Kambari

Orthography Design

Janie P. Stark

SIL International
2010
SIL e-Books

©2010 SIL International
ISBN: 978-1-55671-245-6
ISSN: 1934-2470

Fair Use Policy
Books published in the SIL e-Books (SILEB) series are intended for scholarly research and educational use. You may make copies of these publications for research or instructional purposes free of charge (within fair use guidelines) and without further permission. Republication or commercial use of SILEB or the documents contained therein is expressly prohibited without the written consent of the copyright holder(s).

Series Editor
George Huttar

Volume Editor
Mary Huttar

Managing Editor
Bonnie Brown

Compositor
Karoline Fisher
DEDICATION

To

My husband, John, without whom I wouldn’t have finished this book.
Editor’s Note

This work is a slightly revised version of the author’s 2000 Ph.D. dissertation, University of Ilorin, Nigeria. Since the initial version of this book was written, many important works on orthography design have appeared (e.g., *Handbook of Orthography and Literacy*, ed. by R. Malatesha Joshi and P.G. Aaron, Routledge, 2005). Due to circumstances beyond the author’s control none of these recent works are included in the references.

The reader should keep in mind that this work describes only one case of writing system development. Several principles of orthography design are discussed in detail. One principle which is left implicit is that community participation is a key factor in the development of an acceptable writing system.
FOREWORD

This study deals with designing writing systems: it goes beyond basic principles of orthography development to develop a series of principles intended to guide orthography decisions in situations where the basic principles fall short. Seventeen principles, useful to anyone involved in orthography design and/or reform are delineated. Examples are taken primarily from the three Kambari languages of Nigeria, where the author and I were technical advisors to the Kambari Language Project staff. These examples illustrate the application of advanced orthography principles where the basic principles either were not applicable or produced conflicting results.

Some areas discussed in detail are: dealing with the language’s area of complexity, morphophonemic processes, word-division decisions, and the strategic use of spaces to lessen ambiguity in written forms. Preserving the constant visual image of morphemes, an aid to experienced readers, is emphasized in contrast to orthographic representation of allomorphic variation. Throughout the discussion the primary role of an orthography, i. e., to convey meaning to a mother-tongue speaker, is kept in the forefront.

The approach presented here advocates that the orthography developers consider multiple options for any decision throughout the developmental period. Each option is to be considered in light of its implications for the overall system; thus, every decision affects and is affected by every other decision. The nature and role of consistency in orthography development is expanded beyond the basic phoneme-grapheme correspondence to embrace system-wide and applicational consistency.
The author’s specific expertise and experience in African languages results in an African focus in the discussion and illustrations. By doing so, she could insure the accuracy and the pragmatism of the data and principles. And further, Africa represents an overwhelming number of languages in need of orthographies. Those working in other areas of the world will find the principles surmount area restrictions and provide universal guidelines for decision making. The readers will easily see the parallels between the African illustrations and their own situations.

John E. Stark, June, 2002
ACKNOWLEDGEMENTS

This work began as my Ph.D. dissertation at the University of Ilorin, Ilorin, Nigeria. As a result, much of the credit goes to Professor O. Awobuluyi and the faculty of the Department of Linguistics and Nigerian Languages at that university.

In addition I am grateful to the Chairman of the Kambari Language Project (KLP) Board of Trustees, Dr. I. Majanbu, who has joined with the Chairman of the KLP Executive Committee, Mr. J. S. Jatau, to undertake both project development and project recognition for KLP throughout the Kambari-speaking area. I am also grateful to the members of the Board of Trustees and Ex-Co for their dedication and input to the project.

I acknowledge the immense contribution of the KLP staff, who are dedicated to work day after day, year after year, for the betterment of their people and for the glory of God. Truly, without their input, this study would have never come about.

I am grateful to my family and friends in the U. S. whose encouragement and support have made my stay in Nigeria possible.

With heartfelt thanks I acknowledge my debt to my husband, John, for the many, many hours of discussion from which these thoughts and concepts emerged. He was always ready to help and support me in a thousand ways throughout the process.

In conclusion, whatever mistakes there are in this book, mea culpa.

Janie P. Stark

June 2002
Contents

Dedication
Editor’s note
Foreword
Acknowledgements
Index of tables
Index of Figures
Abbreviations/conventions
Abstract

1 Basic Orthography Design: Where It Works, Where It Falls Short
  1.1 The foundations of orthography design
     1.1.1 Language/area-specific articles
     1.1.2 Theoretic articles
     1.1.3 “How-to” works
  1.2 Summary of the foundations
  1.3 A philosophy of orthography design
  1.4 The simplest description of orthography design
     1.4.1 Sound/symbol correspondence
     1.4.2 Adequate clarity
     1.4.3 Social sensitivity
  1.5 The second level of orthography design
     1.5.1 Tone
     1.5.2 Word-division decisions
     1.5.3 Standardization
  1.6 An illustration: the various Kambari orthographies
     1.6.1 The Crozier orthography
     1.6.2 The Carroll orthography

2 About the Kambari Languages
  2.1 Introduction
  2.2 Scope of research
     2.2.1 Terminology
     2.2.2 The Kambari cultural setting
     2.2.3 Sociolinguistic setting
     2.2.4 Lexical comparison
     2.2.5 Semantic dissimilarity
     2.2.6 Agwara Area Kambari
     2.2.7 Auna Area Kambari
     2.2.8 Salka Area Kambari
     2.2.9 Kambari Language Project
  2.3 Structure of Kambari
     2.3.1 The phonology of Kambari
     2.3.2 Kambari morphology and syntax
     2.3.3 Summary

3 Development of the KLP Orthography
  3.1 Language complexity
3.2 Kambari areas of low complexity
   3.2.1 Letters corresponding to English and Hausa
   3.2.2 Letters corresponding to Hausa
   3.2.3 Consonants unique to Kambari
3.3 Kambari areas of high complexity: orthographic difficulties
   3.3.1 Decisions regarding graphemes
   3.3.2 Decisions regarding phonological processes
   3.3.3 Variations by speakers
   3.3.4 Decisions regarding morphemes
   3.3.5 Kambari power syllables
3.4 Conclusion

4 Insights on Orthography Principles
4.1 Advanced principles
   4.1.1 The convey meaning principle
   4.1.2 The consistency principle
   4.1.3 The optimal simplicity principle
   4.1.4 The limited ambiguity principle
   4.1.5 The preference for independent words principle
   4.1.6 The underlying form principle
   4.1.7 The stable form principle
   4.1.8 The easiest-for-users principle
   4.1.9 The acceptability principle
   4.1.10 The familiarity principle
   4.1.11 The proliferate options principle
   4.1.12 The weigh advantages and disadvantages principle
   4.1.13 The teaching load principle
   4.1.14 The act-the-same, write-the-same principle
   4.1.15 The clarification from co-languages principle
   4.1.16 The evaluation of conflicting parameters principle
   4.1.17 The positive criterion principle
4.2 Interaction of advanced principles
4.3 Standardization, harmonization, and conformity
   4.3.1 Standardization and harmonization in Kambari
   4.3.2 Conformity in the Kambari setting
4.4 Simultaneous consideration of co-languages
4.5 General insights
   4.5.1 The role of tradition
   4.5.2 The network model
   4.5.3 It has to be taught
   4.5.4 Appearance matters
   4.5.5 Orthography development: progression toward the ideal
   4.5.6 Anything is possible
   4.5.7 The role of testing
   4.5.8 The central role of meaning
   4.5.9 The role of the community
   4.5.10 The influence of the LWC
   4.5.11 It is worth doing
4.6 Conclusion
Appendix A Sample text
   English source
   Agwara Kambari
   Auna Kambari
   Salka Kambari

Appendix B Orthographies
   Agwara orthography
      Introduction
      The philosophical basis of the KLP orthography
      The current alphabet
      Vowels
      Consonants
      Problems still to be considered
      A sample story in the Agwara Kambari orthography
   Auna orthography
      Introduction
      The philosophical basis of the KLP orthography
      The current alphabet
      Vowels
      Consonants
      Problems still to be considered
      A sample story in the Auna Kambari orthography
   Salka orthography
      Introduction
      The philosophical basis of the KLP orthography
      The current alphabet
      Vowels
      Consonants
      Problems still to be considered
      Sample text in Salka orthography

References
Index of tables

Table 2.1: Kambari vowels
Table 2.2: Kambari consonant phonemes
Table 2.3: Suffix addition in Auna Kambari
Table 2.4: Salka vowel copying
Table 2.5: Vowel copying skipping high vowels
Table 2.6: NCMs of U-class nouns
Table 2.7: Noun classes
Table 2.8: Kambari verbal elements
Table 3.1: Morphemes to the left of the verb root
Table 3.2: Subject pronouns
Table 3.3: Morphemes to the right of the verb
Table 3.4: Kambari power syllables
Index of Figures

Figure 2.1: Kambari intelligibility and identity
Figure 2.2: Map of Nigeria
Figure 2.3: Map of the Kambari area
Figure 3.1: Glottal syllable decision flow chart
Figure 3.2: Relation between pronoun sets 1 and 2
Figure 4.1: Word profiles
Figure 4.2: Standardization and harmonization
**Abbreviations/conventions**

AGW  Agwara Area Kambari. This convention (along with AUN and SAL) is primarily used with data examples from a specific Kambari that are distinct from the other co-languages.

AUN  Auna Area Kambari

KLP  Kambari Language Project

LWC  Language of Wider Communication, used here to include trade and official/national languages. In the Kambari area, there are two LWCs, Hausa and English.

MT  Mother Tongue. A MT speaker is the same as a native speaker. In the Kambari area, the MT is Kambari, i.e. specifically Agwara Kambari, Auna Kambari, or Salka Kambari.

SAL  Salka Area Kambari

**Phonetic/phonemic conventions**

/y/ as in the English ‘yet’

/j/ as in the English ‘jest’

/c/ as representing the sound ‘ch’ in the English ‘chest’

The schwa is represented phonemically /ə/ and orthographically <a>

The glottal stop is represented phonemically /ʔ/ and <ʿ> elsewhere

**Conventions used in glossing examples:**

CMP  Completive action suffix

IT    Iterative suffix

NCM  Noun class concord marker

NCMP  Non-completive action found in Set 2 Pronouns

PERF  Perfective aspect

REP   Repeated action marker

[s.o.]  Square brackets surround information implicit in the context. S.o. is the abbreviation for ‘someone’.

sth.  Something
This study is concerned with the orthography development of three Kambari languages: Agwara Kambari, Salka Kambari, and Auna Kambari. It documents the standardization of the languages as implemented by the Kambari Language Project. It provides the linguistic and sociolinguistic data that influenced the orthographic decisions, and in a number of instances articulates various options, examining the arguments for and against them.

The foundational principles of orthography design are well established by the work of earlier scholars. Those foundational principles are summarized, and another set of principles proposed to guide orthography decisions in instances where the foundational principles are silent or conflicting. The principles are seen to be interactive, clustering to provide preference for one decision over another. One of the unique features of this work is the extension of simultaneous orthography development beyond the boundaries of dialects and into the realm of co-languages within a language cluster.
1
Basic Orthography Design:
Where It Works, Where It Falls Short

1.1 The foundations of orthography design

People have developed writing systems for unwritten languages for centuries, and continue to do so. These developers often proceed with no specialized knowledge beyond basic literacy in their own language or with only basic literacy in a trade language. The former occurs if they are working to develop a language other than their own, the latter when their mother tongue is the subject of development. Sorting through the complex systematic relationships of sounds and meanings which comprise a single human language is not a task to be undertaken lightly nor without special preparation. With the world-wide promotion of education, speakers of unwritten languages clamor, ‘What about my language?’ These languages deserve the best possible help in creating writing systems that are optimally efficient at communicating information between speakers of the language.

Many books and articles have been written through the years to give aid to those who are developing orthographies. These can be broadly grouped into three classes: language or area-specific suggestions, theoretical concerns, and how-to approaches. Looking at these works provides a basis for the extended discussion on orthography which forms the body of this book.

1.1.1 Language/area-specific articles

Typically literature of the language/area-specific type of articles contains a brief discussion of orthography principles, a detailed accounting of which characters to use for an extensive list of phonemes in a given language or area, and a sketchy paragraph saying that
in-depth understanding of the linguistic and sociolinguistic factors will affect orthography decisions.

In 1965, Swadesh presented a paper of this nature, ‘Possible devices for alphabetic simplification’ at the University of Ghana. He lists three basic principles in orthography development: 1) ‘adequacy’, i.e. significant sounds of the language differentiated; 2) ‘mechanical efficiency and economy’, i.e. the letters should be easy to write both by hand and with typewriters, and 3) ‘harmony’ with co-existing systems, which he limits to previous orthographies of the same language. He then catalogues sounds and possible representations. At the end he says,

Suitable simplified alphabets generally can be devised only after careful advance studies, by capable personnel, of all factors; the nature of the phonemic system, its relation to the form structure (morpho-phonemics), relative frequencies, previous practices in writing the language, facts about other co-existing languages, attitudes and opinion [sic] of the people. (1965:9)

Someone looking to this type of article for guidance in orthography design may collect general hints about how to proceed, and some specific help if the article deals with languages from the same linguistic family, but will be left with many questions and few answers.

1.1.2 Theoretic articles

Articles which deal with the theoretical principles of orthography and use data to illustrate the theoretical points can also be found.

Smalley edited a volume on different aspects of orthography development (1964a), and summarized the basic principles in the first chapter (1964b:1-17). His main point is that since writing is different from speech, a writing system that is merely phonetic is ultimately unsatisfactory. He describes different kinds of orthographies: phonetic, i.e. rigid conformity
to exact speech; phonemic (his ideal: writing the significant sounds of the language in a consistent way); syllabic, i.e. each grapheme represents a pronounceable bit of speech, not a phoneme; and morphemic, i.e. the grapheme represents the meaning, not the sound. He also enlarges on the pitfalls of overdifferentiation and underdifferentiation, and states:

The person who devises a writing system takes an enormous responsibility in deciding for underdifferentiation and must base it upon the most careful analysis of the phonemic structure and of the functional load of the phonemes which are not going to be fully represented. A limited amount of underdifferentiation is possible because of language redundancy, that is because every language has the important characteristic of using a sound structure and a grammatical structure which are complex, more complex than would seem necessary for minimal efficiency. (Smalley 1964b:11)

Here again those actually trying to devise a writing system will now have more questions than answers.

1.1.3 “How-to” works

A third type of writing on orthography presents a basic how-to approach.

In Williamson (1984) and in Barnwell (1998), the basic principles for orthography development are described more or less step-by-step. Someone interested in developing an orthography should definitely start with these principles which are divided into two categories. The first two principles, accuracy and consistency, deal with strictly linguistic factors such as the phoneme inventory and its representation. The last three principles, convenience, harmonization, and acceptability, deal with sociological elements, e.g. the levels of bilingualism and literacy in the speech community, and also the availability of written materials and the means to produce them.
**Accuracy**

According to Williamson (1984:7–8) and Barnwell (1998:74), accuracy is based on the phoneme-grapheme correspondence, i.e. as a result of the phonological analysis the orthographer assigns written symbols to each phoneme in the phonemic inventory. Thus, if a language has thirty phonemes, there are thirty letters in the alphabet. One seeks accuracy in this sense as the mechanism for eliminating phonic ambiguity, i.e. the full specification of phonemic features eliminates the confusion that would result if there were graphemic overlap of certain phonemic features (two or more alphabet characters that represent the same phoneme).

An additional point to accuracy according to these authors is that allophonic variations should not be included. The idea here is that if a given sound in the inventory of phones for a language exists only as an allophone, it will not be represented in the orthography. Often, if such a sound is present as a phoneme in surrounding languages, speakers of the developing language will want a letter for that sound in their writing system. This is one area of conflict in orthography design that the basic principles do not address.

**Consistency**

The word ‘consistency’ is used in multiple ways in the literature. Most commonly the meaning is that each orthographic symbol represents one sound (grapheme-phoneme correspondence). Luelsdorff (1989:126) demonstrates that a one-to-one phoneme-grapheme correspondence and a one-to-one grapheme-phoneme correspondence (each phoneme is represented by one grapheme, and each grapheme represents one phoneme) is far easier to read than a system that does not adhere to these principles. He gives the example of the English “mat” as a pattern more quickly learned than “mate” where the phoneme /e/ needs two graphemes <a-e> to prompt the correct pronunciation and meaning. Barnwell (1998:74) expands: a) an orthographic symbol represents one and only one sound, and b)
the inclusion of orthographic symbols which have no phonetic reality should be avoided (no ‘silent’ letters).

Another area of consistency is in word division. Division parameters must be established and maintained consistently, e.g. if the first-person subject pronoun is written as a word, all other subject pronouns should be written as words.

The various ways orthographies can be consistent often conflict, and the conflicts caused by applying different parameters ‘consistently’ will create dilemmas in orthography decisions. For instance, a single sound, such as [v], may exist as both a phoneme in its own right and as an allophone of /f/ in certain environments, making one-to-one sound/symbol correspondence and one-to-one phoneme/symbol correspondence impossible. Another common conflict is brought about by morphophonemic variation. In context, the sounds of adjacent words affect each other, and if strict sound/symbol correspondence is maintained, then the same word will have various spellings based on the words that surround it. These are also areas of difficulty in orthography design that are not addressed by the basic principles.

**Convenience**

The symbols chosen must be both easy to write and easy to type. Special characters that are not available on a typewriter or at a local printer should be minimized, if not avoided altogether. Barnwell (1998:74) suggests that special characters be limited to less frequently used phonemes. Williamson (1984:9) prioritizes these basic principles by saying it is more important to be ‘accurate’ than to be ‘convenient’. She gives the example of marking tones. Tones are difficult to write using an ordinary typewriter; therefore, they are ‘inconvenient’. Though the elimination of tone marks may be appealing to the language developers (principle of convenience), if it results in an ambiguous orthography, tone marks
must be included (principle of accuracy). Although most writers on orthography
development touch on the principle of convenience, Smalley, for one, argues against the
tyranny of the typewriter, which he calls ‘mechanical imperialism’ (1964b:14). Many
languages use phonetic symbols not found on an ordinary typewriter; other languages find
ways to write those phonemes without using phonetic symbols. Both systems work, i.e.
people can be taught to read using either system. The expanding influence of computers into
the remote regions of the world demands adequate symbols for clear communication while
respecting the principles of accuracy. Working within the international standardized
systems, which are today much broader that in 1964, yields immeasurable benefits in ease of
production for written materials.

Convenience most often creates conflict in orthography design by compromising the
communication potential, i.e. not encoding enough detail to capture all the meaning of the
original thought in the language. Convenience also brings about other types of orthography
design dilemmas. One-to-one correspondence may not be possible if the phonemic inventory
of the developing language is larger than the available symbol set in an area. Digraphs used
for a single phoneme violate the one-to-one correspondence, at least in its purest sense. For
instance, if the letter ‘e’ is chosen to represent [e] and the digraph ‘ae’ to represent [æ], then
the grapheme ‘e’ carries two roles in the orthography.

Harmonization

Barnwell (1998:74) has summarized this principle with these four points: a)
conformity to government policies (where applicable); b) conformity to the LWC in the area;
c) if the new orthography cannot conform to the national language or the language used in
higher education, it should not conflict, i.e. it should not use a grapheme to represent one feature in the mother tongue that is used to represent a different feature in the LWC; d) conformity to regional languages, as far as possible (see also Williamson 1984:10).

The area of conformity can bring about conflicts in orthography design. What if the national language and the regional LWC are using different symbols for the same phoneme? What if government policy, if followed, would compromise the ability of the orthography to communicate clearly? These are still more conflicts between the basic principles of orthography development.

**Familiarity/acceptability**

Familiarity (Williamson 1984:10–11) and acceptability (Barnwell 1998:74) mean an orthography that is acceptable to the language community and matches their expectations as to what an orthography should be like. They have these expectations because they are familiar with the orthographies of other languages.

Concerning acceptability, Simons (1994:LL97930) states, ‘In a very real sense, social acceptability is the overriding principle of all…where there is no acceptability the solution cannot be tolerated, regardless of its linguistic or pedagogical desirability.’

Writing about standardization Hatfield (1991:253) states that, ‘If a written code exists for a language but is not acceptable to the speakers of that language, then, essentially, standardization of the written form has not been achieved.’

Acceptability can bring about conflict in orthography design. Things that are not familiar, especially in an area where second-language literacy is widespread, can cause

---

1 At the basic level of orthography design these serve as two terms for the same principle. In section 4.1.9 I will distinguish between them. (see page 159 and page 160)

2 In lieu of a page number, LL followed by a number refers to the electronic access record number of the LinguaLinks library CD (1998).
people to reject the orthography of the developing language as too difficult, while the reality may be that an orthography without those symbols may fail to represent sufficient meaning information. A methodology for resolving such conflicts is lacking in the basic principles of orthography design.

1.2 Summary of the foundations

These five principles: accuracy, consistency, convenience, harmonization, and acceptability are the foundation, the beginning point, of orthography development. These starting points can be summarized as follows. Accuracy: an orthography must adequately reflect the structures of the language. Consistency: an orthography should behave in a predictable manner. Convenience: writers and publishers should be able to easily produce documents. Harmonization: where established written codes for neighboring languages deal with similar linguistic phenomena, that method of representation can be used. Acceptability: if the people like it and use it, it is acceptable.

Recognizing that the basic principles fall short of providing all the answers to orthography design in no way diminishes their value or effectiveness. They will indeed form the basis for all the early decisions in orthography design, most of which will stand unchanged in the ultimate orthography. However, it would be naïve to think that simple conformity to these five principles is all that is necessary for orthography development. We have seen that the principles contradict one another. It is at this point that further principles come into play. Languages are uniquely complex; each language has its own complexity that will require special care in developing the working orthography. It may be that the proposed solution for that complexity violates one or more of the above general principles. Such violation is inevitable in real-world applications, and the additional principles presented
here will help the orthographer resolve these conflicts in a manner maximally beneficial to the mother-tongue users of the writing system under consideration.

1.3 A philosophy of orthography design

While working on a new writing system, the people involved need a philosophy of orthography design which can help to decide between conflicting options.

Gordon (1986:66–67) describes two extremes in orthography design. The first is the maximally precise transcription of sounds. This may be either phonetic or phonemic. The other extreme would be a morphemic orthography, i.e. a purely meaning-based system in which every morpheme would have its unique orthographic representation. Between the extremes he posits morphophonemic orthographies which ‘offer consistency of morpheme shape along with a relatively high degree of phonemic regularity’ (1986:67). This is what I have described as morpheme-driven philosophy of orthography design. Gordon is addressing the issue that in a strictly sound-based orthography, the spelling of a word is not dependent on its meaning, but on its context. A single morpheme may have several orthographic representations, and the recurrent pattern of the morpheme is obscured. This makes rapid reading nearly impossible, and every word must be subvocalized to discern its meaning.

Marlett, in describing the ‘economy principle’ (similar to section 4.1.3 the optimal simplicity principle), writes that ‘an optimal orthography is not characterized by unnecessary variation in the shapes of the morphemes’ (1999:221).

Dawson enlarges on this theme by stating,

Phonemic spelling helps the beginner but morphophonemic spelling helps the fluent reader. What we need to aim for in a practical orthography is one which is sufficiently accurate phonemically that decoding is possible, but one which keeps the meaning units in their basic form wherever possible (morphophonemic spelling) and which makes the adjustments toward community acceptability, which are
generally dictated by conformity to the national language.
(1989:LL86415)

In evaluating an orthography, Dawson stresses acceptability and morphophonemic representation geared toward fluent readers. Simons provides a different evaluation parameter, that of overall cost of achieving fluency.

When given a number of alternative solutions to an orthography problem, the solution which promises the overall least effort is to be preferred. Overall effort is measured by the amount of time required for an illiterate to become fluently literate. Once a reader has become fluent, there is no effort involved in an orthography...The less time and effort required to gain mastery, the greater the chances that the individual student will succeed. (1994:LL979884)

Using Simons' parameter, the morphophonemic approach, which maintains the constant visual image\(^3\) of words, allows a learner to reach fluency with less overall effort than a strictly phonemic approach. Simons' “overall effort” as stated above refers to the length of time and degree of teaching required to learn to read. It fails to address an even more significant issue, the amount of effort a maximally fluent reader must expend to comprehend a written passage. Despite Simons' claim that “Once a reader has become fluent, there is no effort involved...”, there is in fact effort expended by every reader every time he or she reads. The degree of effort expended by a maximally efficient reader is determined largely by the design of the orthography. Reducing the overall effort required of skilled readers is the goal of a good orthography. If the system requires little effort once learned, then people are more likely to read often and for a variety of purposes. If the system requires extensive effort, even after it is fully learned, people will read only that which has some form of high external motivation, such as class-room assignments. The ease of use for the

---

\(^3\) Constant visual image refers to the consistent representation of a word, or the root of a word so that a reader becomes familiar with the entire word and no longer processes the individual letters or syllables.
fluent reader is the guiding principle in the morpheme-driven philosophy of orthography design.

All these factors combine to provide a philosophy of orthography design as follows: the goal of an orthography is to accurately convey the intended meaning of the writer to the reader, to do so in such a way that minimizes the effort required of the fluent reader and results in materials that are used by the speakers of the language.

1.4 The simplest description of orthography design

In the process of orthography design, the practitioner soon discovers it is a cycle in which step one must be repeated based on what was decided in step two, and then both steps one and two need to be revised based on what was decided in step three, and so on. Orthography design is also a web, in which every decision pulls and pushes on all the other decisions without any reference to a specific sequence. The five principles of Barnwell and Williamson, the assorted theoretical points scattered about in the literature, and the various language-specific illustrations found therein can be distilled and condensed to a three point procedure: 1) develop a symbol set that 2) encodes the meaning 3) in such a way that people will use the system.

1.4.1 Sound/symbol correspondence

Sound/symbol correspondence in its simplest form involves carrying over the sound/symbol correspondence of regional written languages into the target language where the two systems have identical or similar sounds. In instances where the target language has sounds not represented in the developed languages of the region, Swadesh (1965), Van Dyken and Kutsch Lojenga (1993), and similar works for various regions of the world provide a starting point. The decisions of sound/symbol correspondences require a phonetic
inventory, a phonemic inventory, and a statement of the primary phonological processes present in a language.

### 1.4.2 Adequate clarity

Adequate clarity means that an orthography should encode the information needed for the mother-tongue reader to discern the meaning of the text. The fuller definition states that an orthography should encode all and only the information needed by a mother-tongue reader to enable proper understanding of the text. The “only” in the fuller definition of clarity means that decisions in this area serve to reduce ambiguity and to keep the writing system simple. A syntactic description of the language is necessary to enable the orthography designers to utilize word breaks in a way that enhances clarity.

### 1.4.3 Social sensitivity

Social sensitivity means that the writing system needs to be designed in such a way that people can and will use it. Conformity to area norms, acceptability in the eyes of the end users, ease of production of handwritten and machine produced text, feasibility of local commercial printing, and similar considerations come into play in this area of orthography design.

### 1.5 The second level of orthography design

Three crucial areas lie outside of the issues covered in the basic elements of orthography design: dealing with tone, word divisions, and bringing together variant forms of the same language.

#### 1.5.1 Tone

In addition to the basic principles outlined above, tone needs to be considered in African languages and elsewhere around the world. Tone decisions will vary from language to language based on the application of good orthographic principles to the tonological
patterns of individual languages. Tone was not ultimately of serious concern in the
development of the Kambari orthographies, the languages in which my most extensive
experience was gained. The literature discussed here, from Williamson and others, provided
the analytical background for the decisions in the Kambari orthographies.

Williamson (1984:41–42) states that if ambiguities exist when tone is not marked,
tone must be marked regardless of ‘convenience’ or ‘harmonization’. This is an essential
position based on the principle of adequate clarity.

Mfonyam (1990:29–30) in his article ‘Tone analysis and tone orthography’,
summarized the principles of tone in orthography:

1. The target of orthography development is the mother-tongue speakers of that
   language.
2. Tone principles must be both systematic and consistent, i.e. the settings where
certain tone marks occur should always be predictable.
3. ‘Surface tones should be marked rather than underlying tones.’ That is to say, if a
tone results from the proximity of two tones in a specific environment, the
resulting tone should be marked, not the two that make the environment.4
4. An orthography should not be cluttered with too many tone marks.5
5. The more stable level tone should be marked.
6. In a tone system containing three or more tonemes, start by marking Low and High
tones.

However, when considering phonemes rather than tonemes, much of the literature recommends the underlying
rather than the surface form, especially when the process occurs across morpheme or word boundaries. I would
anticipate places where such would be the case with tone marking as well.

I imagine that in seeking to avoid too many tone marks, he is thinking of ‘unmarked’ as a specific toneme.
7. and 8. Here Mfonyam states principles that apply generally for Bantu languages:

7) Low tone is more stable than High tone, and

8) Low tone is more easily perceived by new readers.

9. Whichever tone is most easily perceived, that is the one that should be written.

10. When choosing which tone to mark, do not mark the tone that occurs most frequently.\(^6\)

Mfonyam also describes a scale for marking tones in an orthography, from zero (no marks) to full (every tone marked every time). He advocates the ‘minimal representation’ which falls in the middle of the scale, but is still variable from language to language. He states, ‘The orthographic minimal representation consists of the representation of the minimum number of tones…that are required to make the necessary meaning distinctions in the language.’ (Mfonyam 1990:24).

Because of the strong emphasis in the literature on phoneme-grapheme correspondence, and the strong injunctions against underdifferentiation, one is led to believe that any deviation from ‘one phoneme equals one grapheme’ is a serious violation. The tone marking for Trique, a language of Mexico, illustrates that an exact toneme to tone mark correspondence is not always practical, and that a compromise solution may be the most functional.

Longacre (1964) describes the Trique tone system as having five level tones plus glides that are sequences of tones on a single syllable, so that, if fully specified, each syllable could have any one of twenty different marks. As Longacre points out, ‘This was adjudged to

---

\(^6\) My experience is that unmarked tone indicating the most frequent tone value, and marking on the less frequent tones, conforms to both the simplicity principle in that there are fewer things to write, and to the acceptability principle in that a page of text in the new language will more closely resemble a page of text in the other regional orthographies.
be too complex to be practical’ (1964:133). The solution proposed was to use diacritics to represent tone, but that each diacritic would represent a group of tones. Such a system significantly reduces the ambiguities in comparison to a non-tone marked text, though some ambiguities remain because of the range of tones allowed by each mark. This solution was employed rather than marking tone on a specified list of words because the functional load of tone in Trique is very high.

Consideration of the Trique and similar situations led me to the advanced orthography principle I call “proliferating the options” (see section 4.1.11). The inventory of tone-marked words is a solution if the list contains a relatively few number of words. If there are many such words, the teaching load becomes easier if all instances are marked (perhaps in a limited context, e.g. all verbs) than for learners to memorize a long list of marked words. As seen in the Trique example, if the full specification of certain phonemic features is too unwieldy, the grouping of features represented by a single orthographic convention may be the most practical solution.

1.5.2 Word-division decisions

The preponderance of the literature on orthography development deals with the sound-symbol correspondence and therefore does not help the language developers in deciding what things in the language stand as independent words, and what things are affixes. In oral language, there are no ‘words’ per se, only utterances, and utterances in no way parallel the written word. Utterances can be made up of more than one word, a single word, or even less than a word. For example, in answer to ‘I still didn’t hear if you said ‘contact’ or ‘contract’?’, one could respond simply ‘tract’, which is a morpheme, but not a word in this context. The following quote illustrates the dilemma of word-division decisions, i.e. understanding which parts of a multimorphemic utterance are ‘words’.
We transcribed word lists and linguistic data from over fifty different languages [of southern Sudan]. We soon discovered that gathering such information was not always easy. Many people, when asked for a certain word, would reply with a sentence. The word we wanted would be included in that sentence, but we had no way of knowing the phonetic parameters of that word and were unable to isolate it from the surrounding grammar. The informant was often of little help since he or she could not conceive of an isolated word out of context. We soon learned how to get the words for physical objects since they could more easily stand in isolation and needed less context. However, getting words such as verbs, adverbs, and connectors in isolation always remained difficult. Sudanese who had been to school and learned to read could isolate words easily since they had been taught to do so, but people who had not learned to read spoke in full sentences, since language is used to communicate experience, not words. (Arensen 1992:49–50)

Van Dyken and Kutsch Lojenga (1993) outline principles for making word divisions, and acknowledge the reality of conflicting parameters. They define ‘word’ as a) a written form with the same meaning as the oral form; b) a unit surrounded by spaces or punctuation; c) consistently spelled; and d) a smaller unit than a sentence. They then describe the psychological and sociopolitical influences and list the linguistic criteria subgrouped in the areas of semantic, grammatical, phonological criteria with a final section dealing with interaction among the criteria.

These points are treated under the heading of semantic criteria: REFERENTIAL INDEPENDENCE, i.e. the word has meaning in the mind of the mother-tongue speaker when said in isolation; CONCEPTUAL UNITY, i.e. when compound words and reduplicated words form a single unit of meaning that is different from phrases made up of those words; and MINIMAL AMBIGUITIES, i.e. the insertion of spaces can differentiate what are minimal pairs in the orthography as in ‘green house’ (a house painted green) and ‘greenhouse’ (a building where plants are grown).
The points of the grammatical criteria mostly deal with functors: 7 MOBILITY, i.e. if a functor can occur in different word-order positions, it should be considered a separate word. SEPARABILITY, i.e. a functor should be considered a word if another word can enter between it and what generally follows it. SUBSTITUTABILITY, i.e. if what is already established as an independent word can replace the functor in that position, the functor is also an independent word.

The next three points deal with the phonological criteria: PRONOUNCEABILITY in isolation; PHONOLOGICAL UNITY, i.e. if there are features such as vowel harmony which are contained within a word; and PHONOLOGICAL BRIDGING, i.e. if phonological processes link it to another word.

Sometimes in trying to make word-division decisions, the application of some of the criteria creates one set of results, but the application of other criteria creates the opposite results. The last section gives pointers to how to deal with this problem: CONSISTENCY, i.e. formulate and apply the word-division rules in a predictable manner; REDUNDANCY, i.e. often several criteria point to the same solution; and AGREEMENT, i.e. in the three sets of criteria, the semantic, grammatical, and phonological, look for consensus between two sets. When that has been done, prioritize the decisions so that a) semantic considerations should be given preference over grammatical ones; b) give preference to grammatical considerations over phonological ones; c) establish which features apply within a word, and which apply across word boundaries; d) hyphenate the ambiguous morphemes; and e) avoid contracted forms.

---

7 Functor is defined as ‘used in word classification for words and bound morphemes whose role in the language is largely or wholly grammatical’ (Crystal 1997:162). It is perhaps more common in literature on literacy and reading theory than in formal linguistic studies. It is used here to avoid implying word-division decisions and to allow a single term to cover what would otherwise need the entire phrase ‘grammatical words and bound grammatical morphemes.’
Linguists have long divided morphemes into words, affixes, and clitics. Words are those things that are clearly independent. Affixes are those that are clearly bound to another morpheme. Clitics do not function independently, and are often phonologically linked in the same way as affixes. At the same time, they have wide distribution, often ‘attaching’ to several categories of words, which affixes do not. This mixed behavior of clitics makes them the bane of theoretical linguistics and its need for clean division of categories, but the concept can be very helpful in practical orthography design.

Something can display the phonological properties of an affix and still show the syntactic properties of a word. Marlett provides a discussion on ‘phrasal affixes’, i.e. bound morphemes functioning at the phrase level, not the word level. He gives the example of the English possessive in these phrases, *the Queen’s hat*, and *the Queen of England’s hat*. (1999:48) In the latter, the hat is the property of the Queen, not of England. He goes on to state, ‘in the past undue attention has been given to the question of whether a morpheme may constitute an utterance in and of itself. This criterion may provide a good test for what may be an independent utterance (naturally), but not for what may be a word’ (1999:50).

An example of the problems of making consistent word divisions is documented by Beck (1964:156–158). This article is especially pertinent to the Kambari orthography development, since several Kambari functors function in a similar way. Beck uses examples from the Inamwanga New Testament (East Africa), which he condemns as ‘artificial and inconsistent’. To illustrate what he means, the following data presents, in this order: the published form, Beck’s ‘improved’ form, and the underlying morphemes:

- *ni vintu* (Mk. 7:13) for *nivintu*; actually *na ivintu* ‘and the things’
- *na makasa* (Mk. 7: 5) for *namakasa*; actually *na amakasa* ‘and the hands’
- *nu mulilo* (Mk. 9:44) for *numulilo*; actually *na umulilo* ‘and the fire’.
There are several points to consider:

1. I agree with Beck’s statement that the printed option is poorly done. It does not preserve the content words in an isolatable form, i.e. if asked what *vin*tu means, a speaker of that language would probably not know, since the underlying form is *iv*in*tu*.

2. The underlying forms presented here demonstrate the widespread principle of vowel elision, i.e. that when there are two consecutive vowels, only the second one is pronounced. At no point does Beck acknowledge that readers can learn to pronounce *na ivin*tu as [nivin]tu].

3. Beck presents his own solution, *nivin*tu, as the only option. He argues for joining the grammatical morpheme to the content words. He objects to writing them as separate words in this way:

   We will have to be careful that words do not become confusingly long, but we should on the other hand be free from the misunderstanding that a multitude of partial words, some of them not amounting to more than one letter, are easier to read than fewer words which are a little longer. It is a generally made observation, which is not ignored in primary education, that even in the first stages of learning the pupil is apt to read not letter by letter, but rather by catching with the eye the picture of the whole word as it appears in print, especially if it is only short.

   (1964:156)

   In summary he says: a) avoid excessively long words, b) avoid partial words and strings of one or two letters, and c) maintain a ‘picture of the whole word’—what I have described as the word profile (figure 4.1 in chapter 4). However, Beck does not see the inconsistency in writing `<ivin*tu` ‘things’, `<nivin*tu` ‘and the things’, `<mwivin*tu` ‘in the things’, `<pivin*tu` ‘from the things’. With the Beck system, the reader must catch a different ‘picture of the whole word’ for every different manifestation of a prefix plus *vin*tu. Thus instead of one ‘picture’ for each noun, the reader must have a ‘picture’ of maybe six or ten
manifestations of that noun. Languages have a core vocabulary\(^8\) of about five hundred words that account for ninety percent of communication. Most primers are designed to expose readers to this core vocabulary as early as possible. With the Beck system, instead of a basic vocabulary of five hundred words, the new reader needs exposure to several thousand words, each many times in order to automatically recognize the word profile or to establish a pattern in their minds of a prefix followed by a noun.

### 1.5.3 Standardization

Ansre (1971:680–682) defines standardization as ‘the process by which a specific variety of a language emerges as the preferred variety of a speech community….Strictly speaking, language standardization is an intralanguage phenomenon.’ According to Ansre, it includes but is not limited to orthography and terminology development.

Johnson (1990:5–6) defines standardization as ‘a process by which one form of a language becomes accepted in the whole community as the norm for writing their language. Variation in the spoken form of the language will likely remain, but the written form would be formally codified. This may be true even to the point that the varieties of the spoken form are not immediately intelligible with each other.’

Hatfield (1991:252–254) lists three types of language development: codification (orthography and supporting materials), modernization (terminology development), and standardization. She defines standardization as ‘the development of a written form of a language which covers as broad a grouping of dialects of that language as possible and which will still be comprehensible to its speakers.’ She lists the benefits of standardization (as opposed to codifying a language variety without consideration of the sociolinguistic factors) as helping governments achieve educational goals, enhancing of ethnopolitical

---

\(^8\) The concept of core vocabulary is presented with supporting data from Kambari in section 4.1.
identity, and enabling a greater number of people to use written materials. According to her, factors affecting the development of standardized languages include the political situation, degree of language awareness of MT speakers, the size of the language group, existing unifying infrastructure, and other sociolinguistic factors.

In writing about standardization, Miller (1991:260) states that the main factors affecting language use are identity and utility, i.e. usefulness. She cautions language planners to consider well the non-linguistic factors involved in standardization, i.e. to undertake the task of standardization with careful planning and effort, to make decisions based on the feelings of the whole language community, to avoid disparagement of dialect variants, to allow creativity and change, and to recognize that the standardization process takes time (1991:263–264).

Simons (1994:LL97905), in his introduction to ‘Principles of multidialectal orthography design’, states ‘dialect variation may be so great as to prevent speakers of two dialects of the same language from understanding each other.’ It must be assumed by this that he has moved into the realm of co-languages not dialects, since the commonly used test of dialects is that they are mutually intelligible. He does address the co-language issue by stating, ‘When dialect variation is so great as to make one orthography for many groups impractical, dialect comparisons should still be made in order to make the orthographies as compatible as possible’ (1994:LL97915).

He advocates diligent comparison of the varieties, and that they are to be represented at whichever level they are most alike, be it phonetic, morphophonemic, or fast speech. ‘When such a solution is possible, the result is an orthography which is both multidialectal and the way everybody speaks’ (Simons, 1994:LL97919).
Both Capo (1989) and Simons (1994) propose that a single grapheme may have different phonetic realities in different dialects; for example, the grapheme $<s>$ could be pronounced as $[s]$ in dialect A and $[ʃ]$ in dialect B. Capo deals only with dialects that are mutually intelligible, and it is assumed that only one set of literacy materials would be needed to serve all. He explains four kinds of correspondences phonemes may manifest in different dialects: a) the same phonemic realization in all dialects, e.g. both dialect A and dialect B have the phoneme /d/; b) dialect-specific phonemes in regular correspondence, e.g. where dialect A has the phoneme /c/, dialect B has the phoneme /ts/; c) when the correspondence between phonemes is not predictable on the basis of dialect, e.g. in both dialects A and B there are some words that have /l/ in one language and /ɾ/ in the other, but there is no pattern as to which dialect uses which phoneme. In this case, Capo advised introducing a third grapheme to account for the overlapping words. d) One dialect has a phoneme that does not exist in the other dialect (1989:38–45).

Both Capo and Simons function almost exclusively in the realm of phonology, in that a codification of sequences of graphemes in such a way that MT readers can interpret them is all that is needed in a multidialectal orthography. Neither addresses the lexical/semantic divergence of dialects/co-languages. At the dialect level, often there are two words for the same thing, but one is used by speakers of a restricted locality. In this situation, the word with the broadest range of distribution is the one to be used in written materials that serve the language community. At the co-language/language level, the lexical/semantic differences are more severe, and it is less likely that one set of written materials can serve the cluster of co-languages.

Simons (1994:LL97930) addresses the issue of speakers of dialect B who are already literate in the LWC which has both /s/ and /ʃ/, and reject the orthographic $<s>$ when the
sound is [ʃ] in their own language. As was quoted above, acceptability is a strong, even
overriding factor in the development of an orthography.

In addition to the linguistic factors of standardization, the sociolinguistic factors
must be taken into account. Sadembouo (1989:23) advocates the creation of a language
committee to oversee the development and spread of the standardized MT. He defines
standardization as including the development of the writing system, dialectal unification, i.e.
that all the dialects agree to use the same writing system, and harmonization by which he
means sensitivity on the part of language developers in creating an acceptable standard.

1.6 An illustration: the various Kambari orthographies

1.6.1 The Crozier orthography

In Alphabets of Africa (Hartell 1993:234), Crozier proposed an orthography
consisting of twenty-eight consonants and nine vowels. He did not specify if this
orthography was intended to serve all the Kambari co-languages or specifically one, i.e.
Salka Kambari. Tone would be written with an acute accent mark [´], representing ‘high’
tone where needed for clarity, ‘but this is restricted to certain grammatical functions since
the lexical function of tone is rather light’ (1993:234). He did not discuss word spacing. By
looking at the presentation of data in his dissertation (Crozier 1984), one would assume he
intended to join most functors to content words. Judging by the data in both the dissertation
and the Alphabets of Africa, one would also assume that he intended to write every feature,
such as nasalization, fully every time it occurs phonetically, i.e. on every syllable.\footnote{In general, the KLP orthography (the general points that are the same for all three co-languages) separates functors from content words and does not fully specify features when they are predictable or are not needed for comprehension. KLP orthography allows for phonetic spread of nasalization (it is marked once per word, not on every syllable), and for readers to supply length, nasalization and tone as part of natural, intuitive pronunciation.} No books or materials were prepared using the Crozier orthography. The proposals for treatment of vowels, length, and nasalization merit further examination.

**Crozier proposed vowels**

Crozier lists nine vowels: /a, e, i, o, u, Ñ, Ñ, n, Ó/, suggesting the schwa symbol <œ> or a dot under the <a>, for /œ/.

Concerning [e, œ], neither Auna Kambari nor Agwara Kambari has [e, œ] in its phonetic inventory, only Salka Kambari does. Careful phonological analysis shows that in Salka Kambari, [e, e] and [o, œ] are allophones, and in some environments function in free variation. Crozier’s [i] is a conditioned allophone of /u/. (See section 2.3.1.1 under vowel length for phonological details). Including these symbols in orthography is over-representing the sound system beyond what is needed by a mother-tongue reader for adequate information retrieval, a violation of the principle of simplicity.

**Vowel and consonant length**

Crozier proposed writing length with a double letter, indicating length wherever it occurred phonetically. Vowel length in Salka Kambari plays a significant role in grammatical constructions, but does not distinguish a large number of pairs of lexical items. To treat all phonetic vowel length as essential in the orthography is another form of over-representation. (See Appendix B for a detailed discussion of Salka Kambari vowel length.)
**Nasalization**

Crozier advocates writing `<n>` to denote nasalization following every nasalized vowel. However, nasalization spreads, especially through continuant consonants and the glottal stop, and therefore only needs to be written once, generally word finally. Using the example from Salka Kambari `[ʔỳùwà]` ‘refuse’, Crozier would write `<'yinwan>`; for `[܃ùlù]` ‘leaves’; Crozier `<əvun’un>`. The sequence `<n'>` seems to be unnecessarily difficult in the Crozier system, requiring the reader to maintain two separate pronunciations for word-medial `<n>`.

**1.6.2 The Carroll orthography**

The Reverend Father Timothy Carroll lived and worked among the Agwara Kambari people for a number of years until the mid 1990s. Though he has not presented an orthography statement, he has prepared several books in Agwara Kambari, including primers and some Bible portions. This section cites examples from these materials and draws the conclusions regarding the underlying philosophy from an extensive examination of those materials.

In general, the Carroll orthography uses special (IPA) characters to represent the schwa and nasalization. The Carroll orthography uses `<d>` and `<b>`, also widely used in Hausa materials in the area, and writes length on both vowels and consonants. It is not clear what criteria was used in selecting the words to be marked for segment length, as Agwara Kambari speakers have reported a lack of agreement between their perception of length and what is indicated in Carroll’s materials. On selected words, tone marks appear in this orthography.

In the instances mentioned above, the features mentioned are non-contrastive or redundant, and hence unnecessary in conveying meaning. In many instances, the Carroll
orthography writes double vowels where a portion of the population pronounces a sequence of vowel-glottal stop-vowel. (See section 3.3.3 for a discussion on this feature).

The Carroll orthography joins all noun class markers to content words and treats other functors both as words and affixes, i.e. in one environment the functor will be written independently, and in a different environment the same functor will be written as an affix. The clusters of letters appearing between spaces often cannot be identified in isolation, e.g. <notona> in isolation is meaningless to a Kambari speaker; in fact, it is two morphemes with a phonetic adjustment for the following word <ani> which also is made up of two morphemes. This inconsistent representation of morphemes is a violation of one of the basic applications of the principle of consistency. An examination of four publications reveals that the morpheme ‘with’ is written in the Carroll orthography as a prefix before words beginning with a vowel, as an independent word if the following word begins with a consonant, and as a suffix with specific morphemes. The orthographic representations (marked by < >) in the following examples are in the Carroll orthography.

1)  /N usana/ [nù.sà.nà] <nusana> ‘with morning’

2)  /N kula/ [n.kù.là] <n kula> ‘with name’

3)  /kòbolo N vangu ni/ [kò.òbò.lòn vànqùní] <kòbolun vangun ni> ‘together with his younger brother’

The Carroll orthography is based on context-specific phonetics as in the orthographic <ə> at the end of <notona> to mark the elision between the final /i/ of /otoni/ and the following NCM [ə], or the <n> at the end of <vangun> which represents a morphophonemic blending of /vangu ni/. The result is an orthography that is very difficult to learn, especially to write.
If the production of written material is too difficult, an orthography will not be fully accepted by the people. This system creates an orthography in which even the maximally efficient reader is required to expend a large amount of energy to understand what is written due to the inconsistencies in patterns. This too makes it unlikely such an orthography will receive wide acceptance.¹⁰

¹⁰ The literacy work of the Roman Catholic Church in the Agwara speaking community occurs as dry season Catechist training, and has been a great help in that young people gain the basic concepts of how to read, and can later apply those concepts to reading Hausa and English as they enter the federal school system.
2
About the Kambari Languages

2.1 Introduction

This chapter presents the sociological and linguistic details of the Kambari language data that forms the primary basis for illustration of advanced principles of orthography design proposed later. The reader may chose to go directly to chapter 3 ‘Development of the KLP Orthography’ and refer back to this section for the supporting linguistic details. Readers interested in the structures and linguistic milieu of African languages will find the data and analysis presented here helpful in various types of linguistic inquiry.

2.2 Scope of research

The primary fieldwork presented here is concerned with orthography development in three of the languages of the Kambari Language cluster: Salka Area Kambari (Tsishingini), Auna-Wara Area Kambari (Tsikimba), and Agwara Area Kambari (Cishingini). These three co-languages are being developed by the Kambari Language Project (KLP). As technical advisor to KLP, I was responsible to guide the linguistic analysis and the orthography design process. The final decisions as to which options to utilize in the standardized written form of the languages were in the hands of the KLP staff and representatives of the community. What follows is a critical analysis of the structural and sociological factors present in this language development effort.

2.2.1 Terminology

Commonly, speech forms are divided into two categories: dialects and languages. Dialects are very similar speech forms whose speakers understand each other, i.e. they are mutually intelligible. Languages are speech forms which diverge to a greater extent, and whose speakers do not understand each other. It is usually assumed that speakers of different
dialects belong to the same ethnic group, and speakers of different languages belong to different ethnic groups. In the Kambari cluster this generalization does not hold. Although there is a sense of ethnic unity and the different speech forms are related linguistically, the people of one speech form do not understand the others. In both linguistic analysis and day-to-day usage patterns, each speech variety is a unique language. Outsiders, especially those who distribute funds, see the ethnic unity and relegate the speech forms to insignificant ‘dialects’. Dialects are mutually intelligible; the different Kambari languages are not. Because a binary division of speech forms fails to adequately designate the communication reality of the area, I devised a four-cell matrix to clarify this issue.

<table>
<thead>
<tr>
<th>+ Ethnic identity(^a)</th>
<th>+ Mutual intelligibility</th>
<th>- Mutual intelligibility</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dialect</td>
<td>Co-language</td>
</tr>
<tr>
<td>- Ethnic identity</td>
<td>Exolect</td>
<td>Language</td>
</tr>
</tbody>
</table>

\(^a\)Ethnic identity in the minds of the MT speakers may be manifested as the linguistic awareness that the two speech forms are/are not the same language.

**Figure 2.1: Kambari intelligibility and identity**

The terms dialect and language retain their common definitions. I have added co-languages, which are sister languages of a language cluster whose speech forms are not mutually intelligible. This is an extension of the use of ‘co-’ as in co-wife. Each wife in a polygamous household is a co-wife. Co-wives share the identity of the husband, but each has her own place in the household. The three Kambari forms in this study are co-languages.

Exolects are speech forms that are mutually intelligible, but whose speakers deny any common ethnicity. Chambers and Trudgill (1980:4) described the Scandinavian situation. Linguistically Norwegian, Danish, and Swedish are one language since they are mutually intelligible; however, the three have distinct writing systems and the speakers
consider them different languages. Because of this situation, Norwegian, Danish, and Swedish are what I have termed exolects. According to Crozier and Blench (1992:64 and personal communication), there is a speech form of Kambari, Baangi, that is linguistically part of the cluster, but whose speakers deny being ethnically ‘Kambari’. Thus, Baangi is an exolect.

This study is about three co-languages simultaneously developing three distinct orthographies. Appendix A is a sample text written in English by Kambari speakers, then translated from the single English source into the three Kambaris. Even engineering sameness in this way, the three are seen to be quite different.

2.2.2 The Kambari cultural setting

The Kambari people live in the western section of Nigeria’s Middle Belt, surrounding Lake Kainji.11 Historically they lived in walled cities, the walls of which are still visible from the air. They fled from these cities at the time of the Hausa-Fulani jihads, and now are found mostly in villages or small, extended-family settlements of fewer than one hundred people. The small settlements are formed and abandoned as the people migrate looking for fertile farmland. Kambaris are farmers whose preferred food is millet, although they also grow cash crops.

There is no central Kambari leader; each settlement operates under the ‘village head’ system. The emirs in the area are not ethnically Kambari, though some have maternal connections. In politics, the people prefer a candidate who is Kambari, even if from a different area, to a non-Kambari.

---

11 See maps in fig. 2.2 and enlarged inset in fig. 2.3.
Each of the language areas contains practitioners of three religions: Christianity, Islam, and African traditional religion, though the relative proportions of each religion differ among the areas. The literacy level, especially outside the main towns, is low.
2.2.3 Sociolinguistic setting

Nearly every Kambari person is at least market-level fluent in Hausa. Every settlement of more than one hundred people is multiethnic and Hausa is the common language of all. The basic domains of language usage are Kambari in the home and monolingual interactions (when all participants in a conversation speak the same variety of Kambari). Hausa is used in commerce, in Christian churches, and if one or more participants in a conversation are not of the same Kambari co-language. Whenever two Kambaris from
different areas meet, they converse in Hausa. English is used among the educated in discussions about government, education, or when a participant in the conversation does not speak Hausa.

The Kambari cluster is comprised of seven to ten languages. Crozier and Blench (1992:118) give the Kambari cluster the following language family classification: Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Kainji, Western, Kambari. The languages under consideration here are the three westernmost co-languages of the cluster. Of the three, Agwara Area Kambari is the most geographically widespread, with numerous small settlements in both the Salka and Auna areas, and is possibly the most populous.

2.2.4 Lexical comparison

I have compared and contrasted the three Kambari languages in the areas of phonology, morphology, and syntax. My conclusion is that the three co-languages are structurally very similar in these areas. The differences which merit their distinction as separate languages are more apparent semantically, lexically, and in the prevalence of selected linguistic features. To simplify the later discussion on orthography, I have carefully selected examples that were the same or very similar in the three co-languages. The similarity in the examples should not lead readers of this work to conclude that the differences among the three co-languages are merely dialectical. Agwara Kambari and Auna Kambari, for example, differ in the vowels in these words: AGW mogono ‘king’, AUN magono ‘king’. If those kinds of small differences accounted for the vast majority of the language data, of course the two would be considered dialects of the same language. Both linguistic and sociolinguistic characteristics establish the three as separate languages. The lexicostatistic results from comparisons based on a 228-word list are as follows: Agwara-Auna 56%, Agwara-Salka 39%, and Auna-Salka 31%.
The cognate data given above accounts for the differing communication potential of various cognate levels. A perfect cognate, with identical form and meaning, received a cognate value of 3 when computing the cognate percentage. Cognates that are recognizable by mother-tongue speakers after a small exposure to the second variety received a value of 2. Linguistic cognates discernable by analysis but not easily seen by untrained speakers received a value of 1. Elements that show cognate characteristics, but impede effective communication, were assigned a negative 2. These elements are often referred to as *faux amis* or *false friends* (Hayward and Moulin 1983) and (Malone 1985) or, as I will use, false cognates.

False cognates are particularly dangerous in co-languages, as the cultural backgrounds of the speakers involved are the same. In both Agwara and Salka Kambari, a woman can *wo'o* a baby; however, in Agwara she breast feeds it and in Salka she carries it on her back. In Auna Kambari, there is *gebe* ‘push’. In Salka Kambari *gebe* is ‘find a fiancée’ and *kebe* is ‘push’. In Auna Kambari *kebe* is ‘help support’ or ‘buttress’, *kebe* in Agwara Kambari ‘hold in the hands.’ These are but a few examples of words in the Kambari cluster in which a linguist can recognize a classical cognate, i.e. two words whose etymology traces back to a single protoform and which in everyday communication lead to misunderstandings between speakers of the various varieties of Kambari.

The following miscommunication actually occurred. A Salka Kambari baby was crying and disturbing a meeting. An Auna Kambari person told the baby’s mother to shush it (*pada*). The Salka Kambari mother was instantly outraged, ‘What do you mean, ‘slaughter’ (*pada*) my baby!’ Fortunately, a bilingual Kambari person sitting nearby was able to sort everything out.
The false cognate examples illustrate that the Kambari languages are “the same, but different.” This characteristic is present in grammatical morphemes of the languages as well. An excellent example of the same-but-different reality of the three Kambari languages is the third person plural set 2 pronoun. The meaning of this set is derived from the same underlying morphemes in all three co-languages: /a-a\[12\] ‘3P + NCMP (non-completive aspect)’ (see figure 3.2 in section 3.3.5.5). Despite the identical shape of the underlying forms, the three diverge significantly as to how the morpheme is realized in speech. Auna Kambari inserts a ‘k’ to separate the morphemes /a-ka/ written <a ka> to make it visually distinct from <aka> ‘women’. Agwara Kambari replaces the first /a/ with the phoneme /œ/ (a phoneme used nowhere else in the language, although it is present in the other Kambaris): <'ya> ‘3P-NCMP’, which is in direct contrast to the Agwara ‘2P-NCMP’ <ya>. Salka uses a geminate vowel morpheme <aa>.

Early researchers believed that Agwara and Auna Kambari were dialects of one language. The statement was apparently based on the observation that Auna Kambaris understood Agwara Kambari (not vice versa). This one-way intelligibility phenomenon is due to the settlements of Agwara Kambaris in the Auna area, not on their linguistic similarities, i.e. some Auna Kambaris have learned to understand Agwara Kambari because they interact frequently.

The Agwara Kambaris who do not live near Auna do not understand Auna Kambari. This was vividly demonstrated in the early days of the Kambari Language Project. A speaker and fluent reader of Agwara Kambari spent several hours at a review session of a story in Auna Kambari. The story was examined and discussed entirely in Auna Kambari. The next day the Agwara Kambari man read the same story but could not answer questions about the

\[12\] In Agwara and Salka Kambari these morphemes are both V\[a\] and subject to vowel copying. The /a/ is used here to make the comparison of the three co-languages more readily apparent.
content of the story. One of the many failures in comprehension in this instance is that the Agwara man did not make the cognitive leap from kuwa ‘house’ (in Auna Kambari) to his own kpa’a ‘house’, yet these two words are classic cognates in the sense that they have descended from a single protoform. The consonant pattern of [k] followed by labial consonant, the vowel elision from the first syllable in the Agwara form, then the conversion of the [kC] sequence to the phoneme /kp/, the doubling of the final vowel with glottal insertion to maintain two syllables, are all phenomena evidenced elsewhere in the languages.

2.2.5 Semantic dissimilarity

Not only is there lexical dissimilarity, there is also a surprising amount of divergence among semantic sets. A semantic set is a group of words that function in a certain domain, e.g. cooking terms. Farming is one of the central domains in a Kambari’s life. The three Kambari areas have similar terrain, climate, crops, and farming methodology. Yet in spite of the similarities, they do not have the same concepts, much less the same words. The differences among the co-languages are not in the nature of different labels on the same files, but more like different filing systems. Each Kambari language has two or three names for the mounds/ridges used for planting. Of the seven words among the three languages, only one pair is, to the linguist, recognizably cognate. Agwara has the word kagu for the kind of mound used in planting yam and cassava, and the word volu for the kind used in planting everything else. Salka Kambari has the word agjili for the kind used in planting yam, and the word kedu for everything else. Auna Kambari has katuru for the kind used in planting yam, gbangulu for cassava and uyolo for everything else. It is the plural forms AGW yolu and AUN iyolo that are the only cognates, though in Agwara Kambari the word includes the kind of ridge for planting cassava, but in Auna it does not.
To me, the most unexpected area of semantic dissimilarity is that of the semantic set for kinship, especially that of siblings. In European languages, we find that they have a two-word set of sibling terms distinguished by the gender of the sibling, i.e. *brother-sister*, *hermano-hermana* (Spanish), *bruder-schwester* (German), *frere-soeur* (French), *fratello-sorella* (Italian) *brat-sestra* (Russian transliterated). Not only do the Kambari languages have a different number of terms for siblings, matching semantic constituents reveals very little cognate overlap. Even more troublesome, both for analysis and day-to-day communication, in some cases the word in Auna Kambari and its cognate in Agwara Kambari can be either cognate in meaning or a false cognate, depending on the context. ‘Sibling’, when referring to Kambari languages, is used here in the African sense of patrilineal kin, a usage made necessary by the semantic groupings of the kinship terms of the Kambari language systems. Salka Kambari has a two-part semantic distinction: *zakpara* ‘elder sibling’, and *vawun* (used by older people) or *zawa’a* (used by young people) ‘younger sibling’. I have one elder sister; she is my *zakpara*. My husband has one elder sister; she is his *zakpara*. One must know the age of the sibling in relation to ego (the referent, i.e. ‘me’ if I am speaking of ‘my elder sibling’ or ‘that man’ if I am speaking of ‘that man’s elder sibling’).

Auna Kambari has three overlapping semantic parameters for identifying siblings: one dealing with the age relationship between ego and sibling (elder-younger), one dealing with the gender relationship between ego and sibling (same-opposite), and one dealing with the gender of the sibling when it is the same as ego (man’s brother-woman’s sister). These parameters are realized in four terms: *makoshi* ‘elder male sibling when ego is male’, *mgku* ‘elder female sibling when ego is female’, *vangu* ‘younger sibling (gender unspecified)’, *utaku* ‘sibling of opposite gender to ego (age relationship unspecified)’. There is a fifth word, *tavu*,

---

13 Patrilineal kin here refers to extended family based on male descent of a common male ancestor X number of generations ago. It includes the English idea of cousin, but only from the male line.
which is a rarely used synonym of *utaku*. In Auna Kambari, my elder sister is my *mgku*, and my husband refers to his elder sister as his *utaku*. If we each had an elder brother, my husband would refer to his brother as his *makoshi*, and I would refer to mine as my *utaku*.

Agwara Kambari’s sibling terms take into account the age relationship between ego and sibling (elder-younger) and the gender relationship between ego and sibling (same-opposite). There are three words: *mokoshi* ‘elder same gender’, *vangu* ‘younger same gender’ and *taku* ‘opposite gender (age relationship unspecified)’. In addition, Agwara Kambari has a set of kinship terms used in conversation to refer to the kin of the addressee, e.g. ‘I saw your sibling yesterday. Your sibling gave me this letter.’ There are two words for these siblings: *ambu* ‘younger than addressee’ and *tavu* ‘opposite gender of addressee’. In Agwara Kambari, my elder sister is my *mokoshi*, and my husband’s elder sister is his *taku*. If we each had an elder brother, he would be my *taku*, but my husband’s *mokoshi*. If my husband’s sister were referring to him, he would be her *taku*. If my sister were referring to me, I would be her *vangu*. If an Agwara Kambari were speaking to my sister about me, he would refer to me as her *ambu*.

Among the three co-languages, lexical dissimilarity, semantic variation, variety in the application of phonological processes, preferred grammatical strategies, and sociological factors all indicate that the three are distinct languages.

### 2.2.6 Agwara Area Kambari

Agwara Area Kambari people live on both sides of Lake Kainji. West of the lake they live primarily within the area bounded by the northern border of the Kainji Game Reserve, the Benin border, to the northern border of Niger state, and east of the lake, from Yelwa area south to Nasko town. There are also significant settlements of Agwara speakers in Auna and Salka areas and west of Mokwa to New Bussa. Early work in Agwara Kambari was done by
Mierau who left an unpublished draft of his language analysis (1967). No books were
published in Agwara Kambari during this time. Reverend Father Timothy Carroll of the
Roman Catholic Church lived in the area for many years and published some books, in
Agwara Area Kambari including primers and the gospel of Mark. For further history of
Agwara Area Kambari, see *Kingdoms at War* (Capro 1995) and Yusuf (2000).

The Agwara Kambari people trace their recent history as a migration of part of the
speakers radiating from the Ngaski area (north of Nasko and east of Lake Kainji) across the
river to Agwara and south and west from there. Generally, Agwara Area Kambaris refer to
their language as *Cishingini*. The people who speak Cishingini are called *Ashingini* (*Kashingini
is the singular form*). Older Kambaris refer to their language as *Tsukuku* and the people
*Akuku* based on the root *kuku* ‘old/original’. The people of the Agwara area also refer to
anyone who is ethnically Kambari (speakers of Auna-Wara Area Kambari, Salka Area
Kambari, and others) as *Ashingini*. Kambaris from other areas refer to Agwara Area Kambaris
as *Nwunci* ‘the uncooked (uncivilized) ones’. Agwara Kambaris, naturally, dislike this term.
Therefore, to avoid confusion, the language will be referred to as Agwara Kambari. Agwara
Kambari has two major dialects, *Cishingini* (spoken around Agwara town) and *Tsuqashi
(spooken around Rofia town). Agwara data presented is from the Cishingini dialect.

### 2.2.7 Auna Area Kambari

*Tsikimba*, the language of the Auna area, is found surrounding the towns of Auna
and Wara in Niger and Kebbi States. The people who speak *Tsikimba* are called *Akimba
(*Kakimba* is the singular form). The Akimba people refer to anyone who is ethnically
Kambari (speakers of Agwara Area Kambari, Salka Area Kambari, and others, including
themselves) as *Ashingini* (plural of *Kashingini* ‘Kambari person’). *Tsishingini*, to the Akimba,
refers to the Kambari language cluster. Auna Kambari was recognized as a language distinct
from Agwara Kambari in the thirteenth edition of the *Ethnologue* (Grimes 1996). Auna Kambari has three major dialects: *Tsikimba* (spoken around Auna and Wara towns), *Tsishen* (spoken along the eastern edge of Lake Kainji south of Wara), and *Tsugaunshi* (spoken around Auna town). Because the autonymic terms are fluid, meaning different things to different people, the data based on the Tsikimba dialect will be referred to as Auna Kambari.

Historically, the Auna area people lived in the lowlands near the Niger River and were relocated at the building of Kainji Dam. The clan names, retained by the people of the Auna area, refer to geographical points that are now under water.

No linguistic analysis of Auna Kambari had been done prior to the work of the Kambari Language Project. The formal analysis (see also J. E. Stark 2000) is presented in this study.

### 2.2.8 Salka Area Kambari

The people of Salka Area refer to their language as *Tsishingini*. The people who speak Tsishingini are called *Shingini* (*Shingini* is the singular form). The people of the Salka area also refer to anyone who is ethnically Kambari (speakers of Auna-Wara Area Kambari, Agwara Area Kambari, and others) as *Ashingini*. The people of Salka town refer to the language of Ibeto town as *Tsuvadi*, a recognizable dialect of Tsishingini. The people of Auna area refer to the language of the people of Salka town as *Tsuvadi*. The people of the area around the road from Kontagora to Rijau also refer to their own language as *Tsuvadi*. Though a member of the Kambari cluster, this last *Tsuvadi* mentioned is a different language from that spoken in Salka area and is not included in this study. In this work, the language of the Salka area will be referred to as Salka Kambari.

---

14 The name Kainji comes from the Kambari word for ‘hair’ (AUN *kenji*, AGW *kanji*) referring to the hair of explorer Mungo Park who died in the Kambari area. A swatch of his hair was taken to the local emir.
The majority of the Salka Kambari people live in Magama Local Government of Niger State, in the area between Ibeto and Nasko. Other Salka Kambari speakers live in Mashegu and Borgu Local Governments of Niger State. Missionaries did early language work in Salka Kambari in the 1920s and 1930s when portions of the Bible were translated. Reverend Durkee of the United Missionary Church was the most active in this area of work. Only one or two copies of these materials are extant today. For more history of the early missionary work in Salka, see Mamman (1998). Hoffmann wrote articles based on Salka Kambari (1963, 1964, 1965). Crozier researched and wrote a dissertation on Salka Area Kambari (1984). He also proposed an orthography which was published in the book *Alphabets of Africa* (Hartell 1993:234), but no books were published using it.

**2.2.9 Kambari Language Project**

The Kambari Language Project was formed in 1992 with the goal of developing Kambari writing systems and producing useful literature for the Kambari people. The executive committee and a board of trustees consist of members from each of the participating languages (at this time Salka, Auna, and Agwara Kambari). The project has area-specific subgroups and coordinators. KLP has a full-time staff and two technical advisors (my husband John and myself).

KLP demonstrates the realities of the Kambari social setting (one ethnic identity/many languages) in that one project entity produces literature in three languages. KLP shares facilities, financial resources, and personnel based on respective need regardless of the source of the funds or other resources.

**2.3 Structure of Kambari**

If an example in this chapter is specified only as ‘Kambari’, that example is identical in all three co-languages. Unless otherwise stated, points made apply equally to all three co-
languages. If language examples are Auna Kambari (hereafter AUN) will be used. A principle indicated by examples labeled ‘AUN’ applies equally to all three co-languages, but the specifics of the example are unique to Auna Kambari. For instance, in describing the phoneme /p/ in Kambari there is the word /àpàrà/ ‘dishes’ (all three are identical), and in AUN /kàpòdò/ ‘toad’ (‘toad’ in AGW is /kòpòdò/ and in Salka is /òopòdò/).15 The word ‘toad’ in the three languages is similar, but not identical; and since the point of the example, /p/, is not in the first syllable, the AUN example serves for all. Where Agwara Kambari (AGW) and Salka Kambari (SAL) diverge in significant manner, these areas will be noted and examples given.

### 2.3.1 The phonology of Kambari

#### 2.3.1.1 Phonemic inventory

Development of an inventory of graphemes that is adequate to capture and express the full expressive power of the language is essential before advanced decisions such as affixation, word divisions, punctuation, and paragraphing can be considered. Furthermore, surface variation in the spoken form of morphemes must be accounted for by an understanding of phonological processes before one can arrive at a satisfactory graphemic inventory and the corresponding orthographic representations of the morphemes of the languages.

**Vowel phonemes**

There are six basic vowels in each of the three Kambari languages.

---

15 The morphemic structure of Kambari makes it more appropriate to represent vowel length by double vowels in the underlying forms. In phonemic transcription, adjacent identical vowels with tone on the first vowel only represent vowel length. When tone is marked on both vowels, each vowel represents a full syllable with independent tone and syllable timing.
Table 2.1: Kambari vowels

Vowel length

There is phonetic vowel length. This can be analyzed in two ways: 1) as another phoneme, increasing the phonemic vowel inventory to twelve, the six short ones each with a long counterpart, or 2) as a sequence of two short vowels. As the orthographic implications are identical with either analysis, and the focus of this study is orthography, not phonological theory, the distinction between the two is moot for the purpose at hand.

Salka Kambari uses vowel length as a productive strategy in differentiating grammatical morphemes. In the noun class with the greatest number of members, singular nouns are marked by vowel length in the initial syllable while plural nouns are marked by a short vowel.

4) SAL /òkòwò/ ‘egg’ /òkòwò/ ‘eggs’
   /àbànà/ ‘mouse’ /àbànà/ ‘mice’
   /ëtëlë/ ‘bone’ /ëtëlë/ ‘bones’

Vowel length in the three languages has a low functional load in distinguishing lexical items. Relatively few words contain phonetic vowel length, fewer still are in direct contrast to a minimal partner with a short vowel, and almost none of the words making the minimal pairs would occur in a similar enough context to cause ambiguity in communication.

5) AUN [káká] ‘cake’ [kákā] ‘chameleon’
Hoffmann (1965), later cited by Crozier (1984), posits nine basic vowels in the phonemic inventory of Salka Kambari.

In addition to the six listed in the chart above, Hoffmann and Crozier cite [ε, ɔ, i].

The phones [ε, ɔ] are allophones of /e, o/ in a form of limited free variation. Neither Auna Kambari nor Agwara Kambari has [ε, ɔ] in its phonetic inventory, only Salka Kambari does. Since the phonologies are very similar, this is an indicator that an allophonic pattern should be sought. In Salka Kambari, there is a natural vowel-quality progression from beginning of a word to the end of a word. The phones [ε, ɔ] can occur word initially: [ɛlɛlɛ] 'stoves', [ɔkɔwɔ] 'eggs', while [e, o] never do. [e, o] occur in one-syllable words16 [dɛ] 'call' and [lɔ] 'drive away'; [ε, ɔ] never do. A word spoken very carefully and consciously will preserve [ε] or [ɔ] all the way through, but if spoken rapidly or casually, the vowel becomes [e] or [o] toward the end of the word; thus /ɛlɛlɛ/ 'stoves' may be rendered phonetically [ɛlɛlɛ] ~ [ɛlɛlɛ] ~ [ɛlɛlɛ]. There are no groups of polysyllabic words that contain only [e] or [o] in contrast to groups of words containing only [ε] or [ɔ], as is true of languages with vowel harmony. There are no pairs of words where the only contrast is this vowel quality difference. The addition of the completive morpheme suffix ‘-i’ to [ɛnɛ] ‘see’ produces predictable phonetic raising [ɛnɛi ~ ɛnɛi], and [ɔnɔwɔ] ‘become’ + -i produces [ɔnɔwɔi ~ ɔnɔwɔi]. This is evidence of allophonic variation, not vowel harmony. Native speakers respond to word-initial [e] and [o] as bad pronunciation, not as signaling different morphemes. For these reasons, the [e, ε] and [o, ɔ] pairs are best analyzed as allophones.

---

16 Salka Kambari has sixteen one-syllable words containing [e] or [o].
Concerning the [i], this is an allophone of /u/. The [i] never occurs word initially nor word finally, though all of the other vowel phonemes do. It only occurs in predictable positions, i.e. replacing the final vowel of the verb root before a suffix when the labial feature is not present in adjacent syllables, the initial syllable of a reduplicated verb in Salka Kambari (again without a labialized environment) and following certain consonants, the /ts/ of the noun class prefix being the most common environment. Thus the phoneme /u/ is manifested as [i], in the distribution pattern described, when there is no labial feature in either the syllable immediately before or the syllable immediately after the [i]. If there is a labial feature in either of the syllables creating the environment, the phoneme is realized as [u]. This labial feature can be either consonant or vowel.

6) /tsùgònô/ [tsùgònô] ‘kingship’ (/o/ is labial)
   /tsùgbàin/ [tsùgbài] ‘bigness’ (/gb/ is labiovelar)
   /tsùsùmbà/ [tsùsùmbà] ‘locust bean pod cluster’
      (/u/ is labial)
   /tsùsàlàn/ [tsìsàlà] ‘okra’
   /tsùlàn/ [tsìlàn] ‘flight’
   /tsùgèbè/ [tsìgèbè] ‘courtship’ (the /b/, though labial, is too distant to influence the vowel of the first syllable)

Diphthongs

[ai, oi, au] etc. can be analyzed in various ways in Kambari. The phonological possibilities include allowing a VV rhyme in a single syllable, introducing a set of diphthong vowels into the phoneme inventory, or allowing a semivowel syllable coda (syllable shape CVS).

These sound segments are nearly always formed by the addition of a morpheme, most common is the /-i/ CMP suffix on verbs.
7) AGW /báná -ɨ/ → [báná] go-CMP ‘went’

Only a handful of monomorphemic words in each language contains a diphthong sequence. The sounds have a single tone, and receive the timing value of a single syllable. There are no long diphthongs, i.e. *[aa]' does not exist, neither does contrast exist between [aɨ] and [a.i].

8) [bá.nà]' go-CMP ‘went’, not *[ba.na.i]

The CVS analysis would require the CMP morpheme to be a single /-y/, as in 9).

9) /báná -y/ → [báná] go-CMP ‘went’

The preferred treatment is to allow the morphemes to remain vowels at the underlying level, and a process of syllabification in the derivation fits the [+hi] vowels into a syllable coda semivowel slot in appropriate environments such as word final or before a consonant. This also allows the underlying form to carry a tone, which undergoes a $T_1 \rightarrow T_2$ tone elision when the vowel segments convert to semivowels.

In the phonemic data presentation, the underlying /-i/ is represented without superscript, i.e. /bánā/. This not only conforms to the preferred phonological analysis, it also anticipates the ultimate orthographic form.

**Consonant phonemes**

There are twenty-six phonemic consonants, shown in table 2.2.

---

17 Tone will not be marked on * (not acceptable) data as there is no viable source for the information.
Table 2.2: Kambari consonant phonemes

For the phoneme /z/, some speakers pronounce [z] and [dz] in free variation, while others pronounce only [z]. Phonetically, some of the non-labial consonants allow allophonic palatalization, e.g. [kʰ, lʰ]. Palatalization occurs when the consonants /s, z, ts/ precede front vowels producing [ʃ], [j], and [c]. This palatalization also occurs across morpheme boundaries. In (10), the vowel copy rule applies first, changing the suffix /-sV/ to [-se], which creates the environment for the palatalization process.

10) AUN /wènè/ + /-sV/ → [wènijè] see-IT ‘see over and over’

The [i] before [-jc] is the allophone of /u/ discussed earlier. In certain contexts, and among certain speakers of Auna and Salka Kambari, this palatalization process is optionally applied. For example, AUN /kàkùmbisi/ ‘climber’ and /tsi/ ‘the particular thing (which is a member of the Ts- noun class)’ remain unpalatalized for most speakers. Agwara Kambari applies the palatalization process more completely, as in /tsiŋini/ ‘Kambari language’, where the underlying /ts and /s/ appear on the surface as their palatal variants [ciŋini].
Nasals

The phonetic inventory of nasals includes nasal consonants [m, n, ŋ, ง], syllabic nasals [m, น, ง, งญ], and nasalized vowels.

Nasal consonants

The nasal consonants [m, n] are phonemes. Either can occur at the beginning of words, and also syllable initially or syllable finally. The phoneme /m/ can occur at the end of words, /n/ cannot. The velar nasal phone occurs only syllable finally when the onset of the next syllable is an oral velar consonant. In the same way, the contrast between /n/ and /m/ is neutralized in the syllable-final position, since they are homorganic to the following consonant.

11) AUN [kùmbá] /kùNbá/ ‘climb’
[kàkìndá] /kàkìNdá/ ‘market stall’
[kûryôŋgò] /kûrióNgò/ ‘playing’
[bôlôŋmgbônó] /bôlôNgbônó/ ‘gather’

This homorganic assimilatory process crosses morpheme boundaries.

12) AGW [m bànà] ‘I go’
[õ dànà] ‘I say’
[ij kônà] ‘I catch’

Syllabic nasals

Syllabic nasals only occur word initially preceding a syllable with a consonant onset and are homorganic. In stream speech they lose their syllabicity and function as the coda of the preceding syllable to that consonant.

13) SAL /tsùsàʔàn tsù Ngonò/ [tsì. săʔ tson. gò.nò]
okra NCM king-PL ‘the okra of the kings’
Nasal vowels

There is anticipatory and residual nasalization on vowels immediately adjacent to nasal consonants. This articulatory phenomenon is active across word divisions as well as within words. It has no bearing on either the phonemic analysis or the orthography.

The distribution of the phonemes /n/ and /m/ indicates a need for analysis. There is no restriction to the distribution of the two except word finally. Only [m] occurs word finally, [n] never does. However, in the word-final position there are phonetically nasalized vowels. Based on the gap in the distribution, and the fact that there is no contrast between CVn and C followed by a nasal vowel, all nasal vowels are interpreted as a sequence of /Vn/.

Nasalization spreads left through continuants and the glottal stop, as in these examples,

14) AUN /jàn/ [yà] ‘do’
    /ùwòvòn/ [ùwòvòn] ‘fear (noun)’

15) AGW /ńrèʔè ñ lè/ [ńrèʔèn lè]
    PL-antelope NCM theirs
    ‘their antelopes’

In the preceding example, the single-syllable morpheme /n/ at the phrase level becomes the coda of the preceding syllable and activates nasalization spread. Nasalization does not spread through non-glottal stops, as in (16).

16) AGW /kádàn/ [kádà] ‘fish’

2.3.1.2 Cluster restriction

Kambari has a rule prohibiting more than two vowels or two consonants in sequence (with some minor exceptions). For example, in Auna and Agwara Kambari the conjunctive morpheme meaning ‘and’ is an unspecified for place nasal, /N/. If this precedes a word beginning with a syllabic nasal, a cluster of three consonants results: /ń ŋáŋgà/ ‘and trees’.
However, because of the cluster restriction, both languages insert a vowel following the ‘and’ morpheme, as in (17).

17) AUN /nɔ ñdɔŋɔ/ ‘and trees’
   AGW /nɔ ñdɔŋɔ/ ‘and trees’

In Salka Kambari, the same consonant sequence prohibition occurs, though rarely, because the equivalent ‘and’ in Salka Kambari is /nɔ/ not a syllabic nasal, and Salka Kambari has significantly fewer words with syllabic nasals than the other two. Only in the environment of a single nasal consonant noun concord marker (hereafter NCM) preceding a noun beginning with a syllabic nasal, does this prohibition apply. This environment occurs infrequently, and there is much phonetic variation among speakers as to which strategy to use: a) extra length on the syllabic nasal, b) the use of [mu], already in use with other nouns of the same class (see table 2.7 below), or c) /mV[146]/ which is the same as the NCM of the singular form.

18) SAL /Ngbele M Ngono/ [ŋm̩gbɛlɛ ɲ ɡɔnɔ] ~ [ŋm̩gbɛlɛ m̩ ɡɔnɔ] ~
    [ŋm̩gbɛlɛ mɔ ɡɔnɔ] goats NCM (of) kings
    ‘the goats of the kings’
   /mɛgbɛlɛ mɔ ɡɔnɔ/ goat NCM kings (singular form)
   ‘the goat of the kings’

In the third option, neutralization of the plurality of the NCM (mɔ in the example above) is insignificant because of the fully specified plurality value of the head noun.

Clusters of vowels are somewhat more complex. Since most words end with vowels and a good number of words begin with vowels, and some of the most common functors in the languages consist of a single vowel, the clustering of vowels is a common occurrence. If the three vowels are identical, they blend to the point the native speaker is not aware that there are in fact three vowels.
19) AUN /ànákà àcìmbi/ ‘cows of farmers’ → [ànáká:̀cìmbí]

If the vowels are not identical, different vowel processes can take place to reduce the cluster effect. These vowel processes will be discussed later.

A further strategy of reducing vowel clusters is the alternation between i ~ y and u ~ w. The underlying form of both the pronoun ‘you-PL’ and NCM for nouns beginning with /i-/ is /i/. The underlying form of the pronoun ‘3S (hereafter ‘he’)’ and the NCM for the nouns beginning with /u-/ is /u/. In Kambari, before words beginning with consonants, these are realized by the same vowel as in the underlying forms:

20) /ikèbè i vɔ/ [ikèbìwɔ] money NCM my ‘my money’
/ùnɔ ú vɔ/ [ùnwɔwɔ] mouth NCM my ‘my mouth’

However, if the initial letter of the following word is a vowel and an underlying cluster of three vowels is created, Kambari compensates by transforming /i/ or /u/ into a syllable onset semivowel, [y-] or [w-], respectively.

21) /ikèbè i àkàyà/ ‘money NCM grandparent’
(note the cluster e-i-a) becomes
[ikèbè yàkàyà]

/ùnɔ ú àkàyà/ ‘mouth NCM grandparent’ becomes
[ùnɔ wàkàyà].

2.3.1.3 Vowel processes

Vowel co-occurrence restriction

Kambari has the restriction that non-identical non-high vowels cannot occur in the same word. The non-high vowels are /a, ò, e, o/, and a word can have any number of identical non-high vowels, but cannot mix them within the same word, e.g. no word can have both /ò/ and /o/. There is no restriction in the distribution or frequency of high
vowels. Hoffmann (1965) described this phenomenon in Salka Kambari, and the same holds true for the other two languages with an additional corollary in Auna Kambari. In Auna Kambari, this co-occurrence operates at the level of noun roots. The class identifying prefixes for two noun classes maintain the vowel quality of /a/, e.g. /ka-/ (and its plural form /a-/) and /ma-/. They do not undergo any vowel changing phonological process and are realized in the surface structure as [ka-], [a-], and [ma-] regardless of what other non-high vowels are present in the noun root. For Salka and Agwara Kambari, the vowels in the corresponding noun class prefixes are not underlyingly /a/ but are represented by an underspecified non-high vowel (/V[-Hi]/) that copies the features of the non-high vowel present in the root.

22) AUN [mà-gònò] ‘king’
SAL and AGW [mò-gònò] ‘king’

AUN [à-tèlè] ‘bones’
SAL and AGW [è-tèlè] ‘bones’

An additional exception to the prohibition in Salka Kambari is the suffix /-to/, which increases transitivity and can be attached to any verb.

23) /daŋwá/ ‘fear’ /daŋwàtò/ ‘frighten’
/cipá/ ‘come down’ /cipòtò/ ‘unload (bring sth. down)’
/rèwéc/ ‘be wet’ /rèwètò/ ‘wet sth.’

The other two languages have a cognate suffix /-tàngú/. Salka Kambari [-to] can be explained by applying the velar consonant deletion rule (described later in this chapter) so that /-tàngú/ becomes [–táú], and the vowel coalescence rule ([–táú] becomes [–tò]). The phoneme /n/ is fluid among the three Kambaris, i.e. often cognate words will differ only by the presence or absence of /n/.
Vowel copying

In Kambari, certain suffixes contain vowels that derive their pronunciation from nearby vowels. In this process, the right-most vowel of the root acts as the source or control of the sound of the vowel of the suffix. This is vowel copying to the right, i.e. the source for the vowel features is to the left of the target vowel. In table 2.3, the left column shows the unaffixed verb root, and the right column shows the phonetic outcome of the addition of the suffix to the verb root. These examples are from Auna Kambari.

<table>
<thead>
<tr>
<th>Root</th>
<th>Suffix Addition</th>
</tr>
</thead>
<tbody>
<tr>
<td>/dăná/ ‘say’</td>
<td>[dănįsá] /dăná + sV/ ‘say-IT’</td>
</tr>
<tr>
<td>/cipǐ/ ‘come down’</td>
<td>[cipùsó] /cipǐ + sV/ ‘come-IT’</td>
</tr>
<tr>
<td>/kècè/ ‘count’</td>
<td>[kècįjé] /kècè + sV/ ‘count-IT’</td>
</tr>
<tr>
<td>/sòró/ ‘pound’</td>
<td>[sòrúsó] /sòró + sV/ ‘pound-IT’</td>
</tr>
<tr>
<td>/sùkù/ ‘send’</td>
<td>[sùkùsù] /sùkù + sV/ ‘send-IT’</td>
</tr>
</tbody>
</table>

Table 2.3: Suffix addition in Auna Kambari

Salka and Agwara Kambari also have vowel copying to the left, although Auna Kambari does not. The vowel quality to be copied (the source vowel) is the first non-high vowel (a, o, e, or o) of the preceding morpheme. This process occurs throughout the two languages, and can copy from any word, regardless of the part of speech: noun, verb, adjective, pronoun, etc. In table 2.4 are examples from Salka Kambari of vowel copying to the left, i.e. the vowel in the NCM is copied from the following word.
<table>
<thead>
<tr>
<th>Noun ‘king’</th>
<th>NCM</th>
<th>Word from which NCM is copying its vowel features</th>
</tr>
</thead>
<tbody>
<tr>
<td>mògònò</td>
<td>mà</td>
<td>kàsá ‘separates’ (verb)</td>
</tr>
<tr>
<td>mògònò</td>
<td>mà</td>
<td>ṭyòwá ‘goes’ (verb)</td>
</tr>
<tr>
<td>mògònò</td>
<td>mè</td>
<td>jé?én ‘dances’ (verb)</td>
</tr>
<tr>
<td>mògònò</td>
<td>mò</td>
<td>sòró ‘pounds’ (verb)</td>
</tr>
<tr>
<td>mògònò</td>
<td>mà</td>
<td>vò ‘my’ (possessive pronoun)</td>
</tr>
<tr>
<td>mògònò</td>
<td>mè</td>
<td>lè ‘their’ (possessive pronoun)</td>
</tr>
<tr>
<td>mògònò</td>
<td>mà</td>
<td>gbàin ‘big’ (adjective)</td>
</tr>
<tr>
<td>mògònò</td>
<td>mè</td>
<td>kè?èn ‘small’ (adjective)</td>
</tr>
<tr>
<td>mògònò</td>
<td>mè</td>
<td>mòkòdfi ‘short’ (noun)</td>
</tr>
<tr>
<td>mògònò</td>
<td>mò</td>
<td>màgòlà ‘hitting’ (noun)</td>
</tr>
</tbody>
</table>

**Table 2.4: Salka vowel copying**

If the first vowel of the word containing the source vowel is a high vowel (u or i), the copying mechanism will skip over that vowel, seeking the first non-high vowel of the word (a, ã, e, or o), even if that is the second or third vowel of the word, as seen in table 2.5.
<table>
<thead>
<tr>
<th>Verb</th>
<th>Noun NCM Verb</th>
<th>3P-pro Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>ciğá 'want'</td>
<td>möğònò mà ciğá 'king wants'</td>
<td>à ciğá 'they want'</td>
</tr>
<tr>
<td>gùbá 'herd'</td>
<td>möğònò mà gùbá 'king herds'</td>
<td>à gùbá 'they herd'</td>
</tr>
<tr>
<td>ciğó 'come down'</td>
<td>möğònò mò ciğó 'king comes down'</td>
<td>à ciğó 'they come down'</td>
</tr>
<tr>
<td>lùwó 'drive'</td>
<td>möğònò mò lùwó 'king drives'</td>
<td>ò lùwó 'they drive'</td>
</tr>
<tr>
<td>rító 'learn'</td>
<td>möğònò mò rító 'king learns'</td>
<td>ò rító 'they learn'</td>
</tr>
<tr>
<td>püró 'wait'</td>
<td>möğònò mò püró 'king waits'</td>
<td>ò püró 'they wait'</td>
</tr>
<tr>
<td>[no verbs with the vowels 'i-e']</td>
<td>möğònò mè cinderè 'seventh king'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>müñò nè yitèle 'king of stories'</td>
<td></td>
</tr>
<tr>
<td>pírí 'watch'</td>
<td>möğònò mò pírí 'king watches'</td>
<td>à pírí 'they watch'</td>
</tr>
<tr>
<td>tsùn 'pour'</td>
<td>möğònò mò tsùn 'king pours'</td>
<td>à tsùn 'they pour'</td>
</tr>
</tbody>
</table>

Table 2.5: Vowel copying skipping high vowels

If there is no non-high vowel in the source word, the vowel a is supplied as the default vowel.

Vowel coalescence

Auna and Salka Kambari share the process of vowel coalescence. Agwara Kambari does not. Vowel coalescence is the process whereby two vowels come together (the first is a and the second is either i or u), and instead of either vowel or a diphthong, a third vowel is realized. Hoffmann (1965) referred to this as vowel contraction when he described the phonetics of vowel interaction across morpheme boundaries in Salka Kambari. If the sequence is /a/ followed by /i/: /ai/, [e] will be pronounced. If the sequence is /a/ followed by /u/: /au/, [o] will be realized.
24) SAL /isùlà i rò/ ‘another wall’ is realized [isùlérò]
   /à ìsùlà/ ‘at wall’ is realized [ësùlà]
   /à ǜ?wá/ ‘at house’ is realized [ò̀wá]
   /ǜwá ù tsú/ house NCM our ‘our house’ is realized [ù̀wò̀ tsú]

In Agwara Kambari, instead of vowel coalescence, the /a/ is elided.

25) AGW                  /à ǜnjós/ [ùnjós] ‘at mouth’
    AUN and SAL       /à ǜnô/ [ònô] ‘at mouth’

Vowel elision

There are three kinds of vowel elision in Kambari. The first is the process where
when two vowels come together across morpheme boundaries, the first vowel is deleted, and
only the second vowel is realized. For example (from all three Kambaris):

26) /ikèbè i lè/ money NCM their ‘their money’ is realized [ikèbîlè]

Another instance of vowel elision in Kambari across morpheme boundaries\(^{18}\) is when
the consonant before the final vowel of the first morpheme is a sonorant (l, r, n, w, y, s) and
the second morpheme begins with a non-sonorant consonant, usually with the same point of
articulation as the sonorant. In this setting, the final vowel of the first morpheme deletes and
the final consonant of the first morpheme becomes the coda of the syllable preceding it.

27) AUN /ù bànlà tò/ he go PERF ‘he had gone’ is phonetically [ù bân tô]
    /ù wàlà tò/ he walk PERF ‘he had walked’ is phonetically [ù wàltô]

The third kind of vowel elision is the deletion of a vowel between identical
consonants, or between consonants with the same point of articulation, leaving either a long

\(^{18}\) This elision can also occur within morphemes, but there is great variation among speakers as to whether the
elision is applied fully, partially, or not at all.
(double) consonant, or even a single consonant. Since most of the nouns in Kambari are three syllables, i.e. the noun class prefix and a two-syllable root, it is possible that any two-syllable nouns in the present day Kambari have undergone this elision process. There are several pointers to this. Agwara and Auna Kambari both have the word /mânù/ ‘bird’. It is a member of the M-noun class, and its plural follows the standard pattern of pluralizing M-nouns, /ûnù/. Salka Kambari’s word for ‘bird’ is /mânûnù/. At some point in history, the /u/ in the middle syllable elided, leaving [mànnù] which became the current form in Agwara and Auna. This is also a process in inflectional morphology. In Auna Kambari, the benefactive suffix is underlyingly /-kV[-Hi]/. When adding the suffix to the verb /rôngô/ ‘pass time’, underlyingly it is /rôngó + -kV/ but it is realized as [rôngkô]. The final /o/ of /rôngô/ was elided between the two velar stops; then the two consonants with the same point of articulation neutralized their distinction, favoring the /k/. Salka Kambari has done something of this sort in one of the verb phrase structures. Both of the other languages have a copular verb construction that is underlyingly a single vowel:

\[
\begin{align*}
28) & \text{ AUN } /û \, ñ \, tô/ \ [wôtô], \text{ AGW } /û \, ã \, tô/ \ [wîtô] \ 'he is' \\
& \quad \text{(structurally: pronoun, copula, PERF)}
\end{align*}
\]

In the negative constructions, only the pronoun and the copula appear. In Salka Kambari, the copula is /ri/.

\[
\begin{align*}
29) & /û \, rî \, nô \, mânûnû \, ñî/ \ 'he is with bird not \ 'he does not have a bird.' \\
& /û \, tô \, nô \, mânûnû/ \ 'he is with bird \ 'he has a bird.'
\end{align*}
\]

The sequence of these morphemes (*ri tô) does not exist in the language; the /ri/ and the /tô/ have undergone vowel elision and consonant neutralization, and ‘he is with
bird’ is underlyingly /ù rì tò nà mànùnù jì/ compared to the AUN /ù ɔ tò n mànù/ and AGW /ù ɔ tò n mànù/. This is supported by the following strings being ungrammatical:

30) SAL */u tɔ nɔ manunu jì/ ‘he is with bird not’
    */u rì nɔ manunu/ ‘he is with bird’

In Salka Kambari, the presence of ‘a’ followed by ‘i’ or ‘u’ is an area where there is much phonetic variation: sometimes it is pronounced as a diphthong [ai] as in the English ‘eye’, sometimes vowel elision occurs, so only the high vowel (i or u) is pronounced, and sometimes vowel coalescence occurs. In one context one speaker may say [ai], in a different context the same speaker may say [i], and in a third context, the same speaker may say [e]. Sometimes, with the same morphemes, both [i] and [e] are used by the same speaker. In addition, speakers from Salka may pronounce it one way, and speakers from Raba (a village 10 kilometers from Salka town, but still part of the Salka Kambari language) may pronounce it another way. This is illustrated by table 2.6 in which U-class nouns with the final vowel /a/ are part of noun phrases with three different modifiers.

19 There is a constraint in all the Kambaris that the perfective marker tə never appears in negative constructions.
Table 2.6: NCMs of U-class nouns

2.3.1.4 Syllable and word patterns

Syllable and word patterns are important in the development of the orthography, because they provide a foundation for word-division decisions. Content words in isolation give the language developers both the internal patterns and the general length of words with unambiguous boundaries.

The Kambari syllable patterns are (examples from AUN):

<table>
<thead>
<tr>
<th>Underlying form: Noun</th>
<th>Surface form</th>
<th>Underlying form: Noun</th>
<th>Surface form</th>
<th>Underlying form: Noun</th>
<th>Surface form</th>
</tr>
</thead>
<tbody>
<tr>
<td>uranà u rè 'day'</td>
<td>[uranaùrò]</td>
<td>uranà u vò</td>
<td>[uranùvò]</td>
<td>uranà u lè</td>
<td>[uranùlè]</td>
</tr>
<tr>
<td>ùcirà u rè 'power'</td>
<td>[ùcirùrò]</td>
<td>ùcirà u vò</td>
<td>[ùciròvò]</td>
<td>ùcirà u lè</td>
<td>[ùciròlè]</td>
</tr>
<tr>
<td>ulingà u rè 'work'</td>
<td>[ùlingùrò]</td>
<td>ulingà u vò</td>
<td>[ùlingòvò]</td>
<td>ulingà u lè</td>
<td>[ùlingòlè]</td>
</tr>
<tr>
<td>ùdàngà u rè 'tree'</td>
<td>[ùdàngùrò]</td>
<td>ùdàngà u vò</td>
<td>[ùdàngòvò]</td>
<td>ùdàngà u lè</td>
<td>[ùdàngòlè]</td>
</tr>
<tr>
<td>utàn u rè 'bow'</td>
<td>[ùtorò]</td>
<td>utàn u vò</td>
<td>[ùtòvò]</td>
<td>utàn u lè</td>
<td>[ùtòlè]</td>
</tr>
</tbody>
</table>

V  [à.mà]  ‘people’
VV  [òò.kò.wò]  ‘egg’ (Salka Kambari only)
N  [ñ.kò]  ‘calabashes’
CV  [mà.kò]  ‘calabash’
CVC  [kà.òám]  ‘fruit, kind’
CVV  [ci.ná-i]  ‘meet-CMP’
(VV here is to be read 'long vowel or diphthong', and is required by the morpho-
phonemic behavior of Kambari)20

The most common word patterns are the following:

V  [à] ‘they’
CV  [kò] ‘dig’
VCV  [àká] ‘women’
NCV  [njìmà] ‘mother’
CVCV  [kãlú] ‘root’
VCVCV  [agãtò] ‘river’
CVCVC  [kãbóm] ‘kind of fruit’
CVCVCV  [kàtèlè] ‘bone’
NCVCV  [njìpásà] ‘blood’
NCVCCV  [njìfãngà] ‘trees’
NCVCCVC  [ňlõngòm] ‘Reshe people’

It is to be noted that the word pattern V contains only grammatical morphemes that
are not readily isolated and identified by the untrained native speaker. Though Salka
Kambari allows the grammatically constructed sequence of two vowels (VV) a wider
distribution than do the other two languages, these are still treated as CV, so the syllable
pattern CV is dominant in all the Kambaris.

One-syllable content words (nouns, verbs, or adverbs) are rare in the languages.
Most verbs have two syllables, and most nouns have three, consisting of the noun class

20 In the discussion of vowel length and vowel co-occurrence restrictions earlier in this chapter, a single syllable
with VV was posited. The occurrences are predominantly created by addition of a morpheme, e.g. [ô-g-ko.wo]
’eegg’ could be analyzed as V (singular prefix) + okowo (noun root intrinsically plural) [ba.na-i.] ‘go-CMP’ would
also fit a VV pattern. Some schools of phonology would classify these as single long vowels and diphthongs. In that
case the VV sequence here can be considered as formal notation for ‘long vowel or diphthong’. 
prefix and a two-syllable root. Auna Kambari is more tolerant of one-syllable verbs and two-syllable nouns than the other two languages.

2.3.1.5 Tone patterns

Lexical tone is demonstrated by pairs of words that differ only in their tonal value. Lexical tone in Kambari carries a very small semantic load.

The following tone data indicates Kambari utilizes two level lexical tones. All the examples are in Auna Kambari, but the patterns are the same in the other two languages unless otherwise noted.

On one-syllable nouns:

One-syllable nouns are very rare in Kambari, i.e. there are none in Agwara Kambari, one in Auna Kambari, and twenty-six in Salka Kambari. The null noun class prefix accounts for the increased number in Salka Kambari (see table 2.7), but that strategy is not used by the other two languages. The small inventory of words makes tone analysis irrelevant for this construction.

On two-syllable nouns:

LL [kàlà] ‘name’

HH No HH nouns occur

HL [kàyà] ‘open ridge’

LH [kàyà] ‘arrow’

On three syllable nouns:

LLL [kàbàcî] ‘split fire wood’

LHL [kàgàlí] ‘clan’

---

Since tone data was elicited in isolation, it is possible that HH words were treated as LL, although native speakers of Auna Kambari responsible for the data have linguistic training and three individuals in consultation concluded there are no HH nouns.
LHH [kàkúlú] ‘stone’

HLL [kàkùlù] ‘raffia palm tree’

HHH [ítýóní] ‘spirit’

No other tone patterns are found on three-syllable nouns.

Salka Kambari has the constraint that no word begins with a high tone. The word initial high tone patterns shown above are not found in Salka data. Therefore, cognate words that have identical phonemic structure may diverge in their tone patterns:

31) SAL [vàlì] ‘man’
    AUN and AGW [vàlì] ‘man’

*On verbs (in the simplest form):*

The citation form of verbs is the imperative. The imperative is nearly always LH.

32) LH [kènè] ‘fetch’
    LH [wòló] ‘shed hair/feathers’

There are some exceptions:

33) LL [ząmà] ‘search’

It should be noted that for each of the Kambari languages, minimal pairs that diverge only by tone are rare. In nearly every case of tone pairs, the context would clarify the meaning, i.e. the two words are in different domains and are not likely to be confused for each other.

34) AGW /cình/ ‘clay’
    /cìnin/ ‘intestine’

In natural communication, as morphemes are assembled in speech the tones present on the individual morphemes may vary from the tones on that morpheme when spoken in isolation. However, in Kambari, these tone changes do not in and of themselves signal any
semantic content. The tone changes are redundant with other morphemic features present in the phrase. The examples below from Salka Kambari show a low tone on the first syllable of wala ‘walk’ in the first example, and a high tone in the second example. This tone change is redundant with the subject pronoun, i.e. the tone on the first syllable of a verb will always be high when following pronouns belonging to pronoun set 2 (see figure 3.2).

35) /à wàlá shì/ they-Set 1 walk not ‘They did not walk.’
    /àa wàlá shì/ they-Set 2 walk not ‘They will not walk.’

Tone processes

There are no semantic elements represented by tone alone. In the cases where underlying tone bearing units are elided or reclassified as consonantal, the tones undergo T₁ T₂ → T₂ tone elision as described in the treatment of diphthongs.

Variations across languages

There are certain phonological variations that are specific to one of the co-languages. If a certain change is applied to words or classes of words, those words become identical with the words of the other two co-languages.

One of these is the velar consonant deletion in Salka Kambari. Instead of the initial V of the Salka VV noun class prefix, the equivalent nouns in the other two languages have /k/.

36) SAL /òopòò/ ‘toad’
    AGW /kòopòò/ ‘toad’

The equivalent instrumental/benefactive suffix is –V (where V is the same as the final vowel of the root) in Salka Kambari and –kV in the other two languages.
37)  /ʊtɔ/ ‘come out’ + suffix → SAL /ʊtɔɔ/,\(^{22}\) AGW /ʊtɔkɔ/

come out with [sth] ‘bring out’

There are instances of this velar consonant deletion pattern elsewhere in Salka Kambari, as in the suffix –to discussed earlier, and some lexical items: SAL /-náa/ ‘cow’ belonging to the noun class of null prefix for singular, and i- prefix for plural, compared to /ká-náká/ ‘cow’ belonging to the kV- class of the other two languages. Another example is SAL /lyùcì/ ‘town’ compared to AGW /likùcì/ ‘town’.

Another language variation is the /y/ of Agwara Kambari and the /r/ of the other two languages. All three co-languages have /y/ before vowels (/ye’ne/ ‘flow’ in all), and all three co-languages have the phoneme /r/ (/rònó/ ‘draw’ in all). However, there is a group of words which would be identical except that Agwara Kambari uses /y/ where the other two use /r/.

38)  AGW /yèvè/ ‘know’, /sòyò/ ‘pound’, /ùyè/ ‘road’
    SAL and AUN /rèvè/ ‘know’, /sòrò/ ‘pound’, /ùrè/ ‘road’

2.3.2 Kambari morphology and syntax

The division of multimorphemic utterances into written words is based on the morphology and syntax of the language. If not for the consideration of the structure of the language, the inclusion of spaces in writing would be either random or absent, i.e. at the whim of each writer or at approximately the paragraph level. Neither of these is satisfactory.

This presentation is not intended to be a complete description of Kambari grammar, but is intended to show aspects of the morphology and syntax that affect orthography decisions. For a description of other aspects of Kambari grammar, see J. E. Stark (2000).

\(^{22}\) See footnote 15 for explanation of tone marking.
There are two areas of morphology and syntax to be considered: a) what constitutes the smallest, independent meaningful utterance, i.e. what is unambiguously a ‘word’, and b) what constitutes the basic pattern of a complete utterance, i.e. a simple declarative sentence. The interaction between these two levels forms the basis of word-division decisions. As with the phonology, the examples given will be in Auna Kambari with reference to the other two languages when they diverge from the pattern. Most of the examples will be based on the text in Auna Kambari that was translated into the three languages from a single source text (see appendix A).

### 2.3.2.1 The word

While nearly everyone can agree on a definition of ‘word’ that produces unambiguous examples, a definition of ‘word’ which can serve as a guide for orthographic decisions is not easily formed. People think of words as units between spaces, but the placement of spaces is largely based on tradition. Another aspect in the definition of word is that it has meaning when said in isolation. However, natural speech is in phrases, not words, and every morpheme, by definition, has meaning even if that meaning is difficult for a native speaker to articulate. Content words are easy to identify; grammatical morphemes are difficult. Thus a definition of ‘word’ based on ‘written between spaces’ is useless in deciding orthographic word breaks, and one based on native-speaker intuition or meaningful in isolation fails to account for structural patterns. My definition of a Kambari word is a unit made up of one or more morphemes that can be pronounced in isolation. One of the ways of resolving the status of grammatical morphemes is to look at the syllable patterns of the unambiguous nouns and verbs. In Kambari, the uninflected verb form is the singular command form (a command to one person). Nouns are made up of a noun-class identifying prefix and a noun root. The noun-class prefix is obligatory for comprehension or
identification of the noun. Noun roots and verbs are generally two syllables. The length of these unambiguous words is important in determining word divisions in areas of the language where it is difficult to establish if a string of morphemes is one word or many. If a basic verb is two syllables, an eight-syllable inflected form of that verb seems excessively long, even though untrained native speakers cannot isolate and identify grammatical morphemes such as the ‘causative’ marker or the object pronouns.

### 2.3.2.2 Sentence

Example 39) shows the basic complete utterance: a simple declarative sentence. Sentences in this section are given in Auna Kambari. In the other two co-languages, the syntax is identical, but the actual morphemes vary slightly. (In the remainder of this chapter, tone is not marked on certain examples to enable the reader to focus on the significant segmental variation.)

39) *Kalobokatsulakanukkamatoku.*

boy NCM buys female NCM chicken

‘The boy bought a hen’

This multimorphemic example, on the basis of some substitution frames, can show that there are three basic parts: Subject Verb Object. The hyphens have been added to isolate the verb root from the subject and object.

40)  


king NCM buys female NCM chicken

b. *Aloba-tsila-kanukkamatoku.*  

boys NCM buy female NCM chicken

c. *Ngonon-tsila-kanukkamatoku.*  

kings NCM buy female NCM chicken

---

24 This sentence is glossed ‘female chicken’ rather than ‘hen’ because it reflects the Kambari structure, i.e. a noun phrase made up of ‘female [animal] NCM chicken’.
d. Kaloboka-denge-kanukkamatoku.
   boy NCM sells female NCM chicken

e. Kaloboka-wene-kanukkamatoku.
   boy NCM sees female NCM chicken

   boy NCM buys female NCM calf

g. Kaloboka-tsila-kagonkamadyondyom.
   boy NCM buys male NCM calf

h. Kaloboka-tsila-makumamatoku.
   boy NCM buys baby NCM chicken

2.3.2.3 Noun phrase

These nouns are isolatable and identifiable from the above examples: /kålóbò/ ‘boy’,
The basic forms of these verbs are also isolatable: /tsùlá/ ‘buy’, /dèngé/ ‘sell’, /wèné/ ‘see’.
There are also the morphemes between the subject noun and the verb: ka, ma, a, n in
examples 39–40c, respectively. They are obviously morphemes, if for no other reason than
that the noun and verb surrounding them can be isolated. Henceforth they will be written in
the examples with spaces, though the spaces are not considered orthographic yet, merely
convenient for identification:

   41) /kålóbò kà tsìlà kànùkkà màtòkù./
   boy NCM buys female-NCM chicken

One also needs to take into the account the ka joined to kanuku ‘female animal’ in
kanuk-ka and to kagono ‘male animal’ in kagon-ka and the ma joined to maku ‘baby’ in maku-
ma. Are these the same ka and ma as the ones following the subject nouns, or are they
phonemically identical though morphemically different? Here they are in various AUN
constructions:
42) *ka/ma* followed by a noun:

{N *ka/ma* modifier-noun}
/kàdèlè kà màkùlà/ gourd NCM palm wine ‘gourd of palm wine’
/màkùlà mà kàlòbò/ palm wine NCM boy ‘the boy’s palm wine’

43) *ka/ma* followed by an adjective:

{N *ka/ma* Adj}
/kàdèlè kà pìgè/ gourd NCM big ‘the big gourd
/màtòkù mà pìgè/ chicken NCM big ‘the big chicken’

44) *ka/ma* followed by a possessive pronoun:

{N *ka/ma* possessive pronoun}
/kàdèlè kà nè/ gourd NCM his ‘his gourd’
/màtòkù mà nè/ chicken NCM his ‘his chicken’

45) *ka/ma* followed by a number (quantity):

{N *ka/ma* number}
/kàdèlè kà tè/ gourd NCM one ‘one gourd’
/màtòkù mà tè/ chicken NCM one ‘one chicken’

46) *ka/ma* followed by verb:

{N *ka/ma* V}
/kàlòbò kà wènèi/ boy NCM saw ‘The boy saw…’
/màtòkù mà wènèi/ chicken NCM saw ‘The chicken saw…’

If the *ka/ma* morphemes are verbal prefixes, they would be inflectional prefixes since both the affixed and unaffixed forms are still verbs. If they were verbal prefixes, then it would follow that the *ka/ma* before possessive pronouns would be prefixes, the *ka/ma* before adjectives would be prefixes, and the *ka/ma* before nouns would be prefixes. Each of these prefixes would be specified for their grammatical category, and the similarities among the prefixes of the various categories would be ignored. On the other hand, the creation of a Kambari-specific grammatical category of noun concord markers (NCM), retains the generalization of their function of joining the noun to other things in the language. Once the
pattern of \{N NCM something\} has been established *kanuk-ka matoku* can be reinterpreted as *kanuku ka matoku* ‘female NCM chicken’, *kagon-ka matoku* as *kagono ka matoku* ‘male NCM chicken’, and *alob-a tsila* as *alobo a tsila* ‘boys NCM buy’.

This structure \{N NCM something\} is the same for all the noun classes in Auna Kambari:

| 47 | 1a. Ka-class: | /kàdle ká nè/ ‘gourd NCM his’ |
| 1b. A-class (PL of Ka-class) | /àdle á nè/ ‘gourds NCM his’ |
| 2a. Ma-class | /màtòkù má nè/ ‘hen NCM his’ |
| 2b. N-class (PL of Ma-class) | /ìtòkù ì nè/ ‘hens NCM his’ |
| 3. U-class | /ùgbòjì ú nè/ ‘wisdom NCM his’ |
| 4a. U-class | /ùjìpà ú nè/ ‘song NCM his’ |
| 4b. I-class | /ìjìpà í nè/ ‘songs NCM his’ |
| (PL of U-class 4a and PL of borrowed words) | |
| 5. Ku-class | /kùmè kú nè/ ‘leg NCM his’ |
| (PL of the Ku-class is either A- as in 1b, or A‘ as in /àìènè á nè/ ‘legs NCM his’, a variant form of the A-class 1b above) | |
| 6. Ts-class (no plural) | /tsùgònò tsú nè/ ‘leadership NCM his’ |
| 7. ‘People’ class | (see following discussion) |

The people of class 7 are terms that can only be used for humans, especially kinship, but do not include nouns describing the role of a human as *kacimbi* ‘farmer (member of Ka-class)’ or *magono* ‘king (member of Ma-class)’. The members of class 7 are without a noun-class prefix and generally do not use NCMs. They form plural with the preposed morpheme *an*, as in *an tata tsunu* ‘PL father (class 7 noun) ours (our fathers)’ compared to *ngono n tsunu* ‘kings (plural of *magono*) NCM ours’. This pluralizing morpheme *an* can also come before a proper name as in *kuwa ku an Gbara* ‘house NCM PL Gbara (house of Gbara and his people)’.

---

25 There are two kinds of U-class nouns in Auna Kambari: derived nouns as in *utsgr* ‘possessions’ from *tsgrg* ‘to get’, and non-derived nouns as in *ung* ‘mouth’. The derived nouns have no plural form. They are presented in two classes in order to harmonize Auna Kambari with the other two co-languages (see table 2.7).
In each of the above classes, the NCM is identical with the noun-class prefix. In Agwara and Salka Kambari, wherever the noun-class prefix is underspecified for the vowel (see table 2.7), the corresponding NCM is also underspecified and copies its vowel features from the following word.

Kambari has many strategies for deriving nouns from verbs. In Kambari, the formation of the participle is one of the more productive strategies, producing U-class nouns (class 3), but derived nouns can be found in any class. In Auna Kambari, as a subset of the Ka- and A- classes, the strategy for deriving the doer of an action from the verb is to use a circumfix /ka-i/ or /a-i/ in which the word-final /i/ replaces the final vowel of the verb as in *kacimbi* ‘farmer’, *acimbi* ‘farmers’ from *cimba* ‘farm’. Agwara Kambari forms these nouns the same way as Auna Kambari, except that Agwara Kambari applies vowel copying to the prefix. In Salka Kambari, the singular form has a null prefix, but the plural is the single non-high vowel prefix (the same as in Agwara Kambari): SAL *cimbi* ‘farmer’, *acimbi* ‘farmers’.

These derived nouns follow the same rules as other nouns.

The other two co-languages have very similar noun classes. Sometimes a cognate noun that is a member of one class in one co-language will belong to a different class in the other co-languages. Occasionally the same noun root can belong to different classes in the same language, giving slightly different meanings to the basic concept:

48) AUN /kà-gònò/ ‘male animal’
   (semantically ‘stud’ or ‘herd leader’)
   /mà-gònò/ ‘king’
   /tsù-gònò/ ‘kingship’

49) /kà-kimbà/ ‘Auna Kambari person’
   /tsù-kimbà/ ‘Auna Kambari language’
From these and other examples, the Kambari noun is seen to be comprised of a noun class indicating prefix and a noun root. The noun root in isolation is not recognized by MT speakers as carrying meaning. Neither does every noun root function in multiple classes as illustrated above. Borrowed words are usually placed in class 4, and in Auna Kambari and Agwara Kambari the singular noun prefix is absent, e.g. AUN *bokiti* ‘bucket’ *ibokiti* ‘buckets’.

Table 2.7 shows a general comparison of the noun classes in the three languages. It should be noted that the noun roots in these examples were carefully selected to be as close to identical as possible. All three co-languages have a residue of nouns that do not exactly match the noun class template, e.g. *mini* ‘water’. *mini* does not have the standard mV- class prefix, but it generates /mV/ as the NCM as in AGW *mini ma gadi* ‘water NCM up (rain)’. Thus, by virtue of its NCM, the noun *mini* belongs to the mV- class.
### Table 2.7: Noun classes

<table>
<thead>
<tr>
<th>Singular prefix</th>
<th>Plural prefix</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AUN</strong></td>
<td><strong>AGW</strong></td>
</tr>
<tr>
<td>1. ka-</td>
<td>1. VV[Hi]-</td>
</tr>
<tr>
<td>kapodo</td>
<td>kopodo</td>
</tr>
<tr>
<td>‘toad’</td>
<td>‘toad’</td>
</tr>
<tr>
<td>2. ma-</td>
<td>2. MV[Hi]-</td>
</tr>
<tr>
<td>magono</td>
<td>mogono</td>
</tr>
<tr>
<td>‘king’</td>
<td>‘king’</td>
</tr>
<tr>
<td>manu</td>
<td>manunu</td>
</tr>
<tr>
<td>‘bird’</td>
<td>‘bird’</td>
</tr>
<tr>
<td>3. u-</td>
<td>3. u-</td>
</tr>
<tr>
<td>ucigi</td>
<td>ucigu</td>
</tr>
<tr>
<td>‘love’</td>
<td>‘love’</td>
</tr>
<tr>
<td>4. u-</td>
<td>4. null-</td>
</tr>
<tr>
<td>uciri</td>
<td>uciri</td>
</tr>
<tr>
<td>‘locust bean seed’</td>
<td>‘locust bean seed’</td>
</tr>
<tr>
<td>5. ku-</td>
<td>5. u-</td>
</tr>
<tr>
<td>kutanu</td>
<td>utanu</td>
</tr>
<tr>
<td>‘razor blade’</td>
<td>‘razor blade’</td>
</tr>
<tr>
<td>kuvon</td>
<td>uvon</td>
</tr>
<tr>
<td>‘war’</td>
<td>‘war’</td>
</tr>
<tr>
<td>6. ts-</td>
<td>6. ts-</td>
</tr>
<tr>
<td>tsugono</td>
<td>tsugono</td>
</tr>
<tr>
<td>‘kingship’</td>
<td>‘kingship’</td>
</tr>
<tr>
<td>7. ‘people’</td>
<td>7. ‘people’</td>
</tr>
<tr>
<td>null prefix</td>
<td>null prefix</td>
</tr>
<tr>
<td>bbgq ‘baby’</td>
<td>bbgq ‘baby’</td>
</tr>
</tbody>
</table>

Note: Tone is omitted to emphasize the segmental elements.

**Other features of the noun phrase**

I have already demonstrated the construction \{N NCM modifier\} in which the modifier slot can be filled with various things. In general, Kambari does not string phrasal
elements, and it is difficult to ascertain whether certain morphemes cannot or do not occur together. If two or more modifiers occur, they are joined to the head noun by noun-class specific NCM.

50) AUN /kànàkà kà pìgè kà màngòñò/
cow    NCM big    NCM king    ‘the king’s big cow’

An optional morpheme occurring finally in the NP is the specifier. Semantically it is the referential ‘the’ (the particular thing already made mention of). In Agwara and Auna Kambari, it is noun class specific, i.e. the initial consonant (or semivowel) of each noun class followed by a vowel, /-a/ in Agwara, and /-i/ in Auna.

51) AUN /kànàkà kà màngòñò mà pìgè kì/
cow    k-NCM  king    m-NCM  big    k-the
‘the [specific] cow of the big king’

In this example, it is not the specific king but the specific cow that is marked. In Salka this morpheme is not noun-class specific but is a syllable made up of a glottal stop and vowel that has copied the features of the vowel preceding it.

52) SAL /mògòñò?ò/  
‘king-the [specific]’  
/lnòli n lè?c/  (UF=/nloli n le-?V/ )  
donkeys [specific] NCM 3PL-the ‘their donkeys’

2.3.2.4 Verb phrase

The next aspect of morphology and syntax deals with the verb phrase. The inventories of morphemes that are part of the verb phrase are nearly identical across the three languages, yet the three languages vary substantially in their preference for certain forms, i.e. one language will use one verbal morpheme frequently, while the other languages seldom use that morpheme.
It is necessary to draw attention to specific morphemes, structures, and processes. The obligatory subject marker may be considered part of the noun phrase or part of the verb phrase. It is better analyzed as part of the verb phrase, because the verb affects it phonologically in Agwara and Salka Kambari, both by vowel copying, and by coalescing with the vowel-initial verbs.

53) AGW and SAL: /V_{[Hi]} bâtsà/ [à bâtsâ] ‘they split’
   /V_{[Hi]} rònò/ [ò rònò] ‘they draw’
   /û ñêç/ [wêç] ‘he sees’
   SAL: /V_{[Hi]} una/ [ôná] ‘they kill’

The fully specified verb phrase (minus the obligatory subject marker) is presented below, though it should be noted that the elements presented here are not yet specified either syntactically or orthographically in regard to their status as affixes or words. The morphemes entered in the cells on the bottom row are the underlying forms for all three languages, and the cells without language data can be filled with a number of different morphemes.

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>repeated action (REP)</td>
<td>verb root</td>
<td>iterative (IT)</td>
<td>transitivity increase morphemes</td>
<td>completive (CMP)</td>
<td>object pronoun</td>
<td>perfective aspect (PERF)</td>
</tr>
<tr>
<td></td>
<td>tsu</td>
<td>sV</td>
<td>i</td>
<td></td>
<td>tã</td>
<td></td>
<td>de</td>
</tr>
</tbody>
</table>

Table 2.8: Kambari verbal elements

In natural speech, it is rare for more than two of these elements, in addition to the verb root, to occur. In the Auna Kambari sample text (see appendix A), there are twenty-five verbs. The most complex verb phrases, two in number, contain only two of these elements in addition to the verb root: bankai /bana-kV (column 4) -i (column 5)/ ‘go-with-CMP [sth.]’;
matsasa tã /matsa-sV (column 3) tã (column 7)/ ‘give birth-IT-PERF’.
1. The first column contains the repeated action marker *tsu*. First, it should be noted that the obligatory subject marker for the Ts-class and the 1P subject pronoun are also *tsu*. Meaning is maintained by the syntax, i.e. that the repeated action marker always follows the obligatory subject marker, e.g. *tsu tsula* ‘we buy’, *u tsu tsula* ‘he REP buys’. Though theoretically possible, *tsu tsu tsula* ‘we REP buy’ is almost never used. Salka Kambari has a reduplicated form of the verb, in which the first syllable of the verb is reduplicated as a prefix, as in *tsu-tsula* ‘REDUP-buy’, *ci-ciga* ‘REDUP-want’. Though highly improbable, it would be possible to produce *Tsu tsu tsutsula* ‘We REP REDUP-buy’.

3. The *sV-* morpheme can co-occur with any other morphemes. It contains the idea of doing an action again and again on a regular basis without interruption. In the Auna Kambari sample text *matsa-sa* ‘give birth-IT’ conveys the meaning that the cow gives birth as often as cows do.

4. Uninflected verb roots are frequently intransitive. The morphemes in cell 4 in all three co-languages strictly follow the vowel co-occurrence rules, copying the non-high vowel from the verb root.

54) AUN

*cidâ/* ‘sink (into the ground)’
*cidângû/* ‘bury’ (sink sth. into the ground)
*wûtâ/ ‘come out’
*wûtâkpâ/ ‘bring out (come out with sth.)’
*lyâ/ ‘eat’
*lyâtângû/ ‘feed’
*tônô/ ‘follow’
*tônômbônô/ ‘follow-reciprocal’
AGW /pècè/ ‘divide (separate part from the rest)’
/pècèkù/ ‘divide [sth.](divide the whole into parts)’ as in
/ù pècèkù tò mûkù/ ‘He divided the children [into groups]’
/pècèkè/ ‘divide-benefactive’ as in
/ù pècèkè tò mûkù/ ‘He divided [sth.] among the children’

In Auna Kambari, certain column 4 elements can be grouped together to give extra information.

[wùtômekpš] UF=/wutà -kV N -kpV/
bring out [sth.] for [s.o.] ‘bring it for XX’

In Salka, the equivalent of the others’ /kV/ ‘instrumental/benefactive’ is a copied V, and one of the transitivity increasing morphemes is -to, which was discussed in the phonology in section 2.3.1. The ordering of these column 4 morphemes is lexically specified, i.e. some of these morphemes precede the column 3 /sV/morpheme, and others follow it. In all the languages, with the addition of column three or certain column four morphemes the final vowel of the verb root changes to /u/ as was discussed in the section 2.3.1.1 vowel phonemes.

5. In Agwara and Auna Kambari, the completive (CMP) morpheme is always i. In both these co-languages, there are certain verb stems ending in /u/, and in these verbs, the completive morpheme /i/ replaces the final /u/. In Agwara Kambari, if the verb ends with a glottal syllable, i.e. the final syllable consists of a glottal stop plus the same vowel as that in the preceding syllable, the /i/ ‘CMP’ morpheme replaces the vowel in the glottal syllable.

AGW /laʔa + i/ → [làʔi] surpass-CMP ‘surpassed’

In Salka Kambari the /-i/ ‘CMP’ morpheme has the allomorph -yi that occurs with verbs ending with high vowels, and with verbs containing the instrumental/benefactive suffix.
58)  SAL  

/salu/ + /i/ → [tänûyi] carry-CMP ‘carried’

/tsula/ + /a/ + /i/ → [tsûlàyi] buy-with-CMP ‘bought for’

as distinct from

/tsula/ + /i/ → [tsûlàyi] buy-CMP ‘bought’

In addition to the completive suffix, another filler of this slot in Kambari is the plural command suffix (command to more than one person), also /i/, but syntax (no overt subject) and context differentiate the meanings.

59)  

/ù wâlai/ He walk-CMP ‘he walked’

/wâlái/ [you-PL] walk-IMP ‘Walk!’

Another filler of this slot, in Auna and Salka Kambari, is the simultaneous-action suffix /i/ which replaces the final vowel of the element immediately preceding it. Thus in Auna Kambari an example is given in (60).

60)  

/wéní/ /wene-i[CMP]/ ‘saw’ as in

/tâtə wéní yì, àkù ù lázái/

father see-CMP him, then he leave-CMP

‘Father saw him, then he left.’

/wéní/ /wene-i[simult]/ ‘is seeing’ as in

/tâtə wéní yì, òù lázái/

father see-simult. him, and (while) he leave-simult

‘Father watches him while he is leaving.’

6. The object pronoun column can be filled by pronouns that function as indirect or direct objects. All three languages have the object pronoun le ‘them’, which never copies the vowel of the verb root. To include it as an affix would violate the vowel co-occurrence rule.

7. The perfective marker tə, is one that is difficult for mother-tongue speakers to define or isolate. When occurring as the last morpheme in a sentence, it is always tə, i.e. it does not copy the non-high vowel from the verb root. Thus if it is an affix, it would violate the vowel co-occurrence rule, as would the object pronoun immediately preceding it.
Another filler of this slot in Auna and Agwara Kambari is the morpheme baci that functions here as a conditional morpheme translating as ‘if’ or ‘when’.

61) AUN /kànàkà kà gbònjùrò bàcì, kà màtsàsà tò ànàkà ùjàni./
cow NCM grow if, NCM give-birth-IT PERF cows plenty
‘If the cow grows, it will give birth to many cows.’

Auna and Agwara Kambari also use baci preceding the subject noun to indicate a future condition that must be completed before another action can result. The example below is taken from the Agwara Kambari story in appendix A.

62) AGW /dà bàcì màtòkù mà mà gyà?wàin…/
then if hen [not yet bought] specifier NCM set-CMP
/…dà mùkù mà n gbònjùrói, kòò n ònìjè lc./
then babies the NCM grow-CMP, so-that I sell them
‘When the hen grows and lays eggs, then the babies will grow and I will sell them.’

Salka does not use baci ‘if/when’ in the verb phrase, but instead uses ng preceding the subject;

63) SAL /nò vù úsú…/
‘If you agree’

8. In position 8 de carries the idea of an action taking place before the moment of speech. The morpheme is identical in the three co-languages, but whether it can co-occur with the –ì ‘CMP’ or the tò ‘PERF’ is language specific.

64) AGW /tsù tsùlài dò/
we buy-CMP already ‘We’ve already bought...’

Not included in the chart above because they have a different set of syntactic co-occurrence restrictions, are the pre-posed gerundial morphemes, glossed ING as in ING-go
(going), indicating progressive action. In Auna Kambari this morpheme is \( u \), in Agwara Kambari it is \( ku \), and in Salka it is unmarked (null prefix).

In Auna and Salka Kambari, these gerund-marked forms function as derived nouns. They can occur as the subject of a sentence and are followed by the appropriate NCMs. In Agwara Kambari they can function either as a verb or as a noun. They function as verbs because they fill the verb slot in the construction \( u \) verb \( t\) ‘he verb PERF.

\[
\begin{align*}
65) & \quad /\ddot{u} \text{ cinà ò} / ‘\text{he has met [him]}’ \\
& \quad /\ddot{u} \text{ kúcínà ò} / ‘\text{he is meeting [him]}’
\end{align*}
\]

In Agwara Kambari when functioning as a verb, if preceded by a subject marker, it forms a dependent clause.

\[
\begin{align*}
66) & \quad /\ddot{a} \text{nà à kù-bànà à màkàràntà, dà ù zùwài lè a pí\textisà mèkècù/} \\
& \quad \text{When they ING-go to school, then he put [caused] them they learn reading} \\
& \quad \text{‘When they were on their way to school, he advised them to learn reading.’}
\end{align*}
\]

In a serial verb construction the gerund generally follows either the copula or another verb.

\[
\begin{align*}
67) & \quad \text{AGW} /\ddot{wì} \text{ tò à kù-dànsà à màcì à nì wì tò à kù-bànà à kà\textfìnà/} \\
& \quad \text{he-is at ING-say-IT wives NCM his he-is PERF at ING-go to farm} \\
& \quad \text{‘He is telling his wives every day he will be going to the farm.’}
\end{align*}
\]

The \( ku \)-verb construction in Agwara Kambari accepts any of the postposed morphemes except the \(-i \) ‘CMP’. The second- and third-person singular object pronouns are \( wu \) and \( yi \), not \( nu \) and \( ni \) that follow the \(-i \) ‘CMP’.
In its nominal role, the ku-verb construction can function as the subject of a sentence, generating its own ku NCM, and accepting iterative and transitivizing morphemes. The other two languages function in a similar way regarding nominal usage.

### 2.3.2.5 Serial verbs

Serial verb constructions are prevalent in the languages. They consist of two or more verbs, each of which is generally marked for subject. Any of the verbs in the construction may be affixed. As seen in the Agwara Kambari examples above, the copula precedes any other verbs in the series. Much of adverbial type information is expressed by means of serial verbs.

### 2.3.2.6 Margin elements

Time phrases and discourse level conjunctions are sentence initial. Fronted elements from within the sentence are considered marginal to the sentence. In Agwara and Salka Kambari, they consist of a noun with NCM or emphatic pronoun followed by da ‘equative’.

Auna Kambari has a set of noun class markers and a set of emphatic pronouns that function as both equative and relative markers.
Quantifiers (all, many, plenty, etc) are not noun class specific and may be either noun phrase or sentence final, though some can occur sentence initially. The negative is sentence final. Salka Kambari has two negative morphemes: shi following declarative sentences and wan elsewhere. In direct address, there is also the bracketing construction, she...wan, used for greater emphasis.

2.3.3 Summary

All three Kambari languages are primarily CV in syllable structure, with twenty-four consonant phonemes, six vowel phonemes, and two semivowels, /y, w/. Vowel length and tone play a very small role at the lexical level. Vowel nasalization arises from the leftward spreading of nasality from a word-final /n/ which subsequently deletes. There is a vowel co-occurrence restriction that prevents different non-high vowels from occurring in the same phonological word.

The languages are SVO in clause structure. A fully specified sentence would be:

73) \[ S \rightarrow (\text{Margin-info}) (\text{NP}[su]) \text{ VP (NP}[io]) (\text{NP}[do]) (\text{Margin-info}) \]

Sentence-initial margins are filled with time phrases, conjunctions, or fronted elements. The optional subject noun phrase consists of an explicit noun and any of its modifiers or an emphatic pronoun. Other subject pronouns and the noun-class markers form an obligatory subject marker within the verb phrase. The verb phrase can have both preposed and postposed elements and can be serial. Often objects (both direct and indirect) are implied by the verb. It is very rare that a sentence specifies both the indirect and direct
objects. When specified by a pronoun, the object pronoun precedes the perfective aspect marker \(tg\). When specified by a noun phrase, the object follows the perfective aspect marker \(tg\). Sentence-final margins are filled with locative phrases, quantifiers, and the negative marker.

The noun phrase is made up of a head noun, a noun-class marker (NCM), a modifier, and/or a specifier.

74) \(\text{NP} \rightarrow \text{N (NCM Mod) (Spec)}\)

Modifiers can be a noun phrase, adjective, number, or possessive pronoun. If two or more modifiers occur, they are joined to the head noun by a noun-class specific NCM.

Understanding the morphology and syntax as well as overall patterns of frequency and identical processes in different grammatical roles is one of the bases for orthography decisions, especially in the area of word divisions. The specific issues of Kambari orthography will be discussed in the following chapter.
3 Development of the KLP Orthography

The term ‘orthography’ is generally agreed to mean ‘a writing system’, though the exact boundaries of an orthography are elusive. Can two languages have the same orthography? If two languages have the same alphabet, do they have the same orthography? If two languages have the same alphabet, the same phonological processes, the same word formation/division strategies, and the same punctuation rules, do they have the same orthography? My answer to these questions is pragmatic: if the two languages share a single body of literature, i.e. the same dictionary, the same literacy materials, etc., then yes, they have the same orthography. If however, it is necessary to produce two or more sets of materials, then in spite of massive overlap, the languages have different orthographies. This “body of literature” condition can be seen as an orthographic framing of the question ‘Are these two varieties dialects of one language or different languages?’ The answer given provides a pragmatic test to apply to the dialect question. If you can develop a single orthography, they are dialects, if separate orthographies are required, they are languages.

As the remainder of the discussion concerns orthography, the examples will be given in the Kambari orthography unless indicated by formal linguistic bracketing. Tone carries a very small semantic role in the Kambari co-languages and will not be represented in either the phonemic or the orthographic examples of the remaining sections. The factors regarding writing of tone are discussed in section 3.3.2.3.

3.1 Language complexity

It is helpful in developing and evaluating orthographies to think in terms of specific areas of complexity for a given language. Every language is equally complex, but languages package their areas of complexity in different ways. For example, if you gave one hundred
complexity points to each language, Tyap,\textsuperscript{26} a language with nearly a hundred phonemes but very simple syllables, might give fifty points to its consonants and consonantal processes but one point for syllable structure. Yoruba may give only a few points to consonants and their processes and give the bulk of its complexity points to tone, since the language has both grammatical and lexical tone, and there are thousands of pairs of words and phrases which differ only in tone.

The area of complexity for a given language is the area where most of the orthography discussion will be centered. The tension in these areas of complexity is to develop the simplest way of writing so that the orthography is both accurate and functional. The formulators of the English orthography addressed its area of complexity: homonyms.\textsuperscript{27} English has thousands of homonyms. If English had been written with the one phoneme equals one symbol principle, it would have created mass ambiguity to the point where the reader would struggle and often fail in grasping meaning. Given the appropriate context, it is correct to say any of these three sentences: ‘I have two.’ ‘I have to.’ ‘I have, too.’

Phonetically, ‘two’, ‘to’, and ‘too’ are all [tuː]. However, those three expressions mean very different things. It is because of this factor that Chomsky and Halle (1968:49) state that it comes ‘remarkably close to being an optimal orthographic system for English.’

Orthography development basically follows the scientific method: prepare a hypothesis, test it, and refine the hypothesis into a working theory that may undergo further (and smaller) refinements over time. The first step in the process is that language developers and/or linguistic analysts tentatively propose a way of writing. It sometimes happens that a

\textsuperscript{26} The Tyap (Kataf) language of Kaduna State, Nigeria, can show contrast for about one hundred phonemic consonants but it has a syllable structure that is almost always CV.

\textsuperscript{27} Homonyms are words that have the same form but differ in meaning. Note that the ‘form’ is unspecified. Thus there are two types of homonyms possible once an orthography is established: homophones are different words with the same pronunciation and different spellings, while homographs are words with the same spelling but different meanings (Crystal 1997:185).
group of well-meaning non-linguists get together to ‘write their language’ without taking into account the phonological and grammatical factors in their way of writing. The resulting orthography will be, at best, a harmonization with the closest LWC, and at worst a muddled, unusable hodgepodge. Sjoberg writes,

> Devising efficient orthographies for unwritten languages is a task that requires the services of specialists – of linguists trained to analyze a language into its essential components and to select a minimum set of symbols that will mark all significant contrasts, yet satisfy the demands of practicality in the socio-cultural realm. Only in rare instances has an individual lacking knowledge of these special techniques been able to create a relatively simple and consistent orthography for his language. (Sjoberg 1964:262)

Orthography development should be guided by linguistic specialists who propose a system based on analysis, and undertake systematic testing before decisions are finalized. An orthography is tested in two ways: a) the developers address the consistency issues as they develop materials (see section 4.1.2); and b) the developers get feedback from the community as the materials are distributed. It is a mistake for developers to refrain from public testing in hopes of presenting a ‘perfect’ orthography on the first public exposure. Community feedback is an essential part of the development process. The process of developing a preliminary orthography into a functional working orthography such that future refinements would be minimal takes perhaps five years or more.

Language planners who are developing an orthography should begin with the basic principle that one phoneme will be represented with one symbol. However, they should be aware that this is not always possible, or, in certain cases, even desirable. The advanced principles articulated in chapter four help guide decisions in these areas. With this in mind, I turn to the specifics of the Kambari orthographies.
3.2 Kambari areas of low complexity

3.2.1 Letters corresponding to English and Hausa

When the specific Kambari sounds match the English and the Hausa sounds, the same orthographic symbols are used: <b, d, f, g, h, j, k, l, m, n, p, r, s, sh, t, v, w, y, z>. This is conformity to the major written languages in the area. The Kambari people who are literate in Hausa are accustomed to the letters <j> and <sh>. Phonological processes would make it possible to eliminate these letters and <c> (ch in English), replacing them with <zi>, <si>, and <tsi>. This proposal was rejected from the early developmental stages. While speakers of the languages recognize that a noun beginning with <c>, such as AUN <cigb> ‘okra’ generates the NCM <tsu>, they are not aware of the allophonic relationship between [ts] and [c]. The effort needed to justify to the public, and to teach proper pronunciation for writing AUN *<tsia> instead of <ca> ‘weave’, would be greater than is warranted to maintain phonological elegance. By extension, other palatalized consonants could be written as Cy (<lya> ‘eat’) or Ci (<lia> ‘eat’). Cy conforms to the Hausa palatalized consonants, e.g. kyau ‘good’.

75) AUN  
<lya> ‘eat’  
<ilyuci> ‘town’  
<ikyamba> ‘body’

The other option, writing Ci, has no precedent in nearby languages, and presents no discernable advantage over writing Cy.

The same sociopsychological factors apply to <y> and <w>. The vowel/consonant status of these is fluid, e.g. the NCM /i/ becomes [y] before vowels. Omitting the phonetic transitional [w], e.g. AUN *<buə> instead of <buwa> ‘remain’ is phonologically defensible and in practice possible for the orthography. However, it was not going to be accepted by the Kambari public. This illustrates the distinction between analysis and application. The
most elegant phonology, when reflected in the orthography, produced unacceptable written forms.

3.2.2 Letters corresponding to Hausa

When Kambari sounds match those of Hausa, but not of English, the Hausa symbols are used: <c, ñ, ð, a, e, i, o, u, ts>. The letters <ñ, ð>, though not present on a standard typewriter, are widely used in Hausa materials printed in the area. The Hausa /ts/, as spoken in the Kano variety of Hausa is ejective, but Kambaris do not retain the ejective feature when they speak Hausa. They see no difference between the Hausa <ts> and the Kambari <ts>.

3.2.3 Consonants unique to Kambari

When the Kambari sounds match neither those of English nor those of Hausa, the pattern of other Nigerian languages is followed: <kp, gb>.

3.3 Kambari areas of high complexity: orthographic difficulties

The representation of the letters discussed in the preceding section was practically automatic; any Kambari literate in Hausa or English would probably have made the same choices and rejected any attempts to deviate from them. Conformity and familiarity are the dominant factors in orthography design for areas of low complexity in a language.

The issues presented now do have various options to consider. These issues are examples of the advanced principle of proliferating options (advanced principle 11, section 4.1.11). Weighing the advantages and disadvantages of each option (advanced principle 12, section 4.1.12) clarifies which factors are most important. Options may then be accepted or rejected based on their relative merits or lack thereof.
3.3.1 Decisions regarding graphemes

3.3.1.1 The Kambari /a/

In the non-KLP orthographies presented earlier, both made use of the schwa for this phoneme, although Crozier discussed the possibility of a dot under the letter <a>.

The representation of the Kambari phoneme /a/ presents multiple orthographic options. This decision merits careful attention. Using the set AUN /kali/ ‘snake’, /kãli/ ‘net’, and /keli/ ‘witch’ as examples, the following options exist.

1) Write one symbol that serves two sounds

The /a/ would be written with either <a> or <e> (the one chosen would be used everywhere /a/ occurred). That letter would stand for two sounds.

76) <kali> ‘snake’ <kali/keli> ‘net’ and <keli> ‘witch’

It is an option, but not a good one, since the double duty of one vowel grapheme would create many ambiguous situations. If this option is chosen, it would violate the basic orthographic principle of phonemic accuracy, i.e. each phoneme (especially if it has a high functional load as this one) should be represented with a unique grapheme.

2) Write the IPA character

The /a/ would be written with the IPA symbol <ə>.

77) AUN <kɔli> ‘net’, <kagbɔdagi> ‘baboon’

To its advantage, it is phonetically accurate, and it is in use in some languages of Nigeria. To its disadvantage, many people dislike it (violates acceptability principle). Kambari readers consider the schwa symbol to be difficult to read. At one time I had a book on the coffee table that was written in English with one <ə> in the title (part of a language name). An educated Kambari visitor picked up the book, looked at the title, and said, ‘I cannot read this
book.’ Then he put it down and did not look at it again. The teaching load (advanced principle 13, section 4.1.13) would be greater with this option (overcoming perception of excessive difficulty), especially with Kambaris who are already literate in Hausa which does not use this symbol.

Conformity with Hausa in this area is problematic. The Hausa orthography that is familiar in the Kambari area has five orthographic vowels, but underspecifies phonemic vowel length. Kambaris speaking Hausa correlate the Kambari schwa with the Hausa /a/, and the Kambari /a/ with the Hausa /aa/. Any attempt to distinguish the Kambari schwa ceases to conform to Hausa. In this case, the conformity to Hausa is of less significance than maintaining the accuracy through the unique representation of the Kambari schwa. An additional disadvantage of this option is the difficulty (inconvenience) of making the letter with an ordinary typewriter or in a local printing establishment.

3) Write <a>

The /o/ would be written as <a>.

78)  <kali> ‘net’, <kagbadagi> ‘baboon’

This has the advantages of resembling <a> (a recognizable letter), and being easy to write and type. It is also being used in other languages in Nigeria. Williamson (1984:14) lists it as an alternative to the dot under the vowel currently in use in languages like Yoruba. The <a> has the disadvantages of being so similar to <a> that it would be possible to confuse the two, especially since it is easy to neglect to write the underline. Another disadvantage is the difficulty of using underlining as part of a text, as in a title of a book or a section to be emphasized.

---

28 Hausa-speaking Kambari people are not aware of the recent Hausa orthography changes that distinguish vowel length. Most of the written Hausa available to the people is locally generated.
4) Other possibilities

The schwa would be represented with vowel plus some other diacritic, e.g. a dot under an existing letter such as \( \langle \text{a} \rangle, \langle \text{u} \rangle, \) or with the symbol \( \langle \text{i} \rangle \) or something else. The dotted form has the same advantages and disadvantages as the \( \langle \text{a} \rangle \), except that it is harder to type. The \( \langle \text{i} \rangle \), while easy to type (i backspace hyphen), is generally employed when there are nine or more vowels in a language. In addition to that, of interest to linguists, but perhaps less so to MT speakers, the non-high vowels /a, e, o, \( \langle \text{a} \rangle \)/ function as a natural class; so, to use a symbol from the high vowel series seems to violate the natural grouping of the vowels.

The grapheme \( \langle \text{a} \rangle \) is selected based on the points mentioned above. Readers quickly learn to recognize \( \langle \text{a} \rangle \) and \( \langle \text{a} \rangle \) as different letters when they see them. Both the primer series and the transition primer include lessons in the reading of the two graphemes. In all the books produced by KLP since 1992 (including over two hundred pages of teacher’s guides per language), emphasis could always be indicated with other means than underlining. Of course, with computer-based publishing, compensation with bold and/or italic type is possible. A work-around solution for typewriters or local printers would be to use all capital letters or make the underline of \( \langle \text{a} \rangle \) a half space lower.

3.3.1.2 The Salka Kambari vowel inventory

Crozier lists nine phonemes, and thus requires nine graphemes. Crozier and KLP agree on six phonemes: /a, e, i, o, u, \( \langle \text{a} \rangle \)/. The \([\varepsilon, \partial, i]\) are seen as allophones by KLP.

Concerning \([\varepsilon, \partial]\), neither Auna Kambari nor Agwara Kambari has \([\varepsilon, \partial]\) in its phonetic inventory, only Salka Kambari does. Careful phonological analysis (discussed in section 2.3) shows that in Salka Kambari, \([\varepsilon, \varepsilon]\) and \([\partial, \partial]\) are allophones, and in some
environments function in free variation. There is neither ambiguity nor loss of meaning by representing them with one set of symbols; \(<e, o>\) were chosen for convenience.

Concerning \([i]\), this is an allophone of /u/ based on the analysis of vowel phonemes.
The orthographic considerations of this are treated in section 3.3.2.6.

### 3.3.1.3 The writing of the glottal stop

In the phonological analysis, the glottal stop was considered to be a consonant.

Kambari people who are literate in Hausa and/or English do not think of the glottal stop as a letter of the alphabet but a type of punctuation, since the same symbol (apostrophe) exists as such in written English, and the glottal stop occurs only rarely in Hausa. Most of the Hausa words with the glottal stop are borrowed from Arabic, and even the Hausa name for glottal stop, \(alhamza\), is from Arabic. In the Kambari community, only serious students of Hausa know the name \(alhamza\). Two options exist.

1) **No special mark for the glottal stop would be used**

Kambari could write two vowels consecutively, and teach people that when they see double vowels, there is a break between them. A problem area would be diphthongs, especially the ones formed by adding the suffix /-i/ ‘CMP’ to the end of words.

79) AGW \(<\text{banai}>\)/\(\text{bana-i/ [bå.nài]}\) go-CMP ‘went’
is two syllables,
\(<\text{yain}>\)/\(\text{ya?an-i/ [yà.?in]}\) do-CMP ‘did’
following the affixation rule is two syllables, not one.

2) **Write the glottal stop with the apostrophe, as in Hausa**

The glottal stop written with an apostrophe conforms to Hausa. From this perspective, the decision is not difficult.
The use of the glottal stop in Agwara Kambari, i.e. when to write it, will be discussed later in section 3.3.3.

Phonologically [ʔy] and [ʔw] can be analyzed as either phonemes or a sequence of two phonemes. Either analysis would generate the same output in the orthography. However, there are not many words with the sounds and this low frequency allows orthographic options. Auna Kambari has words with [ʔy] but not with [ʔw]. Salka Kambari has words with both [ʔy] and [ʔw]. Agwara Kambari has words with [ʔw], but only one word with [ʔy]. However, that [ʔy] occurs in the pronoun [ʔya] ‘they-NCMP’, which is in contrast to [ya] ‘you-PL NCMP’ occurring in identical environments. There are two options.

1) No distinction between w and y and their glottalized counterparts

The first option would be to write [w] and [ʔw] both as <w>, and [y] and [ʔy] both as <y>. The contrast between the semivowels and their glottalized counterparts in Agwara and Auna Kambari is not widely utilized, but as the example above illustrates, it is essential. Failure to mark this contrast would violate the necessity to convey adequate semantic information to the reader. In Auna Kambari the best contrast is in (81).

81) <kayanshi> /kayanshi/ ‘one who rubs-down’
    <kayanki> /kayanki/ ‘one who presses’

Nonetheless, for the reasons stated above, the feature must be marked in the orthography.
2) Write the apostrophe before the semivowel

The second option would be to write <'y> and <'w>. Hausa has a few words with <'y> (though no words with <'w>) and represents it as <'y> as in <'ya> ‘daughter’. Thus, in Agwara Kambari the pronouns would be visually different:

82)  <Yiɗai ya kuya'ani?> what is-it you-PL ING-do ‘What are you doing?’
     <Yiɗai 'ya kuya'ani?> what is-it they ING-do ‘What are they doing?’

This is the preferred option. Confirmation for it comes from Salka Kambari, where these sounds in (83) contrast.

83)  <yuwan> ‘do’     <'yuwan> ‘refuse’
     <yon> ‘yesterday’  <'yon> ‘rise’
     <uwa> ‘enter’      <'uwa> ‘house’

A final note regarding the glottal stop in the KLP orthography: KLP publications use a computer type font with a modified form of the apostrophe for the glottal, i.e. <’> instead of <‘>. While not strictly necessary, the modified form helps maintain the distinction between the glottal stop and the symbol marking embedded quotes.

3.3.2 Decisions regarding phonological processes

3.3.2.1 How to write vowel nasalization

In the phonology of the three languages, there are no phonemic nasal vowels. Phonetically nasalized vowels are the result of a phonological process on the phonemic /Vn/ sequence, i.e. nasalization of the vowel, then deletion of the nasal consonant. This is an area where the Kambari languages diverge from both English and Hausa. Both English and Hausa have consonantal /n/ word finally. There are three options.

1) Not to represent nasalization at all

One option was to ignore nasalization completely. This is not a realistic option because there are many contrastive pairs based on this distinction.
2) Write a diacritic

The second option was to write it with a diacritic (tilde) over the vowel.

This, while phonetically accurate, is difficult to type, and the nasalized schwa would have two diacritics, a tilde above and a line below, e.g. \( \tilde{a} \). It would also introduce six new characters to the alphabet, each nasalized vowel. This option is not phonologically accurate, and violates the simplicity principle.

3) Write \(<n>\) syllable finally

Crozier and KLP both advocate writing \(<n>\) to denote nasalization. Crozier proposes marking every syllable that has nasalized vowels. Using examples from Salka Kambari, Crozier would write as in (86).

4) Write \(<n>\) word finally

The fourth option, the one chosen by KLP, was to write \(<n>\) word finally and teach the people that \(<n>\) at the end of words in Kambari is different from \(<n>\) at the end of words in English and Hausa. This option has the advantage of being the easiest to type, is more faithful to the phonological analysis (that nasal vowels are underlyingly /Vn/) and nasalization spreads, especially through continuant consonants and the glottal stop.
Nasalization therefore only needs to be written once, generally word finally. This option introduces no new characters to the alphabet.

An example in the Auna Kambari orthography is:

87)  <kaci> ‘head’
    <kacin> ‘pox’

The following example provides the KLP orthographic form for the words of example (86). The KLP approach simplifies the spelling, especially when the syllable preceding a glottal stop is nasalized:

88)  [ʔyũwɑ̃] ‘refuse’ <'yuwan> 29
    [ɔũũũ] ‘leaves’ <avu'un>

3.3.2.2 How to write long vowels and long consonants

The phonology discussion established the linguistic parameters of length. Using a double letter to represent length is both convenient and conforms to other languages. The question here is when to write length rather than which symbol to use when writing length. There are three options.

1) Never distinguish long vowels or long consonants

All words, without exception would be written with a single vowel or consonant. This option would be conforming to the Hausa tradition on vowel length as it is practiced in the Kambari area. If this were the option chosen, never to represent length, Kambari readers would be taught to pronounce words naturally, i.e. if the word has a phonetic long vowel, they are to pronounce the word with a phonetic long vowel.

29The alternation between <i> and <u> as the vowel of the first syllable is a minor dialect or ideolect variation as to whether the V[+Hi] is influenced primarily by the frontness of [y] or the labialness of [u]. This limited sort of variation occurs in other similar environments in the three Kambaris. In the KLP orthography, the labial feature takes precedence.
2) Write length every time it is present phonetically in the language

This option involves several steps: analysis, consensus on the analysis (that all analysts and influential opinion leaders agree which words do and which do not have length), teaching people to read with correct pronunciation, teaching people to analyze for themselves (hear the distinction in their minds) so that they can 1) write correctly (this assumes a population-wide consistency in following the agreed upon inventory of words with long vowels and consonants), and 2) produce a consistent body of literature using vowel length. This would be a burdensome workload for the clarification of a relatively few instances of potential ambiguity (an example of weighing advantages and disadvantages, advanced principle 12, section 4.1.12).

Another disadvantage is lack of consistent representation of consonant length in locally produced written Hausa. Standardized Hausa represents consonant length, but Hausa as it is written in the Kambari area in letters, invitations, agendas, etc. represents consonant length more or less randomly, i.e. the same writer in the same paragraph may write the same words both with and without length. By extension, Kambari writers would inconsistently represent length in Kambari just as they do in Hausa. Having a standardized system that no one follows is not very different from having no system at all. Tryon (1986:305–306) describing written Bislama, a language of Vanuatu, states,

There is no official spelling for the language, each writer [is] reflecting his own linguistic and educational background in a continuous stream of idiosyncratic spellings. It is not uncommon, for example, to encounter up to four different spellings to designate ‘house’, thus: haus, haos, aus, aos. Until such time as the government addresses itself to this problem, Bislama will continue to be a very largely oral medium of communication.

Writing segmental length with double consonants is possible. The philosophy used to determine the main concern, when to write length, is difficult to determine in the Kambari
setting. Consideration of the problem leads one to mark length on the basis of phonetic perception. This suffers the problems of variation between speakers, of over-representation of phonetic material, and ignoring the native-speaker intuition of the phonological structure of the words.

3) Write only to clarify ambiguities

This option would be to represent long vowels or consonants only when necessary to distinguish one word from another. This is the preferred option. The parameters used to determine ‘when necessary’ are these: a) that there is a minimal pair (one with long and one with short), and b) that the two words occur in similar enough domains to possibly cause confusion. Thus, AUN [kune] ‘leg’, and [kunee] ‘gift’ are both written <kune>, since potential ambiguity is minimal, and context provides adequate distinction in nearly all situations. However, AUN [kàrá] ‘tear’ and [kàrará] ‘mark’ may both occur in the same sentence in such a way that ambiguity occurs if both are written with a single vowel.

89) <A kara (‘tear’) tagaràda, ko a kara (‘mark’) u da wa.>
they tear/mark book, or they tear/mark NCM EQ not ‘They should not tear or mark the book.’

The first <kara> and the second <kara> are indistinguishable. The solution is to represent the length of ‘mark’, so that one would write <kara> ‘tear’ and <kaara> ‘mark’. Of the three Kambari languages, Salka Kambari has the most instances of orthographic vowel and consonant length, especially when making grammatical distinctions. About ten pairs of content words are distinguished by orthographic double vowels or double consonants in Salka Kambari. A further description of Salka grammatical vowel length is given in section 3.3.5.
Adjacent vowel constructions

In Salka Kambari there is a small number of words written with adjacent identical vowels which are pronounced not as long vowels, but as separate syllables. In no instance does a contrastive long vowel form exist. This allows the simplified writing of the forms using the VV convention without any overt syllable marking in the orthography.

\[ /u\ddot{a} + V/ \rightarrow [\ddot{u}.t\dddot{a}.\dddot{a}.] \rightarrow <u\dddot{a}a> \]

*\([u\dddot{a}a] \sim u\dddot{a}a\) does not exist

Many of the nouns in Agwara Kambari that have a vowel-glottal stop-vowel at the end are pronounced in a similar manner in stream speech. This, plus the evidence of the cognate benefactive suffix being /-ka/, points to a consonant onset for the syllable in earlier times.

3.3.2.3 How to write tone

Kambari is tonal, like most African languages, though the burden of comprehension based on tone contrast is slight (see section 2.3.1.5). Three options exist for tone marking in Kambari.

1) Write tone every time

As discussed in the previous section on the representation of length, the marking of tone throughout the language would involve a large workload for the relatively small benefit of clarifying a handful of potential ambiguities.

2) Never write tone

Neither English nor Hausa write tone. Thus if one were to follow the conformity principle, Kambari would not either. However, the conformity principle is useful only to the extent that it allows meaning to be accurately conveyed.
3) Only write tone when needed to clarify ambiguities

This is the option proposed by Crozier in the *Alphabets of Africa* (Hartell 1993:234). However, nine years of field testing materials without tone have shown that there are no instances where the writing of tone is necessary for comprehension. Usually there is some other feature in conjunction with tone that makes the meaning clear.

In the Carroll orthography, words are occasionally marked with tone markings. Mother-tongue readers report that they have no need for the tones, and linguistically trained MT speakers have not been able to identify what information the tone marking is seeking to convey.

Salka Kambari provides a clear example of the Kambari structures that allow an orthography to function without marking tone.

91)  

\[ <a \ t\ a\ w\ shi> \ LLHL \ they \ come \ not \]

‘They did not come.’

\[ <\#a \ t\ a\ w\ shi> \ LHHL \ they-NCMP \ come \ not \]

‘They will not come.’

The second sentence has tone change and also has vowel length to distinguish the pronouns. Thus, the tone feature is redundant, since the long vowel of the pronoun signals the reader as to the tone of the verb.

3.3.2.4 How to represent vowel copying

Vowel copying is a very productive process occurring within and across morpheme boundaries both to the right and to the left. The phonological processes were exemplified in section 2.3.1.3 on the structure of Kambari. The following discussion concerns the orthographic representation of the underspecified vowels in functors in Agwara Kambari, and by extension, in Salka Kambari. The same morphemes in Auna Kambari are always written with <a>.
The options for vowel copying are to a) write the morphemes containing an underspecified vowel with a single default vowel, <a> or <a>, or b) to write the resultant copied vowel (surface form). Regarding option a), the choice of the vowel to be written would be arbitrary, since the indicators in the languages offset each other. The default vowel /a/ surfaces if the following word contains no non-high vowels (favoring <a>). In comparison of co-languages, the equivalent morphemes in Auna Kambari are always <a> (favoring <a>). Analysis of mV-class nouns in Agwara Kambari revealed that if the root contained only high vowels, there was approximately the same number of nouns with the prefix [ma-] as there was with the prefix [ma-], indicating a lexical specification of one form of the vowel in the prefix, but not establishing which is specified and which is default. These mixed indicators require one to consider all the options carefully before selecting one for use in a standardized orthography. The following examples illustrate Agwara Kambari kV-class nouns in which the prefix copies from the non-high vowel in the root.

92) Option a) <kadaku, kadele, kapodo, kagbadagi>³⁰

   Option b) <kadaku, kedele, kopodo, kagbadagi>

   ‘guinea fowl, gourd, toad, baboon’

If one writes the surface form of the copied vowel (option b), the words maintain an orthographic representation of the vowel co-occurrence restrictions of the phonology. This factor contributes heavily to a preference for option b.

From the evidence found in the environment of a single word (the noun-class prefixes) the preferred treatment is determined. Consistent orthographic treatment of the same phonological process requires this copying rule to generalized and applied to such

---

³⁰ This data assumes <a> would be chosen as the default vowel; however, <kadaku, kadele, kapodo, and kagbadagi> would also be possible. The latter three examples are identical with Auna Kambari. (Auna Kambari has the non-cognate word <kakulya> ‘guinea fowl’.)
analogous environments as those involving NCMs, pronouns, and the preposition /V_{[4 Hi]}/
‘to/at’ in Agwara and Salka Kambari, even across word boundaries.

Agwara and Salka Kambari extend the vowel copying process to the preposition ‘to/at.’ Auna Kambari, as described above, does not copy to the left, so in most cases, phonetically the morpheme is [a]. The exception to this phonetic [a] in Auna Kambari is the vowel coalescence that occurs when the preposition precedes words beginning with a high vowel (see section 2.3.1.3).

93) AUN     a ure [ó rè] ‘at road’
       a iyamba [è yám bā] ‘at land’

The prevalence of [a] as the surface form for this morpheme allows the Auna Kambari ‘to/at’ to be written ‘to/at’ as <a> wherever it occurs, even in environments such as those above where it coalesces.

Salka Kambari has the same vowel coalescence process. The orthography of Salka Kambari can extend the orthographic treatment of vowel copying to this morpheme. The coalesced form may be written followed by the complete content word.

94) underlyingly /V_{[4 Hi]} u'wa/ phonetically [ó rèwà] (the V first copied the /a/ from u'wa, then applied vowel coalescence to the a-u sequence) is written <o u'wa> ‘at house’

underlyingly /V_{[4 Hi]} isula/ [è sù:là]

is written <e isula> ‘at wall’

Thus the representation of the preposition /V_{[4 Hi]}/ in Salka Kambari accommodates the processes of vowel coalescence (examples above) and vowel copying, e.g. <a mapara> ‘at dish’, <o mokolongu> ‘at small farm’. In Agwara Kambari (which does not use vowel
coalescence in this place) the preposition may written as the copied form of the non-high vowel in the in (95).

95)  
<o kudolu> ‘at river’
<e kuden> ‘at market’

3.3.2.5 How to deal with vowel elision

This point deals with the writing of the full form of the verb or a contracted form because of vowel elision processes: underlyingly AUN /u rawa tə/ ‘he has arrived’ is phonetically [ʊ́ rə̀ ] tə, and /u bana tə/ is phonetically [ʊ́ bə̀ ] tə. The Stable Form Principle (section 4.1.7) requires that the full form of the verb be written: <u bana tə> ‘he go PERF’, and <u rawa tə> ‘he arrive PERF’. This single visual image allows ease of recognition, and writers are not required to memorize spellings of multiple forms for the a single word, along with rules for when to select which form.

3.3.2.6 How to write sub-phonemic vowels

This point deals with two issues, the elision or partial elision of vowels between a sonorant consonant and another consonant (96), and the representation of the allophone [i].

96)  
AUN <əsəkə> ~ <əskə> ~ <əsikə> ‘let/allow’

The first issue is resolved by maintaining the CV syllable pattern in the language. The complete elision is a fast speech variant, careful speech will have some phonetic reality to the vowel.

The allophonic nature of [i] was discussed in the section on phonology. Following the orthographic principles, it should naturally come about that this vowel would be written <u> and that readers would automatically make the phonetic adjustments. This is the preferred option from both an analytical and a reading theory perspective. However, this
was strongly rejected in test materials. Readers are fully aware of the pronunciation of [u], and they perceive that [i] is not [u]; therefore it must be written differently.

User response cannot be ignored in orthography development. When the developers and analysts perceive a superior solution to the one which seems acceptable to the public, one or the other must give way. If the written form of the language under development will not be used, orthography development becomes meaningless. One method of dealing with the kind of situation described above is to establish a trial period of the theoretically better decision and see if the familiarity factor overrides the objections. An example from the KLP orthography development period illustrates this approach: In 1992, it was suggested by KLP staff that /iNkpa/ ‘fly’ should be written <inkpa> because of the nasalization of the vowel. While it is true that the vowel is nasalized, the representation of nasal features immediately before labial consonants was at that time <m> in all other environments. At the advisors’ recommendation, the staff agreed to try a test period of spelling the word <imkpa>. This recommendation was based on the principle of system-wide consistency, i.e. applying the same orthographic convention wherever the environment occurs. In this case, within words, the nasal before labial consonants is <m>, producing <mkp>. A few years later, after seeing <imkpa> ‘fly’ repeatedly, the same staff declared that it would be incorrect to write it <inkpa>.

---

31 Reading skills in the area are taught as direct sight/sound correlation. In many instances reading is learned first in a language the students do not speak, thus there is a zero relationship between reading and meaning, and readers do not learn to rely on their intuitive knowledge of the language to obtain the correct natural pronunciation of, or the information content of, written materials. This directly affects the functional options at hand for orthography.

32 In personal communication from Professor Oladele Awobuluyi, I have become aware of a similar situation in Yoruba. Awobuluyi writes of the allophones [l] and [n], “They are allophones all right, with [l] occurring before oral vowels, and [n] before nasal vowels. But speakers perceive them as phonetically distinct and therefore take it that they must be written differently. Naïve orthographic tradition going back to the 1840s most regrettably reinforces their contention. The upshot is that no diehard Yoruba linguist has been able to muster enough courage to insist on writing them with the same symbol.”
3.3.3 Variations by speakers

One type of problem that crops up in orthography development occurs when MT speakers pronounce the same lexical item in various ways in identical contexts. This variation can be across the speech community, or it can be variations of the same words by a single speaker. One example of this is found in table 2.6, in which the NCM between /ucira/ ‘power’ and /ro/ ‘another’ is variously [u] or [o]. The NCM could be represented in the orthography with <u> and/or <o>. It is seen in the rest of the examples of table 2.6 that the surface representation is not based on the phonological or morphological properties of either the word before or the word following the NCM, i.e. there is no predictable pattern based on morpheme or environment. Trying to approximate the phonetic realization of these NCMs would mean specifying in the lexicon the NCM for every phrase possible in the language. Even then, it would ultimately be unsatisfactory because of the variations among speakers. The only workable solution that provides a predictable system is to write the underlying form and teach the readers to pronounce it naturally.

Another case of wide variation among speakers is that of glottal syllables in Agwara Kambari. It should be kept in mind that not every glottal stop in the language is questionable orthographically. Some are lexically intrinsic in the word as in <pi'isa> ‘learn’. Some are the benefactive/instrumental allomorph (usually –kV) on verbs with /k/ as the final consonant as in <su'uku> ‘send with/for’ from <suku> ‘send’. These are recognized by all speakers and are represented in the orthography.

The glottal syllables whose orthographic status is in question are word-final syllables made up of a glottal stop followed by the same vowel as that of the preceding syllable. These syllables are most obvious when the word is said in isolation. In fast or informal speech, these glottal syllables are usually not pronounced. There is a wide degree of variation among native speakers as to which words are or are not expanded in such a way.
For most speakers, there is a psycholinguistic reality to the presence of the glottal syllable. Speakers feel that \(<c'a'>\) is the real word. However, they do not retain the glottal syllable in the affixed forms, i.e. \(/c'a' -sV/\) ‘plant-IT’ is pronounced [cəsə]. Fluent oral readers often bypass written glottal stops and at times add unwritten ones, especially paragraph finally. However, the glottal syllable in question here cannot be treated as an utterance-final phenomenon, since this phenomenon is limited to certain lexical items.

Here is the orthographic dilemma: if the glottal syllable is psycholinguistically essential to the lexical meaning of the word, it should be written. If it is merely a rule-generated filler syllable, it need not be written at all. Even speakers who agree it should be written have different inventories of words that accept the glottal syllable.

These are some examples of the perception of native speakers: James states that the word for ‘friend’ is \(/kaja'/\). His full brother Jeremiah states that \(/kaja'/\) is ‘work party’, and \(/kaja/\) is ‘friend’. James replies that both meanings are \(/kaja'/\). Basuna says that \(/kaja'/\) and \(/kaja/\) are in free variation for either meaning. James distinguishes between \(/kusan/\) ‘mountain’ and \(/kusan'/\) ‘soup’. Aaron says that both are identical, \(/kusan/\). James says \(/uye/\) ‘road’. Basuna says \(/uye'e/\) ‘road’. Jeremiah says ‘road’ is \(/uye/\), and \(/uye'e/\) means ‘malted sprouts’.

I arranged a survey to see what sociolinguistic factors triggered glottal syllables. The survey consisted of about thirty pictures; most of these items are listed in the dictionary with glottal syllables. The survey revealed the determining factor to be the age of the speaker. Older people always pronounce the glottal syllables, young people seldom do.

This result brought up the philosophical question whether the orthography should follow the trend away from the glottal syllable or not, i.e. to follow the older people’s example because they have retained original forms of the language, or should it follow the
young people’s example as evidence of language progression? The general consensus among the individuals consulted was for the first option. It was held that the older people speak the ‘real Kambari’. Young people freely mix Hausa and Kambari. Young people will say the non-glottal form first, but once aware of the glottal syllable form, they generally feel that that form is more correct.

With this background established, I proliferated possible solutions for the orthography. These are represented in the flow chart in figure 3.1.

![Glottal syllable decision flow chart](image)

**Figure 3.1: Glottal syllable decision flow chart**
Step A: Is it advisable to write glottal syllables or not? There are three possible choices.

1) Always write the words without glottal syllables (option 1: Never). With this option, <yan> ‘do’ and <kaya> ‘arrow’ would be the correct spellings. This option disregards the native-speaker intuition that these glottal syllables are intrinsically part of the meaning package. It also ignores the perception that the best Kambari speakers include glottal syllables. Ignoring glottal syllables bypasses an opportunity to disambiguate, since words like <uye> ‘road’ and <uye'e> ‘malted sprouts’ would be written the same way.

2) Write either form willy-nilly (option 2). That is to say, words can be written with or without glottal syllables at the whim of the writer. With this option, <ya'an>, <yan> ‘do’, <kaya'a>, <kaya> ‘arrow’ would all be correct spellings. Until a standardized writing system is implemented, random treatment of this feature is what happens in actual practice. From my observation and experience, this uncertainty as to how to write correctly actually deters people from attempting to write in the language. People who are literate in the LWC have learned that there is one correct way of writing, and they want to follow the correct way even in their own language. Thus, option 2 violates the acceptability principle.

3) Write glottal syllables on words in an approved list (option 3: Selected inventory). This is the preferred option. It requires that certain words be written with glottal syllables. In choosing this option, the lexical inventory of words containing glottal syllables must be determined in one of the following ways of step B:

1) Write glottal syllables on every noun and verb containing fewer than the standard quota of syllables (option 3.1) even if no one ever says it that way. In addition to <ya'an> ‘do’ and <kaya'a> ‘arrow’, there would be <kawu'u> ‘lion’ and <kashl'i> ‘eye’, though the latter forms are not used by the speakers of the language. This option avoids arguments
among MT speakers, and thus has certain appeal. However, it creates unnatural, mechanical language and written forms that no one ever pronounces. The acceptability by the language community of this option is unlikely.

2) Write glottal syllables on every word that any speaker anywhere ever pronounces with it (option 3.2). This option involves massive survey and does not address the ambiguity issue. With this option, <kaja'a> would mean both ‘friend’ and ‘work party’, and <uye'e> would mean both ‘road’ and ‘malted sprouts’.

3) Create an approved list of words with glottal syllables and always write those words with them (option 3.3). This list in writing would strive for minimal ambiguity. Using the words from the above examples, <kaja'a> ‘work party’ would be included to distinguish it from <kaja> ‘friend’; uye'e ‘malted sprouts’ (as compared to <uye> ‘road’); <kusa'an> ‘soup’ (<kusan> ‘mountain’). Option 3.3 seeks to create a manageable inventory, striving for maximum disambiguation. It is not necessary to finalize the approved inventory at this point, but choosing this option standardizes the orthographic convention and the representation of the most common words. The consensus among those actively involved in writing the language is that option 3.3 is the most functional. Having the glottal syllable as a mechanism for disambiguation was one of the guiding factors.

Step C (options 4 and 5) arises from the fact that it is possible to write glottal syllables only in certain occurrences of these lexemes. The phonological behavior of these lexemes forces the language developers to consider occasional representation based on context, or obligatory representation based on morphemic consistency.

Option 4 is to write the glottal syllable for certain specified words (determined from one of the subpoints of option 3) when in that context it is pronounced. For example, the unaffixed verb /ya'an/ ‘do’ is written <ya'an>, but its affixed forms would be <kuyan>
‘ING-do’, <yansa> ‘do-IT’, <yanka> ‘do for [s.o.]’, and the phrase <u yan t>a> ‘he do PERF’. For nouns, kaya’a ‘arrow’ would be written <kaya’a> if sentence final and <kaya> if part of a phrase as in <kaya ke te> ‘arrow NCM one’.

Option 5 is to write the glottal syllable every time for certain specified words. For example, ya’an ‘do’ and all of its affixed forms would contain glottal syllables. This preserves the constant visual image of the root: <ya'an> ‘do’, <kuyaa'n> ‘doing’, <ya'an> ‘do-IT’ and <kaya'a ke te> ‘arrow NCM one’. The stable form principle argued against allowing more than one spelling per lexical item. Thus, following the pattern in use elsewhere in the Agwara Kambari orthography, each content word is spelled the same way every time it occurs. Elsewhere in the language, affixed forms of verbs carry the verb root intact. These factors led to the selection of option 5.

If it has been established that glottal syllables should be written, (either from selecting option 2 or option 3 via any of its subpoints), one needs to determine the orthographic convention to be used for representing these syllables. Options 6.1 and 6.2 deal with this issue: Write two vowels in sequence (option 6.1), but teach readers that they can pronounce them smoothly or with a break between them: <ya'an> [yān ∼ yāʔan]. This may have been what the Carroll language developers had in mind when they selected words to represent with double vowels.

Write the glottal stop (option 6.2), but teach readers that they can skip it if they choose to do so: <ya'an> [yān ∼ yāʔan]. It was decided to write the glottal stop (option 6.2) since there are other words in the language written with the glottal stop.

---

33 These were both numbered “6” on the reasoning that no more structural decisions are to be made, the only consideration is the visual representation of the previous decisions.
3.3.4 Decisions regarding morphemes

3.3.4.1 Word spacing

The placement of spaces and affixes is probably the most difficult area of orthography design shared by all languages. In every language, the concrete nouns and verb roots are fairly easily identified as independent words. It is the grammatical information that is difficult to account for in the spacing decisions of orthography. To deal with this area of Kambari orthography, I must anticipate a large number of the advanced principles that are delineated in the next chapter. Especially useful in the area of word spacing are the convey meaning principle, the preference for independent words principle, the stable forms principle, the limited ambiguity principle, and the consistency principle.

3.3.4.2 Kambari spacing decisions

In Kambari, most of the grammatical morphemes are one syllable. In addition, there is huge overlap in underlying forms. This is described more fully in 3.3.5 ‘Kambari power syllables’. In the verb phrase, each morpheme must be carefully considered as to its status as word or affix. In the noun phrase, the specific morphemes under consideration here are the NCMs, the possessive pronouns, adjectives, and the specifier.

Morphemes to the left of the root in the Kambari verb phrase

For the Kambari verb phrase, three possibilities immediately suggest themselves when the ‘proliferate options’ principle (section 4.1.11) is applied: a) all morphemes of the verb phrase are affixed to the verb root; b) all morphemes of the verb phrase are independent words; c) some of the morphemes are affixed, and some are freestanding. Table 3.1, (using the morpheme string explained in section 2.3.2.4), focuses on the verb root and the morphemes to the left of it. The morphemes displayed in the obligatory subject marker column are pronouns and/or NCMs that are homographs with the morphemes in the aspect column.
<table>
<thead>
<tr>
<th></th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>0</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AGW</td>
<td>ku</td>
<td>tsu</td>
<td>ku</td>
<td>vasa</td>
<td>ku tsu vasa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘NCM REP pray’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ku kuvasa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘NCM ING-pray’</td>
</tr>
<tr>
<td>AUN</td>
<td>u</td>
<td>tsu</td>
<td>u</td>
<td>vasa</td>
<td>u tsu vasa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘he REP pray’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>uvasa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘ING-pray’</td>
</tr>
<tr>
<td>SAL</td>
<td>a</td>
<td>tsu</td>
<td>a</td>
<td>vasa</td>
<td>a tsu vasa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘they REP pray’</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>avasa</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>‘about to pray’</td>
</tr>
</tbody>
</table>

Table 3.1: Morphemes to the left of the verb root

Verb root: (Position 0) The uninflected form of the verb is the same as the singular imperative (command to one person). It can be written as a freestanding word since it is both identifiable by MT speakers and can be said as a minimal utterance.

Aspect morphemes: (Position –1) Aspect morphemes immediately to the left of the verb root can be considered as prefixes to the verb. This illustrates how the principle of preferring independent words can be applied to real language data, and in a given context, affixation can be justified. In favor of independent words is the lack of phonological bridging between the aspect morphemes and the verb root in Auna and Agwara Kambari. However, when considering the structure in Salka Kambari, phonological bridging does take place. The primary justification of affixation is based on being able to maintain orthographic distinction between two phonetically identical morpheme strings, e.g. AGW /ku vasa/ ‘NCM pray’, /ku vasa/ ‘ING pray’. If both were spaced or both were affixed, the orthography decision would generate communication-inhibiting ambiguity.
Further support of affixation is that, especially in Agwara Kambari, if a speaker is asked what his word for ‘sit’ or ‘walk’ is, he will answer *kudasangu* /ku-dasangu/ ‘ING-sit’ or *kuwala* /ku-wala/ ‘ING-walk’, respectively. The speakers feel a completeness about the gerundial form that makes it the natural response in the above scenario rather than the imperative. If [kuwala] is said in isolation, MT speakers will invariably understand it to mean ‘walking’, and only understand it to mean ‘it walks’ if an antecedent Ku-noun is present. This argues for the affixation of the gerundial morpheme. The constant visual image of the verbs is maintained in the sense that there is only a single morpheme allowed as a prefix, strictly limiting the possibilities of form at the beginning of the verbs.

*Tsu:* (Position –2) This repeated action morpheme appears between the obligatory subject marker and the verb. There are four options for spacing/affixing: a) everything to the left of the verb root is prefixed; b) the subject marker is freestanding, but *tsu* becomes a verbal prefix; c) the *tsu* is joined to the subject marker, but there is a space between *tsu* and the verb root; and d) the subject marker, *tsu*, and verb root are all spaced. The presence of this morpheme *tsu* is predicated upon the presence of both the subject marker and the verb root. However, the obligatory presence of the three morphemes provides no useful information for deciding which of the four options should be employed.

According to option a, this repeated action morpheme would be an infix: *<utsuvasa>* ‘he-REP-pray’. As was discussed in the morphology section, in Salka Kambari it would be possible to have *<tsutsutsutsula>* ‘we-REP-reduplicate-buy’. Without considering any of the potential affixes to the right of the verb root, with option a, it is possible that words of five or more syllables would be regularly used. That conflicts with conformity to either English or Hausa, both of which only very rarely have five-syllable words. Conflict with conformity to a LWC lowers the acceptability by MT speakers.
Option b: Separate the subject marker to the left of tsu, but join the tsu to the verb root: <u tsuvasa>; <tsu tsutsutsula>. In isolation <tsuvasa> does not invoke the meaning of ‘REP-pray’, but that of an identical morpheme string, ‘we pray’. This is in contrast to the MT speaker’s perception of <kuwala> ‘ING-walk’ in Agwara Kambari mentioned above. The incorrect assignment of meaning of <tsuvasa> makes the ambiguity generated by this option tolerable only if the disadvantages of the other options are insuperable. Another area of generated ambiguity is that of overlap with nouns belonging to the Ts-class.

97) AUN tsugono ‘kingship’, tsu-gono ‘REP-return’
AGW tsugolo ‘horsehair scepter’, tsu-golo ‘REP-break off’

There is the negative factor of obscuring the visual image of the verb instead of following the advanced principle of trying to preserve the beginnings of words in a stable, instantly recognizable form. A positive indicator for affixation is the phonological behavior of this morpheme in Agwara Kambari. It undergoes a palatalization in certain contexts (as does the phonetically identical /tsu/ ‘we’).

98) /tsu ciga/ ‘we want’ → [ci ciga]
/u tsu ciga/ ‘he REP wants’ → [u ci ciga]

Option c joins the repeated action morpheme to the pronoun or NCM, creating another set of subject markers: <utsu vasa>, <tsutsu tsutsula>. A two-morpheme pronoun set is already an active strategy and is discussed in detail later in this chapter. However, among the members of pronoun set 234 there is phonological interaction, but there is no phonological interaction between the subject marker and tsu. There is no positive sense of unity in the minds of MT speakers about <utsu> ‘he-REP’ or <tsutsu> ‘we-REP’. The slight gain by joining the subject marker and tsu in the manner of pronoun set 2 is offset by the

34 In pronoun set 2 the underlying form /u + a/ surfaces as [wa] ‘3P-non-completive (NCMP)."
following factors. First, there is potential ambiguity between AGW <atsu> ‘3P-REP’ and <atsu> ‘1PEmph. Second, a new pronoun set comprised of two morphemes increases the teaching load. Third, following the advanced principle of preference for independent words, there would need to be some positive criterion to trigger the implementation of affixation. There is none, so the principles in favor of spacing lead to option d.

Option d: Write tsu as an independent word: <u tsu vasa, tsu tsu tsutsula>.

Untrained native speakers cannot isolate and identify this tsu; however, following the positive criterion principle, absence of confirmation is not a determining factor. Tsu is easy to teach, and once writers are aware of its function, they have no difficulty in writing it correctly.

Obligatory subject marker: (Position –3) These consist of pronouns or NCMs of either set 1 or set 2 (see figure 3.2).

There is almost complete overlap of the forms of the pronouns and NCMs, e.g. ma may be either the first-person non-completed pronoun (set 2) or the NCM of the mV-class. The NCMs were discussed in the Kambari morphology and syntax section (2.3.2) where it was seen that these morphemes occur in a number of other constructions. The subject markers undergo the same phonological processes as the NCMs in other constructions. The advanced principle of ‘act the same, write the same’ applied to this group of morphemes indicates that however these are written, they should follow the same strategy throughout the language. Limiting the number of permitted prefixes on the verb root allows the verb to be easily recognized in varying environments. Affixation of subject markers would create longer words than readers are accustomed to seeing in the LWC. There is no clear-cut justification for making them affixes, but the syntactic behavior of associating with multiple word classes leads toward independence. The primary, decisive factor in favor of spacing the
subject markers is that by doing so strings of phonemically identical morphemes are disambiguated. The <tsu> ‘we’, <tsu> ‘REP’, and <tsu-> ‘noun class prefix’ were discussed above. The subject markers, which are phonemically identical with other functors, if joined to the following morpheme, often increase ambiguity by having the same visual image as other content words.

99) AGW  <apasa> ‘PL-water yam’; <a pasa> ‘they cross’
      <mabana> ‘basket’; <ma bana> ‘NCM go’
      <vuca'g> ‘planting [noun]’; <vu ca'g> ‘you plant’

The spacing of these subject markers was not an automatic decision because the untrained Kambari writer tends to join this morpheme to the verb. Also if asked what <u> ‘he’ means, he may not be able to respond. Out of context, <u> is just a letter. In considering this factor, I decided that it would be unusual for an untrained native speaker of any language to be able to define grammatical morphemes. In natural speech, a subject must occur with a corresponding verb. Thus, while content words may be understood in isolation, functors obtain meaning only in the context of a larger segment of speech.

Subject pronouns before vowel-initial verbs

One of the most difficult orthography questions in Kambari is how to write the subject pronoun before verbs, especially before verbs beginning with vowels. Salka and Agwara Kambari exhibit the same linguistic patterns in this context. Auna Kambari does not have verbs beginning with vowels, so beyond the question of whether to space or not to space, this was not a difficulty for Auna Kambari.

Table 3.2 with data from Agwara Kambari, displays the surface and underlying forms of the subject pronouns before a consonant-initial verb and before a vowel-initial verb. The sequence marker, /da/ is included to illustrate the three possible surface realizations of the first-person pronoun ‘I’ which is a single nasal consonant. The three realizations are: as
the coda of the preceding syllable [dam] ‘then I’, syllabic before consonant-initial verbs [m\_banai] ‘I go-CMP’, and as a prefix to vowel-initial verbs [menei] ‘I saw’. In this syntactic environment, the orthographic choice to be made is whether to preserve the phonetic syllables, e.g. <dam banai, menei> ‘then I went, I saw’ or to preserve the identity of the morphemes in some way.

<table>
<thead>
<tr>
<th>Surface form</th>
<th>Consonant initial verb underlying</th>
<th>English gloss</th>
<th>Surface form (phonetic)</th>
<th>Vowel-initial verb, underlying</th>
<th>English gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>m banai</td>
<td>n banai</td>
<td>‘I go-CMP’</td>
<td>menei</td>
<td>m enei</td>
<td>‘I see-CMP’</td>
</tr>
<tr>
<td>vu banai</td>
<td>vu banai</td>
<td>‘you’</td>
<td>venei</td>
<td>vu enei</td>
<td>‘you’</td>
</tr>
<tr>
<td>u banai</td>
<td>u banai</td>
<td>‘he’</td>
<td>wenei</td>
<td>u enei</td>
<td>‘he’</td>
</tr>
<tr>
<td>tsu banai</td>
<td>Tsu banai</td>
<td>‘we’</td>
<td>cenei</td>
<td>tsu enei</td>
<td>‘we’</td>
</tr>
<tr>
<td>i banai</td>
<td>i banai</td>
<td>‘you-PL’</td>
<td>yenei</td>
<td>i enei</td>
<td>‘You-PL’</td>
</tr>
<tr>
<td>a banai</td>
<td>a banai</td>
<td>‘they’</td>
<td>enei</td>
<td>a enei</td>
<td>‘they’</td>
</tr>
<tr>
<td>dam banai</td>
<td>da m banai</td>
<td>‘then I go-CMP’</td>
<td>da menei</td>
<td>da m enei</td>
<td>‘then I saw’</td>
</tr>
<tr>
<td>da vu banai</td>
<td>da vu banai</td>
<td>‘you’</td>
<td>da venei</td>
<td>da vu enei</td>
<td>‘you’</td>
</tr>
<tr>
<td>du banai</td>
<td>da u banai</td>
<td>‘he’</td>
<td>da wenei</td>
<td>da u enei</td>
<td>‘he’</td>
</tr>
<tr>
<td>da tsu banai</td>
<td>da tsu banai</td>
<td>‘we’</td>
<td>da cenei</td>
<td>da tsu enei</td>
<td>‘we’</td>
</tr>
<tr>
<td>di banai</td>
<td>da i banai</td>
<td>‘you-PL’</td>
<td>da yenei</td>
<td>da i enei</td>
<td>‘you-PL’</td>
</tr>
<tr>
<td>da banai</td>
<td>da V_{[Hi]} banai</td>
<td>‘they’</td>
<td>de enei</td>
<td>da V_{[Hi]} enei</td>
<td>‘they’</td>
</tr>
</tbody>
</table>

Table 3.2: Subject pronouns

Six options, loosely organized from underlying form to surface form, can be considered. The one option which should be rejected because it violates word integrity, is to decapitate the verb *<me nei> (underlyingly /M enei/) ‘I saw’. In addition to the violation of the constant visual image of the verb, it creates ambiguities: νeι’ ~ enei is another verb, ‘give-CMP’. This mismatch of syllable and morpheme boundaries is similar to the examples given by Beck (1964:156–158). Beck recognized three options: decapitating the content word, the underlying form, and writing the functor and content word together. He failed to recognize that there are other options, such as those adumbrated here.
1) Write the underlying form

This option would be to write the pure underlying form with no adjustments. In this option (morpheme-driven approach, described as morphophonemic by Gordon (1986:66–67), each pronoun and each verb would have one and only one spelling. There would be a space between pronoun and verb. The abstractness of the orthographic form compared to the surface form is a factor that became apparent in test materials. When seeing <u enei> ‘he see-CMP’, it was difficult for people to pronounce [wenei]. An intuitive awareness of the strong CV pattern in Kambari may have been an influence in this situation.

100) AGW  <da m banai> then I go-CMP ‘I went’  
        <m enei> I see-CMP ‘I saw’  
        <u enei> he see-CMP ‘he saw’

2) Adjust semivowels

This option would be to write the underlying form for all but /u/ ‘he’ and /i/ ‘you-PL’. In those cases, when coming before verbs beginning with vowels, the pronoun would be freestanding as <u> and <i>, but the following verb would begin with the semivowel. This breaks up the possible cluster of three vowels together, /da u enei/ ‘then he see-CMP’, which the Kambari languages try by all means to avoid. However, this violates the advanced principle of trying to maintain the beginning of words in a recognizable form (advanced principle 7, section 4.1.7), and gives three spellings for any verb root beginning with a vowel.

101) AGW  <m enei> I see-CMP ‘I saw’  
        <u wenei> he see-CMP ‘he saw’  
        <i yenei> you-PL see-CMP ‘you saw’

3) Adjust all pronouns

This option (a variety of the morpheme-driven approach) was to retain the isolatable form of the verb with one spelling and adjust the pronouns so that they are easy to
pronounce.\footnote{Emmanuel Dangana from Auna originally suggested this solution during the Introductory Course in Applied Linguistics in 1992.} It limits orthographic variation to a closed set of pronouns, with a maximum of seven or so spellings for each pronoun. The total number of pronouns in the language is fixed, but the number of content words is unlimited since new nouns and verbs can be added at any time. Altering the spelling of content words affects thousands of words. Altering the spelling of pronouns affects six words per pronoun set. The content words retain the constant visual image.

102) AGW  \(<m\text{-en}e\text{i}> I \text{ see-CMP ‘I saw’}\)
\(<v\text{-en}e\text{i}> \text{ you see-CMP ‘you saw’}\)
\(<w\text{-en}e\text{i}> \text{ he see-CMP ‘he saw’}\)

4) Join with hyphen

This option is to use a hyphen between the subject pronoun and the verb. The advantage of this option is that it both retains the isolatable spelling of the verb, and sets the pronoun apart. One disadvantage is that a Kambari text would be cluttered with hyphens that divide the onset and nucleus of a syllable. Another disadvantage is that four of the six forms are unpronounceable in isolation, \(v-, w-, c/\text{ts}-, y-\).

103) AGW  \(<m\text{-en}e\text{i}> I \text{ see-CMP ‘I saw’}\)
\(<v\text{-en}e\text{i}> \text{ you see-CMP ‘you saw’}\)
\(<w\text{-en}e\text{i}> \text{ he see-CMP ‘he saw’}\)

5) Join with apostrophe

This option uses an apostrophe to show the contracted form of the pronouns. It has the same advantages and disadvantages as the hyphen, with the two additional disadvantages. First, with the contraction marked with an apostrophe it resembles the glottal stop and in many cases both occur in the same word. Second, the apostrophe use in
established languages of the area such as English represents an orthographic deletion of an underlying phonological unit, and this proposed use does not conform to that convention.

104) AGW     <m’enei> I see-CMP ‘I saw’
            <n’ya’in> I do-CMP ‘I did’
            <v’enei> you see-CMP ‘you saw’
            <vu’ya’in> you do-CMP ‘you did’
            <w’enei> he see-CMP ‘he saw’
            <u’ya’in> he do-CMP ‘he did’

6) Write the surface form

In the examples of surface forms shown in table 3.2, there is one spelling of <banai>, but six spellings of <enei>. In addition, there are five spellings of /da/ (and more possible). If the orthography consistently represents the surface form (phonemic approach), it loses the identity of the morphemes, and it does not treat the morphemes in a unified way, i.e. that the syllabic nasal pronoun would be suffix, word, and prefix in different contexts. With this option, the syntactic treatment of the morphemes is not consistent. The word is not an identifiable conceptual unit, but simply a cluster of letters between spaces.

105) AGW     a) <menei> I see-CMP ‘I saw’
           b) <m banai> I go-CMP ‘I went’
           c) <dam banai> then I go-CMP ‘then I went’

The difficulties can be seen in a sample lexicon comparing the number of entries if the underlying form was written and the number of entries if the surface form was written. In the first part of the sample, the lexicon based on the morpheme-driven approach contained the six pronouns (<m> and <n> were both listed for ‘I’), four verbs, and the completive marker <-i> occur for a total of twelve items. For the corresponding lexicon of surface forms, there were twenty-eight items. The addition of two verbs to the lexicon based on the morpheme-driven approach added two items. The equivalent addition to the
phonemic-form-based lexicon added fourteen items. Fluency in reading is primarily a function of instantaneous word recognition, which comes through seeing the same word many times. Word-attack skills (sounding out words) play a central role in the learning process, but are used by experienced readers only for the infrequent encounters with unfamiliar words. Writing the surface forms means that readers would be required to use the word-attack skills much more frequently, as the number of forms is much greater. Word recognition of a core vocabulary would require a much higher number of learned units with a surface form orthographic approach.

Option 3, in which the pronouns are freestanding and adjust their spelling for easy pronounceability is highly recommended. The NCMs can follow the same rules of writing before vowel-initial verbs as the pronouns. This provides orthographically consistent pronouns and NCMs, and also follows the orthographic treatment of vowel copying processes used elsewhere. It has been in use in Salka and Agwara Kambaris with widespread public acceptance.

Auna Kambari, which does not implement right to left copying, presented a different situation. The recommendation for Auna is to always write the underlying form of the pronoun, as explained in option 1. This has proven effective in Auna Kambari materials.

Morphemes to the right of the root in the Kambari verb phrase

The morphemes to the right of the verb root are shown in table 3.3: -sa ‘iterative’(+1), various transitivizing morphemes(+2),36 -i ‘CMP/imperative’(+3), object pronoun(+4), tā ‘PERF’(+5), de ‘already’(+6).

36 These morphemes are lexically specified as to their position before or after the iterative morpheme. The double-headed arrow in the chart represents the variation of order that occurs. For purposes of reference, the column number will be used to refer to the morpheme regardless of its sequence in a given realization.
Table 3.3: Morphemes to the right of the verb

As with the morphemes to the left of the verb root, there are three basic options:

a) affix everything: <tsulasakaletde> ‘buy-IT-with/for-them-PERF-already’, b) space everything: <tsula sa ka le tā de>, or c) affix certain things and space the rest.

In seeking to prevent long words, option a may be rejected. However, option b would create long strings of one and two letter words with purely grammatical functions, which, while not impossible, feel very choppy. In taking into account MT speaker intuition, <sa> in isolation has no meaning; <ka> in isolation either has no meaning or too many meanings. On the other hand, considering <tsila> ‘buy’ and <tsilasa> ‘buy IT’, both are identifiable in isolation. In a given context, the two words are not interchangeable, and MT speakers would reject the insertion of the wrong form as inappropriate. When comparing the two words, the same MT speakers cannot articulate the difference between them, and may even say they have the same meaning. The same speakers cannot articulate the meaning of /tā/ ‘PERF’. Following the advanced principle that the absence of identification by untrained native speakers is not a determining factor, other linguistic reasons and principles are needed for deciding on either spacing or affixing.
The most relevant factors for Kambari are, a) vowel copying, b) separability, c) ambiguous situations, and d) syntactic consistency.

Factor a) looks at vowel copying from the verb root, which extends to the right through the iterative (+1) and transitivizing (+2) morphemes, but does not affect object pronouns or any morpheme further to the right. If affixed, any morphemes containing non-high vowels not affected by the vowel copying would violate the vowel co-occurrence restrictions discussed in section 2.3.1.3.

106) AGW *<u kəŋale> ‘he catches them’
(a and e are restricted from being part of the same word)

Factor b) separability clearly establishes a space between object pronouns and tə by the insertion of an independent word.

107) AUN: <I so yi tani tə uvarangusu.>
NCM [evil spirit] drink [habitually] him that-one PERF ING-throw-IT
‘It was always throwing HIM to the ground.’

The word tani ‘that-one’ is restricting the pronoun to the person in focus. It can occur in various places in the sentence, and native speakers recognize it as an independent word. Because of this separability, tə is spaced and everything to the right of it as well.

Factor c) seeks to avoid the creation of ambiguous situations because of adjacent optional morphemes that have the same spelling. It is optimally simple to employ spaces to establish meaning where phonemically identical morphemes would be ambiguous if treated the same in regard to independence or affixation. The potential ambiguity is displayed in the examples in table 3.3. In Salka Kambari, a space between the ‘CMP/imperative’ morphemes and the object pronoun disambiguates these strings of morphemes.
108) \(<\text{tanu yi}>\) carry-CMP ‘carried’
\(<\text{tanu yi}>\) ‘carry him’

To illustrate further, I will present another ambiguous situation occurring in both Auna and Agwara Kambari, although with different surface forms for the morphemes. In speech these two phrases are very similar.

109) AUN \(<u\text{tsila ka}>\) ‘he buys [it] for [s.o.]’
\(<u\text{tsila ka}>\) ‘he buys it [that thing that belongs to the Ka-class]’

If both were affixed or both were spaced, ambiguity would result. Thus, the noun class object \(ka\), filling the same slot (+4) as the object pronouns, is spaced and the \(-kv\) ‘instrumental/benefactive’ (+2) is joined to the verb root. This is further confirmed by the vowel co-occurrence rule: the affix \(-kv\) is subject to vowel copying, i.e. \(<u\text{kanaka}>\) ‘he catches with/for’, but the noun class object is not, i.e. \(<u\text{akan ka}>\) ‘he catches it’.

Factor d) deals with syntactic consistency. Applying the ‘act the same, write the same’ principle at the morphemic level, morphemes of the same grammatical category should be written using the same strategy. It was decided earlier that subject pronouns would stand independently. This then is an argument in favor of making the object pronouns freestanding as well. Another factor to be considered is that there is considerable overlap between object pronouns and possessive pronouns. Thus, freestanding subject and object pronouns indicate that possessive pronouns should be freestanding as well. On the basis of these factors, it was concluded that morphemes to the right of the verb root up to and including \(-i\) ‘CMP/imperative’ (+3) would be affixed, and the object markers (+4) and any morphemes to the right would be freestanding.

**Word spacing decisions in the noun phrase**

The Carroll orthography joins all NCMs to content words and treats other functors both as words and affixes, i.e. in one environment the functor will be written independently,
and in a different environment the same functor will be written as an affix. The clusters of letters appearing between spaces often cannot be identified in isolation, e.g. <noton> in isolation is meaningless to a Kambari speaker; in fact it is two morphemes with a phonetic adjustment for the following word <oni> which also is made up of two morphemes. Example 110) shows the morphemes and the corresponding construction in the KLP orthography.

110) /n oni a ni/37 <n oni a ni> with followers NCM his ‘with his followers’

It would be preferable to preserve the syntax and the constant visual image of content words by careful word division practice. As shown below, this is possible in Kambari.

The head noun

The components of the noun phrase were described in section 2.3.2.3 with examples of different types of modifiers and noun classes. The head noun is clearly an independent word by all criteria. The discussion that follows assumes this and focuses on the remaining morphemes of the noun phrase.

The NCM factors

In the noun phrase, the NCMs linking the noun to its modifiers have the same surface form as the subject markers in the verb phrase. In addition to that, modifiers may consist of nouns including proper names, adjectives, demonstratives, numbers, and possessive pronouns. Syntactic association with multiple word classes is a positive criterion for separation. The noun class factor in the language is important for readers and especially

---

37 The underlying phonological form is /n V-toni V ni/ in which the V of the noun prefix copies the features of the /ɑ/ in the root. The V preceding /ni/ becomes a since there are no non-high vowels in /ni/ from which it can copy.
writers to be aware of, and it is easier if the orthographic rules are the same whether these morphemes occur in the verb phrase or in the noun phrase. Thus, the NCMs in the noun phrase are spaced to make consistent with the NCMs that occur in the subject agreement position in the verb phrase. This is an application of the act the same, write the same principle. What sounds like circular reasoning is actually a demonstration of the cyclical nature of developing an orthography network. In a sense, everything affects everything else, and to make one decision is to make all decisions at once.

With NCMs now spaced, it follows that possessive pronouns should also be spaced. Their independence thus parallels that of subject and object pronouns. The main advantage to the three-word noun phrase \{N NCM Mod\} compared to a two-word noun phrase \{N NCM-Mod\} is that the content words (nouns and modifiers) have a constant visual image. In support of this are the two phrases in Agwara Kambari that are disambiguated by consistent use of spaces:

111) \(<u\ singai>\ ‘NCM good’  
     \(<usingai>\ ‘right [not left]’\)

Adjectives

The majority of modifiers are clearly independent words, such as other nouns in a possessive construction. These present no particular orthographic challenges once the decision to space the NCM is made. In the case of adjectives, the situation is somewhat different.

Like many African languages, true adjectives in Kambari are a very limited set, the most common members being ‘red, black, white, big, small, new’. In isolation, a nominal form is used, e.g. /ushili/ ‘red-one’, but when used as a modifier of a particular noun, the /u-/ prefix is replaced with the NCM, e.g. AGW /malala ma shili/ ‘goat NCM red’. The root
form of the adjective /shili/ is not identifiable by untrained speakers in isolation, so the
orthographic quandary exists whether to join the NCM to the adjectival root since lexically it
requires a preposed morpheme <mashili>, or to be consistent with the orthography patterns
of the rest of the language, i.e. to always space NCMs <ma shili>. Systemic consistency and
reduced teaching load indicate treating all NCMs as freestanding, e.g. SAL <usheli u shili>
‘girl NCM red’, allowing the ‘unrecognizable’ <shili>. However, as has been shown with
Kambari publications, increased familiarity with written materials allows readers to easily
identify <shili>, even in isolation. The nominal form is written with the prefix /u-/ , e.g.
SAL <usheli u tu ushili> ‘girl NCM is red-one’.

The specifier
Another word/affix situation to be considered is that of the specifier. Syntactically
the specifier (‘the particular something’) is the same in all three Kambari languages. It is an
optional phrasal suffix (see Marlett 1999:48). The lexical and phonological behavior of the
specifier is one area in which the three co-languages diverge. In Agwara and Auna Kambari
this morpheme is specified as to the noun class of the head noun it modifies. In Salka
Kambari, it is not noun class sensitive. In Auna Kambari the specifier consists of the
consonant of the noun class plus /i/ as ki, mi, tsi, etc. with both A-class nouns and I-class
nouns sharing ‘yi’ and U-class nouns utilizing ‘vi’. In Agwara Kambari the specifier consists
of the noun class consonant plus /a/ as in ka, ma, a, etc. In Salka Kambari this morpheme
has a different shape. It is a glottal initial syllable followed by the same vowel as the final
vowel of the word immediately preceding it.

38 The ‘vi’ is a historical remnant from a time when these U-class nouns were part of the Vu-class. This analysis of
Auna U-class nouns is based on the semantic overlap between them and the Agwara Vu-class nouns (see table 2.7).
Syntactic factors (phrasal affix status) indicate writing the specifier as an independent word in all three languages. Noun class factors support writing the specifier as an independent word in Agwara and Auna Kambari, i.e. for systemic consistency, any grammatical morpheme specified for noun class is written as an independent word. The phonological factors in Agwara and Auna Kambari, i.e. the specifier in these languages is not affected by the phonological properties of the adjacent words, also argues for writing the specifier as an independent word.

Not only is the shape of the specifier in Salka Kambari different, it is not noun class dependent. Therefore, the noun class factor weakens the argument for writing the Salka specifier as an independent word. The specifier in Salka Kambari behaves differently in regard to phonological interaction. Its shape is determined entirely by the final vowel features, including nasalization, of the preceding word. Therefore, phonological factors in Salka Kambari argue for writing it as an affix. Thus, language-specific factors dictated differing spacing decisions regarding the specifier among the three co-languages. In the following examples, I have selected orthographically and semantically identical nouns from the three co-languages to emphasize the variations in the specifiers:

112) AUN <apara yi>, AGW <apara a>, SAL <apara a>
   ‘dishes the’ (the Agwara and the Salka forms have identical pronunciation)
   AUN <agali yi>, AGW <agali a>, SAL <agali i>
   ‘gaps from tooth loss the’

3.3.4.3 How to write the conjunction meaning ‘when’

The underlying form of this word is either /anV/ with an unspecified vowel, or /an/ ‘when’. It is not easy to determine its precise orthographic representation because it would never occur in isolation, and what is more, the following word would trigger vowel elision if in fact the underlying form of the morpheme contains a final vowel. Thus, <an u sumai>,

<anu u sumai>, and <ana u sumai> would all surface as [anusumai] ‘when he ran’. This morpheme is the conjunction in a tail-head linkage in a narrative. Three spellings are equally possible, <ana, anu, an>. This illustrates how, when insufficient linguistic evidence for a decision is available, any of the available options may be chosen provided that the chosen form is consistently written for the morpheme at every occurrence.

The spacing of this morpheme, like that of the preposition, is reinforced by the fact that its scope is always the entire following phrase.

3.3.4.4 How to write /N/ ‘with’

This section deals with the writing of the conjunction which is a syllabic nasal in Agwara and Auna Kambari and is variously glossed ‘with’, ‘and’, or ‘in’. Before words beginning with consonants, the syllabic nasal is homorganic. The orthographic form could therefore be either <m> or <n> since both /m, n/ exist as phonemes.

The Carroll orthography evidences an inconsistent treatment of this morpheme. An examination of four publications reveals that the morpheme ‘with’ is written in the Carroll orthography as a prefix before words beginning with a vowel, as an independent word if the following word begins with a consonant, and with specific morphemes as a suffix. The orthographic representations in the following examples are in the Carroll orthography.

113) /N usana/ [nü.sä.nä] <nusana> ‘with morning’

114) /N kula/ [nü.kü.lä] <n kula> ‘with name’

115) /koôolo N vangu ni/ [köô.öô.lön vänqüini] <köôolun vangun ni> together with younger-brother his ‘together with his brother’

39 However, Kambari orthography uses <m> before bilabials and <n> before both alveolar and velar consonants as there is no velar nasal phoneme in phoneme inventory.
It would be better if the orthography preserves the integrity of the content words, i.e. each content word if pulled out of a sentence and looked at singly would be recognizable as a word, i.e. a constant visual image. The Carroll orthography is based on context-specific phonetics as in the orthographic `<a>` at the end of `<notona>` to mark the elision between the final /i/ of /otoni/ and the following NCM [ə], or the `<n>` at the end of `<vangun>` that represents a morphophonemic blending of /vangu ni/. An orthography based on the underlying syntax and morphology with concessions in the functors to the phonetic surface form is more consistent. The orthographic form of content words is consistent, based on the form in isolation. The orthographic form of functors is allowed a limited systematic variation to approximate phonetic realization.

The intriguing orthography question is in the area of writing this conjunction before words beginning with vowels. Phonetically the syllabic nasal loses its syllabicity and functions as the onset of the following syllable in the same way the pronoun 'I' did before vowel-initial verbs in the previous section. The following examples are identical in the three co-languages.

116) /n usana/ [nû.sà.nà] ‘with morning’
   /n acimbi/ [nà.cím.bi] ‘with farmers’

There are five options. The one option rejected out of hand is truncating the content word:
*<nu sana> ‘with morning’, because *<sana> is meaningless in isolation.

1) *Just write /N/*

With this option, the conjunction would always be written as a freestanding `<n>` or `<m>`. 
2) Join to noun

With this option, the conjunction would be a prefix of the following word. This option doubles the constant visual image of the inventory of nouns, each having a form in isolation and a prefixed form.

3) Join with hyphen

With this option, a hyphen would be inserted between the conjunction and the noun.

4) Join with apostrophe

With this option, an apostrophe would be inserted between the conjunction and the noun. The hyphen and apostrophe options have the same advantages and disadvantages as were discussed earlier in the sections on NCMs and subject pronouns.

5) Adjust the spelling of /N/

This option is to always write the conjunction freestanding, but before words beginning with vowels, copy the initial vowel, and add it to the nasal consonant to form a syllable. This is the same strategy as option 3 of the subject pronouns.
121) *<nu usana>* ‘with morning’
   *<ni ikebe>* ‘with money’
   *<na acimbi>* ‘with farmers’

Since these options roughly parallel the options concerning subject pronouns, to
maintain a consistency of pattern application throughout the language, the same strategy
would be used in both environments be it space, hyphen, or whatever. Arguments for the
subject pronoun lead one to chose spacing the pronoun, so then consistency dictates the
‘with’ be spaced also, selecting option 1. Salka Kambari does not have quite the same
situation since the corresponding conjunction is not a syllabic nasal, but underlyingly /na/.
In Salka Kambari, before words beginning with vowels, this conjunction is written the same
as in option 5: *<nu usana>* , which reflects the presence of the underlying vowel and the
phonological interaction.40

3.3.4.5 Phonetically unperceived morphemes

This section deals with the writing of grammatical morphemes, which are present
syntactically but are unperceived by MT speakers due to the proximity of identical
phonemes. It is only in certain phonetic environments that these morphemes are
inaudible/unperceived. In other syntactically identical but phonetically different contexts,
they are audible and obligatory. The point in question is whether or not to write them in the
contexts in which they are silent.

---

40 Option 5 *<nu usana>* ‘with morning’ was originally chosen by Auna and Agwara Kambari. In 1996, Auna
Kambari readers voiced a preference to write ‘with’ simply as a freestanding *<n>* or *<m>* (option 1 *<n usana>*).
In 1997, Agwara Kambari followed the lead of Auna Kambari and began writing the conjunction as a freestanding
*<m>* or *<n>*. In 1999, after discussion with the Auna Kambari literacy teachers, the consensus was to drop the
*<m>* and write only *<n>* , even before words beginning with labial consonants. This is an important historical
element demonstrating that, with increased fluency, readers/writers prefer a morpheme-driven approach.
The orthographic question to be solved is whether to write example b) as <ili i na> or as <ili na>. Preserving the syntactic structure in the orthography by writing grammatical morphemes in all environments is an application of consistency at the level of syntax. This type of consistency forces a violation of phonetic consistency. The opposite choice would, obviously, violate syntactic consistency. The semantic information encoded in these structures is critical, and should be retrievable by the reader. This factor argues conclusively for writing such grammatical morphemes in all environments regardless of their audible properties. Such a decision leads to a commitment to teach grammatical constructions as part of literacy training, since MT speakers who learn to ‘write by ear’ would never represent these inaudible morphemes.

3.3.5 Kambari power syllables

In an idealized situation, there would be a unique one-to-one correlation between phonetic shape, written representation, and meaning. Since linguistic reality often violates this idealization, orthography decisions are not simple. The extent to which one phonetic shape has more than one meaning, or that one meaning has more than one phonetic shape, is an area in which a language may be complex. The English homophone example cited earlier in section 3.1 exhibited one phonetic shape, [tuː], with three meanings and three written representations. English also has words with one meaning spelled different ways, e.g. ‘a’ and ‘an’ are both the indefinite article and are used predictably based on the consonant or
vowel at the beginning of the following word: *a car, an automobile*. In this example, a single morpheme has two phonetic shapes and two written representations.

### 3.3.5.1 Overview of power syllables in Kambari

The area of complexity in Kambari is that a given grammatical word usually has more than one phonetic representation, creating multiple orthographic forms for the same morpheme. Taken individually, any one of these orthographic forms may represent multiple morphemes. For example: The Kambari functors *<e>* and *<o>*\(^{41}\) may have the same meaning (allomorphs) in different contexts, or the functors *<e>* and *<o>* may mean ‘they’ and ‘to/at’, respectively in one sentence and ‘to/at’, and ‘they’, respectively, in the next. In this section, I am going to ignore the few functors that have a one-to-one phonetic and semantic correlation and focus on the ones with the most overlap. I have called the latter group of functors ‘power syllables’; ‘power’ because they encode a wide range of semantic and syntactic information, and ‘syllables’ because the set is comprised of monosyllabic morphemes, some of which are affixes while others are independent words. They are the twelve most frequently used syllables in the language, occurring in every sentence in the language (except simple imperatives). The number (twelve) refers to an underlying phonological form regardless of meaning or phonetic permutation. The power syllables are these shown in (123).

123) *N* (syllabic nasal consonant), *ma, vu, va, u, wa, tsu, tsa, i, ya, and a*

Auna and Agwara Kambari have the power syllable *ka*, which appears in Salka Kambari as a double vowel. In the above syllables containing *a*, the *a* in Agwara and Salka Kambari is in fact an underspecified non-high vowel which copies its features from the following word. Agwara and Salka allow vowel processes to move both to the right and to

---

\(^{41}\) *e* and *o* are copied forms, and only occur in Salka and Agwara Kambari.
the left. Therefore, for example, the power syllable $mV$, written $<ma>$ in Auna Kambari may be written as $<ma, m, me, mi, mo, mu>$ in the other two. These syllables accommodate both the vowel copying and the replication of the vowel in the NCM occurring before vowel-initial words, e.g. /m enei/ $\rightarrow$ $<me enei>$ ‘I see-CMP’. The semantic and syntactic components of these power syllables are virtually identical across the three co-languages. Semantically these power syllables are the underlying forms of the basic and most common subject pronouns. They are also the noun class prefixes and NCMs for ninety per cent of the nouns. In addition, they have many other functor roles including affixes.

3.3.5.2 The example of the power syllables /i/ and /ya/

Table 3.4 illustrates overlapping relationships between meaning and form. For instance, a single meaning can have up to seven forms as the second-person plural subject pronoun does in Agwara and Salka Kambari. A single phonemic unit can carry seven or more distinct meanings.
<table>
<thead>
<tr>
<th>Power syllable meaning</th>
<th>Auna</th>
<th>Agwara</th>
<th>Salka</th>
</tr>
</thead>
<tbody>
<tr>
<td>Second-person plural subject pronoun</td>
<td>i</td>
<td>i, ya, ye, yi, yo, yu</td>
<td>i, ya, ye, yi, yo, yu</td>
</tr>
<tr>
<td>Noun class plural prefix</td>
<td>i-</td>
<td>i-</td>
<td>i-, yi-</td>
</tr>
<tr>
<td>NCM</td>
<td>i</td>
<td>i, ya, ye, yi, yo, yu</td>
<td>i, ya, ye, yi, yo, yu</td>
</tr>
<tr>
<td>NCM without overt noun (its noun, ili ‘things’ is understood)</td>
<td>i</td>
<td>i, ya, ye, yi, yo, yu</td>
<td>i, ya, ye, yi, yo, yu</td>
</tr>
<tr>
<td>Non-completed second-person plural subject pronoun (NCMP) (underlyingly i + a ‘2P pronoun + NCMP’)</td>
<td>ya</td>
<td>ya, ye, yo</td>
<td>yaa, yee, yoo (more on long vowels later)</td>
</tr>
<tr>
<td>Copula (Agwara only)</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completive action suffix (CMP)</td>
<td>-i</td>
<td>-i</td>
<td>-i, -yi</td>
</tr>
<tr>
<td>Plural imperative suffix</td>
<td>-i</td>
<td>-i</td>
<td>-i, -yi</td>
</tr>
<tr>
<td>Simultaneous action suffix (Auna and Salka, replacing the final vowel of the verb stem)</td>
<td>-i</td>
<td>-i</td>
<td>-i</td>
</tr>
</tbody>
</table>

Table 3.4: Kambari power syllables

In addition to these forms, there are language-specific morphemes that, while having no phonologically conditioned variants themselves, share the surface form of one of the variants listed above. For example, yi in all three languages is also one form of third-person singular object pronoun. In Auna Kambari, yi is also the specifier for A-class and I-class nouns. In Agwara Kambari, yi is also the question word ‘what’, and the compound word ‘you-PL are’ (underlyingly i ‘you-PL’ + i ‘is’). In Salka Kambari, yi is also the third-person possessive pronoun. Ya in Agwara and Auna Kambari can also mean ‘who’. In Agwara Kambari ya is the specifier for I-class nouns. In Salka Kambari, yg is also the preposition ‘to’ when the goal is a person.

---

42 In Salka Kambari the third-person singular object pronoun is always yi. In Auna Kambari, both yi and ni exist as third-person singular object pronoun with ni appearing as the object of a command: Nkki ni. ‘(you-PL) give him [it]’; U nkki yi ‘He give-CMP him [it]’. In Agwara Kambari both yi and ni exist as third-person singular object pronoun. The use of the proper form depends on the verbal aspect: Wi ty e kuneke yi ‘He-is at ING-give him [it]’; U nekei ni. ‘He give-CMP him [it]’. 
124) *U tanu ta nron yg Lumamu.*

he carry PERF PL-yam to Lumamu
‘He has carried yams to Lumamu.’

This *yg* also has variant forms if the following word begins with a vowel. *Yo* in Salka Kambari also means ‘snake’. It is no surprise that each of the other power syllables also has a long list of functor roles.

**3.3.5.3 Phonologically induced variants**

The variants are accounted for by two phonological processes. First is vowel copying to an underspecified lexical vowel. Second, when the power syllable *u* or *i* appears before a vowel-initial word, it becomes the corresponding semivowel. For Kambari, the preferred syllable structure is CV. Converting high vowels (*i* and *u*) to their respective semivowels (*y*, *w*) is one strategy for reducing strings of vowels.

Underlyingly, the NP ‘farmers’ money’ is

125) /ikebe i acimbi/

| VCVCV | V | VCVCV |

By converting the NCM to a semivowel, the CV structure is maintained:

126) *ikebe y acimbi

| VCVCV | C | VCVCV |

However, <y> as a consonant cannot appear alone as a word. It is unpronounceable in isolation and is therefore unacceptable as a freestanding word. A solution to this problem is to allow the NCM to be written as <y> plus the same vowel as that beginning the following word. The addition of an orthographic vowel to the NCM would aid to easy reading, since reading does not require a pause at spaces.

127) <ikebe ya acimbi>

| VCVCV | CV | VCVCV |
Consistency dictates this strategy should be generalized to the NCM of U-class nouns as well as to pronouns before verbs beginning with a vowel: <we enei> ‘he see-CMP’, <ye enei> ‘you-PL see-CMP’.

In Agwara Kambari, the distinction between the pronouns which are underlyingly /i + a/ ‘you-PL + NCM’ and /i/ ‘you-PL’ forms, both of which in certain contexts can surface as <ya, ye,> etc. is preserved by the grammatical constraint by which the non-completed forms must occur with verbs prefixed by ku- ‘progressive aspect’.

128) /i ene tâ / → <ye ene tâ> ‘you-PL have seen’
/i + a ku-ene/ → <ye kene> ‘you are seeing’

In Salka the distinction is preserved through the use of a double vowel for the non-completed forms, which will be discussed later.

**3.3.5.4 Syntactic role of power syllables**

The syntactic factor is very important in discerning the meaning of the power syllables. For example, the syntax for the two sentences: ‘The farmers went to the widows’ farms,’ and ‘The boys returned to the herbalists’ alcoves,’ is the same. The power syllables are marked with capital letters:

PL-farmer the 3P went to PL-farm NCM PL-widow  
‘The farmers went to the widows’ farm.’

O-lobo A O bonoi O O-bogozu O O-boci.  
PL-boy the 3P returned to PL-alcoves NCM PL-herbalist  
‘The boys returned to the herbalists’ alcoves.’

Thus in formulating the orthographies, it was essential to pay strict attention to a functor’s syntactic role and preserve the meaning distinctions via consistent spacing rules. Even with that, it is impossible to generate a strategy that visually differentiates each
meaning without moving into an English-like disregard for the sound-meaning correspondence.

Salka Kambari has four functors which may appear following a verb stem as <yi>, i.e. the completive aspect marker, the plural imperative marker, the third-person singular object pronoun, and a directional preposition ‘to’ if the following word begins with /i/. Thus the morpheme string /tanu-yi/ could have four possible interpretations. The preposition, not being part of the verb phrase, stands independently. That leaves the first three meanings, all of which are part of the verb phrase. If all of them were affixes, or if all were independent, there would be ambiguity for the reader. However, if the pronoun stands independently, and the other two are mutually exclusive affixes whose meanings are easily discerned from the context, the level of ambiguity is minimized. In this way, there are two orthographic representations which have contextual clues as to meaning.

130) <tanuyi>
   ‘(you-PL) carry [imperative]’ and ‘carry-CMP’
   <tanu yi> ‘carry him’ and ‘carry to [followed by person’s name beginning with /i/]’

This example serves to illustrate the use of orthographic conventions to assist the reader in arriving at the desired meaning, while allowing a degree of ambiguity for the sake of orthographic simplicity. These factors affected the orthographic decisions of all power syllables as their behavior in specific contexts is similar to that of yi in Salka.

3.3.5.5 The example of the power syllables before verbs in Salka Kambari

The most complex example of differentiating power syllables in a given environment is that of the set of functors in Salka Kambari whose underlying form is either a short or long non-high vowel immediately preceding a verb. There are nine such morphemes. It is theoretically possible to formulate a system in which all nine functors have unique
representations incorporating non-pronounced graphemes that are not in use or infrequently used elsewhere in the writing system, e.g. (<x>, <h>) as the last letter of the functor. Other orthographic devices, such as a functor-initial glottal stop, the use capitalization, diacritics, or even a system of superscript numbers corresponding to the semantic role of each functor, could provide orthographic distinction. The principles of convenience, teaching load, acceptability, advantages and disadvantages, optimal simplicity, and seeking what will be easiest for users, all work together to eliminate the above options without serious consideration. A system entailing nine unique forms would be too cumbersome to be practical. However, by grouping meanings and maintaining strict spacing and double vowel rules, the Salka Kambari orthography can be a maximally practical and minimally ambiguous system.

**Functor 1: 3P PRO (Set 1)**

Functor 1 is the third-person plural subject pronoun. This pronoun is a member of pronoun set 1. The orthographic treatment of this morpheme was detailed in the discussion of the verb phrase.

**Functor 2: an NCM (Set 1)**

Functor 2 is the NCM for the noun class whose prefix is a short non-high vowel (plural). Contrast between this NCM and the third person-plural pronoun is neutralized in the subject position. In other parallel instances, the distinction between NCM and pronoun is realized by separate morphemes. For example, in the subject marker position, the noun <mogono> ‘king’ may be followed by either <mV> ‘NCM’ or <u> ‘third-person singular pronoun’.
In the case of \( V_{\text{[hi]}}, \) class nouns, the pronoun and the NCM are phonologically identical.

The following could be either the pronoun or the NCM. The distinction between subject pronouns and NCMs is a discourse level feature, concerning the importance of the character and the relation of the action to the matrix events, pronouns being more prominent. If the distinction between NCM and pronoun were critical, the discourse features in the context would disambiguate it.

**Functor 3: an NCM (Set 1)**

Functor 3 is the NCM for the noun class \( V_{\text{[hi]}} \) whose singular prefix is a long non-high vowel. If singular NCMs of the class \( V_{\text{[hi]}} \) are written with a single vowel, the plural/singular distinction reflected by other NCM sets is not captured. However, NCMs only occur in reference to a particular noun. That noun (the antecedent) is always specified for number (singular/plural).

Ambiguity caused by underspecifying length in the NCM would only occur if a singular and a plural noun belonging to this class were referred to only by their NCMs. Already inherent in the language is the potential ambiguity of referring to two nouns of the
same number and class only by their NCMs. Thus, ambiguity caused by underdifferentiating vowel length in NCMs hardly occurs and is easily resolved by restating the noun. For these reasons, the ambiguities resulting from neutralizing the orthographic distinctions between functors 1–3 are seen as minimal. Furthermore, creating an orthographic distinction for these NCMs would greatly increase the teaching load, and would create more orthographic problems when dealing with functors 4 to 6 of this group. This functor is also written with a single non-high vowel spaced from the verb <a>.

**Functor 4: 3P PRO (Set 2)**

Functor 4 is the 3P plural subject pronoun of pronoun set 2 which is a combination of two morphemes, ‘3P pronoun + NCMP’ is underlyingly VV[-hi]. Figure 3.2 shows the relationship between pronoun sets 1 and 2.

**Salka Pronoun Morphology**

![Diagram of Salka Pronoun Morphology](image)

**Figure 3.2: Relation between pronoun sets 1 and 2**

In careful speech, if the functor 4 is sentence initial, and if the verb begins with a consonant, the pronoun is discernibly long. However, if the following verb begins with a vowel, and the pronoun, now having copied its features from the verb, follows a word ending with the same vowel, discernment of length is nigh impossible. Thus, a Salka Kambari writer cannot depend on sounding out the phrase to determine if the pronoun he
wants to write is functor 1 or functor 4. Thus, identification of these morphemes depends on syntactic and contextual factors.

*Functor 5: an NCM (Set 2)*

Functor 5 is the set 2 form of functor 2.

*Functor 6: an NCM (Set 2)*

Functor 6 is the set 2 form of functor 3.

Functors 4 to 6 undergo the same patterns of neutralization discussed for the first three functors. Functors 4 to 6 include the non-completive morpheme which itself is $V_{\text{Hi}}$. If functors 2 (\texttt{a}) and 3 (\texttt{aa}) were distinguished using a double vowel for functor 3, then functor 3 and functor 5 would both be written using double vowels \texttt{aa}, with the important aspectual meaning distinction neutralized. In addition, if functors 3 and 5 were distinguished by double vowels \texttt{aa}, functor 6 would be distinguished with the use of a triple vowel \texttt{aaa} or some other mechanism that was discussed earlier. A word consisting of three consecutive identical vowels was deemed unacceptable. Therefore, following the pattern of functors 1, 2, and 3, orthographically functors 4, 5, and 6 are written as a double non-high vowel that is spaced from the verb \texttt{aa}.

*Functor 7: simultaneous action*

This morpheme, underlyingly a non-high vowel, denotes simultaneous action. This use may be another sense of the preposition ‘at’. This morpheme occurs before nouns.

134) \texttt{ele a atsuma a mmalu}  
‘they while in the middle (lit. ‘stomach’) of walking’

It occurs before verbal nouns that are clearly nouns.

135) \texttt{ele a nlangi}  
‘they while fighting’
It also occurs before prefixed verbs\(^{43}\) (in (136)) or before verbs that probably belong in the previous category, but whose prefix is a null marker as in (137).

136) \(<\text{ele e me-bece}>\)
   ‘they while ING-search’

137) \(<\text{ele a } 'yəwə}>\)
   ‘they while going’

Because of the variety of word classes that can follow this morpheme it is best to make it an independent word. Thus, functors 1, 2, 3, and 7 are all written as a single non-high vowel spaced from the following word. In isolation \(a 'yəwə\) could mean either ‘they go’ or ‘while going’, but testing has shown that in context, there is no confusion between the two meanings.

**Functor 8 : a nominializer**

This morpheme, underlyingly a non-high vowel, is a derivational prefix. It is lexically restricted, only occurring before certain verbs. \(Lyuna\) is the verb ‘herd’. In the following example, \(a-lyuna\) is a noun.

138) \(<\text{ Alyuna a da a yuwusan.} >\)
    herding it is they do-IT  ‘They go herding every day.’

As it creates a noun that generates NCMs, this functor behaves as any other noun class prefix, and therefore is subject to vowel copying. In certain contexts, it is difficult to determine if functor 8 is present or not. For example, the phonetic string \([wənəyɨtəlyûnə]\) ‘He saw him herding,’ could be interpreted three ways depending on the underlying morphemes which comprise the [alyuna] ~ [aalyuna] segment.

\(^{43}\) Hyphens in these examples are not part of the orthography, but are used to isolate the prefixes under discussion.
139) a) <We ene yi tã a alyuna.>
   /a a-lyuna/
   (at-functor 7 herding-functor 8)

140) b) <We ene yi tã alyuna.>
   /alyuna/
   (herding-functor 8)

141) c) <We ene yi tã a lyuna.>
   /a lyuna/
   (at-functor 7 null prefix-herd)

In deciding how to write this phonetic string, option b is rejected because this construction is ambiguous with the construction created with functor 9 ‘about to-herd’. Both options a and c are orthographically distinct from the construction using functor 9. There is conceptually a slight semantic difference between options a and c. Option a represents the occupation: ‘he saw him doing the occupation of herding’. Option c represents the activity: ‘he saw him doing herding’. Both options a and c are viable orthographically.

*Functor 9: an inflectional morpheme*

Functor 9 is an inflectional morpheme denoting impending action.

142) We ene tã kaka a-karga o una.
   ‘He saw grandmother about to-go to the bush.’

All verbs can accept this morpheme. Orthographically, it is treated as a prefix. The vowel of the prefix undergoes vowel copying. In the case of vowel-initial verbs, the initial vowel of the verb is doubled, e.g. <ene> ‘see’, <eene> ‘about to see’.

Thus, the nine functors are written in the language using three strategies, a single vowel spaced from the verb <a lyuna> ‘(functors 1, 2, 3, and 7) herd’, a double vowel spaced from the verb <aa lyuna> ‘(functors 4, 5, and 6) herd’, and a single vowel prefix <alyuna> ‘(functors 8 and 9) herd’. All nine are subject to vowel copying. The three
strategies proposed here provide adequate orthographic distinctiveness among the nine morphemes to satisfy the primary function of orthography: to convey meaning.

### 3.3.5.6 Other uses of VV in Salka Kambari

The final example included in this chapter illustrates the use of an orthographic convention to distinguish set membership or some feature of the language that is too important to be ignored but too infrequent for readers to easily assimilate. The orthographic solution in the Salka Kambari example below is analogous to a decision to write tone on every verb rather than only on verbs with a certain feature or on a list of verbs to be memorized. A decision of this nature is predicated on the principle that marking every occurrence of every member of a set of grammatical elements raises the users’ awareness of that grammatical structure. Where a particular member of the set reflecting critical grammatical information occurs infrequently, the user is accustomed to processing orthographic indicators for that set of features.

By far the most difficult area for native speakers of Salka Kambari to analyze and write correctly is the set 2 third-person plural subject pronoun or NCM where double vowels indicate non-completed, often future, action (functors 4–6 discussed above). For the other set 2 subject pronouns, this distinction is marked by the inclusion of a vowel-copying sensitive low vowel: Set 2: [ma, va, wa, tsa, ya], instead of set 1: [n/m, vu, u, tsu, i] ‘I, you, he, we, you-PL’ (see figure 3.2).

However, pronouns before vowel-initial verbs create another level of complexity for orthography decisions. The option of always writing the underlying form of set 1 pronouns before vowel-initial verbs was rejected by MT speakers as too abstract, therefore violating the acceptability principle. Thus the accommodating strategy employed for set 1 pronouns <me ene, ve ene, we ene> ‘I see, you see, he see’ was developed. In careful speech, the
morpheme string of a set 2 pronoun and vowel-initial verb (\(<\text{wee ene}>\ '3S\ set 2 see'\)) has greater length than the string of the set 1 pronoun and the same verb (\(<\text{we ene}>\ '3S\ set 1 see'\)). This allows utilization of the functor 4–6 strategy, doubling the vowel for set 2 pronouns when occurring before vowels. However, the orthographic forms of set 1 before vowel-initial words duplicated the orthographic forms of set 2 before consonant-initial words, i.e., \(<\text{we}>\ '3S'\) was a member of both pronoun sets: \(<\text{we ene}>\ '3S\ set 1 see', \(<\text{we kene}>\ '3S\ Set 2 fetch'\). The set 2 pronouns thus had \(<\text{we}>\) before consonant-initial verbs and \(<\text{wee}>\) before vowel-initial verbs. Pronoun set 1 would have seven allomorphs for the 3S pronoun. Pronoun set 2 would have eight allomorphs for the 3S pronoun, of which four were identical with those of set 1. Practically, there is substantial teaching load in making readers and writers aware of the contexts in which double vowels were written so that these structures are used consistently.

At this point in the orthography development, the orthography could accurately convey meaning. This was achieved by the complex decoding strategy of recognizing that \(<\text{we}>\) is set 1 before vowels, and \(<\text{we}>\) is set 2 before consonants. The system proved functional everywhere except the two problematic areas of 3P of set 2 pronouns and all set 2 pronouns occurring before vowel-initial verbs. In these instances, MT authors consistently failed to appropriately indicate the distinction between set 1 and set 2 pronouns. In order to maintain accurate meaning in written material, an efficient way of enabling MT authors to correctly utilize set 2 pronouns was needed. Of importance to orthography considerations is that once a commitment to the teaching of set 2 recognition was made, there was no longer a need for the orthographic ambiguity of overlapping members of pronoun sets 1 and 2. As an orthographic convention, all set 2 pronouns are written with double vowels, reducing the allomorphs of each pronoun in set 2 to four. Set 1 remains with the seven forms, but none of them overlap with the orthographic forms of set 2.
This avoids the difficult task of teaching which form of the pronoun to write based on sound and phonetic timing. Instead, writers are encouraged to make the direct connection between the concept and the orthographic form.

The orthographic convention of writing the set 2 pronouns and NCMs with double vowels is an example of the consistent use of an orthographic convention to represent a semantic concept. The fact that every 3P pronoun and any pronoun before vowel-initial verbs must be analyzed made the frequency of the analysis process quite high. When every member of the pronoun set is written with a double vowel regardless of its phonetic quality, there is no overlap in the pronominal spellings of the two classes.

3.4 Conclusion

At the time of this writing, the orthographies of Salka, Auna, and Agwara Kambari had been in public use for six years, and more than 20,000 people had been through literacy classes based on the orthography. Over 100 different titles have been published in each of the three languages using this orthography. The first drafts of dictionaries have been produced and distributed in limited quantities for public feedback prior to the full printing of a dictionary for each language. Materials written in the orthography, such as choruses for church use, personal letters, and stories submitted to the Kambari Language Project for publication are a common sight. This widespread and growing acceptance is the final test of an orthography. Nonetheless, there are remaining orthography decisions, and this will continue throughout the life of the written form of the language. Orthography standardization never completely stops. However, the decisions that remain will be
comprised of adjustments in the spelling of words and other small changes. These minor refinements will come as the result of widespread, long-term mother-tongue literacy and are the ultimate statement of success, as they indicate continuing interest and on-going mother-tongue involvement in language development.
4 Insights on Orthography Principles

4.1 Advanced principles

These principles set forth here are ‘advanced’ in the sense that they are primarily useful in dealing with the problem areas that remain after the low complexity phoneme-grapheme decisions have been completed. The basic principles presented in chapter 1 will only take orthography development to a solid starting point, i.e. to the point of codification of content words in isolation, but they leave many unanswered questions. These advanced principles are meant to provide decision-making guidelines in the difficult areas.

The idea of a core vocabulary is important to many of the advanced principles of orthography. In spite of the thirty thousand or more words in a language’s lexicon, in normal communication a small number of words are used over and over. One of the skills of experienced readers is that of instant recognition of familiar words by their word profile. A word profile is an imaginary outline delineating the tall, rounded, and tailed letters that make up a word. Adams (1990:18) states,

...skillful readers seem to recognize familiar words as wholes. At the same time, skillful readers visually process virtually every individual letter of every word as they read, and this is true whether they are reading isolated words or meaningful connected text. To be sure, readers do not necessarily notice misspellings and misprints at a conscious level. But, conscious or not, studies show that even the slightest missprint, tucked deep within a long and highly predictable word, is often detected by the visual systems of skillful readers.

This detection is based on a changed word profile as in the examples with the and its misprint teh:
I did a core vocabulary study of 105 short stories (35 in each of the three Kambari co-languages). Anything preceded and followed by a space or punctuation counts as a *token*, a term used in corpus linguistics to describe the total word count of a body of texts. The corpora from the three co-languages averaged 5,000 tokens. Another factor to take into account is the number of unique *words*, i.e. *dana* ‘say’ counts as 1 *word*, but 38 tokens (occurrences), and *danai* ‘say-CMP’ also counts as 1 *word* with 76 tokens in the Agwara Kambari corpus. In number, the words in the three corpora range from 648 (Agwara Kambari) to 803 (Salka Kambari). For this analysis, I considered core vocabulary to be any word that occurred more than once. There were 395 words in the core vocabulary (61% of total words) in the Agwara Kambari corpus, 411 (54%) in the Auna Kambari corpus, and 434 (54%) in the Salka Kambari corpus. These accounted for between 92% and 94% of the tokens, i.e. approximately 94% of the total corpus is processed automatically by experienced readers, and only 6% to 8% of the words need decoding skills such as sounding out syllable by syllable. Extrapolating from this, it would seem that the core vocabulary in the orthographies of other languages is 500 words or less. Therefore, the reader needs to learn to recognize only 500 words by their constant visual image to be able to process 94% of any page of text. This small inventory of stable word forms makes fluent reading easy to achieve.

I did two further checks with the corpora. First, I compared the recommended orthography for the Agwara Kambari corpus to a morphemically regularized orthography in which the grammar words were represented with a single spelling, e.g. all occurrences of the NCM *u* are written `<u>`, etc. Grammar words account for 51% of the total number of tokens.
in both formats. I discovered that even though the NCM u has six allographs in the recommended orthography: <wa, wä, we, wi, wo, wu>, all of the allographs except <we> occurred in identical form representing other morphemes in the morphemically regular presentation, e.g. <wa> ‘U-class specifier’, <wu> ‘2S object pronoun’. Thus the morphemically regular orthography did not significantly reduce the number of grammar word forms; it merely redistributed the frequency of the tokens.

The second comparison I did with the recommended orthography in the Agwara Kambari corpus was to compare it to an orthography in which any grammatical morpheme that undergoes phonological processes based on the following word was made a prefix, e.g. <udanai> ‘he said’ (48 tokens), <adanai> ‘they said’ (17 tokens), <vudanai> ‘you said’ (3), <kadanai> ‘NCM said’ (2), <madanai> ‘NCM said’ (2), <nudanai> ‘and he said’ (2), <nadanai> ‘and they said’ (1), <ndanai> ‘I said’ (1). Obviously, this system uses far more words: 1,055 compared to 648 in the recommended orthography. Thus a beginning reader must be exposed to a much larger quantity of written materials before the recognition of such things as <nadanai> ‘and they said’ becomes recognizable via its constant visual image.

Simons (1994) states that in comparing orthography options, the one that provides the easiest path to fluent reading and writing is to be preferred (see section 1.3). This evaluation measure supports the recommended orthography over one in which affixation decisions are based on phonological processes.

4.1.1 The convey meaning principle

The overarching goal of an orthography is to convey meaning. I have had personal experience with a number of developing orthographies. The basic principle of one phoneme/one grapheme is sometimes adhered to to the point that one would conclude that
the purpose of orthography is to represent sounds rather than to convey meaning. Perhaps those developers perceived that the only way of conveying meaning is through the exact representation of sounds. The overt expression of the convey meaning principle at the beginning of orthography development and its frequent reiteration throughout the process will help ensure that the most functional kind of orthography will evolve.

For an orthography to be accurate, it must convey the meaning the author intended. If the readers miss the meaning, assign the wrong meaning, are confused by two possible meanings, or struggle through repeated rereadings to get the meaning, the accuracy of the orthography needs to be addressed. One of the chief ways to convey meaning in an orthography is to make decisions that reduce ambiguity.

An example of conveying meaning in the Kambari orthographies is the spacing of the subject markers in the verb phrase to avoid ambiguity with existing nouns (see section 3.3.4.2). In this way Agwara Kambari readers instantly know that <mabana> ‘basket’ is a noun, and <ma bana> ‘NCM go’ is a verb phrase.

### 4.1.2 The consistency principle

The advanced principle of consistency is that once a system is established, that system is invariably applied, with the result that each written occurrence of the same expression is always written the same way. This is the unstated foundational principle of all orthography development. Any system, regardless of its internal merits or lack thereof can be learned and utilized by readers and writers if it is applied in a uniform manner. The essential thing is that people will know that structure X indicates meaning Y each and every time they see or write structure X.

This advanced principle allows for varying degrees of internal inconsistency within the parameters of ‘consistent’ as it is variously defined. It goes beyond the definition of
consistency that equates one phoneme to one grapheme. It also goes beyond the definition of consistency that requires a single orthographic form for each morpheme. Natural language structures are not in and of themselves consistent enough to allow any of the more simplistic definitions of consistency to function in isolation as an orthographic decision parameter. An orthography that is phoneme-driven seeks to be consistent in the representation of sounds. This often results in the inconsistent representation of morphemes. An orthography that is morpheme-driven seeks to preserve the morphemes in an identifiable form but in doing so the letters on the page do not always reflect the sounds as they are pronounced in that particular environment, especially at morpheme boundaries. ‘Complete consistency’ is a self-defeating concept since absolute consistency at one level conflicts with absolute consistency at the other levels.

In the Salka Kambari orthography, the writing of double vowels for all the set 2 pronouns and NCMs (see section 3.3.5.6) is an example of orthographic consistency in which the decision to consistently represent all the members of a grammatical set with the same orthographic convention resulted in inconsistent representation of the aural vowel length in the individual members of the set. The 3S pronouns in context are phonetically [we ene] ‘3S set 1 see’, [we kene] ‘3S set 2 fetch’, and [wee ene] ‘3S set 2 see’. Orthographically these are now <we ene>, <wee kene>, <wee ene>, respectively.

To write <we/wee> for ‘3S set 2’ depending on the phonetic environment would be inconsistently representing the semantic value of set 2 pronouns. To write <wee> for every occurrence of ‘3S set 2’ is to violate the consistent application of sound/symbol representation. The orthography must be inconsistent one way or the other, but once a convention is selected and utilized, the orthography is consistent under the advanced consistency principle.
4.1.3 The optimal simplicity principle

The advanced principle of optimal simplicity is that the orthography represents only what is needed to convey meaning. Features that are not necessary to convey meaning are not written. In the case of redundant features, the feature that is the easiest to write is the only one to be included. This principle works together with the convey meaning principle to ensure an orthography that is both easy to use and adequate for communication.

To an Agwara Kambari reader <daka> and <dakkà> both convey the same meaning, ‘down’ (see section 3.3.2.2). Since both forms accurately convey meaning, <daka> is preferred as the simpler form. Once a MT reader has grasped the meaning, he automatically pronounces the word correctly.

4.1.4 The limited ambiguity principle

The advanced principle of limited ambiguity is that a certain amount of ambiguity in the orthography is tolerable. Of course, if ambiguity frequently inhibits accurate interpretation of written materials, corrective action is required. However, occasional ambiguity can be tolerated, especially if it occurs only in artificial contexts, such as a single word on a blackboard. Ambiguity may also be tolerated if the context would clarify it in normal communication. For instance, tolerable ambiguity may come about if a feature with a very low functional load shares orthographic representation with another feature. Tolerable ambiguity may occur when homographs which are ambiguous in isolation never occur in the same domain or are not of the same grammatical category. It is always wise to remember that orthographic decisions can increase ambiguity if not enough care is taken in considering the implications of any decision.

In the Kambari orthographies, the power syllables are ambiguous in isolation. Between spaces, the Agwara <i> could be any of the following: you-PL Pro, NCM (acting as
a linker), NCM (acting pronominally), or the copula (described in table 3.4). Because of their syntactic behavior, in context they are not ambiguous. On the other hand, seeking to eliminate all potential ambiguity by making a unique spelling for each morpheme would require an unacceptable level of deviation from predictable sound symbol correspondence and also require a high degree of linguistic sophistication in the users.

4.1.5 The preference for independent words principle

The advanced principle of preference for independent words is that word spacing is used to preserve syntax. Another way to state the principle is that a decision to write a morpheme in isolation is normally preferable to treating it as an affix. Spaces do not indicate pauses in speech; they indicate units of meaning. This fact is often overlooked both in orthography development and in literacy instruction. It is possible, and even desirable in some cases, to have as grammatical words certain morphemes that untrained MT speakers cannot isolate and identify. The fact that such morphemes are non-recognizable in isolation does not preclude independence. This principle stands in contrast to the common practice of assuming affixation whenever phonological processes extend across morpheme boundaries. Even in this environment, I maintain that the default decision should be to space the morphemes, and affixation requires additional supportive factors.

In the early developmental stage of orthography design, it is useful to assume spaces between all morphemes and proceed to justify those morphemes that become affixes. This kind of strategy ensures that word length is moderated. The extremes of word length should be avoided if possible, i.e. strings of words with one or two letters or words with four or more syllables. Appropriate word length provides word profiles that are useful to a reader. If all the words are only one or two letters, there are not enough word profiles to provide help.
If all the words are very long, the profiles become too complex to be of practical use in reading.

In the structure of Kambari, there is conflicting data indicating the spacing or affixation of the morpheme <tsu> ‘REP’. By preferring its status as an independent word, verb length was moderated and potential ambiguity was resolved, e.g. AUN tsugono ‘kingship’, tsu gono ‘REP return’.

4.1.6 The underlying form principle

The advanced principle of the underlying form is that when there is mass variation of certain morphemes across the language community, or even by a single speaker, those morphemes should be written in the underlying form. One example of this situation is when the variants are the pre- and post- forms of a phonological process. Readers reading the pre-process form are free to apply or not apply the process. If the post-process form is written, readers will read only that form. This principle is used primarily to determine the standardized form from among a number of variations occurring in natural speech.

In Salka Kambari, when the NCM /u/ follows a noun ending with /a/, it is unpredictable if the NCM will replace the /a/ (elision) or if it will coalesce with it (see ‘vowel coalescence’ in section 2.3.1.3). In this area of wide variation among different speakers, the retention of the underlying form of the NCM allows readers to apply either process in conformity with their dialect preference.

4.1.7 The stable form principle

The basic principle of stable forms is that the same morpheme is written the same way every time. This is also known as a constant visual image (see footnote 3). The advanced form of this principle is that it is preferable to preserve the stability of the beginnings of words. Suffixes do not obscure the identification of the basic content word;
prefixes can. In general, I believe that prefixed words are ‘stored words’, i.e. that they occur in the lexicon as units, and that suffixed words are ‘built words’, i.e. that the assemblage of the morphemes in the proper order produces the desired meaning. The stable form principle is intended to allow instant recognition of entire words, especially of the core vocabulary as described above.

A corollary to this principle is that if a strict morphemic representation (stable form applied to every morpheme) is rejected as being too abstract, it is better to maintain the stable form of content words and allow orthographic variation of the grammatical words. Variations of content words affect thousands of words, whereas variations of grammatical words apply to a limited set of words. These grammatical words, which have a high frequency of occurrence, can be easily learned.

This principle was illustrated in the Kambari core vocabulary discussion, comparing <danai> ‘say-CMP’, which occurred numerous times in the corpus and is therefore processed automatically by readers, with <nadanai> ‘and-they-say-CMP’ that occurs only once.

According to this principle, the Kambari orthographies’ preservation of content words such as AGW <we enei> [wenei] ‘he see-CMP’, is preferable to the Beck options described in section 1.5.2. Though the verb <ene> ‘see’ has several suffixed forms, by preserving the beginning in this way, it is always readily identifiable.

4.1.8 The easiest for users principle

The easiest for users principle is that orthographies are designed to be most functional for the users. An orthographic decision that is easiest for the users is preferable to one that is theoretically elegant or technically illustrative. An orthography in which readers and writers must regularly consult a ‘glossary or table of notations or conventions’ to decode
the orthography, as Donwa-Ifode (1990:15) recommends does not have everyday usage in mind.

It must be taken into account that there are different groups of users whose orthography needs vary. Beginning readers are largely dependent on sounding out phonemes, while fluent readers rely on instantaneous word and phrase recognition, using the strategy of sounding out words only for unfamiliar ones. Orthographies designed primarily with beginning readers in mind are cumbersome for those who have acquired fluency. Readers find certain orthographic conventions easy to use. Writers find different orthographic conventions easy to use. What is easy for readers may be difficult for writers and vice versa. Beginning readers are guided by classroom instruction; fluent readers function independently. As Dawson stated, ‘Decisions for a popular (or practical) orthography should not be based on which is best for ease of reading acquisition since those choices are not necessarily best for fluent reading and may even hinder fluent reading’ (1989:LL86145).

In applying the easiest for users principle, orthographers need to identify, prioritize, and balance the needs of the various user groups. Preference should be given to an orthography that is easiest for fluent readers and writers rather than one that is targeted for beginning readers. The preference for fluent reader needs is based on the premise that people will read for a lifetime, but only spend a few years (or even less) in learning to read. The preference for the needs of writers is based on the concept that readers need things to read, and that writers will more likely produce things to read if the orthography is designed for ease of writing.

The decision to space the NCM preceding adjectives in the Kambari orthographies (see ‘NCM factors’ in section 3.3.4) even though it functions as an obligatory prefix was
based on the premise that it is easier for users to treat all NCMs in the same manner, than to have different rules based on different grammatical categories.

The decision to write all Salka set 2 pronouns (see section 3.3.5.6) with double vowels was based on the principle that it is easier in the long run to be made aware of the semantic difference of the two sets than it would be to be able to correctly analyze the more infrequent occurrences of the set 2 pronouns requiring double vowels.

4.1.9 The acceptability principle

The advanced principle of acceptability states that an orthography should be designed to invoke a positive feeling of readability upon first exposure. Barnwell (1998:74) and Williamson (1984:10–11) use acceptability and familiarity, respectively, for the reaction of the people to the proposed orthography. I see these as distinct stages, however, in the people’s reaction to orthographies. I define acceptability as the first reaction of an individual or community to the presentation of the orthography of the mother tongue. An orthography that is accepted by the language community will be the one that matches their expectations of what an orthography should be like. They have these expectations because they are accustomed to the orthographies of other languages.

One pitfall in the area of acceptability is the potential of a proposed orthography being rejected because it diverges too much from the orthography of the LWC even though it is a more accurate representation of the structure of the target language. Another pitfall is the insistence of the people on representing the exact surface forms instead of the more abstract underlying forms. This is a problem in that initial evaluation of the orthography by the mother-tongue users cannot be done by people who are experienced readers of that orthography. Observation of various orthography development situations has shown me that experienced readers and writers move away from phonemic representation and toward
morphemic forms. However, instituting morphemic representation in the initial stages may be rejected by untrained readers who insist on the phonemic (or even phonetic) representation. The shift toward morphemic rather than phonemic forms occurred in the representation of the morpheme /N/ ‘with’ in Auna Kambari (see footnote 40).

This feeling of readability should apply both to those who are not literate and to those who are literate in the LWC. For those literate in the LWC, acceptability is directly related to the degree of similarity of the two orthographies, especially in areas of word length, special characters, paragraphing, and punctuation. Another area to consider is that of difficulty factors, i.e. things like frequent diacritics, IPA characters, or even layout considerations such as insufficient white space (margins, spacing between words, lines, and around pictures).

4.1.10 The familiarity principle

The familiarity principle is based on the resulting feeling of correctness that users of an orthography have when they are accustomed to using it. This happened in the Kambari orthographies with the consistent representation of <m> before labial consonants as in <imkpa> ‘fly [verb]’ instead of <inkpa>. Familiarity is, however, both an asset and a liability. It is an asset in that it is essential in community-wide standardization. It is a liability in that it can hinder orthography reform. The decisions made to accommodate acceptability may become fixed features through familiarity. Then, when better solutions are proposed, they are rejected due to the familiarity of the users with the system. I predict this rejection would happen if the Auna Kambari orthography seeks to simplify the representation of the phonetic [i]. Users who are already familiar with the complex system are not likely to respond positively to orthographic reform.
It is important to take into account the level of familiarity of the test subjects when evaluating an orthographic decision. If in following this principle, the optimal orthography is designed for fluent readers, the orthography decisions must be tested by those who are familiar with the orthography, i.e. fluent readers, which is admittedly a problem for new orthographies. However, a high level of mother-tongue speaker involvement in orthography design from the earliest stage will provide at least a small core of people who can perform in this role.

4.1.11 The proliferate options principle

The advanced principle of proliferating options is designed to free creativity and avoid the horns of a dilemma, i.e. orthography decision A is undesirable; therefore decision B must automatically be implemented. The proliferating options principle should be applied in every instance of orthographic decision making. It is exemplified throughout chapter 3. The proliferating options principle comes into play both in deciding how to write a feature and when to write a feature. The orthographic conventions (how to write) deal with things like distinguishing morphemes by double letters, capitalization, hyphenation, diacritics, spaces, etc. When to write specific solutions to orthographic dilemmas may be evaluated along the continuum: Is this feature something that should be written every time it occurs? Should it be written every time it occurs in certain contexts (every verb is marked)? Should it be marked only when in direct contrast with another word and when the difference in meaning could possibly be confused? Should the meaning not be marked overtly, but extra care be given to make sure that when it occurs, no confusion exists? (Possibly through the inclusion of clarifying phrase, e.g. if describing some action in the future, give the sentence an overt time phrase.) Is this feature of so little functional load in distinguishing meaning that it may be ignored? Also included among the possibilities is that of reducing the (impractical) number of representations of certain aspects of the language by grouping them
as in the Longacre example (1964:133) and applied to the Salka Kambari non-high vowel functors before verbs (see section 3.3.5.5) in which ten or more possible solutions were discussed.

4.1.12 The weigh advantages and disadvantages principle

The weigh advantages and disadvantages principle is that each decision has both advantages and disadvantages. A given option may offer enough the advantages to offset the inherent disadvantages of that option. Sometimes a tentative decision that seems advantageous will be reconsidered when the disadvantages are considered. Sometimes the issues are linguistic such as phonological processes or morpheme identity; sometimes they are social such as acceptability or the development of teaching materials.

An example of the application of this principle was seen in the discussion of the orthographic representation of the schwa in section 3.3.1.1. Each option had both pros and cons, and the option chosen, <a>, while having disadvantages, was clearly preferred once all options had been considered.

4.1.13 The teaching load principle

The teaching load principle states that when other factors are more or less equal, choose the option that can be most easily taught. It is based on the premise that everything in the orthography will need to be taught in order for people to read it, and reducing the time and effort needed to achieve fluency will result in more and happier readers. For people who are already literate in the LWC, the teaching load involves only those aspects of the orthography that differ from that of the LWC. For beginning readers, everything needs to be taught. If the underlying form of morphemes is written, the teaching load will be for beginning readers to learn to absorb the meaning and make adjustments for natural
pronunciation. If the surface form of morphemes is written, complex rules must be devised to teach writers correct spelling.

In the decision of how to represent the phonetic [i], the Agwara process of always writing the isolation form of the verb root, then adding the suffix carries a much lower teaching load and is therefore preferable to the strategy adopted by Auna which requires teaching complex and ordered spelling rules.

4.1.14 The act the same, write the same principle

The act the same, write the same principle deals with internal consistency, i.e. consistency of the whole writing system. The same orthographic decision will be made for all morphemes that fill the same syntactic role. The same orthographic decision will be made for all morphemes that undergo the same phonological processes. It is this kind of consistency that enables language planners to summarize the operating philosophy of the orthography in a short paragraph.

This kind of systematized consistency is a result of recursive decisions. A tentative decision is proposed, its implications are studied (how the decision affects other things in the writing system), and perhaps refinements are proposed. Then the implications of the refined decision need to be studied throughout the writing system, and perhaps further refinements will be implemented, and so forth. Thus, orthography development is seen as a network of interconnected lines, not a linear progression.

The representation of vowel copying in the Kambari orthographies is an example of this principle. All the morphemes which exhibit this phonological behavior are represented by the same orthographic convention. The NCMs, pronouns, noun-class prefixes, the preposition ‘to/at’, the Salka Kambari definite article, etc. all vary their orthographic shape to accommodate vowel copying.
4.1.15 The clarification from co-languages principle

The clarification from co-languages principle states that the examination of the language structure of co-languages is beneficial in resolving orthographic quandaries. Often a morpheme will be more phonologically distinctive or more frequently used in one co-language than in another. Insights gained by looking at one co-language are then utilized in making harmonized orthography decisions for all the co-languages.

One example is the treatment of the single morpheme containing ‘y’ in Agwara Kambari, which was validated by the behavior of ‘y’ in Salka Kambari.

4.1.16 The evaluation of conflicting parameters principle

The evaluation of conflicting parameters principle states that when orthographic principles conflict, look to other advanced principles for guidance. Often in the process of applying the basic orthography principles discussed earlier, one finds one principle points to decision X, while another principle when applied to the same situation points to decision Y. In such cases, advanced principles such as systemic consistency, stable forms, insight from co-languages, etc. will enable the language developers to avoid ad hoc decisions.

In the Kambari orthographies, conflicting indicators existed regarding the spacing/affixation of subject markers in the verb phrase. The phonological factors pointed to affixation. The syntactic behavior of the same morphemes in other constructions indicated spacing. The principles of giving preference to independent words and of maintaining stable forms were two of the decisive advanced principles employed at this point.

4.1.17 The positive criterion principle

The positive criterion principle states that if the application of various orthographic principles yields positive results, definite conclusions can be made, while absence of positive results cannot lead to conclusions. For example, if a native speaker is able to isolate and
identify a morpheme, that is a good indicator that the morpheme can stand as an independent word. If however the native speaker is not able to isolate and identify a morpheme, one cannot conclude that that morpheme must be an affix.

An Agwara Kambari MT speaker will recognize <kanaka> ‘cow’ in isolation, providing positive evidence that <kanaka> should be written as a freestanding word. The MT speaker will not be able to provide meaning for <ka> ‘NCM’ in isolation. Applying the positive criterion principle, this negative evidence regarding the status of <ka> cannot be taken as proof that it should be an affix.

4.2 **Interaction of advanced principles**

It should be noted that these advanced principles do not function independently. All are interrelated, though some may provide stronger arguments in specific situations. All of the advanced principles were involved in making the decisions regarding the Kambari verb phrase. Seeking to keep inflected verbs to a moderate length was in consideration of the acceptability principle and the preference for independent words principle. The positive criterion principle was invoked when identification as a word by MT speakers of the aspect-verb root string influenced the decision to make the aspect morpheme a prefix. This is in accordance with the easiest for users principle; they already think of these morphemes as one word, so naturally they write them that way. Following the corollary to the positive criterion principle, when MT speakers failed to isolate and identify grammatical morphemes, other principles were sought for better justification of spacing decisions.

The impetus to space the object pronoun yi ‘him’ in Salka Kambari augmented the implications of the lack of vowel copying from the verb root to le ‘them’ of all the languages and influenced the decision to space object pronouns. This is an example from the clarification of co-languages principle. The act the same, write the same principle was
demonstrated with the subject and object pronouns both deemed independent words since they are part of the same grammatical category. This same principle also affects NCMs occurring in the subject position of the verb phrase and elsewhere. They are deemed independent words and undergo the same phonological processes. Thus, NCMs are written the same way wherever they occur.

In weighing the advantages and disadvantages, long words are a disadvantage, but strings of one or two letter grammatical words are a disadvantage as well. The proliferate options principle was invoked in considering whether to space or affix each pair of morphemes as well as considering whether to space or affix groupings of morphemes. The familiarity principle was invoked in that fluent readers are accustomed to one-syllable grammatical words since they appear frequently, and they are able to quickly recognize content words because of consistent spelling and limited prefixes. The latter factors are also following the stable form principle. The underlying form principle is in operation at the word level, especially in relation to the morpheme  tão. The complete form of the verb is written before  tão even in cases where the final vowel of the verb is elided; <u wala tão> [u wal tão] ‘he has walked’.  tão itself is always written  tão though its vowel elides with the vowel at the beginning of the following word; SAL <u tanu tão amuna> [u tanu tamuna] ‘he carry clothes’.

The optimal simplicity and convey meaning principles were balanced with the development of a system that accurately conveys meaning through the mechanism of spaces. Using one orthographic strategy for pronouns and NCMs regardless of their position in the syntax reduced the teaching load.

The conflicting parameters principle came into play regarding the status of the obligatory subject marker. Its being obligatory indicated affixation. Affixation was also
indicated by the morphemes being phonologically linked to the verb root. Independence was indicated by considering the subject pronouns, object pronouns, and possessive pronouns together, and by considering the various categories of things linked to NCMs. The consideration of co-languages was a factor in that Auna Kambari obligatory subject markers are not phonologically bound to the verb root. The ultimate deciding principle was that of seeking to maximally reduce ambiguity by spacing phonemically identical morphemes differently.

4.3 Standardization, harmonization, and conformity

Various authors have used the terms standardization, harmonization, and conformity in different manners; therefore, it is appropriate to clarify how I will use them, before proceeding to discuss these factors in the Kambari setting.

A standardized language is one that has delineated the linguistic range of the standardized form, i.e. who will be served by the written materials, has defined the alphabet and orthographic rules, and has prepared materials that aid and promote the use of the written language. Co-languages, i.e. sister languages of a language cluster, are standardized independently in that each co-language adheres to its own set of standardized rules and is served by its own written materials.

I diverge from Williamson’s use of the term harmonization (see 1984:10–11). I am using harmonization to refer to the efforts to make the writing systems of co-languages as similar as possible. It is natural to expect that two co-languages should have more points of similarity in their writing systems than either would have with an LWC.

Following Barnwell’s use of the term conformity (1998:74), I use conformity to refer to maximum similarities possible with other languages in linguistic or geographic proximity; i.e. the languages whose writing systems speakers of the developing language are likely to
come in contact with. The theoretical ideal is that all the languages of Nigeria, for example, would be markedly similar in symbols and philosophy. Practically, the greater the conformity to the languages in contact, without doing disservice to the developing language, the easier it will be for readers of the developing language to transfer their reading fluency to the other languages. It will also be easier for speakers of the developing language who are literate in other languages to transfer their reading fluency to their own language.

In any of these three situations (standardization, harmonization, and conformity), there is always a tension between an elegant solution that does not conform and a compromise solution that addresses the reality of other language forms. What is best linguistically may not be the best sociolinguistically, and vice versa.

4.3.1 Standardization and harmonization in Kambari

In consideration of standardization and harmonization, the question arises: are these different languages or dialects of one language? If the latter, there is no harmonization involved; if the former, standardization is undertaken with harmonization constantly in mind. In the case of Kambari, the three Kambari co-languages have been shown to be distinct languages.

Standardization in the Kambari setting deals with the independent standardization of each co-language. Within the co-languages, the choice of the reference dialect, i.e. the form of the language on which the orthography is based, was selected by the Kambari Language Project executive committee who then selected the staff for each co-language in accordance with the committee’s dialect preference. One of the requirements of a new staff member is the rather subjective evaluation of ‘able to speak the language well.’ As part of the training of the staff, they were made aware of dialect variation, and efforts were made by KLP to deal with significant dialect variation. The staff members are from the town that is
the ethnopolitical center. In retrospective analysis, the languages within KLP are center-oriented as is illustrated by the names of the languages, e.g. Salka Area Kambari. The map in figure 2.3 emphasizes how a bordered area, such as is displayed in the language map included in the *Index of Nigerian Languages* (Crozier and Blench 1992), does not capture the sociolinguistic reality. This is demonstrated by speakers of Agwara Kambari, who live in the heart of Salka and Auna areas and have fairly remote ties to Agwara town, and yet request literacy classes using Agwara Kambari materials.

In the standardization process, each co-language is codified individually, i.e. has an established alphabet and rules for writing. Harmonization seeks to make each orthography as similar to the others as possible, taking into account that while the co-languages are similar, they are not the same. Orthographies of co-languages can harmonize in many areas, but divergent language structures dictate differing orthographies. By divergence, I mean structural differences among the co-languages that prevent harmonization. The three Kambari orthographies have harmonized in these areas: identical alphabets, identical rules for spacing of one-syllable functors, and identical rules for affixing inflectional morphemes. In this way, the overall appearance of a page of writing is very similar among the co-languages.

It is my premise that standardization and harmonization are inseparable, and that the orthography of each language is the better for having considered all the languages simultaneously. The degree to which harmonization is achievable is limited by the linguistic structure and social perceptions. Neither harmonization nor standardization can proceed without the other. This is illustrated by the intersecting arrows of figure 4.2. The harmonization factors improved the standardization. The simultaneous standardization improved harmonization. The simultaneous standardization and harmonization of the three
Kambari co-languages was enhanced generally by the daily interaction of the KLP staff and specifically addressing complex issues via group and individual co-language sessions.

By doing language analysis and orthography development in the three co-languages simultaneously, many times a grammatical feature would be more clearly visible in one of the three co-languages, though it was present in all three to some extent. By establishing an orthographic rule for the language with the clearest expression of the grammatical feature, it was then easy to adapt that rule to the other languages.

An example of comparing co-languages concerns the gerundial prefix. In Auna Kambari it is /u-/, but in the early days of analysis, it was difficult to distinguish /u-bana/ ‘ING-go’ from /u bana/ ‘he goes’. In Agwara Kambari, the pronoun is /u/ while the gerundial prefix is /ku-/, and the consonantal onset made it very clear to identify. In parallel sentences in Auna Kambari, the gerund was established, and a substitution test was devised to ascertain if /u/ in that context is pronoun or prefix. In both Agwara and Auna Kambari, the gerundial prefix is universal, i.e. it can attach to any verb. Thus, in the process of standardizing the languages, harmonization occurred in that the morpheme in each language is affixed, but the co-language distinction is preserved in the orthographic representation of the morphemes, i.e. AGW <ku-> and AUN <u->.
When beginning analysis on Salka Kambari, I assumed the language would have an equivalent, universal form. There is no overt morpheme. There is, however, a nominalizing prefix /mV-/ that attaches to most two-syllable verbs. There is also a nominalizing prefix /V-/ (functor 8), but both of these morphemes are lexically specified for certain verbs, and are not universal. The expectation of a universal gerundial prefix, combined with the awareness that Salka makes use of null prefixes (see table 2.7), revealed that the universal gerundial prefix is in fact, null. This is confirmed by irregular verb roots (still without an overt prefix) that follow functor 7 ‘at/while’.

144) SAL
  a. (regular) g 'yqwa 'they go', g 'yqwa 'at/while going'
  b. (irregular) g cu'wgn 'they plant', g ca'qgn 'at/while planting'

If the two occurrences of 'yqwa' in the previous example are morphemically identical, then there is no explanation for the change to ca'qgn. Thus, the pattern established by Agwara and Auna Kambari above was helpful in understanding the structure of Salka Kambari better.

The three Kambari orthographies differ in specific instances where the structures of the languages diverge or if the internal consistency of one co-language’s orthography is in contrast to the harmonized form. The sociological reality of unique group identity among the three groups is reflected in these differing language structures, and in the orthographies also. This phenomenon may be termed disharmonization.

One of the most vivid examples of disharmonization is that of the transliteration of biblical proper nouns. Some of these names are already commonly used in the churches of the areas, but many are obscure. In Bible translation, the task is to transliterate these obscure names to make them pronounceable according to the phonology of the target language, while retaining enough of the original to be recognizable across languages. It
seems axiomatic that the three languages, which have very similar phonologies, working on the same list of names would transliterate them the same way. However, after several lengthy sessions of making conscious effort to harmonize, the lists of names are only about eighty percent the same, e.g. a Greek person (Hellene) in Hausa is *Baheleniya*, AUN *Maheline*, AGW *Mehelene*, SAL *Meheleni*. The /MV-/ prefix is the standard prefix denoting tribal member, e.g. ‘Fulani person’, AUN *Magengeru*, AGW *Mapulatai*, SAL *Mapulata*. One can see disharmonization in both the contrived biblical name and the naturally occurring word.

Harmonization is a goal to be pursued, nonetheless the intuitions and decisions of the MT speakers may limit the degree of or even prevent maximum harmonization.

### 4.3.2 Conformity in the Kambari setting

In the domain of written materials, prior to the work of KLP the only languages Kambari people ever learned to read, or even saw, were Hausa and English. More Kambaris read Hausa than read English, i.e. anyone who reads English also reads Hausa, but there are many readers of Hausa who do not read English. Because of this sociolinguistic situation, I prioritize the conformity in this way: first, conformity to Hausa; second, conformity to English; and third, conformity to other Nigerian languages.

One of the unfortunate influences of Hausa on Kambari stems from the manner that Hausa is written in the Kambari area is not standardized, i.e. any Kambari person writing Hausa spells words any way they land on the paper with no regard for spelling the same word the same way even within the same paragraph. Thus the majority of Kambaris who are literate in Hausa rely heavily on phonetically sounding out the words they read, and never become truly fluent readers in Hausa. They subconsciously transfer this expectation to Kambari and react negatively to meaning-based orthography decisions that will in the long run simplify spelling.
Conformity with English, besides the areas that overlapped with Hausa, was mainly in the domain of punctuation. The symbol set chosen and the function of those symbols is identical with the English punctuation system. The structural factors of Kambari syntax determine the placement of those symbols in Kambari texts. One result of the decision to write each content word in Kambari the same way every time is that Kambari readers learn to associate a single spelling with meaning rather than sound. This enables them to more easily deal with orthographically distinguished homophones in English, e.g. here, hear. Thus at the deepest theoretical level one can say there is conformity between Kambari and English in the area of representation of content words.

4.4 Simultaneous consideration of co-languages

In many language situations around the world, a single co-language is developed without consideration of its sister languages. I strongly advocate simultaneous consideration of co-languages as a strategy for better standardization of each co-language, and also for greater harmonization of the whole package. A single development effort for multiple co-languages is a demonstration of the social reality of a single ethnic identity manifested in different speech forms. The cost in terms of time, effort, and money, while more than that of developing a single language, is less than three independent development projects. Some of the benefits of simultaneous development of co-languages are intangible, e.g. increased language pride, increased motivation for MT literacy due to subconscious competition among co-languages, strong bonds of teamwork and ethnic unity, and the ability to share production facilities and other resources.

In the Kambari setting, one of the most interesting elements has been the interaction of the staff from the different co-languages. Day to day interaction in the office has led to many times when a speaker of one language will seek out his colleague from another
language to help him generate possible answers to the orthographic or other language
related problem he is facing. I believe the regular, face to face interaction of those involved
in making the orthography decisions was a great help in the harmonization aspect of the
project.

4.5 General insights

These general insights represent a summation of the inferences that may be drawn
from a careful examination of the body of this work.

4.5.1 The role of tradition

Orthography is tradition. The more one is familiar with an orthography, the more
that particular orthography ‘feels right’ and the more a decision feels justified as the right
and only choice.

4.5.2 The network model

Orthography development is a network, not a linear progression. In a network each
connection is an integral part of the whole, and any action at one point in the network
affects all the other lines and connections in that network. A decision in one area of the
orthography has implications throughout the language.

Another kind of network concerning orthography is that of users: beginning readers,
fluent readers, writers, preparers of literacy materials, teachers. The question to be
constantly asking is ‘What will be easiest for users?’ keeping in mind the different kinds of
users.

4.5.3 It has to be taught

Teaching writers to ‘sound out’ words is only one way of teaching. Another way is
teaching readers and writers about group membership in the structure of their own
language, i.e. all of these whatnots are doing the same thing and are written following the same pattern. The more things that follow the pattern, the easier it is to teach. Every orthographic decision affects teaching load.

4.5.4 Appearance matters

The visual appearance of a page of type is more important than most people realize. Orthographies can be rejected after only a glance at a page of type. There are two factors affecting acceptability: comparison to a page of type in the target language to one in the LWC and level of perceived difficulty. Someone who says the orthography of the MT is acceptable but chooses to read the LWC and not the MT has in fact rejected the MT orthography.

4.5.5 Orthography development: progression toward the ideal

No orthography is perfect. Not every aspect of orthography development will consume the same amount of time. Certain areas will be decided very quickly. Others will take a long time. Once the basic decisions have been made, one should begin producing materials for field testing the orthography. If an orthography reform effort arises twenty years or so after the introduction of the MT orthography, that reform can be a sign of success for the first orthography, i.e. the orthography has been used enough to make the users themselves see how to make it better.

4.5.6 Anything is possible

Anything is possible. Orthographic principles are guidelines, not laws. Proliferating options in the early developmental stage has the obvious benefit of being more likely to include the best functional option. Developers should list all options, even absurd ones, and prepare sample paragraphs with each one. Being able to articulate why an option is absurd or rejected is an aid in recognizing which factors are important. Once the important factors
are identified, the possible options can be evaluated as to which best represents those factors.

4.5.7 The role of testing

Once options are chosen, a trial period should be instituted to see if, as users become familiar with it, it feels increasingly ‘right’. Jumping from decision to decision leads to frustration on all sides. At some point the orthography should be temporarily frozen so that no changes in the system are allowed for a fixed period of time. At the end of that period, familiarity with the system may have overcome what earlier seemed to be a problem.

4.5.8 The central role of meaning

The purpose of an orthography is to convey meaning. I have said this before. It bears repeating. In any orthography, there is a gap between the visual representation and the exact sound. The reader must interpret a minimal set of symbols intended to create meaning and supply the appropriate phonetic adjustments to produce natural pronunciation. A good orthography represents a distillation of the sounds of natural speech. It provides positive links between sound and meaning, but does not attempt to mirror oral communication.

Focusing on the communicative role of orthography does not mean that it is necessary to completely eliminate ambiguity. Just as spoken communication is at times ambiguous, and context allows proper interpretation, so too an orthography may rely on context and common sense to properly sort out a degree of ambiguity. A good orthography will reduce ambiguity to a level that is unlikely to cause problems in most circumstances.

4.5.9 The role of the community

From the beginning of the developmental stage, written materials need to be distributed widely in the community. The use of low-cost materials and small print runs ensures that the early attempts do not endure into embarrassing perpetuity. A team of
reviewers should be trained to go over each publication before it is released, starting with the first one. Community input is vital, but to have the maximum benefit, it must come from community members who understand the principles involved. Orthography design includes educating a portion of the mother-tongue community as to why certain decisions were made, or why certain options are to be preferred.

Both linguists and preparers of literacy materials need to be involved in the development of the orthography. If the literacy preparer is the primary developer, a linguistic consultant should be called. This is frequently advised in the literature. If the linguist is the primary developer, a literacy consultant should be called. This is rarely if ever mentioned in the literature.

4.5.10 The influence of the LWC

Conformity to the LWC has both advantages and disadvantages. The transfer of skills plays a powerful role in motivating many people to learn to read. Conformity to the LWC may help shift the balance toward a given decision when internal linguistic factors of the MT are ambiguous or conflicting regarding a specific orthographic decision.

The disadvantages come into play when there is linguistic divergence between the MT and the LWC, and the accurate or elegant decision for the MT conflicts with the LWC. The pressure to conform is strong. Orthography developers may find the ‘best’ solution is hampered in implementation by conceptions based on the inadequacy of the orthography of the LWC. There are no easy answers.

4.5.11 It is worth doing

There will never be a perfect orthography. The conflicting needs of readers and writers, new readers and experienced ones; the internal tensions between phonological processes and syntactic behavior, all this and more mean that any orthography is a
compromise. A compromise is always less than perfect from some viewpoint, yet it must be made.

4.6 Conclusion

An opportunity to create an orthography is a tremendous privilege; it is undertaken with fear and trembling. It is also an act of faith, undertaken in the hope that people will accept it on a widespread basis; that books, magazines, and newspapers in the target language will become commonplace; that friends will write letters, lovers write poems, and scholars write books. It is my hope that the information and concepts presented here will help that hope to become reality in more of the languages of the world.
Appendix A Sample text in the three Kambaris

*English source*

Once upon a time there was a poor man who took his gourd full of palm wine to market. He thought, “I will sell it for a lot of money, and I will take the money and buy a chicken with it. When the chicken has set, and the small chickens grow, I will sell them, take the money and buy a calf. When the calf grows and becomes a cow and keeps giving birth, I will have many cows and sell them. Then I will become a rich man.”

The man was very happy, and he started clapping his hands. Then the gourd fell off his head and broke. That was how his dream came to an end.
Agwara Kambari

Kanna ko yoku, vuza vu unambi u dikai kedele ka
Day NCM another, person NCM lacking he pick-CMP gourd NCM

makula a kubanka e kuden. Da u sheshei a kadu
palmwine at ING-go-with to market. Then he plan-CMP at heart

kag ni. Da u danai. "N ciya baci tsulaga, katga n dika
NCM his. Then he say-CMP, "I get if trading, so-that I pick

ikebe ya n tsulaka motoku n i da. Da baci motoku ma
money the I buy-with chicken with NCM it-is. Then if chicken the

ga yawain. Da muku ma ngbonguroi. Katga n denge le,
NCM hatch-CMP, then babies the NCM grow-CMP, so-that I sell them.

kagga n dika ikebe ya n tsula medendem. Da baci
so-that I pick money the I buy calf. Then if

medendem mo gbonguroi, ma yawai kanaka, da u matsai,
calf NCM grow-CMP, NCM reach-CMP cow, then it give-birth-CMP

mi ta o kokpo vuza va anaka n gbundgi, kagga kpamu
I-am PERF at ING_become person NCM cows with plenty, so-that also

mo okpo vuza vu udukuyan.
I become person NCM riches.

Da vuma va u kangai mazanga, da u kang kuabsa ekiye.
Then man the he catch rejoicing, then he start ING-clap hands,

da kedele ka kag fatatsgi de a kaci kag ni. Da ka ko
then gourd the NCM fall-CMP there at head of his. Then the NCM

bosoi. Ta alatani a vuza a zuwai uteku nannai.
break-CMP. Thus dreams NCM person NCM put-CMP end like-that.
Auna Kambari

A'ayin a cau, a yan tə vuma roku, uza da cangai
Time NCM before, they do PERF person another, one that carry-CMP

kadele ka makula ka ne, u bankai a kuden. U danai a
gourd NCM palmwine NCM his, he take-CMP to market. He say-CMP in

katakasuvu ka ne, “N denge baci makula mi n ikebe ushani.
heart NCM his, “I sell if palmwine the with money plenty.

Aku n bidya ikebe yi n tsila kanuku ka matoku. Kanuku ka
Then I take money the I buy female NCM chicken. Female NCM

matoku ka zuwa baci ako, u picuwain mmuku mə ntoku ushani,
chicken NCM lay if eggs, she hatch-CMP babies NCM chickens plenty.

hal a gbonguroi gbə. Aku n denge le n bidya ikebe yi n
until they grow up-CMP all. Then I sell them I take money the I

tsiləka kanuku ka madyondyom. Kanuku ka madyondyom ka
buy-with female NCM calf. Female NCM calf NCM

gbonguro baci, ka matsqsa tə anaka ushani. N denge
grow up if, NCM give birth-IT PERF cows plenty. I sell

baci anaka yi gbə, ma woko tə uza utsəri.
if cows the all, I-CONT become PERF one possessions.

Pini nala, kayanyan ka purə yi, aku u gitə u bəsa
Inside like-that, sweetness NCM pack him, then he begin he clap

akere a ne. Aku kadele ka fətənə di a kaci ka ne
hands NCM his. Then gourd NCM fall-CMP there at head NCM his

she iyamba, ka təsəgi, makula ma wotsəngi.
until ground, NCM shatter-CMP palmwine NCM spill-CMP.

Tyoku da alavutanshi a ne a kotsoi di la vi.
As it-is that dreams NCM his they finish-CMP PLUP there the.
Salka Kambari

A yuwan ṭag lambi a lyuci ro. Ávu urana u lo, u tanu 3P do Perf lacking-one at town another. Then day NCM that, he carry ṣakabu a makula a 'yawo o ukasua. Ávu u uwa majiyan, gourd NCM palmwine at going to market. Then he enter thinking,

u damma, "Nnγ n wina yi ni ikebe lon, ávu n tsula he say, "If I sell it with money plenty, then I buy ṣanwo o moton. Nγ moton u gigi'wan, ávu n tsurγ muton. female of chicken. If chicken it set, then I get chicks.

Nγ muton n yuwan mgbain, ávu n wina le. Nγ n wina le, If chicks NCM do big, then I sell them. If I sell them, ávu n tara ikebe'e, ávu n tsula medendem. Nγ u yuwan then I take money-the, then I buy calf. If it do mgbain, ávu u matsan. Nu u matsan, n tγ tsurγ big, then it give-birth. If it give-birth, I will get inan gbundγ. Ávu n wina le, moo o'wo za vi ikebe. cows plenty. Then I sell them, I-will become one NCM money.

U zanganga tγ majiyan mga yi, sapu u teme maβasa me ekere. He rejoice PERF thinking NCM his, until he begin ING-clap NCM hands.

Reve ṣakabu'u u fatsγ yi. Ávu u buβoso. Reve u Then gourd-the it fall-from him. Then it DUP-break. Then he dada'wan a gtsumγ a alavu a yi. be-interrupted at middle NCM dreams NCM his.
Appendix B Orthographies

Agwara orthography

Introduction

The language of the Agwara area of Niger and Kebbi States in western Nigeria is part of the Kambari cluster and is referred to in the literature as Agwara Area Kambari, and locally as Cishingini. The people who speak Cishingini are called Kashingini (Kashingini is the singular form). The people of the Agwara area also refer to anyone who is ethnically Kambari (speakers of Auna-Wara Area Kambari, Salka Area Kambari, and others) as Kashingini. Therefore, to avoid confusion, in this study the language will be referred to as Agwara Kambari.

Crozier and Blench (1992:118) give the Kambari cluster the following language family classification: Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Kainji, Western, Kambari. Agwara Kambari has two major dialects, Cishingini (spoken around Agwara town), and Tsufashi (spoken around Rofia town). The word “dialect” is used here to refer to two or more speech forms whose speakers recognize differences in pronunciation and vocabulary, but when each speaks his own form, each is understood by the other. Other languages of the Kambari language cluster include: Tsikimba (Auna-Wara area), Tsishingini (Salka, Ibeto and surrounding villages) and Tuva (east of Yelwa, Rijau, north of Kontagora). The languages of the Kambari language cluster share an ethnopolitical identity, but speakers of the different languages of the cluster do not understand one another and use Hausa to communicate. This description is based on the Cishingini dialect as is spoken in Agwara town.

Most of the Agwara Area Kambari people live within the area bounded by (on the west side of Lake Kainji) the northern border of the Kainji Game Reserve, the Benin border, to the northern border of Niger State, and on the east side of Lake Kainji, from Yelwa area south to Nasko town. There many settlements of Agwara speakers in Auna and Salka areas. Early work in Agwara Kambari was done by Mierau (1967) (only an unpublished draft of his language analysis exists, no books published in Agwara Kambari in this time), and later by Father Timothy Carroll, of the Roman Catholic Church who lived in the area for many years and published some books (including primers) and the gospel of Mark. For further history of Agwara Area Kambari see Kingdoms at War (CAPRO, 1995) and Yusuf (2000, long essay).

In 1992 leaders of the Agwara Kambari joined with those of Auna-Wara Kambari and Salka Kambari to form the Kambari Language Project (KLP). James Nlaya Yusuf was
selected to attend the 1992 Introductory Course in Applied Linguistics (ICAL) sponsored by the Nigeria Bible Translation Trust in Jos. In 1998 Basuna John Magaji attended ICAL and joined the KLP staff. John and Janie Stark, who are both reading Ph.D. in Linguistics at the University of Ilorin, have been KLP technical advisors since 1992. A trial orthography has been developed and used by the Kambari Language Project (KLP) to produce written materials, including a primer series to teach those who do not know how to read, a transition primer series for those who can already read in Hausa, several reading books and some Bible passages.

This book follows the outline suggested by Salami (1996).

**The philosophical basis of the KLP orthography**

The KLP orthography divides words into two groups: content words and grammatical words. Content words are nouns, verbs, adjectives, etc., i.e. words which for the most part are easy to isolate, visualize, define, and translate to or from another language. Grammatical words in Agwara Kambari are tense and aspect markers, noun-class concord markers, certain pronouns and prepositions. In the KLP orthography, content words are spelled the same way every time they occur. Grammatical words are allowed to vary in their spelling to more closely approximate their pronunciation in that context. In this way, even with their multiple variations of spelling, there is still a limited set of grammatical words. Most of these are frequently used and therefore quickly learned. On the other hand, content words are an unlimited set (the language can add new words at any time); therefore, to restrict each content word to a single spelling insures that regardless of context, the meaning of that word will be apparent and quickly learned since its spelling never varies. This principle of content and grammatical words is easy to teach since all of the variable grammatical words are a single syllable; whereas nearly all of content words are polysyllabic.

**The current alphabet**

This is the alphabet in its alphabetical order with sounds in the International Phonetic Alphabet (IPA), a sample word in the language and its gloss in English. The English gloss of verbs is given in the infinitive, but the Agwara Kambari form is actually the singular imperative.

<table>
<thead>
<tr>
<th>a</th>
<th>[a]</th>
<th>anaka</th>
<th>‘cows’</th>
</tr>
</thead>
<tbody>
<tr>
<td>ae</td>
<td>[ə]</td>
<td>angku</td>
<td>‘mother’</td>
</tr>
<tr>
<td>b</td>
<td>[b]</td>
<td>bgbg</td>
<td>‘baby’</td>
</tr>
<tr>
<td>b</td>
<td>[ɓ]</td>
<td>batsa</td>
<td>‘to split, to branch’</td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
<td>-------</td>
<td>----------------------</td>
</tr>
<tr>
<td>c</td>
<td>[ɠ]</td>
<td>ciga</td>
<td>‘to want, to love’</td>
</tr>
<tr>
<td>d</td>
<td>[ɗ]</td>
<td>dana</td>
<td>‘to say’</td>
</tr>
<tr>
<td>d’</td>
<td>[ɗ]</td>
<td>daka</td>
<td>‘under’</td>
</tr>
<tr>
<td>e</td>
<td>[ɛ]</td>
<td>edele</td>
<td>‘gourds’</td>
</tr>
<tr>
<td>f</td>
<td>[ɬ]</td>
<td>fodo</td>
<td>‘to remove’</td>
</tr>
<tr>
<td>g</td>
<td>[ɡ]</td>
<td>gaga</td>
<td>‘to pinch’</td>
</tr>
<tr>
<td>gb</td>
<td>[ɡb]</td>
<td>gbama</td>
<td>‘to be strong’</td>
</tr>
<tr>
<td>h</td>
<td>[ɦ]</td>
<td>hawun</td>
<td>‘clear’</td>
</tr>
<tr>
<td>i</td>
<td>[i]</td>
<td>ili</td>
<td>‘thing’</td>
</tr>
<tr>
<td>j</td>
<td>[ɗɮ]</td>
<td>kajiba</td>
<td>‘mat’</td>
</tr>
<tr>
<td>k</td>
<td>[k]</td>
<td>kanaka</td>
<td>‘cow’</td>
</tr>
<tr>
<td>kp</td>
<td>[kɲ]</td>
<td>kpa’a</td>
<td>‘house’</td>
</tr>
<tr>
<td>l</td>
<td>[ɬ]</td>
<td>lapa</td>
<td>‘to beat’</td>
</tr>
<tr>
<td>m</td>
<td>[m]</td>
<td>mogono</td>
<td>‘king/chief’</td>
</tr>
<tr>
<td>n</td>
<td>[n]</td>
<td>nana</td>
<td>‘to deny’</td>
</tr>
<tr>
<td>o</td>
<td>[o]</td>
<td>opodo</td>
<td>‘toads’</td>
</tr>
<tr>
<td>p</td>
<td>[p]</td>
<td>pasa</td>
<td>‘to cross’</td>
</tr>
<tr>
<td>r</td>
<td>[r]</td>
<td>reme</td>
<td>‘to catch’</td>
</tr>
<tr>
<td>s</td>
<td>[s]</td>
<td>saka</td>
<td>‘to observe’</td>
</tr>
<tr>
<td>sh</td>
<td>[ʃ]</td>
<td>shana</td>
<td>‘to fill’</td>
</tr>
<tr>
<td>t</td>
<td>[t]</td>
<td>ta’a</td>
<td>‘to pluck’</td>
</tr>
<tr>
<td>ts</td>
<td>[ʦ]</td>
<td>tsuwa’a</td>
<td>‘starch food (tuwo)’</td>
</tr>
<tr>
<td>u</td>
<td>[u]</td>
<td>unu</td>
<td>‘mouth’</td>
</tr>
<tr>
<td>v</td>
<td>[v]</td>
<td>vara</td>
<td>‘to throw’</td>
</tr>
<tr>
<td>w</td>
<td>[w]</td>
<td>wala</td>
<td>‘to go, to walk’</td>
</tr>
<tr>
<td>y</td>
<td>[j]</td>
<td>yeve</td>
<td>‘to know’</td>
</tr>
<tr>
<td>z</td>
<td>[z]</td>
<td>zuwa</td>
<td>‘to put’</td>
</tr>
</tbody>
</table>
In the Agwara Kambari alphabet there are six vowels and twenty-five consonants, making a total of thirty-one letters. All of the symbols used are common to English and Hausa except for the vowel ‘g’. The symbols shared by Kambari and Hausa, though not English are ‘b, d, and ’ (glottal stop). The letters b and d are the only letters that are not easily made by an ordinary typewriter. The digraphs ‘kp, gb, sh, and ts’ function as a single consonant. The glottal stop [’] (alhamza in Hausa) which looks like an apostrophe, is not included in the set of letters called the alphabet, although it is used as a break between vowels the same way as in the Hausa words addu’a ‘prayer’ shari’a ‘law’. The symbol for the glottal stop [’] is called kapasukpatsu ‘the thing that crosses over’ in Agwara Kambari, and is used as in the word g’isq ‘prayer house’. Agwara Kambari allows ‘y’ to follow certain consonants as in ikyamba ‘body’, gbata ‘cut’, inyama ‘meat’, vudyangi ‘a kind of toad’, and kagulya ‘earthworm’. Hausa has a glottalized y [?] as a separate letter (as in ‘ya ‘daughter’). Agwara Kambari has words with both the ‘y [?] and ‘w [?w] sounds as in ‘ya kubana ba ‘they will not go’ and she’we ‘to do carpentry’.

**Vowels**

The six vowels of Agwara Kambari are as follows:

<table>
<thead>
<tr>
<th></th>
<th>Unrounded</th>
<th>Unrounded</th>
<th>Rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Front</td>
<td>Central</td>
<td>Back</td>
</tr>
<tr>
<td>Close (high)</td>
<td>i</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td>@</td>
<td>o</td>
</tr>
<tr>
<td>Open (low)</td>
<td></td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

**Diphthongs**

Diphthongs are two different vowels next to each other within a word. In Agwara Kambari these are restricted to a low vowel (a, g, e, or o) followed by a high vowel (i or u). Except for a handful of words, diphthongs are formed by adding the –i suffix (meaning either completed action or plural command) to the end of a verb. For example ene ‘to see’ becomes enei ‘saw’, yawa ‘arrive’, yawai ‘arrived’, tgwgi ‘come’, tshwgi ‘came’, rono ‘draw’, ronoi ‘drew’.

**Long vowels**

Long vowels are present phonetically in Agwara Kambari, but there are few pairs of words where the distinction of meaning is based on the presence or absence of vowel length. Even among these pairs, very rarely do both words of a pair belong to the same part of speech or semantic domain. The writing of long vowels involves analysis, consensus of
analysis (community agreement that a certain set of words do in fact have long vowels), writing consistent materials including primers using long vowels, teaching people to read with correct pronunciation, teaching people to analyze for themselves (hear the distinction in their minds) so they can write correctly, and community-wide consistent writing of long vowels. This is a burdensome workload for the alleviation of relatively few instances of ambiguity. In the KLP Orthography, no long vowels are written.

This is an area of divergence from what will be called in this paper the Carroll orthography. The Carroll orthography includes long vowels. Some of the places where the Carroll orthography writes a double vowel, KLP writes a glottal stop between the two vowels. In other instances, because of the low number of pairs of words in direct contrast (minimal pairs whose only difference is vowel length), KLP uses a single vowel.

**Vowel processes**

There are three different ways vowels interact with each other both within words and across word boundaries.

**Vowel copying:**

In Agwara Kambari, certain words and affixes contain vowels which derive their pronunciation from nearby vowels. In this process, a certain vowel acts as the source or control of the sound of other vowels. The source vowel can be either to the left of the target vowel or to the right of the target vowel. The vowel quality to be copied (the source vowel) is the first non-high vowel (a, ã, e, or o) of the root. This process occurs throughout the language, and can copy from any word, regardless of the part of speech: noun, verb, adjective, pronoun, etc. Below is an example of vowel copying to the left.
If the first vowel of the root (the word with the source vowel) is a high vowel (u or i), the copying mechanism will skip over that vowel, seeking the first low vowel of the word (a, g, e, or o) as the source vowel, even if that is the second or third vowel of the word. If there is no low vowel in the root, the vowel ‘q’ is supplied as a default when copying to the left.
The process of vowel copying to the right (the source vowel is to the left of the target vowel), the right-most high vowel will be the source vowel to be copied. The explanation of why the ‘s’ changes to ‘sh’ will be given later.

<table>
<thead>
<tr>
<th>Uninflected verb</th>
<th>Verb plus iterative suffix –sa (IT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>kara ‘to tear’</td>
<td>karasa (kara + sa) ‘to tear-IT’</td>
</tr>
<tr>
<td>kğga ‘to catch’</td>
<td>kğgas (kğga + sa) ‘to catch-IT’</td>
</tr>
<tr>
<td>kece ‘to count’</td>
<td>keceshe (kece + sa) ‘to count-IT’</td>
</tr>
<tr>
<td>soyo ‘to pound’</td>
<td>soyoso (soyo + sa) ‘to pound-IT’</td>
</tr>
<tr>
<td>tsula ‘to buy’</td>
<td>tsulas (tsula + sa) ‘to buy-IT’</td>
</tr>
<tr>
<td>cipa ‘to come down’</td>
<td>cipsg (cipa + sa) ‘to come down-IT’</td>
</tr>
<tr>
<td>suku ‘to send’</td>
<td>sukusu (suku + sa) ‘to send-IT’</td>
</tr>
</tbody>
</table>

**Vowel elision**

Vowel elision is the process where when two vowels come together (usually the last vowel of one word and the vowel beginning the next word), the second vowel “eats” the first vowel, and only the second vowel is pronounced. For example: *Ubuta u na* ‘that place’ is pronounced [ubutu na]. *Ikebi i və* ‘my money’ is pronounced [ikebi və]. In the orthography, the content word is preserved intact, and the NCM stands as an independent word.

Another instance of vowel elision in Agwara Kambari is across morpheme boundaries when the consonant before the final vowel of the first morpheme is a sonorant (l, r, n, w) and the second morpheme begins with a non-sonorant consonant, usually with the same point of articulation as the sonorant. In this setting, the final vowel of the first morpheme deletes, and the final consonant of the first morpheme becomes the coda of the syllable preceding it. For example, the sentence written: *U bana tə* ‘he has gone’ is
phonetically [u ban tã]. *U wala tã* ‘he has walked’ is phonetically [u wal tã]. The full form of the content word is written.

**Vowel co-occurrence**

Kambari (any of the varieties, including Agwara Kambari) does not have vowel harmony in the same way Igbo does. Igbo has two sets of vowels, and the vowels within a word can only draw from the vowels belonging to one set or the other, but never from both sets within the same word. Kambari restricts certain vowels from occurring within the same word as other vowels. The pattern is this: any word can have any number of the same low vowel (*a, ã, e, or o*), but no two different low vowels can occur in the same word. That is to say, a word can have one ‘e’ as in *de* ‘there’, two ‘e’s as in *ene* ‘to see’, three ‘e’s as in *kedele* ‘gourd’, etc., but no word will mix ‘e’ with ‘a’, ‘e’ with ‘ã’, ‘e’ with ‘o’, or ‘a’ with ‘o’, etc. There are no restrictions on the occurrence or distribution of the high vowels ‘i’ and ‘u’.

Hausa and English both allow vowels to be mixed within words with no restrictions. This has implications when Kambari borrows words from English and Hausa. *Bokiti* ‘bucket’ is an example of word borrowing which conforms to the Kambari rule of co-occurrence.

**Consonants**

As was mentioned earlier, there are twenty-five consonants in the alphabet, with the addition of the glottal stop, there are twenty-six. Labialized consonants (*'w*) and palatalized consonants (*ly, ky, dy, ty, gy, 'y*) have been discussed earlier. Thus the consonants are as follows:

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labio-</th>
<th>Alveolar</th>
<th>Post-</th>
<th>Palatal</th>
<th>Velar</th>
<th>Labio-</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p</td>
<td>t</td>
<td></td>
<td>k</td>
<td>kp</td>
<td></td>
<td>?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>d</td>
<td></td>
<td>g</td>
<td>gb</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implosive</td>
<td>ð</td>
<td>ð</td>
<td></td>
<td>h</td>
<td></td>
<td>j</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f</td>
<td>s</td>
<td></td>
<td>v</td>
<td>z</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td>ts</td>
<td>c</td>
<td></td>
<td>j</td>
<td></td>
<td>y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td>ð</td>
<td></td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Long consonants

In Agwara Kambari, long consonants have the same status as long vowels. They exist phonetically, but there is hardly any contrast between words with long consonants and words with short consonants. For example [dakka] ‘under/south’ has a long consonant, but there is no word in Agwara Kambari *[daka] to contrast. The few pairs that do contrast do not occur in the same part of speech or the same semantic domain. In the KLP orthography, few long consonants are written, e.g. *kanna ‘day’, *nannai ‘like that’.

Palatalization

In addition to the ‘y’ following certain consonants as described earlier (ikyamba ‘body’, kagulya ‘earthworm’ etc.), there is an additional palatalization process. In certain contexts, when ‘s’, ‘ts’ or ‘z’ is followed by either ‘i’ or ‘e’ (the front vowels), they become ‘sh’, ‘c’, or ‘j’, respectively. For example, keceshe /kece + sa/ ‘to count-IT’, and *aza e re [eje re] ‘two [somethings]’.

Nasalized vowels

Because /m/ can appear at the end of a syllable or word, but /n/ never does, in the KLP orthography, what sound like nasalized vowels (a vowel where part of the sound comes through the nose) have been interpreted as a sequence of a vowel followed by /n/. In reading Cishingini, ‘n’ at the end of a word is pronounced through the nose instead of as the Hausa nan ‘here’ or the English ‘man’. For example:

- *kaya’an [kajija] ‘termite’ (in contrast to kaya’a [kaja] ‘arrow’) katsgn [kaitsa] ‘bellows’ (in contrast to katsa [kaitsa] ‘knock’)
- *kuntsun [kuntsu] ‘forest’ (in contrast to kuntsu [kuntsu] ‘sore’)
- *mashin [maʃi] ‘pestle’ (in contrast to mashi [maʃi] ‘crying’)
- *kojon [kodjo] ‘bog’ (in contrast to gbojo [gbojo] ‘become wise’)
- keven [keve] ‘corpse’ (in contrast to yeve [yeve] ‘know’)

This vowel nasalization is always represented in the orthography by an ‘n’ word finally. If the root of a word ends with this kind of nasalization, and a suffix is added, the ‘n’ is still written word finally. For example: *iwan + i → *iwain [iwaan] ‘refused’.

Syllabic nasals and nasals before consonants

In Agwara Kambari, at the beginning of words, ‘m’ and ‘n’ coming before another consonant acts like a syllable. (mogono ‘king’ has three syllables; ngono ‘kings’ also has three syllables.) This ‘m’ or ‘n’ is written as ‘m’ before labial consonants (consonants made with
the lips: m, p, b, ɓ, gb, and kp) and ‘n’ before all other consonants. This pattern of ‘m’ before m, p, b, ɓ, gb, and kp, and ‘n’ before all other consonants is also used within a word, and if the ‘n’ or ‘m’ is a single word which is spaced from the following word:

- mpsa ‘blood’
- ndanga ‘trees’ (syllabic nasal)
- cimba ‘to farm’
- banka ‘to take to someone’ (nasal within a word)
- m batsai ‘I split’
- n danai ‘I said’ (freestanding pronoun)

**Tone**

Agwara Kambari has a small number of minimal tone pairs (under 20). Most of the minimal pairs are in different domains, so there is little possibility of confusion when reading. For example ci\text{\textacute{\textl}}n [ci\text{\textacute{\textl}}n] ‘clay’ and ci\text{\textacute{\textl}}n [ci\text{\textacute{\textl}}n] ‘intestines’.

Reader tests of tone in marked and unmarked texts have shown that no information is lost with the unmarked texts. Furthermore, there is reader resistance to tone marks, shown as a reluctance to try to read tone marked texts. If reading tests showed significant amounts of information were confusing without marked tone, KLP would propose including tone marks. Such is not the case; therefore, tone is unmarked in the Agwara Kambari orthography.

**Word spacing**

As a written form of a language is introduced to the community, the transfer from a complete thought to a linear sequence of individual words is not intuitive. Consistency in writing involves spelling the same word the same way every time and spacing between words in a consistent manner. The Agwara Kambari orthography operates on the principle that the ‘big words’ (content words: nouns, verbs, adjectives, adverbs) are spelled the same way every time. The ‘little words’ (functors: pronouns, prepositions, verbal morphemes) alter their spelling to allow the written form to more closely reflect the oral form. One of the ways of ensuring consistency in the big words is to always leave a space between the little words and the big words:

1. *U walai.* ‘He walked.’
2. *U tsu wala tg.* ‘He has always walked.’
3. *U danai le.* ‘He told them.’
4. *U dana le tg.* ‘He has told them.’
5. *M bana tg.* ‘I have gone.’ (This is visually different from *mbana* ‘baskets’.)
6. *Wi tə n kedele kə vg.* ‘He has my gourd.’
   Literally ‘He is with gourd of mine.’

7. *Wi tə n ikebe i vg.* ‘He has my money.’

8. *Wi tə m mbana mə vg.* ‘He has my basket.’

9. *Wi tə nu mbana n vg.* ‘He has my baskets.’

10. *Wi tə n ikebe ya Agbaru.* ‘He has Agbaru’s money.’

11. *Wi tə n umolu wi llisingai.* ‘He has Ilisingai’s guitar.’

   In sentences 6–11, notice that the possessive is made up of a noun, a NCM (noun concord marker, e.g. the *kg* following *kedele* ‘gourd’ in sentence 6), and the possessive pronoun or noun (*kedele kg vg*). There is a space between each of these. Notice also that the ‘with’ is represented by *m* before consonants made with the lips (*m mbana*), *n* before vowels (*n ikebe*) and other consonants (*n kedele*) and *nu* before syllabic nasals (*nu mbana*). Readers are taught to blend the vowels into one syllable (*n ikebe* sounds like *[nikebe]*). Notice the NCMs in sentences 10 and 11 (*‘ya’* and *‘wi’*). If the word following the NCM begins with a consonant, the NCM would be ‘i’ and ‘u’, respectively (*ikebe i Cindanai, umolu u Cindanai*). When the NCM comes before words beginning with vowels, the ‘i’ and ‘u’ become ‘y’ and ‘w’ and form a syllable with the vowel of the following word.

**Problems still to be considered**

The multiplicity of functor shapes (grammatical words). For example, *tsu* in Agwara Kambari has seven meanings/functions: first-person plural subject pronoun, first-person plural object pronoun, first-person plural possessive pronoun (the corresponding first-person singular pronouns are */N/*, *mu*, *vg*, respectively), repetitive action marker (*U tsu wala tə.*) ‘He has always walked.’), noun-class prefix (*tsu-kulu* ‘raffia palm tree’, *tsu-kpe* ‘okro’), NCM (*tsukpe*e tsu gbain ‘big okro’), and a morpheme denoting comparison/manner (as/how). As of now, most of the functors in the language have between four and seven orthographic representations; for example, the subject pronoun *u* ‘he’ may appear in texts as *u, wa, vg, we, wi, wo* or *wu*. If the following verb begins with a consonant, the *u* form is used (*u ronoi* ‘he drew’); if the verb begins with a vowel, the *u* changes to *w* and the vowel of the verb is copied backwards to make a pronounceable word; for example, *we ecei* ‘he asked’, *wo o’oi* ‘he tired’, *wu uma* ‘he killed’. Would it be beneficial in the long run to only use the underlying form (in the above example, *u*), and diligently teach natural pronunciation? Thus the above examples would be written *u ecei* ‘he asked’, *u o’oi* ‘he tired, *u uma* ‘he killed’.
Note: This orthography write-up was based on seven years of using a trial orthography, and every new book was passed by a committee of reviewers who had been trained in both how to read the orthography and what is expected of a reviewer. At the time of the orthography submission to the Language Development Center, the orthography will have been approved by the KLP Committee, by various influential people within each language area, and by a combined (representatives from all three languages) session of an Orthography Approval Committee. On April 17, 1999, the KLP Committee approved the Agwara Kambari Orthography. A booklet, *Uye u Kuya'an Idani i Cishingini* (*Presenting the Cishingini Orthography*) containing stories, a pictorial alphabet, and a brief explanation of some general orthography principles, is being circulated to the community leaders starting in May 1999.
A sample story in the Agwara Kambari orthography

Mامةي му Ulyuwa


U yaŋә bacy tuwu'a, u даңɡu bacy a kulyә'a. U lyuku bacy unә u nam pa u nan de u dana аyi u ciya ba. U lyuku bacy u nam pa u nan de u dana аyi u ciya ba.

Ayn tutu ta u ci yongo a kuya'an nan. Катьа mуku n kана kubana mi indәй ni. Катьа u kана kukiyangу le.

The Greedy Old Woman

Once there was an old woman. She was a widow and all alone. Every day she would be looking for food for herself. Even if she got plenty of food, it would not be enough. That is because she had many mouths. If she cooked tuwo, she would sit and be eating it. If she put a bite in this mouth, that mouth would say it didn't get any. If she put a bite in another mouth, a mouth there would say it didn't get any.

Every day she would spend the day doing like that. Then the children started going to look at her. Then she would get up and chase them away.
**Auna orthography**

*Introduction*

The language of the Auna area of Niger and Kebbi States in the west of Nigeria's Middle Belt is part of the Kambari cluster and is referred to in the literature as Auna Area Kambari, and locally as Tsikimba. The people who speak Tsikimba are called Akimba (Kakimba is the singular form). The Akimba people refer to anyone who is ethnically Kambari (speakers of Agwara Area Kambari, Salka Area Kambari, and others, including themselves) as Ashingini (plural of Kashingini ‘Kambari person’). Tsishingini, to the Akimba, refers to the Kambari language cluster. Because these terms are fluid, meaning different things to different people, the Tsikimba language will be referred to as Auna Kambari in this book.

Crozier and Blench (1992:118) give the Kambari cluster the following language family classification: Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Kainji, Western, Kambari. Auna Kambari was recognized as a language distinct from Agwara Kambari in the thirteenth edition of the *Ethnologue*, a listing of all the languages of the world (Grimes 1996). Auna Kambari has three major dialects, Tsikimba (spoken around Auna and Wara towns), Tsishen (spoken along the edge of Lake Kainji south of Wara), and Tsugaunshe (also spoken around Auna town). The word “dialect” is used here to refer to two or more speech forms whose speakers recognize differences in pronunciation and vocabulary, but when each speaks his own form, each is understood by the other. Other languages of the Kambari language cluster include: Cishingini (Agwara area on the west side of Lake Kainji and from Yelwa to Nasko on the east side of Lake Kainji), Tsishingini (Salka, Ibeto, and surrounding villages), and Tsuvadi (east of Yelwa, Rijau, north of Kontagora). The languages of the Kambari language cluster share an ethnopolitical identity, but speakers of the different languages of the cluster do not understand one another and use Hausa to communicate. This description is based on the Tsikimba dialect as is spoken in Auna town.

In 1992 leaders of the Auna Kambari, Agwara Kambari, and Salka Kambari formed the Kambari Language Project (KLP). Emmanuel D. Dangana was selected to attend the 1992 Introductory Course in Applied Linguistics (ICAL) sponsored by the Nigeria Bible Translation Trust in Jos. In 1998 John Isaiah attended ICAL and joined the KLP staff. John and Janie Stark, who are both reading Ph.D. in Linguistics at the University of Ilorin, have been KLP technical advisors since 1992. A trial orthography has been developed and used by the Kambari Language Project (KLP) to produce written materials, including a primer series to teach those who do not know how to read, a transition primer series for those who can
already read in Hausa, several reading books and some Bible passages. Until KLP, no
language analysis or development had taken place.

This book follows the outline suggested by Salami (1996).

**The philosophical basis of the KLP orthography**

The KLP orthography divides words into two groups: content words and
grammatical words. Content words are nouns, verbs, adjectives, etc.; words which for the
most part are easy to isolate, visualize, define, and translate to or from another language.
Grammatical words in Auna Kambari are tense and aspect markers, NCM (noun concord
markers linking the head noun to its modifier), certain pronouns and prepositions. In the
KLP orthography for Auna Kambari, each content word and each grammatical word has only
one spelling. The KLP orthography is based on *meaning*, and the reader automatically
supplies the phonetic alterations, e.g. The phrase written *ikebe i le ‘their money (literally
‘money, NCM, their’)*’ is pronounced [ikebile]. The phrase written *ikebe i acimbi ‘farmers’
money (literally money, NCM, farmers)*’ is pronounced [ikebe yacimbi].

**The current alphabet**

This is the alphabet in its alphabetical order with sounds in the International
Phonetic Alphabet (IPA), a sample word in the language and its gloss in English. The English
gloss of verbs is given in the infinitive, but the Auna Kambari form is actually the singular
imperative.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Sound</th>
<th>Word</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>[a]</td>
<td>anaka</td>
<td>‘cows’</td>
</tr>
<tr>
<td>ə</td>
<td>[ɔ]</td>
<td>agata</td>
<td>‘river’</td>
</tr>
<tr>
<td>b</td>
<td>[b]</td>
<td>baba</td>
<td>‘baby’</td>
</tr>
<tr>
<td>b ̣</td>
<td>[ɓ]</td>
<td>batsa</td>
<td>‘to split, to branch’</td>
</tr>
<tr>
<td>c</td>
<td>[tʃ]</td>
<td>ciga</td>
<td>‘to want, to love’</td>
</tr>
<tr>
<td>d</td>
<td>[d]</td>
<td>dana</td>
<td>‘to say’</td>
</tr>
<tr>
<td>d ̣</td>
<td>[ɗ]</td>
<td>madanga</td>
<td>‘tree’</td>
</tr>
<tr>
<td>e</td>
<td>[c]</td>
<td>kadele</td>
<td>‘gourd’</td>
</tr>
<tr>
<td>f</td>
<td>[f]</td>
<td>fofo</td>
<td>‘to remove’</td>
</tr>
<tr>
<td>g</td>
<td>[g]</td>
<td>gebe</td>
<td>‘to push’</td>
</tr>
<tr>
<td>gb</td>
<td>[ɡb]</td>
<td>gbamgbara</td>
<td>‘cassava’</td>
</tr>
<tr>
<td>h</td>
<td>[h]</td>
<td>hoga</td>
<td>‘hello (greeting)’</td>
</tr>
<tr>
<td>i</td>
<td>[i]</td>
<td>ikani</td>
<td>‘chains’</td>
</tr>
<tr>
<td>j</td>
<td>[dʒ]</td>
<td>jila</td>
<td>‘sew’</td>
</tr>
<tr>
<td>k</td>
<td>[k]</td>
<td>kakamba</td>
<td>‘bush’</td>
</tr>
<tr>
<td>kp</td>
<td>[kp]</td>
<td>kpatsu</td>
<td>‘boat’</td>
</tr>
<tr>
<td>l</td>
<td>[l]</td>
<td>lapa</td>
<td>‘to beat’</td>
</tr>
<tr>
<td>m</td>
<td>[m]</td>
<td>magono</td>
<td>‘king/chief’</td>
</tr>
<tr>
<td>n</td>
<td>[n]</td>
<td>namba</td>
<td>‘to lack’</td>
</tr>
<tr>
<td>o</td>
<td>[o]</td>
<td>apofo</td>
<td>‘toads’</td>
</tr>
<tr>
<td>p</td>
<td>[p]</td>
<td>pasa</td>
<td>‘to cross’</td>
</tr>
<tr>
<td>r</td>
<td>[r]</td>
<td>ranqa</td>
<td>‘to cook’</td>
</tr>
<tr>
<td>s</td>
<td>[s]</td>
<td>sala</td>
<td>‘to shout’</td>
</tr>
<tr>
<td>sh</td>
<td>[ʃ]</td>
<td>shana</td>
<td>‘to fill’</td>
</tr>
<tr>
<td>t</td>
<td>[t]</td>
<td>teli</td>
<td>‘six’</td>
</tr>
<tr>
<td>ts</td>
<td>[ts]</td>
<td>tsuwa</td>
<td>‘starch food (tuwo)’</td>
</tr>
<tr>
<td>u</td>
<td>[u]</td>
<td>ugeve</td>
<td>‘kind of antelope’</td>
</tr>
<tr>
<td>v</td>
<td>[v]</td>
<td>vara</td>
<td>‘to throw’</td>
</tr>
</tbody>
</table>
In the Auna Kambari alphabet there are six vowels and twenty-five consonants, making a total of thirty-one letters. All of the symbols used are common to English and Hausa except for the vowel Ñ. The symbols shared by Kambari and Hausa, though not English are ɓ, ɗ, and ḥ (glottal stop). The letters ɓ and ɗ are the only letters which are not easily made by an ordinary typewriter. The digraphs kp, gb, sh, and ts function as a single consonant. The glottal stop [‘] (alhamza in Hausa) which looks like an apostrophe, is not included in the set of letters called the alphabet, although it is used as a break between vowels the same way as in the Hausa words addu ‘prayer’, shari ‘law’. The symbol for the glottal stop [‘] is called ma’imkpatsu ‘thing used for jumping’ in Auna Kambari, and is used as in the word a’aci ‘heads’. Auna Kambari allows ‘y’ to follow certain consonants as in ikyamba ‘body’, kakulya ‘guinea fowl’, kadfa ‘dove, makparyangi ‘donkey’, udyoku ‘neck’, kagyara ‘a kind of grass’, katya ‘garden egg’, and unyama ‘meat’. Hausa has a glottalized y [‘j] as a separate letter (as in ‘ya ‘daughter’). Auna Kambari also has words with the ‘y [‘j] sound as in ‘yuwan ‘to refuse’.

Vowels

The six vowels of Auna Kambari are as follows:

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unrounded</td>
<td>Unrounded</td>
<td>Rounded</td>
</tr>
<tr>
<td>Close (high)</td>
<td>i</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td>a</td>
<td>o</td>
</tr>
<tr>
<td>Open (low)</td>
<td></td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

Diphthongs

Diphthongs are two different vowels next to each other within a word. In Auna Kambari these are restricted to a low vowel (a, ɑ, e, or o) followed by a high vowel (i or u). Except for a handful of words, diphthongs are formed by adding the –i suffix (meaning either completed action or plural command) to the end of a verb. For example wene ‘to see’ becomes wenei ‘saw’, rawa ‘arrive’, rawai ‘arrived’, tuw ‘come’, tuwi ‘came’, rono ‘draw’, ronoi ‘drew’.
Long vowels

In the KLP orthography for Auna Kambari, the only word containing a long (double) vowel is *kaara* ‘to make lines’ in contrast to *kara* ‘to tear’.

Vowel processes

There are five different ways vowels interact with each other both within words and across word boundaries.

Vowel copying

In Auna Kambari, certain suffixes contain vowels that derive their pronunciation from nearby vowels. In this process, right most vowel of the root acts as the source or control of the sound of the vowel of the suffix. In the right column, the change in the vowel preceding the ‘-sV’ ‘iterative (IT) will be discussed in the next section. The explanation of the ‘-s’ changing to ‘sh’ will be given later.

| dana ‘to say’ | dangsa (dana + sa) ‘to say-IT’ |
| tuwa ‘to come’ | tuwasa (tuwa + sa) ‘to come-IT’ |
| kece ‘to count’ | kecishe (kece + sa) ‘to count-IT’ |
| soro ‘to pound’ | soruso (soro + sa) ‘to pound-IT’ |
| suku ‘to send’ | sukusu (suku + sa) ‘to send-IT’ |

Unstressed /u/

When a suffix is added to the root, the last vowel of the root is replaced with /u/. In this position, the vowel does not receive the same length and “importance” as the same vowel in other positions. In Auna Kambari, rejecting a single representation of /u/ and its allophone [i], they have chosen to write the orthographic vowel in the highest position of the column of the vowel chart where the vowel of the following syllable occurs, i.e. if the vowel of the following syllable is ‘i’ or ‘e’ (both front vowels), write ‘i’; if a central vowel (a, a), write ‘a’, and if a back vowel (u, o, or labial consonant), write ‘u’ (see examples in the right column above).

Vowel elision

Vowel elision is the process where when two vowels come together (usually the last vowel of one word and the vowel beginning the next word), the first vowel deletes, and only the second vowel is pronounced. For example: *ikebe i le* ‘their money (literally money, NCM,
their)’ is pronounced [ikebile]. In the orthography, the content word is preserved intact, and the NCM stands on its own.

Another instance of vowel elision in Auna Kambari is across morpheme boundaries when the consonant before the final vowel of the first morpheme is a sonorant (l, r, n, w) and the second morpheme begins with a non-sonorant consonant, usually with the same point of articulation as the sonorant. In this setting, the final vowel of the first morpheme deletes, and the final consonant of the first morpheme becomes the last letter of the syllable preceding it. For example, the sentence written: *U bana t* ð ‘he has gone’ is phonetically [u ban tå]. *U wala t* ð ‘he has walked’ is phonetically [u wal tå]. In the orthography, the full form of the content word is written.

**Vowel coalescence**

Vowel coalescence is where when two vowels come together (the first is *a* and the second is either *i* or *u*), instead of either vowel, a third vowel is pronounced. If the sequence is ‘*a*’ followed by ‘*i*’, ‘*e*’ will be pronounced. If the sequence is ‘*a*’ followed by ‘*u*’, ‘*o*’ will be pronounced. Though the orthography retains both vowels, readers are taught to pronounce naturally. For example: *a ikyamba* ‘at the body’ is pronounced [ekyamba]; *a uyamba* ‘at the land’ is pronounced [oyamba].

**Vowel co-occurrence**

Kambari (any of the varieties, including Auna Kambari) does not have vowel harmony in the same way Igbo does. Igbo has two sets of vowels, and the vowels within a word can only draw from the vowels belonging to one set or the other, but never from both sets within the same word. Kambari restricts certain vowels from occurring within the same word as other vowels. The pattern is this: any word/root can have any number of the same low vowel (*a*, *g*, *e*, or *o*), but no two different low vowels can occur in the same word/root. That is to say, a word can have one ‘*e*’ as in *de* ‘there’, two ‘*e*’s as in *wene* ‘to see’, *webele* ‘to look at’, etc., but no word will mix ‘*e*’ with ‘*a*’, ‘*e*’ with ‘*g*’, ‘*e*’ with ‘*o*’, or ‘*a*’ with ‘*o*’, etc. In the case of nouns, a noun is made up of a noun class prefix and a noun root (*ka-dele* ‘gourd’ *ma-gono* ‘king’). These noun class prefixes are always ‘*ka*’ and ‘*ma*’, and the vowel co-occurrence rule applies to the root only. There are no restrictions on the occurrence and distribution of the high vowels ‘*i*’ and ‘*u*’. The writing of words like *dangsa* ‘say over and over’, though an apparent contradiction of the vowel co-occurrence rule, are actually a pointer to the reader that the ‘*g*’ is phonetically different (it is an orthographic approximation of the sound of the allophone [i]).
Hausa and English both allow vowels to be mixed within words with no restrictions. This has implications when Kambari borrows words from English and Hausa. *Bokiti* 'bucket' is an example of word borrowing which conforms to the Kambari rule of co-occurrence.

**Consonants**

As was mentioned earlier, there are twenty-five consonants in the alphabet, with the addition of the glottal stop, there are twenty-six. Palatalized consonants (ly, ky, dy, dy, ty, gy, 'y) have been discussed earlier. Thus, the consonants are as follows:

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Alveolar</th>
<th>Post-alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Labio-velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td>kp</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>d</td>
<td>g</td>
<td>gb</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implosive</td>
<td>ð</td>
<td>ð</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f</td>
<td>s</td>
<td>j</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>v</td>
<td>z</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td>ts</td>
<td>c</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>j</td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td>ř</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
</tr>
<tr>
<td>Lateral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Long consonants**

In the KLP orthography, long (double) consonants are written only in the following word: *kulla* ‘eight’.

**Palatalization**

In addition to the ‘y’ following certain consonants as described earlier (*ikyamba* ‘body’, *kakulya* ‘guinea fowl’ etc.), there is an additional palatalization process. In certain contexts, when ‘s’, ‘ts’, or ‘z’, is followed by either ‘i’ or ‘e’, they become ‘sh’, ‘c’, or ‘j’, respectively. For example, *kecishe* (*kece + sa*) ‘to count-IT’, *laji* (*laz* (from *laza*) + *i*) ‘departing (simultaneously with another action)’.
**Nasalized vowels**

Because /m/ can appear at the end of a syllable or word, but /n/ never does, in the KLP orthography, what sound like nasalized vowels (a vowel where part of the sound comes through the nose) have been interpreted as a sequence of a vowel followed by /n/. In reading Tsikimba, ‘n’ at the end of a word is pronounced through the nose instead of as the Hausa nan ‘here’ or the English ‘man’. For example:

- *karan* [kārā] ‘termite’ (in contrast to *kara* [kara] ‘to tear’)
- *kutsan* [kuʦₐ] ‘yearning’ (in contrast to *kutsₐ* [kuʦₐ] ‘to plait’)
- *kalen* [kalē] ‘usefulness’ (kale [kale] ‘sheanut tree’)
- *kacin* [kaici] ‘kind of pox’ (kaci [kaici] ‘head’)
- *kakoton* [kakotō] ‘snail’ (kakoto [kakoto] ‘back of head’)
- *makun* [makū] ‘cap’ (maku [maku] ‘child’)

This vowel nasalization is always represented in the orthography by an ‘n’ word finally. If the root of a word ends with this kind of nasalization, and a suffix is added, the ‘n’ is still written word finally. For example: *yan + i* → *yain* [jain] ‘did’.

**Syllabic nasals and nasals before consonants**

In Auna Kambari, at the beginning of words, ‘m’ and ‘n’ coming before another consonant acts like a syllable (magono ‘king’ has three syllables; ngono ‘kings’ also has three syllables). This ‘m’ or ‘n’ is written as ‘m’ before labial consonants (consonants made with the lips: m, p, b, ɓ, gb, and kp) and ‘n’ before all other consonants. This pattern of ‘m’ before m, p, b, ɓ, gb, and kp, and ‘n’ before all other consonants is also used within a word. If the ‘n’ means ‘I’, ‘with’, or is a NCM, it is spaced from the following word and always written ‘n’, even before labial consonants:

- *mpasa* ‘blood’
- *ndanga* ‘trees’ (syllabic nasal)
- *cimba* ‘to farm’
- *banka* ‘to take to someone’
  (nasal within a word)
- *n batsai* ‘I branched’
- *n danai* ‘I said’ (freestanding)

**Tone**

Auna Kambari has a small number of minimal tone pairs (under 20). Most of the minimal pairs are in different domains, so there is little possibility of confusion when reading. For example *kaya* [kàyā] ‘arrow’ and *kaya* [kàyā] ‘open ridge’, or *kagbodo* [kàgbòdò] ‘bicycle hub’ and *kagbodo* [kàgbòdò] ‘hand-woven cloth’. Tone is not written.

Reader tests of tone in marked and unmarked texts have shown that no information is lost with the unmarked texts. Furthermore, there is reader resistance to tone marks, shown
as a reluctance to try to read tone marked texts. If reading tests showed significant amounts of information were confusing without marked tone, KLP would propose including tone marks. Such is not the case; therefore, tone is unmarked in the Auna Kambari orthography.

**Word spacing**

As a written form of a language is introduced to the community, the transfer from a complete thought to a linear sequence of individual words is not intuitive. Consistency in writing involves spelling the same word the same way every time and spacing between words in a consistent manner. The Auna Kambari orthography operates on the principle that each word has its one correct spelling. One of the ways of ensuring consistency in spelling is to always leave a space between the little words and the big words.

1. *U walai.* ‘He walked.’
2. *U tsu wala.* ‘He always walked.
3. *Wa wala ta.* ‘He will walk.’
4. *U tonuko le.* ‘He told them.’
5. *U tonuko le ta.* ‘He has told them.’
6. *N bana ta.* ‘I have gone.’ (This is visually different from *mbana* ‘baskets’.)
7. *U banai a kashina.* ‘He went to the farm.’ (*u* spaced from the verb is the pronoun ‘he’.)
8. *Alobo a kagai ubansa ara ne.* ‘The boys kept going to where she was.’ (*u* joined to the verb is continued action—similar to the –ing of ‘going’ in English.)
9. *Wg ta n kadele ka vg.* ‘He has my gourd.’ Literally ‘He is with gourd of mine.’
10. *Wg ta n ikebe i vg.* ‘He has my money.’
11. *Wg ta n mako ma vg.* ‘He has my drinking calabash.’
12. *Wg ta nko n vg.* ‘He has my drinking calabashes.’
13. *Wg ta n ikebe i Awaca.* ‘He has Awaca’s money.’
14. *Wg ta n ugeve u Awaca.* ‘He has Awaca’s antelope.’

In sentences 9–14, notice that the possessive is made up of a noun, a NCM, and the possessive pronoun or noun (*kadele ka vg*). There is a space between each of these. Notice
also that the ‘with’ is represented by a freestanding ‘n’. Readers are taught to blend the vowels into one syllable (n ikebe sounds like [nikebe]). Notice the NCMs in sentences 13 and 14 (‘i’ and ‘u’). When these NCMs come before words beginning with vowels, though written ‘i’ and ‘u’, the ‘i’ and ‘u’ sound like ‘y’ and ‘w’ joined to the vowel of the following word (i Awaca [yawaca] and u Awaca [wawaca]).

Problems still to be considered

The verb ‘is’ is the single vowel á, or in the past tense gri. Should the pronouns be attached to the verb, making ‘is’ have multiple spellings, or should they be separated with a space, thus retaining the unique spelling, but appearing awkward to read: (see next page)

<table>
<thead>
<tr>
<th>Pronoun joined</th>
<th>Pronoun spaced</th>
<th>Pronoun joined</th>
<th>Pronoun spaced</th>
</tr>
</thead>
<tbody>
<tr>
<td>mg ‘I am’</td>
<td>m á ‘I am’</td>
<td>mgri ‘I had been’</td>
<td>m gri ‘I had been’</td>
</tr>
<tr>
<td>wg ‘you are’</td>
<td>vu á ‘you are’</td>
<td>wgri ‘you had been’</td>
<td>vu gri ‘you had been’</td>
</tr>
<tr>
<td>wá ‘he is’</td>
<td>u á ‘he is’</td>
<td>wári ‘he had been’</td>
<td>u gri ‘he had been’</td>
</tr>
<tr>
<td>tsá ‘we are’</td>
<td>tsu á ‘we are’</td>
<td>tsári ‘we had been’</td>
<td>tsu gri ‘we had been’</td>
</tr>
<tr>
<td>i’gá ‘you PL are’</td>
<td>i á ‘you PL are’</td>
<td>i’gári ‘you had been’</td>
<td>i gri ‘you had been’</td>
</tr>
<tr>
<td>a’gá ‘they are’</td>
<td>a á ‘they are’</td>
<td>a’gári ‘they had been’</td>
<td>a gri ‘they had been’</td>
</tr>
</tbody>
</table>

KLP recommends the ‘pronoun joined’ option. The verb ‘is/to be’ constitutes one of the most irregular elements of nearly every language. In Tsikimba it is the only verb beginning with a vowel. By joining the pronoun to the verb, a “compound word” is formed. These words form a small set. In all other places, the pronoun or NCM would be spaced from the following word.

Note: This orthography write-up was based on seven years of using a trial orthography, and every new book was passed by a committee of reviewers who had been trained in both how to read the orthography and what is expected of a reviewer. At the time of the orthography submission to the Language Development Center, the orthography will have been approved by the KLP Committee, by various influential people within each language area, and by a combined (representatives from all three languages) session of an Orthography Approval Committee. On April 17, 1999, the KLP Committee approved the Auna Kambari Orthography. A booklet, Ure u Ukorongusu u Tsikimba (Presenting the Tsikimba
Orthography) containing stories, a pictorial alphabet, and a brief explanation of some general orthography principles, is being circulated to the community leaders starting in May 1999.
A sample story in the Auna Kambari orthography

Kazāi ka Aminya

Kazāi ka aminya ka roku ka pini a Ndazu. La u wusha bacī aminya a ama, aku u bana agatā u ba u zāṣa. U zā bacī, aku u wecikpe a da pini a agatā yi. U laza u aşakā a da pini, bawu uza da wa wundika yi. Kain dem nala.

Kain ka te, an u wecikpe u așakai a da, u gono u rawa u cinai a pārā yi a da de. Pini nala, u uwa ubolo ili i 'ya wa tonuko ama a da a nākā yi kuzā ki.

The Clothes Washer

There was once a clothes washer in Ndazu town. He would receive people’s clothes, then he would go to the river and wash them. When he had washed them, then he would spread them beside the river to dry. He would leave them there without anyone to watch them for him. Every day he did like that.

One day, when he had spread the clothes out and left them, he returned to find that someone had stolen them. At that point he started trying to find something to tell the people who had given him their washing.
**Salka orthography**

*Introduction*

The language of the Salka area of Niger State in western Nigeria is part of the Kambari cluster and is referred to in the literature as Salka Area Kambari, and locally as Tsishingini. The people who speak Tsishingini are called Ashingini (Shingini is the singular form). The people of the Salka area also refer to anyone who is ethnically Kambari (speakers of Auna-Wara Area Kambari, Agwara Area Kambari, and others) as Ashingini. Therefore, to avoid confusion, in this book the language will be referred to as Salka Kambari.

Crozier and Blench (1992:118) give the Kambari cluster the following language family classification: Niger-Congo, Atlantic-Congo, Volta-Congo, Benue-Congo, Kainji, Western, Kambari. Salka Kambari has two major dialects, Tsishingini (spoken around Salka town) and Tsuva$h_i$ (spoken around Ibeto town). The word “dialect” is used here to refer to two or more speech forms whose speakers recognize differences in pronunciation and vocabulary, but when each speaks his own form, each is understood by the other. Other languages of the Kambari language cluster include: Tsikimba (Auna-Wara area), Cishingini (Agwara, Nasko and west of Yelwa area), and Tsuva$d_i$ (different from that of Ibeto; this Tsuva$d_i$ is spoken east of Yelwa, Rijau, north of Kontagora). The languages of the Kambari language cluster share an ethnopolitical identity, but speakers of the different languages of the cluster do not understand one another and use Hausa to communicate. This description is based on the Tsishingini dialect as is spoken in Salka town.

Most of the Salka Kambari people live in Magama Local Government of Niger State, Nigeria, in the area between Ibeto and Nasko. Other Salka Kambari speakers live in Mashegu and Borgu Local Governments of Niger State. Early language work in the Salka Area Kambari language was done by missionaries in the 1920s and 1930s when portions of the Bible were translated. No copies of that work are in existence today. For more history of the early missionary work in Salka, see the long essay by Wakaso Amos Mamman, (1998). In the 1960s, Hoffmann wrote some articles based on Salka Kambari (1963, 1964, 1965). Apparently there are still some texts that Hoffmann transcribed at the University of Ibadan. In the early 1980s Crozier researched and wrote a dissertation on Salka Area Kambari (1984). He also proposed an orthography which was published in the book *Alphabets of Africa* (Hartell 1993:234), but no books were ever published using it. Essentially, the orthography proposed by Crozier and the orthography now used by KLP are the same, and the differences will be discussed in the relevant sections. In 1992 leaders of the Salka Kambari joined with those of Auna-Wara Kambari and Agwara Kambari to form the Kambari Language Project (KLP). In 1993, the Salka group selected Wakaso Amos Mamman as the
person to focus on language development for Salka Kambari. In 1998 Danladi Matthew attended ICAL and joined KLP staff. Both Wakaso Mamman and Danladi Matthew received training at the Introductory Course in Applied Linguistics offered by Nigeria Bible Translation Trust, Jos. John and Janie Stark, who are both reading Ph.D. in Linguistics at the University of Ilorin, have been KLP technical advisors since 1992. A trial orthography has been developed and used by the KLP to produce a primer series to teach those who do not know how to read, a transition primer series for those who can already read in Hausa, several reading books and some Bible passages.

This book follows the outline suggested by Salami (1996).

**The philosophical basis of the KLP orthography**

The KLP orthography divides words into two groups: content words and grammatical words. Content words are nouns, verbs, adjectives, etc.; words which for the most part are easy to isolate, visualize, define, and translate to or from another language. Grammatical words in Salka Kambari are tense and aspect markers, noun class syllables, certain pronouns, and prepositions. In the KLP orthography, content words are spelled the same way every time they occur. Grammatical words are allowed to vary in their spelling to more closely approximate their pronunciation in that context. In this way, even with their multiple variations of spelling, there is still a limited set of grammatical words. Most of these are frequently used and therefore quickly learned. On the other hand, content words are an unlimited set (the language can add new words at any time); therefore, to restrict each content word to a single spelling insures that regardless of context, the meaning of that word will be apparent and quickly learned since its spelling never varies. This principle of content and grammatical words is easy to teach since all of the variable grammatical words are a single syllable; whereas nearly all of content words are multisyllabic.

**The current alphabet**

This is a current alphabet in its alphabetical order with sounds in the International Phonetic Alphabet (IPA), a sample word in the language and its gloss in English. The English gloss of verbs is given in the infinitive, but the Salka Kambari form is actually the singular imperative.

<table>
<thead>
<tr>
<th>a</th>
<th>[a]</th>
<th>agada</th>
<th>‘lizards’</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>[ɔ]</td>
<td>gmli</td>
<td>‘deities’</td>
</tr>
<tr>
<td>b</td>
<td>[b]</td>
<td>babs</td>
<td>‘baby’</td>
</tr>
<tr>
<td>b</td>
<td>[ɓ]</td>
<td>batsa</td>
<td>‘to split, to branch’</td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-------</td>
<td>----------------------</td>
</tr>
<tr>
<td>c</td>
<td>[ʧ]</td>
<td>ciga</td>
<td>‘to want, to love’</td>
</tr>
<tr>
<td>d</td>
<td>[d]</td>
<td>damma</td>
<td>‘to say’</td>
</tr>
<tr>
<td>d'</td>
<td>[d̟]</td>
<td>udolu</td>
<td>‘river’</td>
</tr>
<tr>
<td>e</td>
<td>[ɛ]</td>
<td>elele</td>
<td>‘stove’</td>
</tr>
<tr>
<td>f</td>
<td>[f̥]</td>
<td>faba</td>
<td>‘to beat’</td>
</tr>
<tr>
<td>g</td>
<td>[g]</td>
<td>igagu</td>
<td>‘termites’</td>
</tr>
<tr>
<td>gb</td>
<td>[ɡ̟]</td>
<td>gbagba</td>
<td>‘duck’</td>
</tr>
<tr>
<td>h</td>
<td>[h]</td>
<td>hawun</td>
<td>‘clear’</td>
</tr>
<tr>
<td>i</td>
<td>[i]</td>
<td>ili</td>
<td>‘thing’</td>
</tr>
<tr>
<td>j</td>
<td>[dʒ]</td>
<td>jon</td>
<td>‘cricket’</td>
</tr>
<tr>
<td>k</td>
<td>[k]</td>
<td>kgara</td>
<td>‘to depart’</td>
</tr>
<tr>
<td>kp</td>
<td>[kp]</td>
<td>kpakpa</td>
<td>‘climbing rope’</td>
</tr>
<tr>
<td>l</td>
<td>[l̥]</td>
<td>lolo</td>
<td>‘marriage’</td>
</tr>
<tr>
<td>m</td>
<td>[m̥]</td>
<td>mogono</td>
<td>‘king/chief’</td>
</tr>
<tr>
<td>n</td>
<td>[n̥]</td>
<td>nama</td>
<td>‘meat’</td>
</tr>
<tr>
<td>o</td>
<td>[ɔ̊]</td>
<td>opodo</td>
<td>‘toads’</td>
</tr>
<tr>
<td>p</td>
<td>[p̥]</td>
<td>papura</td>
<td>‘to rub’</td>
</tr>
<tr>
<td>r</td>
<td>[r̥]</td>
<td>aralu</td>
<td>‘roots’</td>
</tr>
<tr>
<td>s</td>
<td>[s̥]</td>
<td>sapa</td>
<td>‘to inject/stab’</td>
</tr>
<tr>
<td>sh</td>
<td>[ʃ̥]</td>
<td>shipuru</td>
<td>‘mosquito’</td>
</tr>
<tr>
<td>t</td>
<td>[t̥]</td>
<td>tawa</td>
<td>‘to come’</td>
</tr>
<tr>
<td>ts</td>
<td>[t̥s̥]</td>
<td>tsuwa</td>
<td>‘starch food (tuwo)’</td>
</tr>
<tr>
<td>u</td>
<td>[u̥]</td>
<td>umuna</td>
<td>‘cloth’</td>
</tr>
<tr>
<td>v</td>
<td>[v̥]</td>
<td>vali</td>
<td>‘man’</td>
</tr>
<tr>
<td>w</td>
<td>[w̥]</td>
<td>wala</td>
<td>‘to go, to walk’</td>
</tr>
<tr>
<td>y</td>
<td>[j̥]</td>
<td>yomo</td>
<td>‘monkey’</td>
</tr>
<tr>
<td>z</td>
<td>[z̥]</td>
<td>zuwa</td>
<td>‘to put’</td>
</tr>
</tbody>
</table>
In the Salka Kambari alphabet there are six vowels and twenty-five consonants, making a total of thirty-one letters. All of the symbols used are common to English and Hausa except for the vowel ‘ğ’. The symbols shared by Kambari and Hausa, though not English are ‘b, d, and t’ (glottal stop). The letters ñ and ò are the only letters which are not easily made by an ordinary typewriter. The digraphs ‘kp, gb, sh, and ts’ function as a single consonant. The glottal stop ‘[’ (alhamza in Hausa) which looks like an apostrophe, is not included in the set of letters called the alphabet, although it is used as a break between vowels the same way as in the Hausa words addu’a ‘prayer’ and shari’a ‘law’. The symbol for the glottal stop is called iliyaizuva ‘thing that is up’ in Salka Kambari, and is used as in the word ura'atsu ‘large water pot’. Salka Kambari allows ‘y’ and ‘w’ to follow certain consonants as in mikyon ‘smoke’, utyo ‘end’, kingyo ‘kneel’, odyo ‘a kind of fruit’, ilyalya ‘food’, and ukwan ‘skin’. Hausa has a glottalized y [ñ] as a separate letter (as in ‘ya ‘daughter’). Salka Kambari has words with both the ‘y’ [ñ] and ‘w’ [ñw] sounds, but the KLP orthography treats them as the iliyaizuva (’) followed by ‘y’ or ‘w’. Examples of these sounds are ‘yqwa ‘to go’ and u'wa ‘house’.

### Vowels

The six vowels of Salka Kambari are as follows:

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unrounded</td>
<td>Unrounded</td>
<td>Rounded</td>
</tr>
<tr>
<td>Close (high)</td>
<td>i</td>
<td>a</td>
<td>u</td>
</tr>
<tr>
<td>Mid</td>
<td>e</td>
<td>ã</td>
<td>o</td>
</tr>
<tr>
<td>Open</td>
<td>a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In this book, “high vowels” are ‘i’ and ‘u’. “Low vowels” are composed of the mid and open vowels, ‘e, ã, o, a’. It is to be noted that in Salka Kambari, the orthographic symbols ‘e’ and ‘o’ are pronounced [e] and [o], respectively, especially in careful speech. In fast speech, toward the ends of words, and when next to the vowels ‘i’ and ‘u’, the vowels are pronounced [e] and [o]. There is no contrast, i.e. no phonemic distinction between [e] and [e] or between [o] and [o]. They are allophonic, i.e. two pronunciations of the same letter in the same way in English the plural ‘s’ can be pronounced [s] or [z] as in [kats] ‘cats’ and [dogz] ‘dogs’. This is one of the differences between the KLP orthography and the Crozier orthography. Crozier listed both ‘e’ and ‘o’ as well as ‘e’ and ‘o’. In the KLP orthography, only the more common symbols ‘e’ and ‘o’ are used.
Another difference between the Crozier orthography and the KLP orthography is that Crozier included an additional vowel ‘i’ (often called ‘barred i’), but that vowel is not an additional letter in the alphabet (i.e. phoneme), but a special pronunciation of ‘u’ following certain consonants, or in certain positions such as immediately before a suffix. In the KLP orthography, this ‘barred i’ will be written as ‘u’ if the following consonant is labial (m, p, b, f, gb, kp as in tsugbain ‘bigness’), or if the vowel in the next syllable is ‘o’ or ‘u’ (tsugono ‘kingship’). The ‘barred i’ will be written ‘a’ if the vowel in the following syllable is ‘a’ or ‘e’ (tsgson ‘okro’, sgsg ‘to sift’). It will be written ‘i’ if the vowel in the following syllable is ‘i’ or ‘e’ (tsisewun ‘rivalry/hatred’).

**Diphthongs**

Diphthongs are two different vowels next to each other within a word. In Salka Kambari these are restricted to a low vowel (a, g, e, or o) followed by a high vowel (i or u). Except for a handful of words, diphthongs are formed by adding the –i suffix (meaning either completed action or plural command) to the end of a verb. For example ene ‘to see’ becomes enei ‘saw’, rawa ‘arrive’, rawai ‘arrived’, tawg ‘come’, tawgi ‘came’, rono ‘draw’, ronoi ‘drew’.

**Long vowels**

There are pairs of words in Salka Kambari in which the only difference is whether a vowel is short, or whether it is dragged out (takes longer to say). Hausa has this distinction, but it is not preserved in the Hausa orthography, as in gari. When it means ‘town’ it is pronounced differently than when it means ‘flour’. In Salka Kambari, only when there are pairs of words which might confuse the meaning intended by the writer are long vowels written with double vowels. As of now, there is a restricted set of lexical items, each with a short vowel counterpart.

<table>
<thead>
<tr>
<th>Long vowel</th>
<th>Short vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>aapa ‘to suck breast’</td>
<td>apa ‘here’</td>
</tr>
<tr>
<td>goon ‘only’</td>
<td>gon ‘fishhook’</td>
</tr>
<tr>
<td>jaara ‘to be insane’</td>
<td>jara ‘to smash’</td>
</tr>
<tr>
<td>kggra ‘to draw a line’</td>
<td>kgra ‘to depart’</td>
</tr>
<tr>
<td>laamba ‘number’</td>
<td>lamba ‘to lack’</td>
</tr>
<tr>
<td>paala ‘to cheat’</td>
<td>pala ‘to cover’</td>
</tr>
<tr>
<td>shaama ‘to vomit’</td>
<td>shama ‘to rot’</td>
</tr>
</tbody>
</table>
Also in the language, there are prefixes and suffixes whose presence is indicated by doubling the vowel next to it. In certain nouns this occurs at the beginning of a word and signifies singular or plural.

<table>
<thead>
<tr>
<th>Long vowel (singular)</th>
<th>Short vowel (plural)</th>
</tr>
</thead>
<tbody>
<tr>
<td>aakala ‘cake’</td>
<td>akala ‘cakes’</td>
</tr>
<tr>
<td>gggaga ‘mouse’</td>
<td>gggaga ‘mice’</td>
</tr>
<tr>
<td>eeteje ‘rope’</td>
<td>eeteje ‘ropes’</td>
</tr>
<tr>
<td>ookowo ‘egg’</td>
<td>ookowo ‘eggs’</td>
</tr>
</tbody>
</table>

In verbs this occurs at the end of the word and indicates an action ‘with something’.

<table>
<thead>
<tr>
<th>Short vowel</th>
<th>Long vowel (‘with’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tsula ‘to buy’</td>
<td>tsulaa ‘to buy with’</td>
</tr>
<tr>
<td>tgwq ‘to come’</td>
<td>tgwq ‘to bring (come with)’</td>
</tr>
<tr>
<td>e ‘to see’</td>
<td>ene ‘to see with’</td>
</tr>
<tr>
<td>rono ‘to draw’</td>
<td>ronoo ‘to draw with’</td>
</tr>
</tbody>
</table>

Another area where long vowels are used to distinguish meaning is that of pronouns denoting future action. The usual pronouns (not future) never have long vowels. The future pronouns are always written with a double vowel.

<table>
<thead>
<tr>
<th>Usual pronoun</th>
<th>Future pronoun</th>
</tr>
</thead>
<tbody>
<tr>
<td>a kamba ‘they help’</td>
<td>aa kamba ‘they will help’</td>
</tr>
<tr>
<td>g tgwq ‘they come’</td>
<td>gg tgwq ‘they will come’</td>
</tr>
<tr>
<td>e dewe ‘they call’</td>
<td>ee dewe ‘they will call’</td>
</tr>
<tr>
<td>o rono ‘they draw’</td>
<td>oo rono ‘they will draw’</td>
</tr>
<tr>
<td>u kamba ‘he helps’</td>
<td>waa kamba ‘he will help’</td>
</tr>
<tr>
<td>u tgwq ‘he comes’</td>
<td>waa tgwq ‘he will come’</td>
</tr>
<tr>
<td>u dewe ‘he calls’</td>
<td>wee dewe ‘he will call’</td>
</tr>
<tr>
<td>u rono ‘he draws’</td>
<td>woo rono ‘he will draw’</td>
</tr>
</tbody>
</table>
The pronouns vu, vaa ‘you singular’ and m, maa ‘I’ work the same way as the tsu, tsaa pronouns above.

Where there is no possibility of confusion, all other long vowels are written with a single letter, and readers are taught to pronounce the words naturally.

**Vowel processes**

There are four different ways vowels interact with each other both within words and across word boundaries.

**Vowel copying**

In Salka Kambari, certain words and affixes contain vowels which derive their pronunciation from nearby vowels. In this process, a certain vowel acts as the source or control of the sound of other vowels. This vowel copying occurs both to the left and to the right, i.e. the source vowel can be either to the left of the target vowel or to the right of the target vowel. The vowel quality to be copied (the source vowel) is the first non-high vowel
(a, a, e, or o) of the root. This process occurs throughout the language, and can copy from any word, regardless of the part of speech: noun, verb, adjective, pronoun, etc. Below is an example of vowel copying to the left.

<table>
<thead>
<tr>
<th>Noun</th>
<th>NCM</th>
<th>Word from which NCM is getting its vowel pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>mogono ‘king’</td>
<td>ma</td>
<td>kasa ‘separates’ (verb)</td>
</tr>
<tr>
<td>mogono</td>
<td>ma</td>
<td>‘yqwq ‘goes’ (verb)</td>
</tr>
<tr>
<td>mogono</td>
<td>me</td>
<td>je’en ‘dances’ (verb)</td>
</tr>
<tr>
<td>mogono</td>
<td>mo</td>
<td>soro ‘pounds’ (verb)</td>
</tr>
<tr>
<td>mogono</td>
<td>ma</td>
<td>va ‘my’ (possessive pronoun)</td>
</tr>
<tr>
<td>mogono</td>
<td>me</td>
<td>le ‘their’ (possessive pronoun)</td>
</tr>
<tr>
<td>mogono</td>
<td>ma</td>
<td>gbain ‘big’ (adjective)</td>
</tr>
<tr>
<td>mogono</td>
<td>me</td>
<td>ke’en ‘small’ (adjective)</td>
</tr>
<tr>
<td>mogono</td>
<td>mo</td>
<td>mokodi ‘short’ (noun)</td>
</tr>
<tr>
<td>mogono</td>
<td>ma</td>
<td>magala ‘hitting’ (noun)</td>
</tr>
</tbody>
</table>

If the first vowel of the root (the word with the source vowel) is a high vowel (u or i), the copying mechanism will skip over that vowel, seeking the first low vowel of the word (a, a, e, or o), even if that is the second or third vowel of the word from which the vowel quality is derived. If there is no low vowel in the root, the vowel ‘a’ is supplied as a default when copying to the left.
The process of vowel copying to the right (the source vowel is to the left of the
rightmost vowel) is slightly different. The vowel copied to the suffix –sV (-sa in the examples
below, meaning ‘iterative’ (IT)) is the right most vowel of the root. The vowel immediately
preceding the suffix in the right column becomes /u/ before the suffix, but if its allophone
[i] surfaces, the root vowel of the verb is preserved in the orthography, e.g.
[kec] is written keceshe. Explanations of why the ‘s’ changes to ‘sh’ will be given later.

<table>
<thead>
<tr>
<th>Uninflected verb</th>
<th>Verb plus iterative suffix -sa</th>
</tr>
</thead>
<tbody>
<tr>
<td>kamba ‘to help’</td>
<td>kambusa (kamba + sa) ‘to help-IT’</td>
</tr>
<tr>
<td>'yagw ‘to go’</td>
<td>'yagwusg ('yagw + sa) ‘to go-IT’</td>
</tr>
<tr>
<td>kece ‘to count’</td>
<td>keceshe (kece + sa) ‘to count-IT’</td>
</tr>
<tr>
<td>soro ‘to pound’</td>
<td>soruso (soro + sa) ‘to pound-IT’</td>
</tr>
<tr>
<td>tsula ‘to buy’</td>
<td>tsulusa (tsula + sa) ‘to buy-IT’</td>
</tr>
<tr>
<td>cipa ‘to come down’</td>
<td>cipusg (cipa + sa) ‘to come down-IT’</td>
</tr>
<tr>
<td>piri ‘to watch’</td>
<td>pirishi (piri + sa) ‘to watch-IT’</td>
</tr>
<tr>
<td>tsun ‘to pour’</td>
<td>tsusun (tsun + sa) ‘to pour-IT’</td>
</tr>
<tr>
<td>ushi ‘to receive’</td>
<td>ushishi (ushi + sa) ‘to receive-IT’</td>
</tr>
</tbody>
</table>

**Vowel elision**

Vowel elision is the process where when two vowels come together (usually the last
vowel of one word and the vowel beginning the next word), the first vowel is deleted, and
only the second vowel is pronounced. For example: Urana u ro ‘another day’ is pronounced
[uranu ro]. Ikebe i va ‘my money’ is pronounced [ikebi va]. In the orthography, the content
word is preserved intact, and the NCM stands on its own.

Another instance of vowel elision in Salka Kambari is across morpheme boundaries
when the consonant before the final vowel of the first morpheme is a sonorant (l, r, n, w)
and the second morpheme begins with a non-sonorant consonant, usually with the same
point of articulation as the sonorant. In this setting, the final vowel of the first morpheme
deletes, and the final consonant of the first morpheme becomes the coda of the syllable preceding it. For example, the sentence written: *U wawa tə* ‘he saved [it]’ is phonetically [u wau tə]. *U wala tə* ‘he walked’ is phonetically [u wal tə]. The full form of the content word is preserved in the orthography.

### Vowel coalescence

Vowel coalescence is where when two vowels come together (the first is *a* and the second is either *i* or *u*), instead of either vowel, a third vowel is pronounced. If the sequence is ‘*a*’ followed by ‘*i*’ (*ai*), ‘*e*’ will be pronounced. If the sequence is ‘*a*’ followed by ‘*u*’ (*au*), ‘*o*’ will be pronounced. If the ‘*a*’ is the final vowel of a content word, the content word retains the ‘*a*’ at the end. For example, *ida i ro* ‘another land’ is correctly written, but is pronounced [idee ro]. If the ‘*a*’ is part of a variable grammatical word, and the ‘*i*’ or ‘*u*’ belongs to the following content word, to aid pronunciation, the ‘*a*’ is changed to ‘*e*’ or ‘*o*’ while the content word remains unchanged. For example: *e ida* /a ida/ [eda] ‘at the land’, *o uwa* /a uwa/ [o'wa] ‘at the house’.

The presence of ‘*a*’ followed by ‘*i*’ or ‘*u*’ is an area where there is much phonetic variation, i.e. the exact string of morphemes may be pronounced by the same speaker in various ways: sometimes it is pronounced as a diphthong [ai] as in the English ‘eye’, sometimes vowel elision occurs, so only the high vowel (*i* or *u*) is pronounced, and sometimes vowel coalescence occurs. In one context one speaker may say [ai], in a different context the same speaker may say [i], and in a third context, the same speaker may say [e]. In addition, speakers from Salka may pronounce it one way, and speakers from Raba (10 kilometers away, but still speak Salka Kambari) may pronounce it another way.

### Vowel co-occurrence

Kambari (any of the varieties, including Salka Kambari) does not have vowel harmony in the same way Igbo does. Igbo has two sets of vowels, and the vowels within a word can only draw from the vowels belonging to one set or the other, but never from both sets within the same word. Kambari restricts certain vowels from occurring within the same word as other vowels. The pattern is this: any word can have any number of the same low vowel (*a, g, e, or o*), but no two different low vowels can occur in the same word. That is to say, a word can have one ‘*e*’ as in *de* ‘there’, two ‘*e*’s as in *ene* ‘to see’, three ‘*e*’s as in *medele* ‘gourd’, etc., but no word will mix ‘*e*’ with ‘*a*’, ‘*e*’ with ‘*a*’, ‘*e*’ with ‘*o*’, or ‘*a*’ with ‘*o*’, etc. There are no restrictions on the occurrence and distribution of the high vowels ‘*i*’ and ‘*u*’. An exception to this is the suffix ‘-to’ which can be attached to any verb: *da'wan* ‘to fear’ *da'waton* ‘to frighten; *cipa* ‘to come down’, *cipato* ‘to unload (bring something down)’, *re'we* ‘to be wet’, *re'weto* ‘to wet something’.
Hausa and English both allow vowels to be mixed within words with no restrictions. This has implications when Kambari borrows words from English and Hausa. Bokiti 'bucket' is an example of word borrowing which conforms to the Kambari rule of co-occurrence.

**Consonants**

As was mentioned earlier, there are twenty-five consonants in the alphabet, with the addition of the glottal stop, there are twenty-six. Labialized consonants (kw, gw, 'w) and palatalized consonants (ly, ky, dy, dy, ty, gy, 'y) have been discussed earlier. Thus the consonants are as follows:

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Labiodental</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Labiovelar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plosive</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td>kp</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>b</td>
<td>d</td>
<td>g</td>
<td>gb</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implosive</td>
<td>ð</td>
<td>d</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f</td>
<td>s</td>
<td>ñ</td>
<td>h</td>
<td>v</td>
<td>z</td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>c</td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td>ɾ</td>
<td>y</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Long consonants**

There are very few contrastive long consonants in Salka Kambari. Where there might be confusion, they are written as double consonants. As of now there is a restricted set of words written with a double consonant to distinguish them from their short consonant counterparts.

<table>
<thead>
<tr>
<th>Long consonant</th>
<th>Short consonant</th>
</tr>
</thead>
<tbody>
<tr>
<td>accaka ‘sticks’</td>
<td>acaka ‘trousers’</td>
</tr>
<tr>
<td>addama ‘trouble’</td>
<td>adama ‘because’</td>
</tr>
<tr>
<td>akkama ‘food stuck in pot’</td>
<td>akama ‘headache/migraine’</td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>amma ‘but’ (borrowed from Hausa)</td>
<td>ama ‘people’</td>
</tr>
<tr>
<td>assala ‘mouthful’</td>
<td>asala ‘vultures’</td>
</tr>
<tr>
<td>damma ‘to say’</td>
<td>dama ‘worry’</td>
</tr>
<tr>
<td>adamma ‘about to say’</td>
<td>adama ‘because’</td>
</tr>
<tr>
<td>kucci ‘sacrifice’, ‘nine’</td>
<td>kuci ‘silently’, ‘always’</td>
</tr>
</tbody>
</table>

In all other places, where there is no confusion, they are written as single consonants: vuma [vumma] ‘person’.

**Palatalization**

In addition to the ‘y’ following certain consonants as described earlier (mikyôn ‘smoke’, lyuwa ‘to eat’ etc.), there is an additional palatalization process. In certain contexts, when ‘s’, ‘ts’ or ‘z’ is followed by either ‘i’ or ‘e’, they become ‘sh’, ‘c’, or ‘j’, respectively. For example, keceshe (kece + sa) ‘to count-IT’, pirishi (piri + sa) ‘to watch-IT’, and ejere ‘two [somethings]’ (underlyingly aga g re [literally ‘ones NCM two’] undergoes both vowel copying and palatalization).

**Nasalized vowels**

Because /m/ can appear at the end of a syllable or word, but /n/ never does, in the KLP orthography (and phonology), what sound like nasalized vowels (a vowel where part of the sound comes through the nose) have been interpreted as a sequence of a vowel followed by /n/. In other words, ‘n’ at the end of a word is pronounced through the nose instead of as the Hausa nan ‘here’ or the English ‘man’.

- tawan [tāwā] ‘to shoot a bow’ (in contrast to tawa [tawa] ‘to remove’)
- shuwâ [ʃuwa] ‘to pass gas’ (in contrast to shuw [ʃuwa] ‘to transplant’)
- je’en [dʒe’en] ‘to dance’
- ušhin [uʃi] ‘twenty’ (in contrast to ushi [uʃi] ‘to receive’)
- lon [lɔ] ‘abundance’ (in contrast to lo [lo] ‘there’)
- tsun [tʃu] ‘to pour’ (in contrast to tsu [ʃu] ‘we’)

This vowel nasalization may spread through certain consonants (’, y, w, s, sh, and v), and it is always represented in the orthography by an ‘n’ wordfinal. If the root of a word ends with this kind of nasalization, and a suffix is added, the ‘n’ is still written word finally.

- yuwan [jùwà] ‘to do’  
  yuwan + i → yuwain [jùwài] ‘did’
tewen [tēwē] ‘to shoot [arrow]’
tewen + i → tewein [tēwē]n ‘shot [arrow]’
gqvun ‘leaf’
gqvun ‘the particular leaf’
ukwan ‘skin/leather’
ukwa’an ‘the particular skin’
moten ‘chicken’
moto’on ‘the particular chicken’

Syllabic nasals and nasals before consonants

In Salka Kambari, at the beginning of words, ‘m’ and ‘n’ coming before another consonant acts like a syllable (mogono ‘king’ has three syllables; ngono ‘kings’ also has three syllables). This ‘m’ or ‘n’ is written as ‘m’ before labial consonants (consonants made with the lips: m, p, b, f, gb, and kp) and ‘n’ before all other consonants. This pattern of ‘m’ before m, p, b, f, gb, and kp, and ‘n’ before all other consonants is also used within a word, and if the ‘n’ or m is a single word which is spaced from the following word.

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>mpasa</td>
<td>‘blood’</td>
</tr>
<tr>
<td>nruja</td>
<td>‘fishing’</td>
</tr>
<tr>
<td>cimba</td>
<td>‘to farm’</td>
</tr>
<tr>
<td>gandala</td>
<td>‘to roll’</td>
</tr>
<tr>
<td>m batai</td>
<td>‘I branched’</td>
</tr>
<tr>
<td>n dammai</td>
<td>‘I said’</td>
</tr>
</tbody>
</table>

Tone

Lexical tone: Salka Kambari has a small number of minimal tone pairs (under 20). Most of the minimal pairs are in different domains, so there is little possibility of confusion when reading. For example [uná] ‘to kill’, [unà] ‘bush’.

Grammatical tone: An example of grammatical tone in Salka Kambari occurs in the verb form. In writing, the imperative is always part of a direct quote and has no overt subject, while other verb modes require an expressed subject, so there is no confusion.

“Wala!” ‘Walk!’ [wàlå]
U wala shi. ‘He did not walk.’ [ù wàlá shå]
Waa wala shi. ‘He will walk go’ [wà wålå shå]

Reader tests of tone in marked and unmarked texts have shown that no information is lost with the unmarked texts. Furthermore, there is reader resistance to tone marks, shown as a reluctance to try to read tone marked texts. If reading tests showed significant amounts of information were confusing without marked tone, KLP would propose including tone marks. Such is not the case; therefore, tone is unmarked in the Salka Kambari orthography.

Word spacing

As a written form of a language is introduced to the community, the transfer from a complete thought to a linear sequence of individual words is not intuitive. Consistency in writing involves both spelling the same word the same way every time, and spacing between words in a consistent manner. The Salka Kambari orthography operates on the principle that
the ‘big words’ (content words: nouns, verbs, adjectives, adverbs) are spelled the same way every time. The ‘little words’ (functors: pronouns, prepositions, verbal morphemes) alter their spelling to allow the written form to more closely reflect the oral form. One of the ways of ensuring consistency in the big words is to always leave a space between the little words and the big words.

1. U walai. ‘He walked.’
2. U tsu wala. ‘He always walked.’
3. U wala tg. ‘He has walked.’
4. U damma nle. ‘He told them.’
5. U damma le tg. ‘He has told them.’
6. N lyungu tg. ‘I sent.’ (This is visually different from nlyungu ‘shoulder blade’.)
7. U tg ng medele mg vg. ‘He has my gourd.’ Literally ‘He is with gourd of mine.’
8. U tg ni ikebe i vg. ‘He has my money.’
9. U tg ng matada mg vg. ‘He has my soup dish.’
10. U tg ng ntada n vg. ‘He has my soup dishes.’
11. U tg ni ikebe ya Awanca. ‘He has Awanca’s money.’
12. U tg nu umuna wi ikebe. ‘He has expensive cloth.’
13. U tg ni ikebe yu umuna. ‘He has money for cloth.’

In sentences 7–11, notice that the possessive is made up of a noun, a NCM (noun concord marker linking the head noun and its modifier, e.g. mg following medele ‘gourd’ in sentence 7), and the possessive pronoun or noun (medele mg vg ‘gourd of mine’). There is a space between each of these. Notice also that the ‘with’ is represented by ng before consonants (ng matada), and n with the same vowel as the initial vowel of the following word (ni ikebe, nu umuna). Readers are taught to blend the vowels into one syllable (ni ikebe sounds like [nikebe]). Notice that the noun phrase structure of sentences 12 and 13 follows the same pattern as that of the possessive: umuna wi ikebe ‘cloth of money, i.e. expensive cloth’, and ikebe yu umuna ‘money of cloth, i.e. the money needed to buy cloth’. The noun
phrase structure is noun, NCM, and modifier. The clause structure when a subject noun is specified is noun, NCM, verb.

**Problems still to be considered**

The multiplicity of functor shapes and meanings. For example, *tsu* in Salka Kambari has six meanings/functions: first-person plural subject pronoun, first-person plural object pronoun, first-person plural possessive pronoun (the corresponding first-person singular pronouns are /N/, *mu*, *v* respectively), repetitive action marker (*U* *tsu* *wala* ‘He repeatedly walked.’), noun class prefix (*tsu-kutu* ‘leprosy’, *tsu-sumba* ‘locust bean fruit’), and NCM (*tsusumba* *tsu* *gbain* ‘big locust bean fruit’). As of now, most of the functors (grammatical words) in the language have between four and seven orthographic representations; for example, the subject pronoun *u* ‘he’ may appear in texts as *u*, *wa*, *wq*, *we*, *wi*, *wo* or *wu*. If the following word begins with a consonant, the *u* form is used (*u* *ronoi* ‘he drew’); if the word begins with a vowel, the *u* changes to *w* and the vowel of the verb is copied backwards to make a pronounceable word; for example, *we* *ecei* ‘he asked’, *wo* *o’woi* ‘he became’, *wu* *una* ‘he killed’. Would it be beneficial in the long run to only use the underlying form (in the above example, *u*), and diligently teach natural pronunciation? Thus the above examples would be written *u* *ecei* ‘he asked’, *u* *o’woi* ‘he became’, *u* *una* ‘he killed’. If so, each functor would have only one spelling, though the overlapping of different meanings with the same spelling would still exist.

Note: This orthography write-up was based on six years of using a trial orthography, and every new book was passed by a committee of reviewers who had been trained in both how to read the orthography and what is expected of a reviewer. At the time of the orthography submission to the Language Development Center, the orthography will have been approved by the KLP Committee, by various influential people within each language area, and by a combined (representatives from all three languages) session of an Orthography Approval Committee. On April 17, 1999, the KLP Committee approved the Salka Kambari orthography. A booklet, *Ure wi Idana i Tsishingini* (*Presenting the Tsishingini Orthography*) containing stories, a pictorial alphabet, and a brief explanation of some general orthography principles, is being circulated to the community leaders starting in May 1999.
The Blind Man and The Leper

A leper and a blind man were friends. On Friday morning they went begging to the king of farming of Udoro village. Then they were given plenty of guinea corn, enough to fill ten mudus. Then the leper, when he realized the blind man could not see, told the blind man there were only five mudus.

Then they started fighting. Then the leper was beating the blind man. He the blind man, who is without eyes, he said the leper was beating him with planks. He had forgotten the leper had no fingers.
References


1998. LinguaLinks Library ver. 2.5, CD-ROM. Dallas: SIL.


