Précis from the 15th Conference on African Linguistics

UCLA, March 29-31, 1984

STUDIES in AFRICAN LINGUISTICS

SUPPLEMENT 9

Published by the Department of Linguistics and the African Studies Center
University of California, Los Angeles
PREMIS FROM THE FIFTEENTH CONFERENCE ON AFRICAN LINGUISTICS

UCLA, March 29-31, 1984

Edited by
Russell G. Schuh
UCLA

Supplement 9
Studies in African Linguistics
December 1985
The papers in this volume were presented at the Fifteenth Conference on African Linguistics, held at UCLA from March 29-31, 1984. Though the papers here are called "précis", they are really full papers. Authors were given a 2000 word limit, which allowed them to not only write a substantive text, but also to include a generous number of examples, tables, etc.

Inasmuch as the volume does not include all the papers presented at the conference, we have chosen not to call it a "proceedings". However, if the purpose of a proceedings is to reflect the general range and quality of papers presented at a conference, this volume achieves that goal. Of the 89 papers presented at the conference, 65 are included here.

If the papers here are a representative of the current state of African linguistics, it is indeed a dynamic field. Of the language families native to the African continent, only Khoisan is not represented. In addition, there are several papers on European-based creoles which African languages have influenced. Besides papers covering all the major "traditional" fields of linguistics, i.e. phonetics, phonology, morphology, syntax, semantics, and historical language change, there are a number of papers from "hyphenated" fields, e.g. socio-linguistics and psycho-linguistics. Orientation ranges from the purely descriptive to the highly theoretical. Even among the theoretical papers, there is considerable diversity, e.g. in syntax there are papers representing the government binding, relational, and generalized phrase structure approaches, among others.

In short, we believe that for anyone who wants to learn who is doing what in African linguistics today, this volume is a good place to start.

Russell G. Schuh
Los Angeles, December 15, 1985
# STUDIES IN AFRICAN LINGUISTICS

**Supplement 9**  
**December 1985**

## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOREWORD</td>
<td>ii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>iii</td>
</tr>
<tr>
<td>F. Niyi Akinnaso, <em>Some Theoretical Issues in Borrowing and Code-Switching</em></td>
<td>1</td>
</tr>
<tr>
<td>Mohamed Nuuh Ali and Linda Arvanites, <em>The Place of Jiiddu in Proto-Soomaali</em></td>
<td>6</td>
</tr>
<tr>
<td>Qladele Awobuluyi, <em>On the Reality of Vowel Coalescence in Yoruba</em></td>
<td>11</td>
</tr>
<tr>
<td>Hany Amin Azer, <em>The Expression of Negation in Egyptian Colloquial Arabic (ECA)</em></td>
<td>15</td>
</tr>
<tr>
<td>M. Lionel Bender, <em>Gumuz, Koman, Mao, and Omotic</em></td>
<td>19</td>
</tr>
<tr>
<td>Elena Bertoncini, <em>Quantitative Analysis of Swahili Vocabulary</em></td>
<td>22</td>
</tr>
<tr>
<td>Hagit Borer and Laurice Tuller, <em>Nominative/Agreement Complementarity and VSO Order in Standard Arabic</em></td>
<td>27</td>
</tr>
<tr>
<td>Dianne C. Bowcock, <em>African Language Usage in the Classroom, Reported and Observed</em></td>
<td>33</td>
</tr>
<tr>
<td>Vicki Carstens, <em>Wh-Movement in Yoruba</em></td>
<td>39</td>
</tr>
<tr>
<td>Farida Cassimjee and Charles W. Kisseberth, <em>Three Cases of Downstep in Venda</em></td>
<td>45</td>
</tr>
<tr>
<td>Monik Charette, <em>The Appendix in Parametric Phonology</em></td>
<td>49</td>
</tr>
<tr>
<td>Donald G. Churma, <em>Semisegmental Phonology</em></td>
<td>54</td>
</tr>
<tr>
<td>Elizabeth Cowper and Keren Rice, <em>The Destruction of Tonal Structure in Mendé</em></td>
<td>57</td>
</tr>
<tr>
<td>Chet A. Creider, <em>One Father or Two? Polysemy in Kinship Terms</em></td>
<td>63</td>
</tr>
<tr>
<td>Gerard M. Dalgish, <em>Microcomputers and African Research</em></td>
<td>69</td>
</tr>
<tr>
<td>Anthony R. Davis, <em>Syntactic Binding and Relative Aspect Markers in Hausa</em></td>
<td>72</td>
</tr>
<tr>
<td>Katherine Demuth, Nicholas Faracias, and Lynell Marchese, <em>Niger-Congo Noun Class and Agreement Systems in Historical and Acquisition Perspective</em></td>
<td>78</td>
</tr>
<tr>
<td>Shirley Donwa-Ifode, <em>The Intonation System of Isoko</em></td>
<td>83</td>
</tr>
<tr>
<td>David Dwyer, <em>A Segmental, Autome洛ic View of Mendé Tone</em></td>
<td>90</td>
</tr>
<tr>
<td>Carol M. Eastman and Sh. Yahya Ali Omar, <em>Swahili Gestures: Comments (vielezi) and Exclamations (viingizi)</em></td>
<td>95</td>
</tr>
<tr>
<td>Author</td>
<td>Title</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Atteya Yussif El-Noory</td>
<td>EGYPTIAN ARABIC AND ENGLISH: NATIVIZATION PROCESS</td>
</tr>
<tr>
<td>Nora C. England and William A. Ladusaw</td>
<td>QUESTION FORMATION IN KUSAAL</td>
</tr>
<tr>
<td>Nicholas Faraclas</td>
<td>RIVERS PIDGIN (CREOLE) ENGLISH: TONE, STRESS, OR PITCH-ACCENT LANGUAGE?</td>
</tr>
<tr>
<td>Zygmunt Frajzyngier</td>
<td>BORROWED LOGOPHORICITY?</td>
</tr>
<tr>
<td>Gary M. Gilligan</td>
<td>AUTOSEGMENTAL BABANKI</td>
</tr>
<tr>
<td>John Goldsmith</td>
<td>BANTU -ə- : THE FAR PAST IN THE FAR PAST</td>
</tr>
<tr>
<td>H. F. Hailu</td>
<td>CLEFT CONSTRUCTIONS</td>
</tr>
<tr>
<td>Hafedh Halila</td>
<td>SOME SYLLABLE STRUCTURE BASED RULES OF TUNISIAN ARABIC</td>
</tr>
<tr>
<td>Grover Hudson</td>
<td>THE ARABIC DOUBLED VERB CONSPIRACY AND MORPHEME INVARIANCE</td>
</tr>
<tr>
<td>John P. Hutchison</td>
<td>WH-MOVEMENT AND RELATIVIZATION IN TWO CREOLE LANGUAGES: CAPE VERDEAN (CV) AND HAITIAN (H)</td>
</tr>
<tr>
<td>Larry M. Hyman</td>
<td>DEPENDENCY RELATIONS IN SYNTAX: THE MYSTERIOUS CASE OF THE EMPTY DETERMINER IN AGHEM</td>
</tr>
<tr>
<td>Janice L. Jake</td>
<td>RAISING IN FUR</td>
</tr>
<tr>
<td>Camilia N. Keach</td>
<td>OPENTHESIS: A POSITIONAL TREATMENT OF SWAHILI PRONOMINAL CLITICS</td>
</tr>
<tr>
<td>Michael Kenstowicz</td>
<td>ON CORE SYLLABLES IN MODERN ARABIC DIALECTS</td>
</tr>
<tr>
<td>Mairo Kidda</td>
<td>MORPHEME ALTERNATION IN TANGALE: A SYLLABLE STRUCTURE APPROACH</td>
</tr>
<tr>
<td>Alexandre Kimenyi</td>
<td>DOUBLET IN KINYARWANDA: AN INQUIRY INTO THE PROCESS OF SIGN-PRODUCTION</td>
</tr>
<tr>
<td>Karen Lahaie</td>
<td>THE ma- PREFIX IN AFROASIATIC</td>
</tr>
<tr>
<td>William R. Leben</td>
<td>INTONATION IN CHADIC LANGUAGES</td>
</tr>
<tr>
<td>Marshall Lewis</td>
<td>RELATIVE CLAUSES IN ANLO EWE</td>
</tr>
<tr>
<td>Jean Lowenstamm and Jean-François Prunet</td>
<td>ON CERTAIN NOMINAL PATTERNS IN TIGRINYA</td>
</tr>
<tr>
<td>C. T. D. Marivate</td>
<td>THE IDEOPHONE AS A SYNTACTIC CATEGORY IN THE SOUTHERN BANTU LANGUAGES</td>
</tr>
<tr>
<td>Bruce Martin</td>
<td>VOWEL HARMONY IN KIPSIGIS</td>
</tr>
<tr>
<td>P. Mumo Maundu</td>
<td>SOME ASPECTS OF CONSONANTAL SOUND CHANGE IN KIKAMBA</td>
</tr>
<tr>
<td>Girgis Z. Messiha</td>
<td>SOME PHONOLOGICAL PROBLEMS THAT FACE ARAB LEARNERS OF ENGLISH</td>
</tr>
<tr>
<td>M. A. Mohammad</td>
<td>STYLISTIC RULES IN CLASSICAL ARABIC AND THE LEVELS OF GRAMMAR</td>
</tr>
<tr>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>Paul Newman, TONE SPLITTING AND GWANDARA ETHNOKISTORY</td>
<td>233</td>
</tr>
<tr>
<td>Kjell Norlin, ACOUSTIC ANALYSIS OF VOWELS AND DIPHTHONGS IN CAIRO ARABIC</td>
<td>238</td>
</tr>
<tr>
<td>Derek Nurse, THE &quot;INDIGENOUS VERSUS FOREIGN&quot; CONTROVERSY ABOUT THE SOURCES OF SWAHILI VOCABULARY</td>
<td>245</td>
</tr>
<tr>
<td>Dudley K. Nylander, A NEW ANALYSIS OF THE KRIIO CLEFT PREDICATE</td>
<td>251</td>
</tr>
<tr>
<td>David Odden, PROBLEMS IN THE METRICAL REPRESENTATION OF TONE</td>
<td>254</td>
</tr>
<tr>
<td>Annie Rialland, LE FINI/L'INFINI OU L'AFFIRMATION/L'INTERROGATION EN MOBA (LANGUE VOLTAIQUE PARLÉE AU NORD-TOGO)</td>
<td>258</td>
</tr>
<tr>
<td>Ian Roberts, SERIAL VERBS AND GOVERNMENT BINDING THEORY</td>
<td>262</td>
</tr>
<tr>
<td>Douglas L. Sampson, THE PHONOLOGY OF BANDA-TANGBAGO</td>
<td>269</td>
</tr>
<tr>
<td>Paul Sauvageau, THE METRICO-SYLLABIC APPROACH: EVIDENCE IN KINYARWANDA</td>
<td>275</td>
</tr>
<tr>
<td>Thilo C. Schadeberg, THE DEVELOPMENT OF CASE AND FOCUS IN UMBUNDU</td>
<td>282</td>
</tr>
<tr>
<td>Deborah Schindwein, DOWNSTEP IN THE KIPARE VERB COMPLEX</td>
<td>285</td>
</tr>
<tr>
<td>Carol Myers Scotton, LANGUAGE UNIVERSALS AND SYNTACTIC CHANGES IN SWAHILI AS A SECOND LANGUAGE</td>
<td>290</td>
</tr>
<tr>
<td>Aleksandra Steinbergs, LOANWORDS AND MSC'S IN OSHIKWANYAMA</td>
<td>293</td>
</tr>
<tr>
<td>Andrea Tyler, SWAHILI RELATIVE CLAUSES, A GENERALIZED PHRASE STRUCTURE GRAMMAR ANALYSIS</td>
<td>298</td>
</tr>
<tr>
<td>Donna Wagner, OBJECTS IN GOKANA</td>
<td>304</td>
</tr>
<tr>
<td>Wafaa Wahba, SOME ASPECTS OF LF MOVEMENT IN IRAQI ARABIC</td>
<td>309</td>
</tr>
<tr>
<td>Benji Wald, THE INFLUENCE OF ARABIC ON THE SYNTAX OF SWAHILI DISCOURSE</td>
<td>315</td>
</tr>
<tr>
<td>Jennifer J. Yanco, LANGUAGE CONTACT AND GRAMMATICAL INTERFERENCE: HAUSA AND ZARMA IN NIAMEY, NIGER</td>
<td>318</td>
</tr>
<tr>
<td>Eluzai M. Yokwe, THE DIVERSITY OF JUBA ARABIC</td>
<td>323</td>
</tr>
<tr>
<td>Qrê Yusuf, A FUNCTIONAL EXPLANATION FOR THE N1-NP CONSTRUCTION IN YORUBÁ</td>
<td>329</td>
</tr>
</tbody>
</table>
SOME THEORETICAL ISSUES IN BORROWING AND CODE-SWITCHING

F. Niyi Akinnaso
University of Ife, Nigeria

0. Introduction

Drawing upon Yoruba-English bilingual data, this paper provides explanations for lack of rigorous distinction between borrowing and code-switching (§1) and suggests four main ways of establishing the distinction between the two (§2.) In conclusion, an argument is proposed for the relevance of studies of borrowing and code-switching for linguistic theory (§3).

1. Borrowing and Code-Switching

Rather than provide a comprehensive review of previous studies of borrowing and code-switching, emphasis is placed on major shortcomings of previous attempts to distinguish between the two processes. The first shortcoming is that of orientation: no one has explicitly made the distinction between borrowing and code-switching the primary goal of analysis. Rather, only ad hoc distinctions have been made within the context of a general discussion of either borrowing or code-switching, thus making classification a secondary issue [Haugen 1950; Pfaff 1979; Goke-Pariola 1983].

The second shortcoming stems from failure to account for variation in informants' level of bilingual competence. This is an important issue because borrowing and code-switching imply vastly different claims about the competence of the individual speaker: borrowing may occur in the speech of those with only monolingual competence, whereas code-switching implies some degree of competence in at least two languages. Such variation in levels of bilingual competence has serious implications for the rate and extent of morpho-phonological adaptation of borrowed or switched items: morpho-phonological adaptation is obligatory for monolingual speakers, whereas bilingual speakers may or may not adapt borrowed or switched items to the structure of the recipient or matrix language. This calls into question the status of morpho-phonological adaptation of borrowed items as a criterion for distinguishing between borrowing and code-switching.

The third, and perhaps most crucial, shortcoming arises from over-concentration on the surface features of linguistic structure. Many investigators have based the distinction between borrowing and code-switching on the phonology, morphology, and surface syntax of particular utterances considered in isolation, neglecting macro-sociological and socio-pragmatic issues underlying surface linguistic behavior. In this approach, borrowing is viewed essentially as a lexical operation, whereas code-switching is regarded as a syntactic phenomenon [Reys 1974]. There are several loopholes in this approach. First, the association of switching but not borrowing with internal syntactic structure ignores the generally recognized possibility of borrowing not only idiomatic phrases but also full scale morphological and syntactic features from another language [Gumperz and Wilson 1971; Pfaff 1979]. Second, there is no agreement as to the classificatory status of foreign items in speech as they occur in both borrowing and code-switching (see Pfaff [1979] for a review), thus making it difficult to distinguish between the two (see, especially, Goke-
Some further comments should be made about the relationship between borrowing or code-switching and morpho-phonological adaptation. First, morpho-phonological adaptation is mediated by numerous factors, including speaker's level of bilingual competence, interlocutor, topic, and context of speech. For example, the pronunciation of the Yoruba word *burenl* 'bread' varies from the fully adapted form, *[bʊɾɛdl]*, to the unadapted English original, *[bɹɛd]*. The former is found in the speech of nonliterate, monolingual Yoruba speakers, whereas the latter is found in the speech of literate, bilingual Yoruba speakers of English. However, this is not an absolute distinction since literate bilinguals sometimes use the fully adapted form when speaking to non-literate monolinguals, especially in situations where it may be offensive to sound literate.

Second, the degree and perceptibility of morpho-phonological adaptation vary according to the genetic and typological distance between the languages in contact. For example, Hausa loan words require less adaptation than English loan words in Yoruba.

A third comment concerns the extent of morpho-phonological marking required in "language [incorporated] borrowing" as opposed to "speech [spontaneous] borrowing" [Grosjean 1982:308]. It is difficult to predict the degree of morpho-phonological adaptation of loan words largely because the process that leads to language borrowing and the adaptation of borrowed items to the structure of the recipient language is very complex. Thus, although full integration of borrowed items is often accompanied by morpho-phonological adaptation, spontaneous borrowings, too, may be adapted, even on first usage. Besides, there are integrated borrowings that are unadapted [Saville-Troike 1982:67], whereas spontaneous code-switching may be accompanied by phonological and lexico-grammatical adaptation as in Yoruba-English code-switching [Goke-Pariola 1983]. Furthermore, the relationship between morphological adaptation and lexical incorporation is gradient, and depends on the functional load of morphological marking for different syntactic categories. For example, it has been found that there is a higher degree of morphological adaptation for verbs than for other syntactic categories because of the "functional load" and "centrality" of the verb in the sentence [Pfaff 1979:298].

2. Distinguishing Borrowing from Code-Switching

From the foregoing, it seems that the indeterminacy of the distinction between borrowing and code-switching rests on the fact that macro-sociological and socio-pragmatic issues (which often go beyond the particular utterance) are involved. Because of the relevance of degree of bilingual competence to these issues, a necessary preliminary step toward distinguishing between borrowing and code-switching is to define bilingualism. For present purposes, a bilingual speaker might be described as a person who can express himself in spontaneous, intelligible sentences in at least one language in addition to his first language.

The first criterion for distinguishing between borrowing and code-switching rests on the distinction between mono- and bilingualism, and the theoretical question to ask is, "What is the level of bilingual competence required of a speaker before he can use a borrowed item or code-switch?" As indicated above, individual bilingual competence is not required for borrowing to take place,
whereas it is obligatory for code-switching, as in the case of monolingual Yoruba speakers who make use of borrowed items in their speech, whereas only literate, bilingual Yoruba speakers can code-switch, in addition to making use of loan words.

While the above classificatory criterion is targeted at the competence of the individual speaker, the second criterion has to do with the number of grammars involved. The main argument is that borrowing prototypically involves only one grammar (except where the borrowed items are fixed phrases or idiomatic expressions), whereas code-switching involves at least two grammars. A caveat must, however, be added because the distinctiveness of the two grammars is often mediated by several factors, including genetic and typological relationship and the degree of lexico-grammatical adaptation involved in code-switching. Where lexico-grammatical adaptation is minimal as in Spanish-English code-switching [Pfaff 1979], the two grammars are often very distinct, whereas the distinction between the two grammars is often very blurry where there is a great deal of lexico-grammatical adaptation as in Yoruba-English code-switching [Goke-Pariola 1983]. In the latter case, the tendency is to conclude that only one grammar—that of the matrix clause—is involved. However, such a conclusion would be neglecting cases in which code-switching involves juxtaposition of syntactic constructions (complex noun or verb phrases, whole clauses, etc.) whose analysis would require a recourse to two separate grammars. This is not to say that mono-grammar explanations of code-switching are not plausible, but to point out that such explanations often rely on data in which switched items operate largely at a lexical rather than syntactic level. Since dual-grammar explanations of code-switching automatically subsume lexical switching, they seem to account for all the facts without leftovers.

The third classificatory criterion is degree of substitution: given L1 as the indigenous language of a speech community and L2 as the language in contact, it is often possible to render code-switched passages in either of the two languages, whereas equivalent terms are often not available in L1 for items borrowed from L2. Hence, it is often difficult to find lexical substitutions for loan words in the borrowing language, and where substitutions exist, they often post-date the borrowed items in the lexical history of the borrowing language. For example, there are no equivalent terms for the following loan words in Yoruba: yín 'bread', gówù 'gold', sìká 'scarf', wì 'wig', and tàbítì 'table', whereas the code-switched passage, G̲i̲r̲ŷ̲n̲Ō̲̄ḡ̲̄̄ā̲̄̄a, can be rendered in either English or Yoruba:

English: That girl does not make it at all.

Yoruba: Qmóbínrin yén ọ̀ṣe dádáá rárá.

The English and Yoruba sentences are, of course, inherently polysemous, specific meanings being essentially context-dependent. This underlines the fact that the main objective of code-switching is to give socio-pragmatic rather than referential information. This leads us to the fourth criterion.

The final criterion to be discussed here is that of functional classification. The theoretical question to ask is, "What is the primary function of particular L2 items in an otherwise L1 utterance?" The distinction to look for is that between referential and socio-pragmatic meaning. It is suggested that while borrowing serves primarily referential functions, providing labels for concepts, objects, and ideas that have no antecedents in the borrowing language/
culture, code-switching serves primarily socio-pragmatic functions. Thus in the Yoruba sentence, *mo féé jé bùrédi* 'I want to eat bread', the borrowed item, *bùrédi*, refers uniquely to the object 'bread', whereas in the sentence *girl yénnó make è allat*, the switched L₂ items (*girl, make, and all*) may have been introduced for a variety of socio-pragmatic reasons: to show off, to cut off a monolingual participant in the conversation, to specify a particular addressee, to reiterate a message, to deliver an ambiguous message, etc. Viewed in this sense, borrowing is a word or clause level phenomenon, whereas code-switching is ultimately a matter of conversational interpretation, so that the relevant inferential processes are strongly affected by contextual and socio-cultural presuppositions. Furthermore, in code-switching, "it is the juxtaposition of two alternative linguistic realizations of the same message that signals information, not the propositional content of any one conversation-al passage" [Gumperz 1982:84].

3. Theoretical Implications

The theoretical implications of the preceding observations about the relationship between borrowing and code-switching are twofold. First, the distinction between borrowing and code-switching sensitizes us to variations in degree of communicative competence and, especially, the distinction between mono- and bilingualism. Borrowing and code-switching imply vastly different claims about the competence of the individual speaker. A synthesis of research on the relationship between borrowing and code-switching should, therefore, give us further insight into the nature of, and variation in, communicative competence.

Second, syntactic analyses of switched passages often involve at least two grammars, thus extending the possibilities of syntactic knowledge beyond the provisions of the conventional mono-grammar analysis. Two main views have emerged regarding the nature of code-switching grammar (see Akinnaso [1981] for more detailed review).

In one view, represented by Samkoff and Poplack [1980], a code-switching grammar must exist which makes it possible for code-switching to take place as quickly and fluently as it does in spontaneous conversations. It is argued that the constituents juxtaposed in intrasentential switching are too intimately related to be generated separately by the rules from two distinct grammars. This view must, however, be distinguished from the tone that identifies the code-switching grammar with that of the matrix sentence [Goke-Pariola 1983].

In another view, represented by Woolford [1981], it is not necessary to postulate a grammar specific to code-switching. Rather, the two monolingual grammars co-exist, allowing the bilingual speaker to switch back and forth between grammars and to draw phrase structure rules freely from either language. The lexicons and word formation components of the two grammars remain entirely separate from each other. When phrase structure rules are the same in both languages, the nodes may be filled freely from either lexicon; however, when the nodes are created by a rule that exists in only one of the languages, they must be filled from the lexicon of that language.

Borrowing and code-switching are among the major legacies of those who live with two languages. At the bilingual community level, borrowing and code-switching constitute an integral part of the language varieties and communicative norms developed by the speech community. Since bilingualism and language
contact are the rule rather than the exception today, more attention should be paid to the theoretical and practical implications of the major sociolinguistic consequences of language contact, viz., borrowing and code-switching.

REFERENCES


1. The Problem

At the Second International Congress of Somali Studies, C. Ehret and M. Nuuh Ali presented evidence that there is more linguistic differentiation in the Soomaali-speaking area than previously described [Ehret and Nuuh Ali 1983]. They proposed a reorganization of the family tree of Omo-Tana. With the exception of H. Fleming's initial consideration of Bayso [1964], preceding work had accorded that language a unique relationship with the rest of Omo-Tana: it had come to be considered the sole constituent of a Northern branch of Omo-Tana [Heine 1978:8]. By plucking Jiiddu out of obscurity among what have been considered "southern Somali dialects" [Sasse 1979:15,19,34], and putting it with Bayso as one primary branch, as against a second primary branch composed of the remainder of the Soomaali languages, Ehret and Nuuh Ali cut a swathe through the confusion which has reigned with regard to "the full array of Soomaali dialects and languages and...delimiting their territorial extent" [Ehret and Nuuh Ali 1983: Addendum, p.5].

Their grouping of Bayso and Jiiddu has been criticised by M. Lamberti in a paper at that same conference which emphasized the ties of Jiiddu with a southern Somali dialect cluster he labels "Digil" [Lamberti 1983: sections 2, 3.5 and Appendix]. In this response to Lamberti's arguments against the validity of a Bayso-Jiiddu relationship, we examine Lamberti's counter-evidence anew. We find untenable Lamberti's contention that Jiiddu is properly a "Somali" dialect by virtue of its observance of the "older" laws of spirantization and palatalization, both because of inconsistent argument by Lamberti and because there is another older law, "a-raising, which Bayso and Jiiddu share and which feeds the palatalization conditions of Jiiddu. We discuss inconsistent argument in sections 2ff., and "a-raising in section 4.3.

2. Data and Method

In this examination of sound correspondances between Jiiddu and its neighbors (Garree, Tunni) and geographically distant kin (southwestern Ethiopian Bayso, Kenyan Rendille), we ask, "Which is the correct position for Jiiddu: independent language status or southern Soomaali dialect (Lamberti's 'Digil')?" It is important to stress the languages chosen for comparison. Lamberti [1983: section 3.5] has argued for a "Digil" unity, classing Jiiddu with Garree and Tunni. But, in refuting Ehret and Nuuh Ali [1983] he has made reference, not to these two languages, but to Northern Soomaali. This is inconsistent argumentation; two different sets of correspondences are used, the one to explain a "Digil" cluster, the other to refute a Jiiddu-Bayso connection. Lamberti's postulation of a "Digil" cluster points to closer relations between Jiiddu and Garree and Tunni than between Jiiddu and Northern Soomaali. His own criteria seem to demand argument from the standpoint of "Digil" to disprove Ehret and
Nuuh Ali's stand, yet he has not used such data.

We do not, as has Lamberti, use Northern Soomaali evidence to prove the nature of the Jiiddu-Bayso relationship; we prefer to emphasize the nature of the Soomaali I - Soomaali III differences through concentration on three of Lamberti's "Digil" dialects.

3. Sources and Terms

The first description of Jiiddu was accomplished by Moreno [1951]. Bayso has been described by Fleming [1964] and by Hayward [1978-79], which latter is our main source of Bayso data. Rendille data are from Oomen [1981] and Heine [1978]. Our data for Garree and Tunní as well as Jiiddu is from C. Ehret's field notes. We have relied on one of the two Jiiddu dialect forms collected by Prof. Ehret, which appears to differ in some respects not only from Moreno's but also from Lamberti's.

The term *Somali* refers to the nation and people (unless citing others' use, in which case quotation marks identify the form), while *Soomaali* refers to the historical mother language and to its branches and sub-branches. Thus *Soomaali I* is the branch of Omo-Tana containing Bayso and Jiiddu as one sub-branch and what Heine calls Sam as the second sub-branch; *Soomaali II* is Heine's Sam, composed of Rendille as one sub-branch and *Soomaali III* as the second sub-branch. *Soomaali III* contains the remainder of the Soomaali languages and dialects, such as Garree, Tunní, Maay, and the Northern dialects. It is Bayso, Jiiddu, Rendille, Garree, and Tunní which are compared here with a view toward ascertaining the validity of moving Jiiddu out of southern "Somali" dialect status.

4. Subgrouping Evidence

Lamberti ascribed to Jiiddu the status of one of the southern "Somali" "Digil" dialects, because it obeys what he calls "older" sound laws [Lamberti 1983: Appendix A]. These are (i) spirantization of *k* to *h* as in "Somali" and (ii) palatalization of velars before front vowels as in "Somali". We proceed, then, to consider these changes.

4.1. Spirantization. As mentioned earlier, if one describes Jiiddu as a southern "Somali" dialect, then one is constrained to compare it with speech groups from the southern Somali area. Accordingly, we use Garree and Tunní to appraise Lamberti's contentions about the age and scope of spirantization of *k*. If there is such a thing as a "Digil" dialect cluster, then it should be marked by sharing this "older" law. Earlier *k* has spirantized in Jiiddu, except in the environment of another consonant. In the tables to follow, NC means no cognate, dashes indicate no form was collected, and parentheses indicate that seeming cognates are actually loans.

'Sheep's dewlap' hulkul in Jiiddu is a neat illustration of what happens to *k*: initially, this reduplicated syllable shows spirantization, while the second *k*, which is post consonantal, remains a stop. Those *k* 's found in Jiiddu which are not the result of post-consonantal preservation are loans, e.g. kullin 'all', or are reflexes of devoicing of *g*, e.g. 'enek 'breast', dalk 'blood' [Ehret and Nuuh Ali 1983: Addendum]. Neither Garree nor Tunní share the change of *k* to *h*. This difference does not support Lamberti, either with respect to a "Digil" cluster or with respect to the "age" of spirantization. If the "Digil" cluster were a unity at the time of the
Table 1. Spirantization of *k in Jiiddu

<table>
<thead>
<tr>
<th>Bayso</th>
<th>Jiiddu</th>
<th>Rendille</th>
<th>Garree</th>
<th>Tunni</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) initially</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'body'</td>
<td>NC</td>
<td>hor</td>
<td>NC</td>
<td>koor</td>
</tr>
<tr>
<td>b) medially</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'brain/crown of head'</td>
<td>--</td>
<td>masihaa/maska</td>
<td>NC</td>
<td>maskah</td>
</tr>
<tr>
<td>c) finally</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'foot'</td>
<td>lukka</td>
<td>loh</td>
<td>lu₉</td>
<td>NC</td>
</tr>
<tr>
<td>d) no change/ --</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'tooth'</td>
<td>C</td>
<td>0l₉o₀</td>
<td>'el₉a</td>
<td>il₉o</td>
</tr>
</tbody>
</table>

Jiiddu spirantization, then in Table 1 Jiiddu should not be the sole participant of this change. It appears that spirantization was a matter of independent development, then, in Jiiddu and Northern Soomaali: the very limited development of spirantization in Northern Soomaali (in Lamberti's [1983:31] words, "after *e, *a and *ə") cannot be compared to the Jiiddu change, which is not vowel-triggered. Indeed, Banti's caution in characterizing the outcomes of *k in Soomaali is another form of disagreement with Lamberti's formulation. Banti notes that one type of *k change takes place "under conditions that seem to be determined by lexical diffusion rather than by phonological differences", and that *k "may become a pharyngeal fricative, i.e. h ...This is clear in Somali rahn and Rendille rah 'frog' vs. Somali rak-e, Oromo rasæ-a ..." [Banti 1983:15-16]. (NB: rahn is a Northern Soomaali form, rak-e a Tunni or Maay form.)

4.2. Palatalization of velars. Examination of this set reveals that Garree and Tunni do, indeed, share in palatalization with Jiiddu. But before the presence of palatalization in these three languages is construed as support for a "Digil" cluster, note that Rendille also participates in the palatalization of velars before front vowels.

Table 2. Palatalization of velars

<table>
<thead>
<tr>
<th>Bayso</th>
<th>Jiiddu</th>
<th>Rendille</th>
<th>Garree</th>
<th>Tunni</th>
</tr>
</thead>
<tbody>
<tr>
<td>*k 'bird'</td>
<td>kimblr</td>
<td>ʃabberə</td>
<td>cimblr</td>
<td>ʃimmir</td>
</tr>
<tr>
<td>*g 'knee'</td>
<td>glib</td>
<td>jəleb</td>
<td>jilbl</td>
<td>jilbl</td>
</tr>
</tbody>
</table>

Lamberti has correctly pointed out that Bayso does not participate in this "older" sound change. But the Rendille example shows that this change was not limited to Lamberti's "Somali". Moreover, there are cases where Jiiddu lacks palatalization yet the other languages have it. Compare Jiiddu with Rendille in the case of 'water':

*k 'water' bekee *behe bice bյyє (behe)
(That the modern Jiiddu form *ware was preceded by earlier Jiiddu *behe is surmised from the loanword for that meaning in Tunni, which takes the expected Jiiddu form.) Palatalization neither helps to group Jiiddu with Bayso (which lacks it) nor with Lambert's "Digil", since the change is not exclusive to "Digil". As Heine [1978:13-14] has shown, palatalization of velars before high front vowels is a Sam (a Soomaali II) level change in Omo-Tana. While Table 2 shows Jiiddu also has palatalization, it is important to recognize that this palatalization, as 'water' reveals, is not identical to the Soomaali II palatalization, since there are cases in Soomaali II where the affrication of velar stops has applied, which have not been so changed in Jiiddu.

Other sound change evidence further weakening Lambert's arguments for a "Digil" unity and simultaneously supporting a Jiiddu-Bayso relationship is found in *a-raising.

4.3. Raising of *a. In his discussion of pharyngeals, Sasse [1979:36] also described an Omo-Tana level change for Bayso and Western Omo-Tana of *a-raising to either e or i in the environment of a pharyngeal. The following shows that this also takes place in Jiiddu. Bayso and Jiiddu have front vowels from this *a-raising, where Rendille, Garre, and Tunni retain older *a. This shared innovation not only points toward the closer relationship of Jiiddu and Bayso, it further detracts from the claim for palatalization as an "older" law, as we shall see.

Table 3. Raising of *a

<table>
<thead>
<tr>
<th>Baysa</th>
<th>Jiiddu</th>
<th>Rendille</th>
<th>Garre</th>
<th>Tunni</th>
</tr>
</thead>
<tbody>
<tr>
<td>'head'</td>
<td>mete</td>
<td>midj'i</td>
<td>matah</td>
<td>madah</td>
</tr>
</tbody>
</table>

The age of this raising rule explains why the following words are palatalized in Jiiddu, but not in the rest of the languages where palatalization occurred: the original *a had been raised in Jiiddu, but not in the rest.

Table 4. New Jiiddu environments for palatalization

<table>
<thead>
<tr>
<th>Baysa</th>
<th>Jiiddu</th>
<th>Rendille</th>
<th>Garre</th>
<th>Tunni</th>
</tr>
</thead>
<tbody>
<tr>
<td>'name'</td>
<td>mege</td>
<td>miji'</td>
<td>magah</td>
<td>ma'ag</td>
</tr>
<tr>
<td>'hand/arm'</td>
<td>gene</td>
<td>jini'</td>
<td>--</td>
<td>ka'an</td>
</tr>
</tbody>
</table>

5. Conclusion

We have shown that there is a rule older than palatalization in Jiiddu: *a-raising. If this latter were not older than palatalization, the Jiiddu column in Table 4 would have had velars followed by front vowels (from *a-raising), unpalatalized velars in the environment which created palatals! This shared *a-raising in Jiiddu and Bayso counters the case for a "Digil" cluster by demonstrating that both palatalization and spirantization were changes which proceeded in Jiiddu separately from the partially parallel changes in the Soomaali II, or Sam, branch, although areal influence may have helped to motivate in one case or the other.

The Jiiddu palatalization in its particulars (the non-occurrence in 'water'
and the unique occurrences in Table 4) reveal that there is little ground, using this velar palatalization, for Lamberti's placement of Jiiddu in "Digil", since (i) this change is not unique to "Digil", and (ii), given that Jiiddu palatalization was preceded by "a-raising, which fed it, while the Soomaali II palatalization was not so affected, the changes are not identical. In other words, palatalization in Jiiddu had to have occurred after the separation of Jiiddu from the rest of Soomaali.

[Without the generosity of Professor Christopher Ehret, who provided his time as well as his personal field notes, this paper would not have been possible. We thank Prof. Ehret while we hasten to add that any errors to be found herein are wholly our own device.]

REFERENCES


ON THE REALITY OF VOWEL COALESCEENCE IN YORUBA

Oladele Awobuluyi
University of Ilorin

All the cases of vowel coalescence reported in the literature so far follow the same pattern when considered from the standpoint of the outputs of the process: such outputs are always vowels in both the front and back series. Thus, in Swahili the vowels /a/ and /i/ coalesce to produce /e/, a front vowel, while /a/ and /u/ coalesce to produce /o/, a back vowel. Similar occurrences have been reported for Kasem (Chomsky and Halle [1968:358-364], following Callow [1965]), Òwọn Afa [Awobuluyi 1972], and Xhosa [Aoki 1974].

It would appear, however, that that is not the only possibility in language. There is evidence to suggest that Standard Yoruba not only displays vowel coalescence, as recent literature on the phonology of the language fails to indicate, but also features output vowels belonging to the back series only. The relevant data are exemplified below:

(1) i. o ni ikéde → olúkéde
agent possess announcement 'announcer'

ii. o ni iyà → olùyà
agent possess suffering 'ill-starred person'

iii. lòq kò Isòq
saying any saying 'foolish/loose talk'

iv. iwà kò iwà → ìwàkúwà/ìwàkìwà
habit any habit 'bad habits, any habit'

v. ibi lìjokòö → ibùjokòö
place sitting 'seat'

vi. ibi ìgbè → ibùgbè
place habitation 'abode, residence'

vii. a rì ogbò → arúgbò
agent see old-age elderly person'

viii. òjì lé igba → ìjììlùgbà
forty be-in excess two hundred 'two hundred and forty'

ix. ò se ìwòn → ìṣùwòn
agent be measurement 'unit of measurement'

x. jè ìbà → ìjìba
pay homage 'pay homage'

xi. gbọ lìrọó → gbùrọó
hear sound 'hear from someone'

xii. pa irọ → purọ/parọ
tell falsehood 'tell lies'

xiii. sà iré → sûré/sáré
run race 'run a race'
The phonetic mergers encountered in the above data can be summarized as follows:

\[(2)\]

i. \( i + i \rightarrow u \) (in (1), i-vi)
ii. \( i + o \rightarrow u \) (in (1), vii)

(iii) \( e + i \rightarrow u \) (in (1), viii-ix)
iv. \( e + i \rightarrow u \) (in (1), x)

v. \( o + i \rightarrow u \) (in (1), xi)
vi. \( a + i \rightarrow u \) (in (1), xii-xiv)

vii. \( a + o \rightarrow u \) (in (1), xv)

viii. \( a + o \rightarrow u \) (in (1), xvi)
ix. \( i + e \rightarrow o \) (in (1), xvii)

x. \( u n + e \rightarrow o \) (in (1), xvii-xix)
xi. \( u n + e \rightarrow o \) (in (1), xx-xxi)

xii. \( u n + e \rightarrow o \) (in (1), xxii)

xiii. \( a + e \rightarrow o \) (in (1), xxiii)

The phonological process illustrated in (1-2) regularly occurs in every Standard Yoruba utterance meeting the structural descriptions of (a) the agentive construction illustrated in (1,i-ii) and (b) the disapprobative construction in (1,iii-iv). Its next highest degree of occurrence (probably ten to fifteen cases) is in the genitival construction in (1,v-vi). Elsewhere, it occurs
sporadically, but such that the total number of utterances involved may well be over two hundred.

Considerations of the overall size of the relevant data apart, the clear pattern formed by the output vowels seen in (2) suggests that some generalization lies hidden in (1). That generalization would definitely be missed should the utterances exemplified in (1) be regarded as exceptions to be handled with a hodgepodge of ad-hoc devices.

To capture the generalization in question, this paper proposes that Standard Yoruba displays vowel coalescence. The process appears to be governed by the following five related, partly language-specific and partly cross-linguistic, principles:

(3) i. the output vowel must be rounded;
   ii. the output vowel must be phonetically distinct from the input vowels;
   iii. the output vowel copies the functional tongue height of the second input vowel;
   iv. in the event that (3ii) and (3iii) clash, the output vowel assumes the highest tongue height;
   v. the two input vowels must not both be rounded.

There is no viable alternative to vowel coalescence as a means of accounting for the data in (1). Thus, the data cannot be accounted for with vowel deletion, because most of the underlying forms that deletion would require would be either ill-formed, non-existent or semantically distinct from their corresponding surface versions.

Vowel harmony supplemented by total assimilation and vowel deletion cannot be used either. For one thing, that method would require the employment of a rule of total assimilation across a stop which does not otherwise exist. For another, even with that kind of unattested rule plus deletion, the method would only be able to handle four of the twenty-three examples in (1), namely, (1 xviii-xix, xxi-xxii).

Partial assimilation-cum-deletion, the method advocated by Aoki [1974], is of no avail either. In the first place, there is no rule of partial vowel assimilation in Standard Yoruba. Secondly, such a rule would in any case be of no effect whatsoever for eight of the thirteen formulas in (2) above. Thus, there is no way that e and i could partially assimilate to eventually yield u, as in (2 iii).

Stahlke [1976] rejects Aoki's approach to vowel coalescence and endorses instead the Chomsky-Halle [1968:358-364] proposal that vowel coalescence be handled by means of transformational rules. The fact, however, is that such rules can only effect coalescence, miraculously as it were, and without being able to explain precisely how the changes concerned are accomplished. It must be admitted, therefore, that they are unsatisfactory.

Stahlke [1976] makes the interesting observation that the output of coalescence is always a compromise between the two original input segments. If the technical problems involved in formalizing this observation could somehow be overcome, it might provide a viable alternative to conventional transformational
rules as a means of handling coalescence in an explanatory fashion. It turns out, however, that many of the outputs of vowel coalescence in Standard Yoruba can in no way be construed as compromise forms. Thus, it is not clear how \( u \) can be called a compromise between \( a \) and \( i \) in (2) vi. \( a + i \rightarrow u \). Hence, even if Stahlke's observation could somehow be formalized, it would be unable to handle most of the Standard Yoruba data above.

In these circumstances, it is clear that the Standard Yoruba case reported here shows that (vowel) coalescence poses a far greater challenge to phonological theory than was previously thought to be the case. The challenge has to do specifically with (a) explaining in rational and non-metaphysical terms the precise way or ways in which the phonetic changes it involves are accomplished and (b) evolving a rule schema which fully incorporates such an explanation. Standard Yoruba offers no clues whatsoever in this regard, neither do the other languages examined so far, going by the indications in Stahlke [1976]. Perhaps still other languages will, when relevant information becomes available on them.

REFERENCES


In English the semantic load of negation affects either the whole sentence or part thereof. In the case of Egyptian Colloquial Arabic (ECA) the negative particle (the most common of which are mif and ma-...) is realized structurally on the modal (when there is one and the tense is present) or on the tense carrier (the auxiliary kaan when the tense is past). However, it could also be attached to the lexical verb (a syntactic option which is ungrammatical in English). In presence of a modal, if the negativizer appears on the MV that follows, the resulting string is not the negative counterpart of the positive (modal) sentence. Rather, the positiveness of the modal is retained and the structure would then come to indicate that the subject can, would, might... etc. (when he likes) not + MV. In other words, the negative load is both syntactically and semantically externalized. In this case the discontinuous negativizer ma-... could be employed. When the negative appears on the modal itself some will accept the single word negativizer mi and others the discontinuous ma-...

The single word negativizer mi appears before (i) adjectives, (ii) nouns, (iii) lexical verbs (with or without the future prefix ha- attached), and (iv) certain particles notably, ma, and wayya. The discontinuous negativizer, on the other hand, appears on the pronouns. This is only a syntactic option used for the expression of certain attitudes shown later. The more frequent Pron. + (mi + MV) string signifies no overtones.

The examples given below together with their paraphrase gloss establish the fact that the two negative particles are not in free variation. Selectional restriction rules will both delimit and govern their use. In other words, structural variations which appear to be stylistic options to the uninitiated speaker perform different functions, i.e. convey different messages. Any change in the message affects the code (ambiguous sentences excepted) though not vice versa.

Examples of negative structures with mi:

(1) mif mumkin aqiilik bukra
not possible I come to you tomorrow
'I cannot come to you tomorrow'

(2) mif ?ablilin b?d xaalilis
not they accept each other completely
'they totally resent each other'

(3) ?anna mif ?ayyala ?andak
I not I work at you
'I don't work for you'

(4) mif hasma?ak
not I listen to you
'I won't listen to you'
Studies in African Linguistics, Supplement 9

(5) 1īlīfārabiyya di mīʃ sāriifa ʔašlī  pre-adjectival position
the car this not fast very
'this car isn't very fast'

Examples of negative structures with ma-...-

(6) mafīʃ 1uzuum tiigli badri
there isn't need you come early
'there's no need for you to come early'

(7) manīʃ gāyy bukra
I'm not I come tomorrow
'I'm not coming tomorrow', i.e. I insist on not...

(8) magatluf ʿalaawā min sanateen
not come to him increase from two years
'he hasn't had an increase in salary for two years'

Since the grammar of ECA is greatly synthetic, i.e. depending largely on enclitics to map up word relations, word order is relatively loose. The discontinuous negativizer together with appropriate enclitics can generate single word negative sentences like the following:

(9) mafhimtiʃ 'I/you didn't understand'
mabiyeʃrafʃ 'he doesn't know'
mazarunaaʃ, 'they didn't visit us'
maraʃluhumʃ 'he didn't go to them'

Although the single word negativizer is fast becoming standard with independent pronouns, we still hear ma-...-f with these but with special overtones and speaker attitudes towards the communicative value of his utterance. That is, negative pronouns when used in statements indicate something like insistence, warning (or even threat), (self)assertion of authority, whereas in questions (or tags), they indicate something like surprise, resentment, or disagreement. Notice the difference between the following pair:

(10) a. mantiiʃ gāyya maʃaana ↑ 'aren't you coming with us?' you(fem.)are not you coming with us i.e. I'm surprised if you aren't

b. mantiiʃ gāyya maʃaana ↑ 'you're not coming with us' i.e. I insist that you don't come with us

The following are the negative pronouns of ECA:

SINGULAR                    PLURAL
maniiʃ  1st person (masc. & fem.)  mahnaaʃ  1st person
mantaaʃ 2nd person (masc.)         mantuuʃ 2nd person
mantiiʃ 2nd person (fem.)
mahuwwaasʃ 3rd person (masc.)  mahummaasʃ 3rd person
mahiyyaaʃ 3rd person (fem.)

(N.B. Gender distinction is not shown on negative plural pronouns.)
Negation in Egyptian Colloquial Arabic

Use of the discontinuous negativizer ma-.. may result in certain phonological changes exemplified by the following:

(i) In the first person singular /-i/- is used, i.e. maliif, maayiiif, (cf. maniiif above).

(ii) The vowel of the particle is elided in the following: mal(l)haaf, mal(l)kiif, mal(u)humf, mal(u)kumf, mam(a)yiiif, mam(a)kaaf, etc.

(iii) The final /-h/ of the prepositional fiih disappears from the corresponding negative form (unless required by agreement with a preceding noun). Thus,

(11) maliif taalaba filfaal 'there aren't any students in the classroom'

(12) filfaal maliif taalaba 'there aren't any students in the classroom there isn't in it students classroom'

(iv) In a line with a trend that applies to other suffixes, the use of /-i/- implies the following:

(a) occurrence of the extra vowel to avoid a sequence of three consonants, e.g.

(13) lissa marutfi 'I haven't gone yet'

(b) lengthening of a preceding vowel, e.g. makatabuulf 'they didn't write/he didn't write it', maguulf 'they didn't come' (3rd person plural, no gender distinction shown), magaaf 'he didn't come' (3rd person singular masc.).

In negative imperatives the second person imperfect forms are used:

(14) a. matiidxul 'don't go in!

b. matzaa?i kiida 'don't shout so!'

c. matxailihum yintizru 'don't let them wait!'

d. matit?axr? bukra 'don't be late tomorrow!'

other negative structures in the dialect include the following:

(15) laada walaada laakin dukha 'neither this nor this but that'

(16) latxilin wala-rafaya? laakin mutawassii 'neither fat nor thin but average'

(17) lazaa kir wala hayzaa kir 'he has neither studied nor is he going to'

(18) ?irraaqiil da labyesma? wala byetakallim 'this man neither hears nor talks'

Examples (15-18) establish laa .. wala as corresponding to the English 'neither .. nor'. The actual negative particle in the construction is laa- (lengthening of the vowel in example (15) above is due to contraction). wa is
prefixed to every la after the first. Theoretically there is no limit to the number of words which may be preceded by the particle.

Imperfect prefixes, like the future marker ḥa-, are often omitted with verbs other than the first as in

(19) laḥyiili wala yētkallim 'he will neither come nor give us a
he will not come and not he talks ring'

Compare (19) with (17) above.

It is possible for /Ja/ to be suffixed to a verbal form after the first, e.g.

(20) la saʔal Ṣannak wala xaragf 'he has neither asked about you nor not he asks about you and not has he left'
he not leave

It is also quite as frequent to have the first negative of the discontinuous type ma-. For example,

(21) maxaragf yētmajfa wala zakirf druusu he neither went out for a walk nor didn't go out to stroll and not he did he study his lessons didn't study his lessons

Emphatic negatives are those which have ma- without its final -f. This type of negative occurs frequently with a few words commonly associated with negation, e.g. ʕumr and ḥadd, and with oaths. The first of these words ʕumr (lit. life) usually shows as its final sound a pronominal suffix. It comes to mean 'never' when ma- appears on the ensuing verb. Even on the occasion when one hears -f it must be shown on ʕumr after the pronominal suffix has been added. Compare the following:

(22) a. ʕumri maʕuft waḥid zayyu 'I have never seen anyone like him'
in my life I didn't see someone like him

b. maʕumriʕ f i f waḥid zayyu 'never in my life have I seen anyone like him'

Note: the occurrence in (22b) of -li- before the final f.

The second type of emphatic negative is oaths. The common oaths in ECA are wallaahi (lit. and my God), wirabbina (lit. and our Lord) and winnabi (lit. and the prophet), to which may be added ʔin alla (God willing) and in which case a following perfect seems to be a must. Thus:

(23) laa wallaahi maʔdar 'no, by heaven I cannot'
no by God I cannot

(24) ʔinalla maḥaddi xarag 'may nobody ever go out, then!'
even though nobody leaves
GUMUZ, KOMAN, MAO, AND OMOTIC

M. Lionel Bender
Southern Illinois University-Carbondale

1. Introduction

Several languages of the Ethio-Sudan border area have long defied classification. Among these are Gumuz (a single language with many local varieties of which I will deal with Kokit, Sai, Sese, Disoha, and "Hamej" of Ethiopia) and "Mao" (several languages, of which I will deal with Hozo-Sezo, Bambeshi-Diedesa, and Ganza). The major difficulty lies in deciding whether these languages belong to Koman (of the Nilo-Saharan Phylum) or Omotic (of the Afroasiatic Phylum). Partial reconstruction of lexica of these two families now makes it possible to arrive at an unequivocal answer [Bender 1984, forthcoming].

2. Gumuz is Koman

Out of the first 100 basic items on my Koman lists, about 46 can be reconstructed quite firmly to Proto-Koman. The structure of Koman is as follows: I. Komo-Twampa (Uduk), 77% shared basic lexicon; II. Opo (Ansita) vs. I, 62% average shared basic lexicon; III. Kwama vs. II, 42%; IV. Anej (Gule) vs. III, 28%. Note that Anej-Kwama is 18% only, Anej vs. others averages 31%.

Only the least problematical items found in both Anej and the rest of Koman (or strongly in Koman if not found in Anej) will be considered here. A number of other cases of likely comparisons with part of Koman or Anej alone will not be listed here.

<table>
<thead>
<tr>
<th>Item</th>
<th>Common Gumuz</th>
<th>Proto-Koman¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.</td>
<td>bird</td>
<td>mbit</td>
</tr>
<tr>
<td>9.</td>
<td>blood</td>
<td>maxa</td>
</tr>
<tr>
<td>11.</td>
<td>breast</td>
<td>kuwA</td>
</tr>
<tr>
<td>14.</td>
<td>cloud</td>
<td>k'uda</td>
</tr>
<tr>
<td>19.</td>
<td>dog</td>
<td>k'owa</td>
</tr>
<tr>
<td>22.</td>
<td>ear</td>
<td>(t)s'ea</td>
</tr>
<tr>
<td>23.</td>
<td>eat</td>
<td>sa</td>
</tr>
<tr>
<td>31.</td>
<td>foot</td>
<td>cogwa</td>
</tr>
<tr>
<td>32.</td>
<td>give</td>
<td>kYE, tYE</td>
</tr>
<tr>
<td>43.</td>
<td>kill</td>
<td>Sok'w</td>
</tr>
<tr>
<td>49.</td>
<td>louse</td>
<td>sukuna</td>
</tr>
<tr>
<td>56.</td>
<td>neck</td>
<td>-biya</td>
</tr>
<tr>
<td>73.</td>
<td>sleep</td>
<td>Ji?i</td>
</tr>
<tr>
<td>75.</td>
<td>smoke</td>
<td>dukwa</td>
</tr>
<tr>
<td>82.</td>
<td>tail</td>
<td>tsa</td>
</tr>
<tr>
<td>87.</td>
<td>tongue</td>
<td>kwoteta</td>
</tr>
<tr>
<td>89.</td>
<td>water</td>
<td>a(?)Ya</td>
</tr>
<tr>
<td>95.</td>
<td>what?</td>
<td>ins't'e</td>
</tr>
</tbody>
</table>


²Consider reverse order ("metatheses").
The only convincing agreement with Proto-Omotic is 16, 'come': Gumez we, Omotic *wA.

3. "Mao" Languages are Omotic

The structure of Omotic is something like: I. North Ometo-South Ometo, average shared basic lexicon 63%; II. Chara vs. I, 44%; III. Gimira vs. II, 42%; IV. Kefoid (Gonga) vs. III, 38%; V. Janjero (29%) and Dizoid or Majoid (28%) vs. IV; VI. Aroid (South Omotic) vs. V (North Omotic), 19%. Note that Dizoid-Janjero is 18%.

"Mao" languages agree with about half of approximately 58 tentative Proto-Omotic items. The best cases (some having only one of "Mao"-Hozo-Sezo-Ganza represented) are given below.

<table>
<thead>
<tr>
<th>Item</th>
<th>&quot;Mao&quot;</th>
<th>Hozo-Sezo</th>
<th>Ganza</th>
<th>Proto-Omotic</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. bird</td>
<td>ke</td>
<td>E</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>7. burn</td>
<td>taas- *</td>
<td>(§)aTs</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>13. claw</td>
<td>s'uk'um</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>16. come</td>
<td>Gumez</td>
<td>kwa</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>19. dog</td>
<td>kana</td>
<td>----</td>
<td>kana</td>
<td>----</td>
</tr>
<tr>
<td>20. drink</td>
<td>i§</td>
<td>i§</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>22. ear</td>
<td>waale</td>
<td>we(y)</td>
<td>waya</td>
<td>----</td>
</tr>
<tr>
<td>23. eat</td>
<td>mi</td>
<td>ma, me</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>25. eye</td>
<td>aagc</td>
<td>abi, awi</td>
<td>ap</td>
<td>----</td>
</tr>
<tr>
<td>26. fat</td>
<td>k'o(t)s</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>28. fire</td>
<td>kaame</td>
<td>taame</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>31. foot</td>
<td>toge</td>
<td>tuugi</td>
<td>tuku</td>
<td>----</td>
</tr>
<tr>
<td>35. grass</td>
<td>maara</td>
<td>mis('t)</td>
<td>mati</td>
<td>----</td>
</tr>
<tr>
<td>37. hand</td>
<td>kuse</td>
<td>ku(t)s</td>
<td>konso</td>
<td>----</td>
</tr>
<tr>
<td>38. head</td>
<td>tooke</td>
<td>tokl, towil</td>
<td>kwothi *</td>
<td>----</td>
</tr>
<tr>
<td>42. I</td>
<td>tiya</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>44. knee</td>
<td>tukaum</td>
<td>k'umi</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>51. meat</td>
<td>o(t)si</td>
<td>wasi</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>52. moon</td>
<td>aansa</td>
<td>eems</td>
<td>anzi</td>
<td>----</td>
</tr>
<tr>
<td>54. mouth</td>
<td>----</td>
<td>----</td>
<td>nana</td>
<td>----</td>
</tr>
<tr>
<td>59. nose</td>
<td>$inte</td>
<td>$in(t')i</td>
<td>sindi</td>
<td>----</td>
</tr>
<tr>
<td>62. person</td>
<td>esse</td>
<td>----</td>
<td>ciya</td>
<td>----</td>
</tr>
<tr>
<td>75. smoke</td>
<td>ts'ue</td>
<td>s'ubi, s'uwini</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>76. snake</td>
<td>s'oo§A</td>
<td>s'oo§E</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>80. sun</td>
<td>awa</td>
<td>aBi</td>
<td>aba</td>
<td>----</td>
</tr>
<tr>
<td>88. tooth</td>
<td>aats'E</td>
<td>(h)aats'E</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>89. tree</td>
<td>inziE</td>
<td>in(t)s</td>
<td>insa</td>
<td>----</td>
</tr>
<tr>
<td>92. water</td>
<td>haats'E</td>
<td>haan(ts')</td>
<td>ha?a</td>
<td>----</td>
</tr>
<tr>
<td>98. woman</td>
<td>muns'E</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
</tbody>
</table>

1 From work now in process.

2 Possible reversal.
4. Loans Which Deceive

The following are fairly clearcut cases of loans whose status as such becomes clear when tentative proto-languages are taken into account:

17. cut  Kwama k'ofə, "Mao" k'ofi.
19. dog Kwama kəna (< Omotic); Hozo-Sezo w+ʒi (< Amharic).
28. fire Gumuz manja (prob. Nilotic).
29. fish Hozo-Sezo wats'ĩ, Koman wats'ĩ.
40. heart "Mao" ḍẹẹ, Kwama ḍẹẹ.
45. know Hozo-Sezo-'Mao' al-, Kwama aļa, Twampa ari.
64. red Hozo-Sezo tant, Kwama tantaa.
69. see Kwama ʒi, Hozo-Sezo ʒi.
70. seed Hozo-Sezo and Omotic zara (< Amharic); "Mao" ʒok', Kwama ʒoco.
72. skin Kwama ʒoŋk'əo, Hozo-Sezo and Omotic go(ŋ)k.
79. stone "Mao" ʒoọe, Opo ʒawa or jəo, other Koman reversed as (w)ọọ.
82. tail "Mao" yọŋka, Janjero ọọ, Kefa ụnọ, Kwama ọongο(ʔ).

REFERENCES


NOTE

[A full version of this paper entitled "The Limits of Omotic" will appear in R. Hayward, Forth.: Omotic Language Studies.]
QUANTITATIVE ANALYSIS OF SWAHILI VOCABULARY

Elena Bertocnini
Istituto Universitario Orientale, Napoli

It is a well-known fact that an important portion of the Swahili vocabulary consists of Arabic and other borrowings. I tried to figure out the ratios of the Oriental and European loanwords to Swahili words in the Swahili language in my Tentative Frequency List of Swahili Words [1973], distinguishing between single words and occurrences.

In the present analysis I have posed the following questions: in which spheres of the Swahili vocabulary are loanwords most numerous and in which spheres are Bantu words most numerous? The material I used was 40,000 word occurrences taken from 100 sources which were divided into four groups: ancient texts, modern literary texts, journalistic texts, and miscellaneous contemporary texts. In order to embrace the largest strata of the written language, samples of most varied kinds were chosen so that they might represent the greatest possible number of Swahili styles, from classical poetry, chronicles, and fairy tales to modern poetry and short stories, from textbooks and political pamphlets to newspaper articles including political news, reports, crime page, economy, culture, sports, and so on. The complete list of the sources is to be found in Bertocnini [1973]. It must be borne in mind that the material for this analysis was collected in the 2nd half of the Sixties, when the only modern prose writers available to me were Shaaban Robert and the authors of the short stories and essays published in the journal Swahili.

The analysis is based on Hallig & Wartburg's [1952] division of the vocabulary, partly modified according to the needs of Swahili. The whole Swahili vocabulary is divided into 7 sections and several sub-sections (cf. Table I.)

The 1st section, Nature, includes the sky and the air, the earth, plants, and animals. Besides the nouns, there are verbs such as -pambazuka 'become clear', -vuma 'blow', -iva 'become ripe', etc.

The 2nd section, Man as a physical being, contains the following sub-sections: the human body, its organs and their functions (there are such verbs as -konda 'grow thin', -kua 'grow up', -lamba 'lick', etc.); the senses and their activity; sleeping and waking; health and illness, including medical service; movements and positions (besides many nouns and verbs, there are some adverbs such as upesi 'quickly'; further eating, clothing, and housing.

In the 3rd section, Man as a spiritual being, there are the concepts connected with thinking, feeling, willing, and doing (in a general sense, not the actual work of jobs which are in the 5th section; I include here, for instance, -acha 'leave', -gumu 'difficult', tatizo 'problem', mpango 'project', huenda 'maybe', -faa 'be useful', lengo 'aim'); there are also the sub-sections ethics, language, the mass media, the arts and social sciences, and religion.

In the 4th section, Man as a social being, we find existence in general, where I put e.g. kuwa 'to be', ulimwengu 'the world', -enyewe 'self', klumbe 'creature', and kitu 'thing'; there are further the subsections human
life from birth to death; the family and marriage; social relations, pastimes and festivities, and sports.

The 5th section, Social organization and politics, is quite extensive. It concerns communities and countries; state, government and politics; justice and police; defense, war and the army; education; agriculture, hunting and fishing; work and employment with the relative tools and materials; technique, industry, the economy, and the natural sciences; commerce, finance, and property (here we find any concept related with possession, such as -pā 'give', -potea 'get lost', uko-sefu 'lack', mwombaji 'beggar'); the last sub-section is transport, post and telecommunications.

The 6th section, Natural laws, includes properties (with brightness and darkness, sounds, and most adjectives except the evaluating ones like -ema 'good', -baya 'bad', which are put into ethics). The sub-section forms, values, methods, and relations contains such heterogeneous items as mfano 'example', hivi 'like this', kuliko 'more than', mbalimbali 'different', mvingingo 'anything round', -inge 'another', -panga 'put in order', etc. The remaining sub-sections are quantity, order, weights and measures; space; time; causes and conditions (with some adverbs and conjunctions like kwa sababu 'because'); changes and evolution, e.g. -sitawi 'flourish'.

In order to embrace the whole of Swahili vocabulary, I added the 7th section, Pronouns and particles (that is, exclamation, some conjunctions and prepositions).

Often the same item could belong to more than one section or sub-section, e.g. kufa 'die' could be in human life or existence. In this case I have chosen one section in which the word is counted.

I stick to the meaning actually found in the texts, even if it may be different from that given by the dictionaries. In some cases the meaning of a word changes in different groups of texts. For instance, mkutano 'meeting' is put into social relations in Group II (modern literary texts) and into politics in Group III (journalistic texts); upinduzi means "upsetting" (feeling) in Group I (ancient texts), and "revolution" in Group IV (miscellaneous contemporary texts).

I compiled a separate word list and the relative word counting for each of the four groups of texts and then did the same within these groups for each language cluster, i.e. for words of Bantu, Oriental, or European origin. Although each group is based on 10,000 word occurrences, it must be borne in mind that the number of different words is much smaller. (See Table II.)

In all types of texts the words of Bantu origin represent little more than 50%, but there are among them the words with the highest frequency, hence in the count of occurrences the ratio of the words of Bantu origin is much higher, about 70%. (Words from other African languages, e.g. Cushitic, are also included here.)

The ratio of the Oriental loanwords oscillates between 45% in the ancient texts and 30% in the newspapers. The European borrowings, on the contrary, are most frequent in the newspapers (20%) and practically non-existent in the ancient texts.

Another observation we can make from the above table is that the ancient
texts have the richest vocabulary, that is, the highest number of different words, and the newspapers have the poorest vocabulary.

As regards semantic division of the vocabulary into sections, in all groups the less numerous section is the last one, Pronouns and particles, but it contains some Bantu and Oriental items with a very high frequency. On the other hand, the first section, Nature, not only contains few items, but their frequency is also low. Only in the ancient texts, Bantu words in this section are slightly more numerous.

The richest section on the whole is C. Man as a spiritual being, except in the newspapers and in the European cluster, where E. Social organization leads. An important section for the Bantu cluster of the ancient texts is B. Man as a physical being, whereas F. Natural laws is well represented in modern literary texts and in the Oriental cluster as a whole. D. Man as a social being has some importance only in the European cluster and in the newspapers but contains two Bantu words from the top of the frequency list: kuwa 'to be' and mtu 'man'.

Observing Table I, you will see beside each section and sub-section in brackets the total number of items and the percentage of items in the Bantu, Oriental and European clusters respectively. These figures are the sums of the results obtained in each group. If I had made a single list instead of four, the resulting figures would be lower because many items recur in all groups.

Summarizing the results of the analysis, it can be affirmed that most words of Bantu origin are connected with movements and positions. There are also many Bantu words in the sphere of spiritual life (feelings, willing and doing, ethics), but the majority of items in this sphere is provided by Oriental loanwords. The same is true of terms dealing with time and quantity.

Bantu words predominate in zoological and botanical terminology and in the sub-sections concerning the sky and the air, the earth, the human body, sensorial activity, sleeping, food, clothes, and housing as well as in the sub-sections concerning agriculture, the family and marriage, human life, work, industry, social organization, space, and properties. They have a slight majority in the sub-sections on education, language, social relations and pastimes, the mass media, the arts and the social sciences.

Oriental, mostly Arabic, loanwords provide the religious terminology and other terms concerning spiritual activity. Many of them are connected with existence and with natural laws (time, quantity, forms and relations). Even many particles (conjunctions and exclamations) are of Arabic origin. Most of these borrowings entered Swahili long ago; however, Oriental loanwords also predominate in some spheres of modern life, such as commerce, justice, politics, and transport, but in these last two sub-sections all three language clusters are of almost equal importance. European, above all English, loanwords are found especially in the newspapers and concern mostly modern life. They occur most often in the spheres of politics, transport and sports, and also in those of social relations and pastimes. Other terminological sub-sections relatively well represented are commerce, justice, education, work and tools, the arts, the mass media, medical service, food, and clothes. Finally, several European borrowings concern time, weights, and measures.
## Table I. Semantic Division of the Swahili Vocabulary

Figures in parentheses represent total number of items: Bantu %, Oriental %, European %.

### A. NATURE (429: 71, 26, 3)

I. The sky and the air. (48: 65, 27, 8)
II. The earth. (138: 65, 31, 4)
III. Plants. (112: 70, 28, 2)
IV. Animals. (131: 82, 18, 0)

### B. MAN AS A PHYSICAL BEING (1,124: 73, 21, 6)

I a. The human body, its organs and their functions. (147: 89, 10, 1)
b. The senses and their activity. (42: 90, 10, 0)
c. Sleeping and waking. (20: 95, 5, 0)
d. Health and illness; medical service. (86: 51, 31, 18)
e. Movements and positions. (442: 81, 18, 1)

II a. Eating. (121: 63, 22, 15)
b. Clothing. (108: 54, 31, 15)
c. Housing. (158: 62, 30, 8)

### C. MAN AS A SPIRITUAL BEING (2,182: 42, 53, 5)

I a. Thoughts, intellect, aptitudes and inner life. (284: 44, 54, 2)
b. Feelings. (478: 47, 51, 2)
c. The will; general activity, its possibilities and modalities. (786: 40, 58, 2)
d. Ethics and character. (265: 45.5, 54, 0.5)

II a. Language. (256: 44.5, 42.5, 13)
b. The mass media. (65: 45, 29, 26)
c. Literature, music, fine arts, philosophy and social sciences. (130: 45, 35, 20)
d. Religion. (202: 20, 77, 3)

### D. MAN AS A SOCIAL BEING (688: 53, 33, 14)

I. General existence. (127: 47, 52, 1)
II a. Human life from birth to death. (137: 63, 32, 5)
b. The family and marriage. (93: 76, 23, 1)
c. Social relations, pastimes and festivities. (221: 44, 42, 14)
d. Sports.

### E. SOCIAL ORGANIZATION AND POLITICS (1,347: 47, 34, 19)

a. Social organization, communities and countries. (103: 68, 25, 7)
b. State, government and politics. (260: 32, 39, 29)
c. Justice and police. (109: 33, 46, 21)
d. Defense, war and the army. (143: 55, 33, 12)
e. Education. (70: 40, 27, 33)
f. Agriculture, hunting and fishing. (96: 88, 6, 6)
g. Work and employment; tools and materials. (127: 70, 13, 17)
h. Technique, industry, the economy and the natural sciences. (20: 60, 5, 35)
i. Commerce, finance and property. (255: 39, 50, 11)
j. Transport, post and telecommunications. (164: 30, 36, 34)

F. NATURAL LAWS (1,228: 54, 41, 5)
   a. Properties. (207: 68, 30, 2)
   b. Forms, values, methods and relations. (213: 42, 55, 3)
   c. Quantity, order, weights and measures. (245: 45, 47, 8)
   d. Space. (247: 76, 22, 2)
   e. Time. (275: 43, 45, 12)
   f. Causes and conditions. (11: 10, 90, 0)
   g. Changes and evolution. (30: 40, 60, 0)

G. PRONOUNS AND PARTICLES (236: 67, 32, 1)
   a. Pronouns. (99: 89, 9, 2)
   b. Conjunctions, exclamations and some prepositions. (137: 51, 49, 0)

Table II. Number of single words in the sources

<table>
<thead>
<tr>
<th>Category</th>
<th>Total items</th>
<th>Percent of total:</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.   Ancient texts</td>
<td>Total items</td>
<td>1,969</td>
</tr>
<tr>
<td></td>
<td>Bantu</td>
<td>1,070</td>
</tr>
<tr>
<td></td>
<td>Oriental</td>
<td>894</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category</th>
<th>Total items</th>
<th>Percent of total:</th>
</tr>
</thead>
<tbody>
<tr>
<td>II.  Modern literary texts</td>
<td>Total items</td>
<td>1,887</td>
</tr>
<tr>
<td></td>
<td>Bantu</td>
<td>1,040</td>
</tr>
<tr>
<td></td>
<td>Oriental</td>
<td>771</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category</th>
<th>Total items</th>
<th>Percent of total:</th>
</tr>
</thead>
<tbody>
<tr>
<td>III. Journalistic texts</td>
<td>Total items</td>
<td>1,587</td>
</tr>
<tr>
<td></td>
<td>Bantu</td>
<td>790</td>
</tr>
<tr>
<td></td>
<td>Oriental</td>
<td>479</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>318</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Category</th>
<th>Total items</th>
<th>Percent of total:</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV.  Miscellaneous contemporary texts</td>
<td>Total items</td>
<td>1,791</td>
</tr>
<tr>
<td></td>
<td>Bantu</td>
<td>953</td>
</tr>
<tr>
<td></td>
<td>Oriental</td>
<td>625</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>213</td>
</tr>
</tbody>
</table>

REFERENCES


NOMINATIVE/AGREEMENT COMPLEMENTARITY AND VSO ORDER IN STANDARD ARABIC

Hagit Borer
UC Irvine
Laurice Tuller
UCLA

1. Introduction

It has been noted that VSO languages are typically null subject languages. Another property which seems to be a correlate of these is that, in VSO languages, there is a complementary distribution between the appearance of full subject agreement on the verb and an overt subject. That is, if there is full agreement, the subject must be empty, while if the subject is overt, the agreement is empty or incomplete. Standard Arabic (SA) is an example of this situation; basically, the same pattern holds for other VSO languages, such as Breton, Irish, Kwakwa, and Welsh [Borer and Tuller, in preparation]. Consider the sentences in (1) from SA:

(1) a. katabuu 0 t-taqriira
wrote-3pl the-report

b. *katabuu l-wuzaraa' u t-taqriira
wrote-3pl the-ministers the-report

c. kataba l-wuzaraa' u t-taqriira
wrote-3sg the-ministers the-report

In (1a) there is full third person plural agreement on the verb and hence only a null subject is possible. Sentences like (1b), where third person plural agreement occurs with an overt third person plural subject, are ungrammatical. In such cases, the agreement must be incomplete, as in (1c), where the overt plural subject occurs with singular agreement on the verb.

The idea that this nominative/agreement complementarity is related to the VSO property receives further support from the fact that in modern Arabic dialects, the complementarity no longer exists, and the word order has changed to SVO as well.

It is well-known that VSO languages have properties which in other languages are attributed to hierarchical structure. That is, there is evidence that the verb and its direct object have a status that the verb and its subject do not share. This is evidenced in SA by phenomena such as idioms, compositional thematic role assignment, anaphor binding, and Case assignment. Various ways of accounting for the "VP property" of VSO languages have been proposed (see Borer and Tuller [in preparation] for a full review). Here, we will follow work initiated by Emonds [1980] and further pursued by Sproat [1982] and Levin and Massam [1983] who propose that surface VSO order is derived by fronting the verb.

We propose that the VSO word order and the complementarity between agreement and nominative both follow from a single factor, viz. the categorial nature of the AGR(eement) node. Concretely, we propose that AGR in VSO languages is [+N], while in "true" SVO languages, it is [-N]. We further assume
that the [+N] AGR has the properties and distribution of other [+N] elements. We will show that the VSO order, the complementarity, and pro-drop all follow from the interaction of the [+N] node with independently motivated principles of grammar.

2. Nominative/Agreement Complementarity: The Paradigm Case

Following Stowell [1981], we assume that INFL is located in COMP (whether it is base-generated or moved there is immaterial to our analysis).

Both "VSO" and "SVO" languages, then, have the underlying structure in (2). The difference between the two is that while in "VSO" languages AGR is [+N], in "true" SVO languages it is [-N]. [+N] AGR, like all other [+N] elements, may be generated with or without phonological realization. If it is empty, it is subject to the Empty Category Principle, which sanctions the occurrence of empty elements by requiring that they be properly governed (see Chomsky [1981] for definition and discussion). If, on the other hand, AGR is phonologically realized, it, like nouns, is subject to the Case filter [Rouveret and Vergnaud 1980], which requires all N elements to be Case-marked.

The final component of our analysis is the general rule 'Move α', where α may be any category, including V. This will be important since INFL, being affixal in nature, must end up adjacent to the verb. This can happen either if V raises to INFL or if INFL lowers to V (see also Levin and Massam [1983]).

Consider now how these various components of the analysis fit together to produce the complementarity and VSO facts outlined above, looking first at the case in which the [+N] AGR is generated without phonological content. Since it is null, it must be properly governed. As INFL is not a proper governor [Rizzi 1982], the verb must adjoin to COMP in order to properly govern the empty [+N] AGR. An empty subject is impossible, since we assume that null subjects must be identified by an overt AGR element (in the spirit of Taraldsen's generalization and subsequent work on the null-subject parameter).

The result is a VSO structure with a lexical subject, the (c) sentence of (1):

(3)
Consider now the situation where the [+N] AGR is generated with phonological features. Since AGR, in this case, is an overt [+N] element, it is subject to the Case filter, and hence must be assigned Case, in this instance, the nominative Case assigned by INFL. Since nominative Case is absorbed by AGR, the subject NP may not be assigned Case, and hence it may not be overt. The only possible result is therefore a structure with overt agreement and a null subject, the (a) sentence of (1):

(4) s'--------
     |       |
     COMP  S
     |       |
     AGR [+N] NP VP
     |      (overt)
     |      [+NOM]

Notice that verb-fronting is not forced in these structures (AGR is not empty and hence the verb need not move up so as to properly govern it). In principle, then, we can assume that V-fronting is optional. Since the subject is empty, there is no evidence in SA as to whether the verb was fronted. If there is no V-fronting, INFL will lower and affix to the V, so as to satisfy the affixation requirement discussed above.

3. Dislocation and the Complementarity of Nominative Agreement

We would like to consider now some more complex cases, where the complementarity between nominative and agreement appears to break down. Typically, the complementary distribution between nominative Case and overt agreement does not hold in cases of dislocation and movement, as the following cases illustrate:

(5) Left Dislocation
    arrizaalu habuu
    the-men-NOM left-3m-pl
    'the men, they left'

(6) Question
    ayyu rizaalin ahabuu?
    which men-NOM left-3pl
    'which men left?'

Before considering why complementarity is not attested here, let us first distinguish between Left Dislocation constructions and constructions which involve extraction in SA. Based on Ayoub [1981], three pieces of evidence for this distinction can be given. While in Left Dislocation constructions the Case of the dislocated element is always nominative, regardless of the argument position to which it is related (7a), in constructions which involve actual extraction (topicalization, WH-movement) the dislocated element has the Case of its extraction site (7b).

(7) a. arrizaalurafaynaahum
    the-men-NOM saw-we-them
    'the men, we saw them'

b. Zaydan raafa Amrun
    Zaydun-ACC saw Amrun-NOM
    'Zaydun, Amrun saw'
Secondly, Left Dislocated elements must be specific (8a,b), whereas extracted topics may be non-specific (8c):

(8) a. al-kittaabu qaraa-hu Zaydun
    book-NOM read-it Zaydun-NOM
    'the book, Zaydun read it'

b. *kittaabun qaraa(-hu) Zaydun
    book-NOM read-(it) Zaydun-NOM
    'a book, Zaydun read it'

c. kittaaban qaraa Zaydun
    book-ACC read Zaydun-NOM
    'a book, Zaydun read'

Thirdly, while Left Dislocation need not obey subjacency (9a), extraction does (9b,c).

(9) a. Zaydun zaa a r-razulu liatl li ra'aa-hu
    Zaydun came the-man who saw-him
    'Zaydun, the man who saw him came'

b. *al-xayla ta rifu r-razula li li yuhibbu
    the horses-ACC know-you the-man who loves
    'the horses, you know the man who loves'

c. *man ra'ayta l-fataata liatl qaabalat?
    who saw-you the-girl who met
    'who did you see the girl who met?'

Before we proceed to consider the implications that the distinctions between Left Dislocation and extraction have for our analysis, it is worthwhile to note that SA has more than one TOPIC position, as is attested by (10), cited by Ayoub. We thus assume the structure of the TOPIC-COMP complex to be as in (11):

(1). Zaydun abuu-hu yahlumu
    Zaydun-NOM father-his-NOM dreams
    'Zaydun, his father dreams'

(11). 

As will become clear, it is crucial for our discussion that the TOPIC node in (11) head a maximal projection. Concretely, government from the TOPIC position into any position inside S' is not possible.

Returning to the Left Dislocation constructions and the extraction constructions discussed above, it is natural to conclude that while topicalization and questions are derived by actual movement, this is not the case for Left Dislocation constructions. In the latter, a dislocated element is base-generated in position, and a default nominative Case is assigned to it. Note that (5) meets the criteria for a Left Dislocation construction: the dislocated element is nominative, and it is specific. As (5) is a Left Dislocation construction, the absence of complementary distribution between nominative and Agreement is no longer a problem, since the nominative Case assigned to the dislocated element in (5) is not related to the agreement node and is rather a default dislocation Case. S-internally, the sentence in (5) acts precisely as expected: in
the absence of an overt subject in the clause, agreement is overt. The placement of an overt subject in this clause alongside the overt agreement, e.g. a subject pronoun, would, as expected, lead to ungrammaticality.

A more complex issue is presented by the breakdown of complementarity in the cases of real extraction. Consider again the structure in (11). We assumed that when AGR is [e], V moves into the position immediately dominated by the adjoined COMP, so as to properly govern the empty N node. The failure of the V to move results in ungrammaticality. Now consider what happens if AGR is empty and V moves into COMP. The extracted subject now must move into the TOPIC node if extraction takes place, as in (12):

(12)

```
*TOPIC'  S'  
TOPIC   COMP   S   COMP   NP   V (NP)
NP_k  V  COMP  NP  VP  INFL [elk
```

Since S' is taken to be a maximal projection, NP_k cannot properly govern its trace in the [NP,S] position. Furthermore, as INFL is not a proper governor, the empty category in the subject position is not properly governed, and thus a violation of the ECP results.

What happens if the NP_k (instead of V) moves into the adjoined COMP, where it can properly govern its trace? In such a situation, [e] in INFL is not properly governed, resulting in ungrammaticality. (See Borer and Tuller [in preparation] for why NP_k cannot be a proper governor here.)

(13)

```
*S'  
COMP   S   COMP   NP  V (NP)
NP_k  INFL [e]_k  V  (NP)
```

We thus expect any derivation which involves an empty INFL to result in ungrammaticality. The only grammatical derivation of sentences such as (6), (8b) and (9c) involves the following: agreement is generated as an overt element and the extracted subject moves into the adjoined COMP position, from which it may properly govern its trace. The derivation does not involve any V movement (crucially, however, in the phonological component INFL will lower and attach to the verb, in accordance with the analysis sketched in section 1 above). It follows from our analysis that nominative Case may not be assigned in the clause itself. We have, however, established the fact that nominative Case is the default Case assigned in Arabic in dislocated positions. As no Case could be assigned to the extracted subject in the location of its origin, a default nominative Case will be assigned to it following the extraction.

We are now in a position to predict certain asymmetries between subject and object extraction. Note that both V movement and extraction from the [NP,S] position are not possible due to the fact that the subject trace must be
properly governed by its antecedent. Such is not the case for the trace of the
direct object. Thus, while V movement is still necessary in order to properly
govern the INFL node if the latter is [e], such movement will not block the
topicalization of the direct object:

(14)

\[
\begin{array}{c}
S' \\
\text{TOPIC} \\
NP_k \\
\text{COMP} \\
S' \\
\text{COMP} \\
NP \\
V \\
\text{INFL} \\
\end{array}
\]

It thus appears that the structure motivated for the complementizer system in
SA by independent factors accounts naturally for the apparent breakdown of nomi-
avative/agreement complementarity. Upon a closer scrutiny, it turns out that
in fact no such breakdown occurred and that the "redundant" nominative Case has
an independent source which is not related to the AGR node. Once this is estab-
lished, and the conditions under which such independent nominative Case is as-
signed are clarified, it is clear that S-internally, the complementary distri-
bution between agreement and nominative is preserved both in Left Dislocation
structures and under extraction.

REFERENCES

3e cycle, Paris VII [published in *Etudes arabes: Théorie et analyses.*
Département d'arabe, Paris VII.]

Borer, H. and L. Tuller. in preparation. "Nominative/Agreement complementari-
ty and VSO languages." ms., UC Irvine and UCLA.


Research* 1:33-54.


French causatives and conditions on representation." *Linguistic Inquiry*

Sproat, R. 1982. "VSO languages and Welsh configurationality." ms., MIT.

1. Introduction

The Gambia, located in the Sahel of West Africa, is representative of language policy in many African nations in two fundamental ways: (1) it is linguistically complex, and no linguistic group comprises a majority; and (2) it continues to use the former colonial language as its recognized official language. Although the current educational language policy specifies English as the medium of instruction, this paper discusses that in practice the vernacular, especially Mandinka, is used considerably.

Gambian linguistic groups consist of Mandinka (42%), Pulaar (18%), and Wolof (15%), in addition to a variety of others (Jola, Serahulle, Serrer, Aku, Manjago to name a few). Language is in most cases a marker of ethnic identity. Virtually no one speaks English as a first language. Language contact is high, and multilingualism is the norm. Of the 231 primary school teachers surveyed, teachers on the average spoke three languages including English, and three-fourths spoke Mandinka. Excluding the capital city, Mandinka serves as a lingua franca. As the Gambia is 98% Muslim, Arabic used in a classical form functions as a religious language.

Approximately 60% of Gambian children enter primary school, and 40% complete it. Around 15% finish secondary school. The current stated official medium of instruction in schools is English, although the Department of Education has plans to introduce Mandinka, Wolof, and Pulaar into schools. It is still uncertain whether these will be introduced as subjects or used as media of instruction. Presently vernacular reading books are being developed. It is planned that the local languages will be used in some form or other through the fourth grade.

Teacher questionnaires were distributed and classroom observation undertaken in eleven schools in Mandinka-speaking areas of the Gambia. Approximately 95% of these children speak Mandinka, 70% speak it as their mother tongue.

2. Teacher Questionnaires

Teachers were asked to specify the language(s) they used in certain situations or domains. Table 1 indicates the language usage reported in various domains. For example, when speaking with a spouse at home, 38% said they used Mandinka, 35% used Wolof, and 27% used other vernaculars like Pulaar or Serrer. The non-school usage reveals the diglossic situation between the use of the official language at the bank (English) and vernacular usage elsewhere in the less formal domains of home or when shopping. Diglossia is also apparent in the School Non-classroom domain in which teachers use almost 100% vernacular with parents of children and use predominately English with headteachers or other teachers. Parents in most cases do not speak English. All teachers (excluding Koranic teachers) speak English. School classroom usage reveals that teachers
reported using either English or English plus the vernacular. When comparing
the three subjects math, science, and social studies, one notices that the use
of English is less in science and social studies, concomitant with an increase
in the vernacular in these subjects. For social studies approximately 10% of
the teachers answered they use Mandinka only. Note that in giving praise a
high percentage, around 80%, said they use English only, and the remainder re-
ported using English and the vernacular.

Teachers were asked to estimate the percent of the time they spend teaching
in English and the vernacular(s) in one entire teaching day. Presently in
classrooms the vernacular is used only orally and never in a written form.
Teachers had no inhibitions about reporting vernacular usage in spite of an
English-as-a-medium policy.

Table 2 reveals that vernacular estimates range from 43% in first grade and
gradually decrease to 13% in sixth grade. Of the vernaculars, the use of Man-
dinka is the most prevalent. When asked why they used the vernacular, teachers
most often selected two reasons: (1) the child understands the subject matter
better when taught in the vernacular, and (2) it makes learning English easier.

3. Observations

One week of observation at each school was completed in first and second
grades. Every 15 or 20 seconds during teaching time, the observer noted (a)
who was speaking, (b) the language code used, and (c) what was said. These em-
pirical data revealed that essentially four languages were used in classroom
lessons: 60% English, 28% Mandinka, 10% Arabic, and 2% Wolof. Gambian chil-
dren receive 2-3 hours per week of Koranic instruction, which essentially con-
sists of recitation of Koranic prayers or verses and passages from the Koran in
Arabic. A special Koranic teacher teaches these lessons, not the regular class-
room teacher. Table 3 excludes Arabic from the analysis and compares teachers' 
estimates of usage with usage actually observed. Teachers were reasonable ac-
curate in their estimates of their code usage.

The weekly timetable allows for a variety of subjects in varying propor-
tions throughout the week. Table 4 indicates that language usage varies signif-
ificantly with the subject of the lesson. Observed vernacular usage was near 15%
for Language, Reading, and Rhymes and increased to 60–75% for subjects such as
Social Studies, Story Telling, and Health. All of these subjects included the
use of the vernacular in varying proportions plus English. In contrast, the
Koranic lesson (which uses Arabic in a classical form) used Mandinka as its me-
dium of instruction. English was not used at all in Koranic lessons, as these
instructors were not usually exposed to Westernized education, nor were they
competent in English.

In these schools pedagogy takes the form of recitation where the teacher
says a word or phrase, and the child or entire class repeats it. This is due
to a combination of factors such as a dearth of teaching materials, the large
class size, and large numbers of teachers with little formal teacher training,
all of these compounded by a medium of instruction foreign to the students.
The recitation form of teaching is particularly noticeable in Language, Reading,
Rhymes, Math, and Science where English terms are presented, recited, and prac-
ticed orally as a group. Reading and Rhymes as subjects usually consist of re-
citing sentences, paragraphs, or English poems. Math and science at these
first and second grade levels also consist of essentially presenting vocabulary. Social Studies within the Gambian curriculum is culturally based and cannot rely entirely on English as a medium. Mandinka terms and concepts are prevalent in these lessons. For example, teachers may discuss the role of the "Seyfo" (chief) or the "Alkalo" (religious leader) in the community.

Comparison of the types of speech, e.g. command, question, relative to the language used reveals which languages are used in what types of situation. Table 5 reveals that the use of Arabic is concentrated mainly in the recitation, reading aloud category. These usages require little understanding, but rely on rote memory. Arabic here is used in a classical sense and not as a form of interchange and communication. Although English is used in a broad spectrum of ways, it too included recitation, oral reading, and oral spelling. In contrast, Mandinka and Wolof are not used in these forms. One notices also that questions not related to the subject, commands, and personal comments occur slightly more in the vernacular than in English. Direct translation and series explanation occur most frequently in the vernacular, indicating that teachers often present material initially in English and then explain again in the vernacular. Praise seems to be given more often in English, reprimands in the vernacular. This dichotomy of praise and reprimand perhaps reflects and reinforces the status of English and serves as an inducement for the child to learn English.

4. Conclusion

A description of the language usage reported and observed in Gambian classrooms has been presented. In spite of an official stated policy of English in the classroom, the vernacular, Mandinka in particular, is used considerably in first and second grade, especially for culturally-based subjects (Story Telling, Social Studies, Health), and it is the only medium of instruction for Koranic lessons (Arabic). Vernacular usage is also reported up to sixth grade, one might presume with the same subject allocation as apparent here. English usage is high in subjects where English vocabulary presentation and practice are the focus of the lesson. Comparison of reported usage and actual observed usage indicate teachers are aware of the ways they allocate language, whether it be usage among codes, among subjects, or among domains.

As the language policy is changed to include accepted and expanded usage of the vernacular(s), the functions of the vernacular vis-à-vis English will expand. It is possible that this language change may bring with it changes in teaching methods less reliant on English vocabulary presentation, recitation, and rote memory, and more involved in interactive communication.
Studies in African Linguistics, Supplement 9

TABLE 1
LANGUAGE USAGE IN VARIOUS DOMAINS
REPORTED BY TEACHERS

WITH
SPouse
M
W
Other
Vern

SHOpING
M
W
E

AT
BANK
E

WITH
PARENTS
OF
CHILDREN
M
W
E

WITH
HEADTEACHER
E

WITH
OTHER
TEACHERS
E

MATH
E

SCIENCE
E

SOCIAL STUDIES
E

GIVING
PRAISE
E

M=Mandinka
W=Wolof
E=English
V=Vernacular (possibly including M & W)

TABLE 2
TEACHER ESTIMATES OF VERNACULAR USAGE
BY LEVEL

75
50
25

Mandinka Wolof Pulaar

25 36 24 24 19 13
1 2 3 4 5 6
African Language Usage in the Classroom

TABLE 3

VERNACULAR USAGE
REPORTED AND OBSERVED
(level 1 & 2, all subjects combined*)

ENGLISH  MANDinka  WOLOF  PULAAR

*Excludes Koranic lessons (consisting of both Arabic usage and Mandinka usage). Koranic teachers did not complete questionnaires.

TABLE 4

OBSERVED VERNACULAR USAGE BY SUBJECT
### TABLE 5

**OBSERVED CLASSROOM LANGUAGE USAGE**

<table>
<thead>
<tr>
<th>TYPE OF SPEECH</th>
<th>CODE USED (Percent)</th>
<th>ENGLISH</th>
<th>MANDINKA</th>
<th>WOLOF</th>
<th>ARABIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Subject matter</td>
<td>34</td>
<td>35</td>
<td>35</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2. Question re:subject</td>
<td>18</td>
<td>16</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Command</td>
<td>16</td>
<td>21</td>
<td>18</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4. Prayer, recitation</td>
<td>4</td>
<td>-</td>
<td>1</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>5. Reading aloud</td>
<td>12</td>
<td>-</td>
<td>-</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>6. Spelling/alphabet</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>7. Praise</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>8. Reprimand</td>
<td>-</td>
<td>2</td>
<td>2</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>9. Singing</td>
<td>4</td>
<td>3</td>
<td>20</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>10. Comment</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>11. Vernacular prompt</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>12. Direct translation</td>
<td>-</td>
<td>3</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>13. Telling story</td>
<td>-</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>14. Series explanation (teacher</td>
<td>-</td>
<td>6</td>
<td>5</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>explains what was previously</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>said)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of instances 7777 3574 281 1244

*based on 62 hours of actual speech*
WH-MOVEMENT IN YORUBA

Vicki Carstens
UCLA

Yoruba has a focus construction which enables any sentence constituent to appear sentence-initially. Basic word order is SVO:

(1) Ajike ñ ta isu ninu oja 'Ajike is selling yams in the market'
    ASP sell yams in market

The presence of the particle ni following a constituent in sentence-initial position indicates that it is focused:

(2) isu ni Ajike ñ ta 'it's yams that Ajike is selling'
yams FOCUS ASP sell

Wh-questions are formed by obligatory usage of the focus format:

(3) ki ni Ajike ñ ta ninu oja 'what is Ajike selling in the market?'
    what FOCUS ASP sell in market

The structure of (3) will be considered to be as illustrated in (4) below (structure subsequently will be indicated by bracketing).

(4)

```
S
 |   |
S   S
 |   |
NP INFL VP
 |   |
\  /  \\
ki ni Ajike ñ ta [e] ninu oja
```

When a subject is focused or questioned, a third-person singular resumptive pronoun is obligatory:

(5) a. S[Ajike ni S[O ñ ta isu]] 'Ajike is selling yams'
    FOCUS 3s ASP sell yams

b. S[ta ni S[O ñ ta isu]] 'who is selling yams?'
    who FOCUS 3s ASP sell yams

If the locative phrase ninu oja is focused, the particle ti must be introduced between the subject and the verb:

(6) S[ninu oja ni S[Ajike ti ñ ta isu ti]]
    ninu oja FOCUS in market FOCUS 'it's in the market that Ajike is selling yams'
    S[Ajike ti ñ ta isu ti]
    FOCUS 3s ASP sell yams

\( ti \) is obligatory in 'where' questions:

(7) S[nibo, ni S[Ajike ti ñ ta isu ti]] 'where is Ajike selling yams?'
    where FOCUS 'where' FOCUS 3s ASP sell yams
In these sentences ti is semantically empty, and although its presence is clearly motivated by the presence of a locative in COMP, it does not occupy the DS position of that phrase.

Ti also occurs as a "preverb" in the construction exemplified by (8):

(8) Tolu ti oko bo 'Tolu is coming from the farm'
    ASP from farm come

Note the position of the homorganic nasal continuous aspect marker with respect to ti here: their order is the reverse of that in (6). If the object oko is extracted, the order found in (6) is again obligatory:

(9) oko ni Tolu ti m bo 'it's the farm that Tolu is coming from'
    farm FOCUS ASP come

No Yoruba verbs or preverbs other than the three to be discussed in this paper can precede the continuous aspect marker in this way. In cases such as (6), (7), and (9), where the presence of ti is motivated by the presence of a locative phrase in COMP, and it does not inflect or occupy the position in which verbs usually occur, I conclude that it is performing a distinct role and occupying a distinct syntactic position. I will assume the structure of (7) to be roughly the following:

(10)

\[
S \rightarrow \begin{array}{c}
\text{COMP} \\
\text{NP} \\
\text{INFL} \\
\text{VP}
\end{array}
\]

\[
\begin{array}{c}
nibo \_ ni \\
\text{Ajike} \\
ti \_ ti \_ isu [e] \\
\end{array}
\]

Ti is also found in 'how' questions:

(11) S[bawo ni S[Duple ti k' baba re ti]]
    \quad \text{FOCUS} \quad S[\text{greet father 3s poss.}]
    \quad \text{greet her father?'}

Alternatively ŝe can appear on its own or immediately following ti, but one of these options must be exercised or the result is ill-formed:

(12) S[bawo ni S[3s \{ti ŝe \} pa ekun naa? ti]]
    \quad \text{FOCUS} \quad 3s \{\text{kill leopard this} \}
    \quad \text{'how did s/he kill this leopard?'}

(13) *bawo ni o pa ekun naa?

As before, continuous aspect if it occurs is preceded by these elements. It is worth noting that ŝe also serves as a main verb, meaning 'do'.

Embedded 'how' questions, manner relatives, and comparatives are all introduced by the particle bi and must include either ti or ŝe:
(14) mo fé mo [bi i S [o șe pa ekun náà t₁]]
1s want know how 2s kill leopard this
'I want to know how you killed this leopard'

(15) [bi i S [Tolu ti ri t₁]] ya mi lénu'
how S look open me mouth
'the way Tolu looked surprised me'

(16) mo ga [bi i S [Ajike ti ga t₁]]
1s tall how tall
'I am as tall as Ajike is'

Finally in 'why' questions and relative clauses headed by idi ( rè) 'the reason', a particle fi occupies the same syntactic slot:

(17) -nitorikii ni S [Tolu fi ò lọ s'ọja t₁]]
why FOC S ASP go to market
'why is Tolu going to the market?'

(18) NP [idi ti S [ò fi ò lọ t₁]] ni látì ra ìsu reason REL S 3s ASP go be for buy yams
'the reason he's going is to buy yams'

fi has many other uses, of which the most common is as a preverb having an instrumental interpretation:

(19) o ò fi ibon pa ekun
3s ASP gun kill leopard
's/he is killing the leopard with a gun'

When its object is extracted, it usually inverts with the continuous aspect marker:

(20) ibon ni ò fi òpa ekun
gun FOCUS
'it's a gun that she's killing the
leopard with'

In a multi-clause question, the presence of the appropriate verb exclusively in an embedded clause indicates that clause (rather than a higher one) to be the location of the variable bound by the WH-phrase in COMP. In (21) below, nibo can only bind a variable in the embedded clause, i.e. (21) must have the S-Structure (22):

(21) nibo ni Ajike so fun e pe òun ti rì Tolu?
where FOCUS tell you that 3s see
'where did Ajike tell you that she saw Tolu?'
The presence of $ti$ in $S_2$ results in there being just one available interpretation for (21), i.e.

(23) For which place $x$, Ajike told you [she saw Tolu at $x$]

But if $ti$ occurs only in the matrix clause, the question is ambiguous. The question in (24) can have either of the interpretations (a) or (b):

(24) nibo $ni$ Ajike $ti$ so fun e pe oun ri Tolu?

where FOCUS tell you that $3s$ see

where did Ajike tell you that she saw Tolu?

a. for which place $x$, [Ajike told you at $x$] that she saw Tolu
b. for which place $x$, Ajike told you [she saw Tolu at $x$]

This indicates that two $S$-structures may correspond to question (24): that in (22) above and the one below:

(25)

A further complication exists. In addition to the two positions of $ti$ in (21) and (24), it is also possible for $ti$ to appear in each clause. The question in (26) below can have either of the two interpretations available for (24) and additionally for some speakers can be a question involving two places,
for which a third structure must be entailed, in which *nibo* is indexed to two gaps:

(26) *nibo* ni Ajike *ti* so fun e pe oun *ti* ri Tolu  
where tell you 3s see  

a. for which place x, [Ajike told you at x] that she saw Tolu  
b. for which place x, Ajike told you [she saw Tolu at x]  
c. for which places x and y, [Ajike told you at x] that [she saw Tolu at y]  

(A fourth interpretation seems to exist as well, wherein for which x has matrix and for which y embedded scope, but judgements so far have not been entirely clear). The (a) interpretation is unexpected, as it indicates that a *ti* may occur vacuously beyond the logical source of *nibo*.

Although two-story 'where' questions have been used in these examples for ease of exposition, the same facts hold of 'why' and 'how' questions and of three and four clause structures.

These structures were found to obey Subjacency, so I assume a wh-movement derivation for all of them. It is therefore tempting to relate the presence of *ti*, *fi*, and *se* to the ECP. Under standard assumptions, movement would proceed successive-cyclically from COMP to COMP leaving traces, which *ti*, *fi*, and *se* could be thought to sanction in compliance with the Empty Category Principle given in (27a), where proper government is as stated in (27b) (cf. Chomsky [1981], Aoun & Sportiche [1981]):

(27) a. ECP: NP[e] must be properly governed  
    b. A properly governs B if A governs B and  
       (i) A is a lexical category X^0 (lexical government), or  
       (ii) A is coindexed with B (antecedent government)

Minor alteration of the structure I proposed in (10) would allow *ti* to lexically govern a trace dominated by VP. Thus in a single clause sentence such as (7), *ti* could be assumed to properly govern Wh-trace at the source. Likewise in a multi-clause question such as (21), where *ti* is contained in an embedded clause as illustrated below, *ti* could be argued to properly govern NP[e] at the site of extraction:

(28) $S_1[nibo_1 S_1[ti] S_2[t_i]]$  

Consider however the configuration below, representing the structure of (24), reading (b):  

(29) $S_1[nibo_1 S_1[ti] S_2[t_i]]$  

Here *ti* is too remote to lexically govern the extraction site. It might be thought to properly govern the trace in the intervening COMP of $S_2$, which would then antecedent-govern the source, but since *nibo* cannot antecedent-govern the source in a single clause question such as (7), it would be surprising if an intermediate trace could do so. We can see from (24b)/(29) that *ti* is not needed to lexically govern the source, and from (21)/(28) that it is not
required to licence intermediate traces. In short although ti is obligatory, owing to the unpredictability of the clause in which it will appear, either of the two functions which this approach can ascribe to it are demonstrated to be optional, so our understanding of its role is not furthered.

An alternative possibility exists. At the beginning of this paper, I noted that a subject cannot be focused or questioned unless a resumptive pronoun occupies subject position. This suggests that Wh-trace cannot be properly governed from COMP at all in Yoruba. We have seen that a verb's object can be extracted freely, which indicates that verbs can be proper governors. Let us assume that proper government has to be lexical in Yoruba and briefly examine [-N] lexical governors.

All objects in Yoruba must be adjacent to a verb, preposition or preverb. The sentence in (30) illustrates the use of the semantically empty preposition n'i to govern and case-mark the NP onje 'food':

(30) Bola fun m'i n'i onje 'Bola gave me food'
give me PREP food

(*Bola fun m'i onje)

Prepositions cannot be stranded, so are not proper governors in Yoruba. They are further limited in that they cannot take pronominal objects. If an object not governed by a verb is pronominalized, the preverb fi must be introduced for it to cliticize to (note that here continuous aspect would precede fi, as in (19)):

(31) o fi i fun m'i 'she gave it to me'
3s give 1s

This use of fi suggests a similar treatment of ti, fi, and se. If we assume that only verbs and preverbs are proper governors in Yoruba, we can explain the introduction of ti, fi, and se in a way that parallels the insertion of fi above, as a means of overcoming a syntactic restriction, here on wh-movement. Both prepositional phrases and adjuncts lack lexical governors. If antecedent government is not possible in Yoruba, as the evidence from subject focusing suggests, then perhaps before an adjunct or PP can move to matrix COMP it must find a lexical proper governor for its trace. If we assume that movement proceeds through INFL, then perhaps the presence of a trace there triggers insertion of the appropriate verb, in the same way that pronominalization of onje causes fi to be inserted in (31).

The (c) reading of (26) can be derived as mentioned above from a structure including two nibo's, one in each clause, such that at S-structure nibo is indexed to two gaps. Perhaps in cases such as (26b), where ti is included in every clause, the moved phrase has passed through the INFL of each S on each successive-cyclic move. The (a) reading of (26) still fails to be captured, however, and as yet remains a mystery.

REFERENCES


THREE CASES OF DOWNSTEP IN VENDA

Farida Cassimjee and Charles W. Kisseberth
University of Illinois

According to Westphal [1962], Venda, a Bantu language spoken in the Zoutpansberg district of the northern Transvaal in South Africa, displays the phenomenon of downdrift. Each occurrence of a sequence of low-toned vowels between two sequences of high-toned vowels causes the second sequence of high-toned vowels to be realized at a somewhat lower pitch level than the first sequence (where "vowel" = any tone bearing unit, including syllabic nasal, and "sequence" = one or more such units). In addition to downdrift, Venda also displays downstep; that is, there are instances where what is demonstrably a sequence of high-toned vowels is immediately preceded by another sequence of high-toned vowels, yet the second sequence is pronounced at a somewhat lower pitch level than the first. In this paper we examine three instances of downstep in Venda and argue that all of them can plausibly be regarded simply as instances of downdrift but where the low tone that induces the downdrift is a tone that is present in the (autosegmental) tonal tier but unassociated with any vowel.

Before we examine these three cases of downstep, it is necessary to review very briefly the analysis of Venda tonology developed in Cassimjee [1983] where the following tone rules are postulated:

- LOW-DELETION: A Low tone at the beginning of a word is deleted when the preceding word ends in a High tone.
- MEEUSSEN'S RULE: A High tone is changed to a Low tone when immediately preceded by another High.
- HIGH TONE SPREAD: Add an association line between a High and an immediately following vowel (provided that to do so does not result in the crossing of association lines). This rule is applied iteratively from left-to-right.
- CONTOUR SIMPLIFICATION: (a) If a HL sequence is associated with the same pre-penult vowel, disassociate the Low; (b) if a HL sequence is associated with the same ultimate vowel, disassociate the High.

All of these rules are phrase-level rules and they apply in the order listed.

Cassimjee assumes that there are just two underlying tones in Venda, High and Low, and that (in the level of representation that is input to the above rules) each vowel is associated with at most a single tone, i.e. it is not necessary to assume there are lexical contour tones. Furthermore, she assumes that within a morpheme, successive high-toned vowels are all linked to a single High on the tonal tier whereas successive low-toned vowels are linked to separate Lows on the tonal tier. Given these assumptions and the rules mentioned, an extremely complex pattern of alternation in Venda nominals can be accounted for in an elegant fashion.

The first case of downstep: Downstep arises as a consequence of the application of the rule of Contour Simplification (which disassociates the Low in a HL sequence associated with a pre-penult vowel). Consider, for example, the alternation between mu-lambóni 'at the river' (isolation form) and mú-lambilóni
(form used after a word that ends in a High). The symbol "!" marks the fact that downstep occurs between the two sequences of high-toned vowels separated by "!". The post-High form is produced by the following derivation:

1. H #\# L L H L
   \mu-lamboni (input)
   H #\# L H L Low-Deletion
   inapplicable Meeussen's Rule
   H #\# L H L High Tone Spread
   \mu-lamboni H #\# L H L Contour Simplification

Notice that when Contour Simplification disassociates the L from the pre-penult vowel, there is a "floating" Low between the high-toned pre-penult vowels and the high-toned penult vowel. And there is also a downstep at this point. We attribute this downstep to the floating Low. In other words, a floating Low will trigger downdrift just as an associated Low would. All that is required is that this floating Low actually be in the tonal tier at the point where the rule of Downdrift applies.

The preceding example involved a case where Contour Simplification affects a contour tone created by the phrasal application of High Tone Spread. There are also cases where Contour Simplification affects a contour tone resulting from the word-internal application of High Tone Spreading. For example, if we compare the first and third person present tense verbal forms ndi-a-vhona 'I see' and u-a-vhona 'he sees', an analysis immediately suggests itself where ndi- and -a- are Low-toned elements, and u- and -vhona are High-toned elements. We would then obtain the following derivations:

2. L L H H L L H
   ndi-a-vhona u-a-vhona (input)
   inapplicable inapplicable Low-Deletion
   inapplicable inapplicable Meeussen's Rule
   H L H L H u-a-vhona
   inapplicable inapplicable High Tone Spread
   inapplicable u-a-vhona Contour Simplification

In the derivation of u-a-vhona we see that the application of Contour Simplification to the output of the word-internal application of High Tone Spread creates a floating Low tone and that there is, in fact, downstep at precisely the point where this floating Low occurs.
Three Cases of Downstep in Venda

The second case of downstep: Another source of downstep in Venda is a floating Low tone that arises in the morphology (as opposed to the operation of a phonological rule, as in the preceding case). The present tense forms cited above are valid only in cases where the verb word is final in the clause. If another word follows, then the present tense forms lack the /a/ prefix. The "short" forms for first and third person singular present tense are shown in (3):

(3) ndi-vhōná... ū-vhōná...

The proper analysis of these short forms seems to involve saying that whereas the long present tense form has an /a/ prefix and a Low tone that is associated with that prefix, the short form has no /a/ prefix but does have a Low tone located at the point between the subject prefix (ndi- or ū- in the examples under discussion) and the verb stem. The suggested underlying forms, then, for (3) are as in (4):

(4) L L H \_ H L H \_ 

\_ ndi-vhōna... \_ ū-vhōna...

It is this floating Low tone in the short present form that causes the downstep between a High-toned subject prefix like ū- and a High-toned verb stem like -vhōnā.

The third case of downstep: In addition to floating Low tones created by Contour Simplification and in the morphology, we suggest that there are also floating Low tones created by a rule of Tone Metathesis. Consider the following present tense examples:

(5) ndi-a-lįndėla 'I am waiting'

ū-a-lįndėla 'he is waiting'

From the first person singular form it is clear that -lįndėla is a Low-toned verb stem. The problematic form is the third person singular. Why is the High tone on the subject prefix able to spread all the way to the second vowel of the verb stem? Why is there no trace of the Low tone associated with the /a/ prefix? An answer to these questions emerges when one considers the pronunciation of the items in (5) when they follow a word that ends in a High.

(6) ndi-a-lįndėla

ū-a-lįndėla

The data in (6) strongly suggest that there is a floating Low tone in front of the third person form since (a) there is a downstep between the final High of the preceding word and the High of the subject prefix on the verb, and (b) the High of the preceding word has failed to trigger Meeussen's Rule (which ordinarily would change an initial High in a verb to Low after a word that ends in a High).

If we assume that there is, then, a floating Low in front of the third person form, where could it have come from? If we were to say that it originates, via metathesis, from the Low tone associated with the /a/ prefix, then we
would explain both why there is a downstep in front of the verb word and also why the High tone of the subject prefix is able to spread past the /a/ prefix all the way into the verb stem.

The Tone Metathesis rule can be formulated roughly as follows:

\[(7) \text{H L} \quad \text{Verb Stem} \quad \Rightarrow 2 1 \emptyset 3\]

Tone Metathesis is triggered only by Low verb stems, and it applies only when there is a HL sequence in front of the stem. We illustrate the effects of this rule below. (It will be noted that we have assumed the existence of a special rule that deletes the initial Low of a verb stem after a High-toned prefix. This rule has not been fully justified here, but it is important in determining how far into the verb stem a High-toned prefix is able to spread.)

\[(8) \text{H H H} \text{H L L L L} \quad \text{u-a-linda} \quad \text{(input)}\]
\[\text{inapplicable} \quad \text{Low-Deletion}\]
\[\text{H H H} \text{H L L L} \quad \text{u-a-linda} \quad \text{Tone Metathesis}\]
\[\text{inapplicable} \quad \text{Meeussen's Rule}\]
\[\text{H H H} \text{H L L} \quad \text{(initial-Low deletion from verb stem)}\]
\[\text{H H H} \text{H L L} \quad \text{u-a-linda} \quad \text{High Tone Spread}\]
\[\text{inapplicable} \quad \text{Contour Simplification}\]

We have now examined three cases of downstep in Venda and have shown that they all can be accounted for in terms of the principle of downdrift operative in the language as long as we assume the existence of floating Low tones in the tonal tier at the point where Downdrift applies. Furthermore, we have shown that these three cases of floating Low tones have separate origins. In one case the floating Low originates from Contour Simplification, in another case from the morphology, and in a third case from the application of a rule of Tone Metathesis.

REFERENCES


1. Introduction

Vowel shortening processes sensitive to syllabic structure have been treated as specific rules for specific languages in a linear phonological framework. The fact that in different languages the contexts of application of the rules are identical suggests that vowel shortening has to be derived from a universal principle and not by language specific rules which need to be learned independently. Following this line, the difference in the context of application will be explained from the differences in the syllabic structure of languages in question. To account for vowel shortening processes sensitive to syllabic structure, Lowenstamm & Kaye [1982] have proposed a universal principle, the Prosodic Government Principle (henceforth PGP). The principle seeks to explain why we cannot find long vowels in closed syllables in languages where vowel shortening occurs.

In this paper I will discuss an apparent violation of the PGP in Wolof, a West Atlantic language spoken in Senegal. I will show that the violation is indeed apparent because of an independently motivated theory, that of the "appendix". The intersection of these two theories will give precisely the desired results.

2. Prosodic Government Principle

"One could for instance speculate that the Nucleus is the 'head' of the rime constituent. We could then go on to say that this element must in some sense govern its sister constituents of the rime. This relationship would be defined, at least in part, configurationally and could explain why long vowels and [heavy] diphthongs are not found in closed syllables in languages for which such a constraint holds." [Lowenstamm & Kaye 1982: 27]

PGP excludes configurations with a branching nucleus in a branching rime.

(1)

\[ \begin{array}{c}
\ast & \ast \\
O & O \\
R & R \\
N & N \\
C & C \\
a & a \\
b & b \\
c & c \\
x & x \\
x & x \\
(\text{long vowel}) & (\text{heavy diphthong})
\end{array} \]

In such configurations the head of the rime (the leftmost point under the nucleus) does not govern all its sister constituents in its prosodic domain, the rime, since the first branching constituent dominating it is the nucleus.
3. **Yawelmani**

If we consider Yawelmani [Kurada 1967] we see that vowel shortening occurs in all closed syllables, which means that PGP is systematically respected.

\[
\begin{array}{ccc}
\text{stem} & \text{aorist} & \text{passive aorist} \\
\text{gloss}
\end{array}
\]

\[
\begin{array}{cccc}
'\text{ili}' & '\text{ili}+\text{hin}' & '\text{ili}+\text{t}' & '\text{ilet}' \\
& ['\text{ileehin}'] & & '\text{to fan}' \\
'/\text{pana}' & '/\text{pana}+\text{hin}' & '/\text{pana}+\text{t}' & '/\text{panat}' \\
& ['\text{panahin}'] & & '\text{to arrive}' \\
'/\text{aam}' & '/\text{aam}+\text{hin}' & '/\text{aam}+\text{t}' & '/\text{aamit}' \\
& ['\text{aamihin}'] & & '\text{to help}'
\end{array}
\]

The PGP explains why we cannot find long vowels in closed syllables in a language like Yawelmani. By the PGP the branching nucleus of a syllable like -\text{ili} must lose a point of the skeleton when the coda of the rime is occupied: (-\text{ili}+\text{t} + -\text{lit}). If the nucleus does not lose a point of the skeleton, the head of the rime (the leftmost point under the nucleus) will not govern all of its sister constituent in the rime.

\[
\begin{array}{ccc}
\text{imperfective} & \text{inversive} & \text{gloss}
\end{array}
\]

\[
\begin{array}{cccc}
[\text{rOof}] & /\text{rOop+i} & /\text{rOppi} & '\text{to put in/to take out}'
\end{array}
\]

3. **Wolof**

If we now consider Wolof [Diagne 1971; Dialo 1983; Sauvageot 1965], a language in which a vowel shortening process sensitive to syllabic structure is also present, we must explain why PGP is always respected in non-final syllables while it is apparently systematically violated in word-final syllables. Such cases are given in (4) below.
The Appendix in Parametric Phonology

[yEEw] /iEEu+i/ [yEwwi] 'to tie/to untie'
[tEEr] /tEEd+i/ [tEddi] 'to start/to stop a vehicle'

The inversive forms show us that word internally a long vowel shortens when the coda is filled.

(5) /rOOp+i/ [rOpp]

We see in (5) that when the inversive suffix is added to a stem ending in a consonant, this segment is syllabified under the rime which is creating a closed syllable. Since the rime is now branching, the nucleus must lose a skeletal point to prevent violation of the PGP. The skeletal point dominating the consonant could not be syllabified in the following onset since this constituent already dominates a skeletal point. The only possible syllabification is then for the point to occupy the coda position which is not filled. The gemination of the consonant is explained by the fact that the onset of the inversive syllable dominates a skeletal point which has to be associated to a segment.

To explain the fact that in Wolof the PGP is systematically violated in word final syllables, I would like to claim that a word-final syllable containing a long vowel followed by a single consonant is not a closed syllable as seems to be the case, but is rather open. Following this I will suggest that Wolof has an appendix position (a non-rimal constituent) which is, like an extrametrical position, possible at the end of a word or a cycle. This suggests that the syllabic structure of a word like [rOOf] is a syllable composed of a non-branching rime followed by an appendix constituent to which the last consonant is associated. Since the rime is non-branching, there is no reason for the long vowel to be shortened by the PGP.

(6) [rOOf]

Having shown that a word-final syllable containing a long vowel followed by a single consonant is to be analyzed as having a branching nucleus and an appendix, the question is now whether the appendix position is justified elsewhere in the language. The answer is provided by words ending in more than one consonant. We find in Wolof words ending in geminate consonants which are never pre-
ceded by a long vowel. The analysis for such words is one with a branching rime followed by an appendix. Since the coda never branches (because we never find a cluster of more than two consonants word internally and because when a suffix beginning with a consonant is added to a stem ending by a geminate, an epenthetic vowel appears) the appendix position is justified, the two skeletal points dominating the geminate needed to be associated to a syllabic constituent.

(7) bOpp 'head'

![Diagram](image)

It is clear that a final consonant following a long vowel must be analyzed as being in the appendix (because of the PGP) and that the same is true for a word ending in a geminate. But this leaves open the question of the analysis of a final consonant preceded by a short vowel. Is the consonant in the coda or in the appendix? That is, is the appropriate structure (8a) or (8b)?

(8) nak 'cow'

<table>
<thead>
<tr>
<th></th>
<th>a.</th>
<th>b.</th>
</tr>
</thead>
<tbody>
<tr>
<td>σ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>O</td>
<td></td>
<td></td>
</tr>
<tr>
<td>R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>b</td>
<td>O</td>
<td>p</td>
</tr>
<tr>
<td>n</td>
<td>a</td>
<td>g</td>
</tr>
<tr>
<td>n</td>
<td>a</td>
<td>g</td>
</tr>
</tbody>
</table>

The following factors suggest that the syllabic representation of a word containing a short vowel followed by a single consonant has a non-branching rime followed by an appendix dominating the last consonantal segment:

(a) An open syllable is less marked than a closed one.

(b) When a syllable contains a long vowel followed by a single consonant, this consonant is syllabified under the appendix and the syllable is open.

(c) No words end in more than one consonant except those ending in geminates which have the syllabic representation of a single segment associated to two skeletal points, one under the coda and one under the appendix. (The consonant cluster nasal-obstruent is analyzed as the nasal associated to the nucleus and the obstruent under the appendix. No long vowels precede such clusters.)

I consider that this analysis provides a plausible account of the acquisition of syllabic structure. Word-internally and word-final the coda is filled only by a segment associated to another constituent, i.e. only geminates occupy a coda position. There is no evidence that the last consonant of a word such
as nak is under the coda since in no other case would a consonant be uniquely associated to the coda position. Analysing the consonant under the coda would lead us to say that only in word-final position is a single segment syllabified under the coda.

To sum up, my claim is that in Wolof all syllables are open except when closed by the initial element of a geminate.

REFERENCES


In the full version of this paper, I argue for a restrictive theory of tone in what I will refer to as "true tone languages" (TTLs) as contrasted with, say, pitch accent languages. Assuming in broad outlines the theory of autosegmental phonology, I provide evidence that this theory should be restricted, when applied to tone in TTLs, in these ways:

1. Tonal autosegments are at some point in the derivation mapped onto their anchors, after which they behave like any other non-autosegmental feature.

2. Tonal autosegments may be associated with no more than one anchor.

Evidence for (1) is of three types. First, rules that are crucially autosegmental in nature, e.g. deletion of a vowel without concomitant deletion of the tone it bears, universally precede rules that are not, e.g. spreading, in the sense of Hyman and Schuh (cf. Leben [1973], Odden [1980], and Creider [1983]). Secondly, diachronic changes appear never to affect tones in a crucially autosegmental way, as I have argued in Churma [1983, to appear]. This follows from (1) together with the hypothesis that rules are added only at the end of the grammar. Third, no crucially autosegmental rule may be optional, since only possible sound changes may be optional synchronic rules [Churma 1983, to appear].

There are a number of different kinds of evidence which support (2). First, note that it is not necessary to allow violations of (2) in order to allow for rules that affect (apparent) sequences of like tones as if they were a single multiply-associated autosegment, given the existence of the Q-subscript formalism, which is independently required for an account of Kikuyu, where rules affect sequences of like tones even across word boundaries [Clements and Ford 1979].

Furthermore, there are a number of examples for which the unmarked autosegmental case, i.e. multiple association of tones, either requires unwieldy rules or simply gives incorrect results, including the following. First of all, despite the fact that only "stranded" tones would be expected to reassociate with the trigger of the rule that results in the loss of their segmental support [Clements and Ford 1979], in fact any tone whose support is lost reassociates to the trigger [Leben 1978]. Thus, in the often cited case of Lomongo, /bəlóngó bækáké 'his book' + bəlóngbækáké, despite the fact that the high tone on the second o in the input would not be stranded after vowel deletion if it was also associated (in unmarked fashion) with the first o.

Secondly, in some languages, rules that appear to be clearly autosegmental in nature (in the sense that they would be most simply expressed by a crucially autosegmental rule) make it impossible, when treated as such, to state subsequent rules. For example, in Churma [to appear, b] I argue that in Loko there is a rule which converts all tones of the second element of a compound to low (presumably stated most simply as $T_o + L$, with association of all vowels to the
L). However, a later rule must be able to convert the first two Ls (either as simple Ls on vowels of the second element or as part of a word-final falling tone on the first element) to H when they follow a word-final H. The Ls in this case behave as if they were singly associated; however, only the tone of the first syllable is affected if the first element has final falling tone, and only the first two are affected if it has a final H. If multiple association of tones is prohibited, as in (2), this would require that the former rule be revised so that all (singly-associated) tones in second elements are replaced by L, and the latter would be able to apply appropriately.

A third argument in favor of (2) comes from Odden's [1980] analysis of certain forms in Shona, which behave as if they have multiply-associated tones with respect to an early rule (in that indefinitely long sequences of Hs are lowered) but behave otherwise with respect to a later rule, where only a single H of a sequence is lowered. If (2) is adopted, then this difference would be attributed to the presence of a Q-subscript in the case of the former, but not the latter. (These facts could also be taken as support for (1), as in fact Odden argues, with the difference in behavior accounted for in terms of the different behavior of pre- vs. post-mapping rules; however, the putatively post-mapping rule does not appear to have the other kinds of characteristics exhibited by such rules.)

Finally, in Venda, according to Cassimjee [1983], all Hs are multiply associated, but all Ls are singly associated, as indicated by the fact that indefinitely long sequences of Hs are affected by a rule, whereas only the first of a sequence of Ls is affected by another rule. If we assume that there are just 100 lexical items for which the putative difference between singly- and multiply associated tones could be detected from the way in which they are affected by the relevant rules, then the probability of the above distribution occurring by chance is \((1/2)^{100}\), or one in over 1,200,000,000,000,000,000,000,000. An alternative account in which the difference in behavior derives, not from a wildly improbable chance distribution with respect to lexical representations, but from a simple difference in the form of the rules involved (Q-subscript vs. lack thereof) seems clearly preferable.

The converse of (2) is clearly not true, given the existence of contour tones. The different autosegmental tiers behave differently in other respects, as well. For example, while elimination of one of two tones associated with a single vowel is commonplace ("contour-tone simplification"), no language requires that one of two vowels (putatively) associated with a single tone be deleted. Thus, tone in TTLs is more nearly suprasegmental than autosegmental in character—indeed close to being strictly segmental when compared to other prosodic characteristics (as Bell and Hooper [1978] almost suggest).

There are a number of fairly obvious apparent problems for a theory that incorporates principles (1) and (2). Some of these are addressed in my 1983 paper, and I feel that the others will also turn out to be only apparently problematic (some are treated in work in preparation). In any event, the (apparent?) problems facing a theory that does not incorporate the equivalent of these principles are sufficiently severe that the potential explanatory role played by these principles deserves careful attention.
REFERENCES


THE DESTRUCTION OF TONAL STRUCTURE IN MENDE

Elizabeth Cowper and Keren Rice
University of Toronto

In Mende, a Mandan language of Sierra Leone, there are several word formation processes which involve moving or copying tones and introducing low tones that are assigned to all non-initial syllables of the second morpheme of the word. We will claim that in order to account for these processes in the lexical phonology framework proposed by Kiparsky [1982], the rules copying and moving tones and assigning the low tones must destroy previously assigned autosegmental structure.

1. Data

1.1. Reduplication. In Mende, reduplication involves the copying of an entire verb or adjective stem to the right of the base. A final high tone of the base is moved onto the reduplicated element and the remaining syllables of the reduplicated element receive low tone. This is shown in (1).

(1) a. mbélá 'tear, split, break' mbélambélá 'tear to pieces'
    b. hùmù 'murmur in dissatisfaction' hùmùhùmù 'murmur in dissatisfaction'
    c. nyándé 'pretty' nyándenyándé 'very pretty'
    d. fèmbè 'shake' fèmbèfèmbè 'shake repeatedly'

From the form in (1a), it might seem that the high tone is spreading rather than moving onto the first syllable of the reduplicated element. However, the forms in (1c) and (1d) show that in fact the final high is moved. In these forms, the base still carries a low tone which spreads right to associate with the remaining syllables of the base. The inserted low tone then associates with the remaining syllables to the right. In (1b), since rightward movement of the high tone leaves the base with no tone, the high tone spreads back.

1.2. Noun-noun compounding. Noun-noun compounding involves the copying of a final high tone from the first noun onto the first syllable of the second noun. The remaining syllables of the second noun receive low tone. This is illustrated in (2).

(2) a. ngúlu + mítè ngúlümítè 'wooden spoon'
    b. mèndé + hálè mèndénhálè 'Mende medicine'
    c. pùù + hálè pùúnhálè 'European medicine'
    d. mèndé + bèlè mèndébèlè 'Mende trousers'
The word formation processes of reduplication and compounding thus both involve the assignment of low tone to the second member of the word. They differ only in that reduplication involves moving final high tones while compounding involves copying final high tones. In the rest of the paper, our major concern will be how the second element receives its tones.

2. Lexical Phonology and Mende

Within the lexical phonology framework, the organization of the grammar of Mende is as follows.

While the placement of syntax within the lexicon is controversial, we argue for it in detail in Cowper and Rice [1984]. For the present paper, we simply assume the model in (3). The rules we are concerned with here are the tone rules in the lexical phonology and their interaction with morphological processes within the lexicon. Before proceeding with our analysis, we will outline some basic principles of lexical phonology relevant to our analysis.

In the lexical phonology model, there is a cycle back and forth between the phonology and morphology sides of the grammar. While most languages that have been treated in a lexical phonology framework have more than one level, it appears that there is just one level in the grammar of Mende.

The rules in the lexical phonology are sensitive to the internal structure of words; the postlexical rules apply across the board, paying no attention to internal structure. In order to capture this distinction between rule types, the Bracket Erasure Convention is used. This convention erases word internal brackets at the end of each level of the lexicon. Thus, within the Mende lexicon, the internal structure of words is transparent, but postlexically no reference can be made to this structure. The rules of the lexicon are subject to the Strict Cycle Condition; they can only apply in derived environments. In recent work, Kiparsky [1983] proposes that only structure changing rules are restricted by the Strict Cycle Condition to derived environments; structure building rules such as stress assignment and tone association are not so restricted.
and can apply to basic lexical entries.

With this overview of lexical phonology, we will now go on to look at the rule of tone association in Mende.

3. The Place of Tone Association in the Grammar

In arguing for the placement of tone association, we will assume with Leben [1978] that Mende is a language with autosegmental tone and tone spreading. We will show here that (1) tone association occurs in the lexical rather than in the postlexical phonology and (2) it precedes all rules of word formation in the lexical phonology.

We will present two types of evidence that tone association is lexical rather than postlexical. In both cases, if tone association applied postlexically, incorrect forms would be generated. First, consider a case where a polysyllabic stem has a single tone associated with it, as in (4).

(4) a. bulu 'horn' b. bulu + nga 'horns'

If tone association is postlexical, brackets will have been erased, obliterating word internal structure and predicting the tone association in (5).

(5) bulunga

H L

The correct tone association is not (5), however, but (6):

(6) bulunga

H L

This pattern follows only if tone association is sensitive to word internal structure.

A second piece of evidence for tone association being lexical comes from the existence of word internal contour tones. An example is shown in (7).

(7) a. mba 'rice' b. mbanga 'rices'

Again, postlexical tone assignment predicts the incorrect form in (8) rather than the correct (7b) since access to word internal structure is not allowed.

(8) mbanga → *mbanga

L H L L H L
Once tone association is placed in the lexical component of the grammar, the theory of lexical phonology predicts that it will apply immediately to the basic lexical entries. As discussed in Kiparsky [1983], structure building rules can apply to basic lexical entries and are not restricted by the Strict Cycle Condition to derived environments. Tone association is a structure building rather than a structure changing rule and therefore applies as the first rule of the lexical phonology before any phonological or morphological processes have a chance to apply.

4. The Destruction of Assigned Tones

Since tone association can be shown to apply before any other processes, it follows that when word formation processes apply, the component parts of the word already have their tones associated. Thus if any word formation process introduces a new tone pattern, the original associated tones must be deleted. Thus, in a reduplicated form like (9), the tones on the second element must be erased in order to give the correct phonetic form.

(9) mbela mbela
    \[ \text{phonology} \]
    \[ \text{morphology} \]
    \[ \text{phonology} \]

The surface tones in reduplicated forms might be accounted for by claiming that reduplication involves copying merely the segmental tier, leaving the second element toneless. The tone erasure step would then be unnecessary and the derivation would go directly to mbela mbela and then continue as illustrated.

H

\[ \text{tone association} \]

\[ \text{reduplication} \]

\[ \text{tone erasure} \]

\[ \text{insert L} \]

\[ \text{move final H} \]

\[ \text{tone spread} \]

\[ \text{postlexical rules} \]
While such an approach involving copying only the segmental tier will handle the reduplication data, it is unavailable for any word formation process involving elements with lexically assigned tone. An example is noun-noun compounding, as shown in (10).

(10)  
\[
\begin{array}{c}
\text{mende} \\
L \ H \ H \\
\text{hale} \\
\end{array}
\]

\begin{tabular}{ll}
phonology & mende hale \\
& L H H \\
morphology & [[mende][hale]] \\
& L H H \\
phonology & [[mende][hale]] \\
& L H \\
& insert L \\
& L H L \\
& copy final H \\
& L H H L \\
& postlexical rules \\
& L H H L \\
\end{tabular}

tone association
compounding

Since hale with associated H is a basic lexical entry on its own, it will be assigned tones by tone assignment on entering the phonological component of the grammar. Since this item is assigned tones before compounding occurs, the tone erasure step is necessary to eliminate these assigned tones. We thus see that there is no way to avoid erasing autosegmental structure.

To sum up, we have shown that tone association in Mende precedes all rules of word formation. We have also shown that word formation rules involve wholesale changes in tone melody and thus must detach and erase previously assigned structure. What remains to be shown is that the erasure of tones need not be stipulated in the grammar of Mende, but follows automatically from the theory given the existence of the tone insertion, copy, and movement rules. In the derivations shown in (9) and (10), the tone erasure step (in parentheses) need not be stated but follows given the other rules.

As Kiparsky [1983] discusses for English stress rules, structure building rules override existing metrical structure when they apply in derived environments but not when they apply in non-derived environments. This overriding of existing metrical structure is not blocked by the Elsewhere Condition since the string affected by the stress rules is not a substring of any lexical entry. Since the structural description for the stress rule is met and nothing blocks it from applying, the rule must destroy any previously existing structure to be
interpretable: a stress rule cannot apply to stress an already stressed form.

The destruction of autosegmental tone can be viewed similarly: it follows automatically from the theory given the existence of the derived environment low tone insertion rule. This rule is not interpretable unless it automatically erases structure previously assigned since tones cannot be assigned to already toned structure. Since tone melodies in Mende are associated with morphemes rather than syllables, it follows that when tone erasure occurs the tone melody of an entire morpheme disappears.

[The authors' names are given in alphabetical order. No priority of authorship is intended. Thanks to Patrick Conteh for help with the data.]

REFERENCES


ONE FATHER OR TWO? POLYSEMY IN KINSHIP TERMS

Chet A. Creider
University of Western Ontario

1. Introduction

In this paper I propose the use of empirical procedures developed in the study of semantics in linguistics for the resolution of aspects of a long-standing debate in anthropology. All kinship terminologies use a small number of terms. For example, in a Hawaiian-type system, there may be only a handful of terms. The Mbuti pygmies use just five terms [Turnbull 1965:110]: tata 'individual of one generation greater than my parents', epe 'male of my parents' generation', ema 'female of my parents' generation', moko 'individual of my generation', and mlki 'individual of a generation below my own'. Kinship terms which denote a large number of kin types are termed classificatory after Morgan [1871]. Terms which denote a small number of types are called descriptive. These two labels are frequently also used to describe entire terminological systems, although this usage is of course very imprecise.

In the semantic analysis of kinship systems, two general approaches have been taken in the literature. In the first, which I will call the social category approach, it is claimed that kinship terms do not inherently have anything to do with kinship (as biologically defined). So-called kinship terms are in essence terms which express relationships of social category. Relationships which are felt to be ones of kinship in, e.g. English, are "covered" by, i.e. have as functional analogues, terms which are non-biological in basic meaning.

The second general direction for analysis is taken by the "extensionist" hypothesis. Under this hypothesis, kinship terms have core meanings which are biologically defined. These meanings are then "extended" to cover wider ranges of kin types. Thus epe in Mbuti usage would have as its "core" meaning 'father of ego'.

A classical statement of the social category approach was provided by Leach in 1958 in a paper entitled, "Concerning Trobriand clans and the kinship category tabu". A highly polemical reanalysis of the Trobriand data in terms of the extensionist hypothesis was given by Lounsbury [1965], "Another view of Trobriand kinship categories". Since then the debate has raged without, to the best of my knowledge as an impartial outsider, resolution. It has in fact become customary for anthropologists simply to cite their affiliation to one or the other "side" in the debate as though this were now simply a matter of which team one belonged to and had more to do with matters of ascriptive identification than with decidable issues [Kiernan 1977]. In part this is due to the fact that virtually all of the analyses which have been utilized in the discussion have been based on data collected by others. Both Leach and Lounsbury based their analyses on Malinowski's Trobriand data, supplemented by data provided by a missionary. Scheffler and Lounsbury [1971] used data collected by Allan Holmberg, etc. We will see below why this is a self-defeating procedure.
2. A Sketch of the Nandi Kinship Terminological System

Before proceeding further let me outline the principal relevant features of the kinship terminological system of Nandi, a Southern Nilotic language spoken by some million or so individuals in Western Kenya (including closely related dialects). The terminology is a variant of the bifurcate-merging or Iroquois type known as Omaha. The term kwaanta is used in referring to ego's father, his father's brothers and all individuals whom ego's father calls brother. This includes of course all of ego's father's father's brothers' sons since ego's father refers to his father's brothers as kwaanta. In addition one's father's mother's sisters' sons are called 'father' because ego's fathers call the children of their mothers 'brothers and sisters'.

The term eeyoo is the basic word for mother. It is used to refer to one's own mother, to one's father's other wives and to all women whom one's mother calls 'sister'. It also is used to refer to the wives of everyone one refers to with kwaanta.

It should be emphasized that the language does not provide terms to distinguish among these different kinds of 'mother' and 'father'. It is of course possible to use relative clause constructions to identify uniquely the particular individual one is referring to, but in such a case there is no tendency to prefer a kin-based identificational procedure over any other means of uniquely identifying some referent. It is as natural for a Nandi to refer to his FaBr as 'father' as it is for an English speaker to call the same individual 'uncle'. Countless details of usage reinforce the system. To cite just one, the Nandi practice teknonymy, and refer to an individual by the name of his or her (usually first) child. A man without children of his own, but with, for example, an older brother who has had a son named Kibet, will be called 'father of Kibet'.

The term seenke is used to refer to anyone whom ego's father calls 'sister'. These individuals are father's sisters, father's father's brother's sisters and father's mother's sister's daughters.

The term maama denotes anyone that one's mother calls 'brother', i.e. mother's brothers, mother's father's brother's sons and one's mother's mother's sister's sons. This term can also be applied to generations below ego, i.e. to one's MoBrSo and MoBrSoSo. Finally it is a reciprocal term, so it also means 'sister's son', i.e. 'son of anyone ego calls sister'.

This basic system is paralleled on one's own generation by the use of terms for 'sibling', 'brother', and 'sister' for anyone who is the child of someone one calls eeyoo or apa (the address term corresponding to kwaanta 'father'). Similarly, the children of one's sibling are called 'sons and daughters'.

3. The Analysis of Polysemy in Linguistics

George Lakoff [1970], in a well-known Linguistic Inquiry squib, distinguished between ambiguity and vagueness and provided a number of tests which allow linguists to distinguish between the two. The article by Arnold Zwicky and Jerrold Sadock [1975], "Ambiguity tests and how to fail them", provided a convenient summary of the literature together with an evaluation. The basic test I
will use in the next section "passes" according to Zwicky and Sadock and is un­
controversial today.

Our goal will be to determine if the meaning of a Nandi kinship term which is classificatory is polysemous. If it is, then an ambiguity test will allow us to distinguish the different senses which the term has. Let me illustrate with a couple of English examples. First, suppose we wish to determine whether the English term 'father' is ambiguous, i.e. has two senses, as between 'biological father' and 'priest'. Imagine a situation in which there are two indi­
viduals, Bill and John, both of whom regularly attend mass and frequently turn to their respective priests for counselling. In addition both have living fa­
thers. In other words, each of Bill and John has two 'fathers' as shown below.

(1) 1 2 3 4
  biological father priest biological father priest
      John       Bill

Now consider the sentence, 'John called his father, and Bill called his also'. The possible pairings of individuals who are in the 'father of' relationship to John and Bill are 1+3, 2+4, 1+4 and 2+3. It is clear that only the first two are possible readings for the sentence. That is, the construction requires identity of sense and hence distinguishes between two different senses of the same word.

Now consider again the same two individuals. Each of these has two 'uncles' as shown below.

(2) 1 2 3 4
  MoBr FaBr MoBr FaBr
      John Bill

If we look at the sentence, 'John called his uncle, and Bill called his too', it is clear that any of the logically possible pairings of 1+3, 1+4, 2+3 and 2+4 are plausible readings of the sentence. We may conclude that the term uncle, with respect to the distinction between MoBr and FaBr, does not distin­
guish between MoBr and FaBr, i.e. that there is a single sense to the term (no polysemey).

4. The Application to Nandi

Suppose these kinship relations: Arap Suswa is the biological father of Tapsuebi and Arap Barno, now deceased, was one of his brothers. The following sentence is perfectly appropriate in Nandi:

(3) kwaant-aep, Taapsubei Arap Suswa, ak kii Araap Barno korā
father-of Tapsuebi Arap Suswa and Past Arap Barno also

'Arap Suswa is the father of Tapsuebi, and Arap Barno was too'
In this sentence, where Tapsubei has a father (in the English sense) and a paternal uncle (also termed 'father' in Nandi), the two individuals are both given the attribution 'father'. It may be useful to compare an English example. Suppose John is Bill's biological father and Andrew is his priestly father. The sentence, 'John is Bill's father and Andrew is, too', is not possible as a description of this state of affairs. Although syntactically well-formed, it is semantically ill-formed (unless uttered for the sake of a laugh).

As another example, suppose that out of a group of men standing some distance away from the speaker, two men detach themselves and come towards the speaker. If one of the two men is the father of Kiprono, and the other is a paternal uncle of Kiprono, it would be entirely appropriate for the speaker to say,

\[(4)\] kwaanwaakik aap Kiproono chaa pwaaney
fathers of Kiprono those who come (pl.)
'the ones who are coming are fathers of Kiprono'

Compare a situation in English where a pair of eyeglasses and a drinking glass are located some distance from a speaker. It would not be appropriate to say, "The glasses are over there." A somewhat more subtle English example may be constructed. Suppose a similar situation to the Nandi one, but suppose that of the two men who come toward the speaker, one is the brother of the father of John and the other is the brother of the father of John's wife. It would not then be appropriate to utter, "The ones who are coming are John's uncles."

As a final example, suppose we have two friends, Kibet and Kiprono, and (5) is uttered with respect to them.

\[(5)\] kii-kuur kwaanta-nyii Kipeet, ak kokuur nenyiï Kiproono
Past-call father-his Kibet and call his Kiprono
'Kibet called his father, and Kiprono called his'

In this sentence, in a situation where Kibet and Kiprono each have a father (in the English sense) and one or more paternal uncles, i.e. individuals also termed 'father' in Nandi, any of the logically possible combinations of fathers and paternal uncles may be intended.

Analogous sentences with the term eeyoo 'mother, mother's sisters' may be constructed, and the "additional" (from an English-speaker's point of view) sense of 'another of Fa's wives' may also participate fully in the core extension of the term.

What can we conclude from this exercise? At the very least we can claim that from the standpoint of cognitive reality, that is the semantic system operated by an adult speaker of Nandi, there is no polysemy here. The extensionist hypothesis, however attractive it may appear from a formal point of view, sim-
ploy fails to produce an analysis which is cognitively real.

However, as is often the case with the real world, all is not that simple. When we turn to the more extended senses of the Nandi terms, the very same tests clearly distinguish between the above core extensions of the terms and the extended senses of, for example, FaFaBrSo for *kwaanta*. Interestingly, the cross-generational neutralization of the term for MoBr, *maama*, does not survive the test with a unitary sense. The use of the term for MoBr for MoBrSo and MoBrSoSo is criterial for a kinship terminology to be an Omaha type. So this usage is less well-integrated cognitively than the basic Iroquois features.

I mentioned previously that it was unfortunate that the principal articles in the extensionist debate were based on secondary analyses of data collected by others. It is in fact surprisingly easy to simply ask speakers whether there is or is not a unitary sense to a given kinship term. As an English speaker, it is clear to me that 'father-as-priest' and 'father-as-biological father' are different senses of the term 'father', but that the term 'uncle' applies equally well to MoBr and FaBr, i.e. that both are included in the basic extension of the term. In the more subtle case of the use of the term uncle for MoSiHu or FaSiHu, my intuitions are less clear, but I am satisfied with the results of the test given in the discussion of the English analogue to (5). Nandi operate a culture with far fewer straggly pieces and loose ends than ours, and their judgements are much sharper. Their intuitions coincide with the results of the test.

5. Discussion

Despite the fact that at the locus of what might be called primary extensions, the extensionist analysis is not supported in Nandi, it appears to me that it has a very strong basis in the ontogenetic development of the child. A Nandi child first learns to refer to his biological mother and father as *eeyoo* and *kwaanta*. He is then taught to apply these already learned terms to FaBrs and MoSis. My guess is that as a child he operates a system which is best described in the extensionist framework. As an adolescent the child comes into extensive contact with his close relatives, and I suspect that by the time of his initiation in early adolescence, the child has internalized the cognitive system which matches the social reality which is the Nandi social system—one based on the fundamental principle of the equivalence of siblings. The Nandi were traditionally a highly independent, acephalous ethnic group with a high level of military activity. In a society characterized by a high mortality rate for males, the Nandi kinship system, together with associated institutions of teknonymy, the levirate, etc., ensured that children always grew up with someone to inherit from, someone to call father who would arrange their initiations, marriages, provide them with cattle, etc. There were no single parent households in Nandi despite the fact that large numbers of males were lost in military action.

In conclusion, I have suggested new procedures for the semantic analysis of kinship terminological systems. These procedures should be readily adaptable for use in other cultures and languages, and I hope that work of this type will aid in deciding which of several competing analyses has the strongest claim to cognitive validity. This in turn may lead to renewed progress in the area of kinship studies.
[I would like to thank J. T. Creider for discussion of the Nandi examples used in this paper and of the ideas proposed here.]

REFERENCES


MICROCOMPUTERS AND AFRICAN RESEARCH

Gerard M. Dalgish
Baruch College, CUNY

This paper examines some applications of microcomputers in conducting research and analysis in the field of African languages. It describes the minimum capabilities necessary to conduct computerized research and analysis, explores some examples of syntactic, morphophonemic, and historical computerized data-keeping and analysis for Khoisan, Niloctic, Cushitic, and Bantu languages. It concludes with some general comments on the usefulness as well as the limitations of such research and mentions additional applications of interest to Africanists as well as African linguists.

Syntactic computerized analysis is often more complex than the other areas. I worked on a project at the University of Illinois in which sentences from five East African languages—Sandawe, Maasai, Iraqw, Acholi, and Chagga—were entered on the PLATO facility. In the presentation there was a reproduction of a printout for all examples containing WH-Question. One "field"—computer jargon for data that the computer will treat as one group—consisted of the sentences morphosyntactically split; for example, ho-a dlomo ?leng tanga. The second field listed the glosses of each morpheme: WH-ELEMENT+AGENT,BUY,PROG(ressive),MELON. The third field glossed the sentence, in this case, "Who is buying the melon?", and other fields contained "anything of interest" for the particular language, in this case such remarks as "affirmative", "question", etc.

Once the data were entered—a not insubstantial undertaking—one could begin to search for any morpheme, marker, syntactic category, gloss, or combination of any of these. For languages with complicated syntactic word order, such as Sandawe with a practically free word order system, keeping track of the possible permutations is something the computer can do with ease. In fact, one could construct a short computer program to generate possible word order combinations and test such items against the speaker's intuitions for grammaticality, or for semantic/pragmatic focus, etc.

For languages with interesting deictic pronominal processes, such as Maasai, the computer can "find" quickly what might otherwise be disparate data, linking points that might otherwise be missed. The presentation demonstrated a search for the morpheme aa in Maasai; it is used for 1st singular objects and 1st singular subjects under certain circumstances. It is important to note that it is up to the linguist to make sense out of what the computer shows him/her; the computer does not yet do the analysis. But with the computer we have a quick way of bringing all the relevant data together to help us make the judgement.

These initial examples were with PLATO, but many of us do not have access to that facility, and in fact, sometimes mainframes are more trouble than they are worth in terms of convenience and access. But with relatively simple database systems that run on microcomputers, we can begin to conduct research just as easily as with mainframes. I use a very simple database called Incredible Jack which is not without its problems, and in the presentation I provided a
sample template form for syntactic data, and then sentences entered as "rec­
ords" on that template. It is possible even with this limited system to up­
date, i.e. retype or correct, anything you entered, make searches of varying
degrees of complexity, and print out what you need when you need it. A cru­
cial example can be searched for in your file and for some systems, can be
merged into your text without even retyping.

Morphophonemic analysis and historical investigation can also be aided by
the computer. Words or morphemes can be entered along with—as in the case
for Bantu as an example—such relevant information as class membership, any
and all alternations observed, etc., in short, any information that might nor­
mally go on one's index cards, looseleaf binders and file folders. In the
presentation, a sample template for Luyia was given, where the fields include
such information as underlying form, part of speech, gloss, class #, singular,
plural, etc. Once one enters the data, one can examine here paradigms along
any parameter one chooses, e.g. grammatical/semantic paradigms (as infinitive,
imperative, passive, first-person singular, applied, stative, etc. forms); one
can then transfer, say, all the verbs into another file and then enter present
tense forms. One could save a lot of work by being able to look at morphemes
by root-initial or root-final segments, by nasal interactions, vocalic inter­
actions, tonal phenomena, etc., all without laboriously recopying and realpha­
etizing. Relevant alternations need be entered only once, after which para­
digms can be called up by searches of particular fields. For complicated al­
ternations like those observed in certain past tenses of Bantu languages, we
can easily compare paradigms within a morpheme or across morpheme types for a
particular category.

As for historical work, once a list of Guthrie, Meinhof, Bourquin, or any
proposed proto-system of Bantu is entered, one can again compare one's own da­
ta with the proposed proto-forms and match correspondences. One could even
imagine a software package "Guthrie for the Apple Computer" which would ask
for selected items in the Bantu language of your choice, match them semantical­
ly to Guthrie's starred forms, and give you the correspondences (or let you fi­
gure them out.) In the presentation there was a printout of a few morphemes I
entered with their corresponding proto forms, showing where most vowel-initial
and y-initial roots came from. But the bigger one's database becomes, the
more the correspondences could fall out.

Another project that can actually be fun is to program the computer to sim­
ulate the derivations of certain forms. I presented a program written in BAS­
IC that can accept as input two morphemes of a word (verbs mostly) and produce
the correct surface form. This could easily be applied to CAI (computer-as­
sisted instruction) in African languages, where the student can be called upon
to test his knowledge of how to put strings together. It would also be inter­
esting to know what sorts or classes of rules will allow themselves to be ren­
dered in BASIC and which can not. Since the structure of BASIC is well known
to mathematicians, we might find out the structure of certain rules that cannot
be captured in BASIC and test to see how widespread such rules are, or even
whether we should rethink our whole paradigm of rule-writing, rule-ordering,
and the like. At the present, artificial intelligence (AI) research is foc­
ussed primarily on English and a handful of Indo-European languages; African­
ists have long been aware of the dangers of such limitations. As we begin to
attempt to parse our "exotic" languages both morphemically and syntactically,
we may encounter problems that will lead to a rethinking of some AI research.

I conclude with some remarks about some of the advantages and disadvantages of computers in African linguistic research and present some caveats and pitfalls to avoid. Some drawbacks include limitations in memory storage which can slow your work down, especially as you increase the size of your database. This point should not be underestimated. Nor should the value of having good merge capabilities be forgotten—you should be able to switch your data around, both in existing files and in new files, in and out of the databases to a word-processing program.

A much more important potential drawback is the fact that the linguist must have some preliminary analysis in mind prior to deciding on which fields to set up. In some sense, you have to know what's going to be interesting before you go about setting up your schema to enter the data. Obviously, a thorough grounding in the structure of the language family is helpful before one investigates a language within that family. And this problem can also be minimized if one can easily redesign one's file structure, which is possible in most software. Networking with computer-interested Africanists who are attempting this sort of research is also very helpful.

One of the advantages is the sheer speed of finding relevant forms, as well as the possibility of uncovering a correlation that had gone unsuspected. It will also be simpler to comply with requests from colleagues for data; no laborious writing out or spending days at the xerox machine will be necessary. And for monolingual and bilingual dictionaries, word-processing and database systems are perfect for gathering, alphabetizing, searching and replacing, etc. Instant dictionaries could be produced, alphabetized by English or by African language, without retyping.

In short, microcomputers can provide the African linguist with important analytical tools to aid research in and compilation of African language data. Syntax, phonology, historical work, lexicography, CAI, and even AI are all areas that will benefit from the contributions of Africanists with computers.
This paper discusses some features of the verbal aspect system of Hausa and relates them to a class of phenomena known as constituent control, unbounded dependencies, or syntactic binding. Specifically, I will try to develop a unified account of the distribution of Hausa relative aspect markers, within the framework of lexical-functional grammar (LFG). An introduction to LFG is Kaplan and Bresnan [1982] and this paper also incorporates ideas in Zaenen [1983].

The aspect of a Hausa verb is indicated by an aspect marker preceding it, which also indicates the person, number, and gender of the verb's subject, as shown in (1) and (2):

(1) a. Hālīmā tā zō
    H. 3sgf,cmpl come
    'Halima came'

    b. kin zō
    2sgf,cmpl come
    'you came'

(2) a. Hālīmā tana zuwā
    H. 3sgf,cont come
    'Halima is coming/comes (regularly)'

    b. kin zuwā
    2sgf,cont come
    'you are coming/come (regularly)'

There is another form of the completive and continuative aspect markers, known as the relative form, which must be used in the contexts exemplified in (3)-(6):

(3) relative clauses
   a. mutānen dà sukā zō jiyā
      people Comp 3pl,r.cmpl come yesterday
      'the people who came yesterday'
   b. *mutānen dā suñ zō jiyā
      people Comp 3pl,cmpl come yesterday

(4) word questions
   a. ñe ya ganT?
      what 3sgm,r.cmpl see
      'what did he see?'
   b. *ñe yā ganT?
      what 3sgm,cmpl see

(5) cleft and fronting constructions
   a. Kānde cē ta zō
      K. copula 3sgf,r.cmpl come
      'it's Kande who came'
b. "Kânde cê tā zō
K. copula 3sgf,cmpl come

(6) a. cikin mōta (nē) mukā zō 'we came by car'
in car copula 1pl,cmpl come
b. "cikin mōta (nē) mun zō
in car copula 1pl,cmpl come

The relative aspect markers also appear in narrative contexts, as when tell­
ing a story or fable. No attempt will be made here to deal with occurrences in
these contexts, and it is unclear at best whether they are syntactically deter­
dined or related in other ways to those in (3)-(6). This problem aside, the
list in (3)-(6) is insufficient in two respects. First, it doesn't provide a
unified description of the environments in which relative aspect markers occur.
A grammar of Hausa incorporating such a description would be preferable, all
other things being equal, to one that didn't. Secondly, this account isn't
strictly adequate with regard to verbal complements, as (7) and (8) show:

(7) a. mē su kē fātan sun gamā? 'what do they hope they have
what 3pl,r.cont hope 3pl,cmpl finish finished?'
b. "mē su kē fātan sukā gamā?
what 3pl,r.cont hope 3pl,r.cmpl finish

(8) a. yāran dā su kē zāton yā tāfī Kanō
boys Comp 3pl,r.cont think 3sgm,cmpl go Kano
'the boys who thought he went to Kano'
b. "yāran dā su kē zāton ya tāfī Kanō
boys Comp 3pl,r.cont think 3sgm, r.cmpl go Kano

I will now present an outline of some parts of LFG relevant to what follows.
There are no movement or deletion rules in LFG; instead an empty category is
generated in the "gap" and it is anaphorically related to some other constitu­
te. The empty category I will call the bindee, and the corresponding constit­
tuent is its binder. Thus in

(9) Who did you see e?

Who is the binder of the empty category, and both are base-generated by phrase
structure rules. Phrase structure rules in LFG are constrained by functional
equations, which contain information about the grammatical relationships the
constituents bear to one another. As an example here are some of the PSR's
that could be used to generate (9):

(10) S ---› NP Aux NP VP
     (↑ TOP)=↑ +⇒ (↑ SUBJ)=↑ +⇒
     ↓=↓

     VP ---› V NP
     +⇒ (↑ OBJ)=↓
The $\uparrow$ is a variable referring to the functional information at the mother node, while $\downarrow$ refers to the node itself. These equations are subject to certain general principles. The head of a category always bears the equation $\uparrow = \downarrow$; such an equation in effect passes all of the functional information at the current node to its mother node. A maximal category may bear an equation of the form $(\uparrow G) = \downarrow$, where $G$ is a grammatical function such as SUBJ, OBJ, or PRED. Minor categories may only bear the equation $\uparrow = \downarrow$.

In (9) the binding of *who* to the empty category is effected with the $\Lambda$ and $\downarrow$ metavariables. $\Lambda \downarrow$ must be matched by a $\uparrow$. The equation $\uparrow = \downarrow$ is associated with the constituent which is the binder, while the equation $\uparrow = \uparrow$ is associated with the empty node in the gap. Kaplan and Bresnan (1982) proposes conditions on the possible configurations in which $\Lambda$ and $\downarrow$ may appear. Some of these are shown in (11), which also introduces the notion of binding domain:

$$\text{(11)}$$

![Diagram of binding domain structure](image)

In (11), node $y$ is the binder. Node $r$, which is required to be a daughter of $y$'s mother, is the root node of the binding domain. The bindee must be dominated by the root node, as node $z$ is in (11). A further stipulation is that the path from the root node to the bindee may not contain any bounding nodes. Thus, if node $b$ is a bounding node, no node that it dominates, such as $w$, is a possible bindee of $y$. The class of bounding nodes is specified for each language, and it has been argued that $S$ is a bounding node in English. This is the basis of the LFG account of island constraints. For Hausa, I have not assumed that $S$ is a bounding node, as extraction seems to be freer than in English.

We can now define a binding domain as all the constituents dominated by a root node but by no bounding nodes (except possibly the root node itself).

Returning to relative aspect markers, we note that, narrative contexts aside, they seem to occur only within binding domains. We now need some mechanism in LFG which will permit only the relative forms to appear there. Zaenen (1983) proposes a lexical feature BND, which may have the values $+$ or $-$. For Hausa, its value will be $+$ for relative aspect markers and $-$ for neutral ones. In aspects that don't distinguish between relative and neutral forms, BND can be either $+$ or $-$. Also, every phrase structure rule that introduces a root node will associate with it a constraint equation of the form:
(12) \[ S \quad (\uparrow \text{BND}) = c + \]

Unlike the other equations, constraint equations don't contribute any functional information, but require that some other equation in the tree do so. Here, for example, the constraint equation (12) will be satisfied if some constituent of \( S \) bears the equation:

(13) \[ (\uparrow \text{BND}) = + \]

and if the value of \text{BND} can percolate up to \( S \) via a series of \( \uparrow = + \) equations.

Now only one more thing needs to be stipulated—the root node of any binding domain bears the constraint equation (12), while all other \( S \) nodes implicitly bear the equation:

(14) \[ (\uparrow \text{BND}) = c - \]

Now we can give an account of (6) and (7). The structure of (7a) is given in (15).

(15) 

```
S -------
(\uparrow \text{TOP}) = \uparrow \quad \uparrow = \downarrow
(\uparrow \text{BND}) = c +
\downarrow \quad \uparrow = \downarrow
NP -------
\text{mè} -------
\uparrow = \downarrow
VP -------
(\uparrow \text{XCOMP}) = \uparrow
S -------
(\uparrow \text{BND}) = +
\uparrow = \downarrow
\text{Aux}
\text{su kè} -------
\text{fātan}
\uparrow = \downarrow
VP -------
(\uparrow \text{OBJ}) = \uparrow
\text{Aux}
\text{V}
\text{NP}
\text{au}
\text{gama}
\text{e}
```

The root \( S \) node bears the constraint equation (12), as required, and so the verbal aspect marker in that clause must be in the relative form, in order to supply the \( + \) value of \text{BND} which will satisfy the constraint. But in the lowest clause the aspect marker can only be neutral, because the \( S \) node immediately dominating that clause is not a root node. Thus this \( S \) node implicitly bears
equation (14), and the presence of a relative aspect marker in this clause would cause a clash in the value of BND in that clause, accounting for the ungrammaticality of (6)b.

There are two verbs in Hausa whose complements behave somewhat differently, as (16) and (17) show:

(16) a. mutanen da suka riga suka tafi
    people Comp 3pl,r.cmpl have already 3pl,r.cmpl go
    su ne suka saci sanlya
    they copula 3pl,r.cmpl steal cow
    'the people who have already left, they stole the cow'

   b. *mutanen da suka riga sun tafi
    people Comp 3pl,r.cmpl have already 3pl,cmpl go

(17) a. mutanen da suka rikida suka zam kureye
    people Comp 3pl,r.cmpl metamorphose 3pl,r.cmpl become hyenas
    'the people who metamorphose and become hyenas'

   b. *mutanen da suka rikida sun zam kureye
    people Comp 3pl,r.cmpl metamorphose 3pl,cmpl become hyenas

(Example (17) comes from Eulenberg [1972].)

In (16) and (17) both the main verbs and their complements take relative aspect markers when within a binding domain. It is noteworthy that these two verbs also require the subject of their complement to be coreferential to their own. An account of this correlation would be desirable.

One way of doing this is to assume the structure for these sentences shown in (18) on the next page. We have to make riga or rikida a member of * ADJ (adjuncts) to avoid assigning two conflicting values to PRED. This means that these two verbs are treated as a sort of adverbial, and what appears to be the complement is treated as the main verb. Notice that all of the features of both aspect markers will percolate up to the top VP; this means that if they differ in any way there will be a clash in the values of person, number, gender, aspect, or BND, and the structure will be ill-formed.

There is a hitch in this account which is illustrated by (19):

(19) a. mutanen da ba su rigayi suka ga sinemar ba
    people Comp Neg 3pl,cmpl have already 3pl,cmpl see movie Neg
    'the people who haven't already seen the movie'

   b. *mutanen da ba su rigayi suka ga sinemar ba
    people Comp Neg 3pl,cmpl have already 3pl,cmpl see movie Neg

When riga is negated, the aspect marker of the following verb is neutral even in a binding domain, at least for some speakers. One treatment of this problem is to complicate the constraint equations associated with root nodes,
as in (19):

(19) (⊥ BND) = c + or (⊥ NEG) = c +

However, this is clearly ad hoc in the absence of independent motivation.

[I thank Will Leben and Joan Bresnan for their help and encouragement and Bashir Wada for sharing his native knowledge of Hausa.]

REFERENCES


NIGER-CONGO NOUN CLASS AND AGREEMENT SYSTEMS
IN HISTORICAL AND ACQUISITION PERSPECTIVE

Katherine Demuth, UC Berkeley
Nicholas Faraclas, UC Berkeley
Lynell Marchese, San Jose State University

0. Introduction

This paper investigates noun classes and concordial agreement systems in Niger-Congo languages. It examines comparative language data from Cross River and Kru languages, noting overall characteristic directions of noun class and agreement system evolution in both families. It then focuses on problems which first language learners might face in the learning of a full class/agreement system, considering both spontaneous and experimental Bantu language acquisition data as well as evidence from Bantu languages in contact. Historical changes which have resulted in the leveling of gender and agreement distinctions in Niger-Congo in many ways parallel the various stages of acquisition of the system in languages which still maintain these distinctions. Findings from these comparative data sets provide important clues as to the nature of the linguistic change at work in the cases considered, as well as providing evidence for and against various theories of how those changes might have taken place.

1. Cross River Languages

The Cross River subbranch is divided into five groups: Bendi, Upper Cross, Lower Cross, Ogoni, and Central Delta, each of which typifies a different stage in the gradual disintegration of the prototypical Niger-Congo noun class and concordial agreement systems. The most conservative groups (Upper Cross and Bendi) preserve many of the original CV- prefixes as well as the gender and concordial systems which most probably existed in Proto-Niger-Congo. The most innovative groups (Lower Cross, and especially Ogoni) show limited concord and few if any gender distinctions, with prefixes where they persist (many nouns in Ogoni are prefixless) typically taking the form N- or V-. In some groups (Central Delta and Bendi especially) old CV- prefixes are incorporated into noun stems.

In Upper Cross, Bendi, and Central Delta, many languages have complete concordial systems including pronominal, subject-verb, adjectival, determiner-demonstrative, and numeral concord. Numeral concord is, however, most often a relic. In Lower Cross and Ogoni only a few relics of concord have survived, including a few adjectival concords.

From the comparative synchronic data on Cross River languages we find a small range of consistent patterns by which gender/concord loss appears to be taking place. It is not just any gender which persists, but rather the human 1/2 and the non-human 9/10 classes which are the last to be lost. CV- prefixes are not simply omitted, but rather reduced to V- or N- or incorporated gradually into the stem with new prefixes occasionally added. The number of plural classes tends to collapse, becoming distinguished by one overgeneralized marker. Concord appears to be lost first with numerals and adjectives, while it is maintained longest with subject pronouns.
2. Kru Languages

Like Cross River languages, Kru languages represent a reduced version of a more extensive proto-system. While Kru languages are almost exclusively suffixing, noun class suffixes have all but disappeared through loss or coalescence with the noun stem. Again, we find more and less conservative Kru languages, Eastern Kru showing more preservation of the old system, and Western Kru showing more loss. As the result of coalescence in Kru languages, noun systems are becoming more regularized, with singulars being perceived as unmarked and plurals marked, a tendency observed both in Bantu [Stucky 1978] and in Lower Cross.

Despite this phonological coalescence and seeming "loss" of noun class suffixes, concord is extensive in Kru. In Eastern Kru languages, concordial agreement occurs on subject, object, possessive, interrogative, and relative pronouns, as well as on adjectives, demonstratives and definite markers, as seen in the Godié examples below:

(1) πųku πa k δo πa m interleaved nū woto  
man big this drank water cold

(2) o m interleaved λ o o

The gender distinction in subject pronoun systems is maintained longer than any of the other manifestations of concord in Kru. With one exception: all languages still maintain a singular human/non-human distinction throughout the pronoun set.

As class distinctions decrease, there appears to be a corresponding (though not absolutely predictable) decrease in concordial agreement. Though most languages do not retain adjective gender agreement, singular/plural distinctions are maintained on some adjectives. As in Cross River, numbers and definite markers seem to be lost first.

Kru languages thus represent a case where remnants of gender classes are preserved as coalesced nominal suffixes. While the productivity of nominal classification is minimal, the distinctions which are made correspond to the human classes 1/2a and non-human classes 9/10, similar to the most dominant class retentions in Cross River languages. Though the gender system is minimally functional, the agreement system is still productive, indicating a primacy of concord over nominal marking.

3. Bantu Language Acquisition Data

The language acquisition data from Bantu languages stem from a longitudinal study of spontaneous and natural speech from 4 Sesotho speaking children between the ages of 2 and 4 1/2 years old [Demuth 1983], a spontaneous/naturalistic Siswati study of 2 children between the ages of 18 months and 36 months, and an experimental study of 3 children 4 1/2 to 6 years of age [Kunene 1979].

From our current cross-linguistic knowledge of strategies that children em-
ploy when learning a language [Slobin 1984], we have made several predictions about the kinds of phenomena we would expect to find in the acquisition of the Sesotho noun class system: (1) The learning of noun class prefixes might prove problematic as they are word initial, carry little semantic content, are usually found in unlengthened (and unstressed) position, and usually bear low tone. (2) children might try to collapse the number of singular/plural distinctions made, possibly taking the singular form plus prefix as the root and using only one plural marker, probably class 10. (3) class 9 nouns with \( \emptyset \) prefix and monosyllabic nominal stems + prefix would be assigned a singular marker of some sort, making them "fit the paradigm". (4) there would be a somewhat reduced number of gender distinctions in the agreement system initially, perhaps surfacing productively as class 9/10.

Upon considering spontaneous Sesotho language acquisition data we discover that only some of these predictions are upheld. Noun class markers were used with \( \emptyset \text{-}, V\text{-}, \) or CV- markers at the same stage of development, as illustrated below:

(3) (25.0 months):

\begin{align*}
\text{-punko} & \quad -phoko \\
\text{\( \tilde{a}\)-phoko} & \quad \text{\( \tilde{a}\)-phoko} \\
\text{e-punk} & \quad \text{ma-punko} \\
\text{(l\( \varepsilon \)-phoko)} & \quad \text{(m\( \varepsilon \)-phoko)} \\
\text{\'green corn stalk\'} & \quad \text{\'green corn stalks\'}
\end{align*}

Thus, as with the various stages of loss or incorporation of nominal prefixes found in Cross River languages, children also progress through a stage where a single lexical item may be rendered with variable prefix shape, even in consecutive utterances in the same contextual and grammatical environment. Exceptions to this progression of appropriate nominal marking are found in the omission of class 5, 7, 8, and 10 prefixes (Sesotho \( \text{\( \varepsilon \)}, s\varepsilon, (N)\text{\( \varepsilon \)} \text{\( \varepsilon \)} \) and \( (N)\text{\( \varepsilon \)} \text{\( \varepsilon \)} \), when an adjunct (demonstrative, possessive, etc.) follows the noun, as seen below:

(4) \( p\text{\( \varepsilon \)}:\text{\( \varepsilon \)} \) \( \text{l\( \varepsilon \)}\varepsilon \text{\( \varepsilon \)} \) \text{\'that green corn stalk\'}

\( (l\varepsilon\text{-phoko} \) \( l\varepsilon\text{-n\( \varepsilon \)} ) \) green corn stalk that

Omission of these particular prefixes when the noun is used with an adjunct is a phenomenon also found in adult Sesotho speech and may represent an initial stage of prefix loss in progress in Sesotho.

Contrary to our hypothesis, class 9 nouns, which take a \( \emptyset \) prefix, are not overmarked with prefixes from other classes, and nouns from one class are rarely assigned to other classes, i.e. there is no collapsing of plural distinctions in singular/plural markings. This differs from the tendencies observed in Cross River, Kru, and Bantu languages in contact, where plural distinctions are frequently neutralized.

These findings contrast with experimental results from Siswati where children of ages 4 1/2-6 did collapse plural markings when the children were given novel word forms and Siswati nouns out of context and asked to provide the corresponding singular or plural form of the noun. Overgeneralizations occurred (1) toward the use of 9/10 class markings for classes 7/8 and 11/10 and (2) toward the over-
extension of class 2a, the human-relative (family relation) class marker, to classes 2, 4, and the liquid/mass class 14 and the infinitival class 15 (apparently analysed only as stems, as they have no singular/plural alternations). Thus older children tended to classify plural nouns into loose human/non-human classes. Independent evidence from Kituba [Mfwene 1972] and Spoken and Kinshasa Lingala [Bokamba 1983] document this same tendency for overgeneralization of the class 2a human-relation class marker for plural when the noun class system begins to lose gender distinctions. The most interesting result, however, is that these same children did not make these or any overgeneralizations in spontaneous speech, indicating a major difference in the nature of "natural" versus "less-natural" linguistic contexts.

The acquisition of concordial agreement is a gradual process which has already begun prior to 2 years of age, most notably with demonstratives and possessives. It continues till past the age of 3 when most subject concords and object clitics, demonstratives, possessives, relatives, adjectives, and some numerals are used with appropriate gender forms.

4. Conclusion

First language learners appear to focus not simply on nouns, but on entire nominal or verbal phrases where concord is of primary importance and nominal marking only secondary. This would account not only for the relative lack of marking errors for nouns in context on the part of Sesotho speaking children and for the primacy of concord not only in the order of acquisition, but also in its persistence in Kru and Cross River languages where productive gender marking has been lost. Noun and modifier may be conceived of as being a cognitive unit, with concord as the crucial unifying element. It is proposed that young language learners may adopt this unit as a basic learning construct, using concord productively while continuing to experiment with the correspondingly appropriate marking for nouns.
REFERENCES


1. Introduction

Isoko is a Southwestern Edoid language spoken in the Isoko Local Government Area of the Bendel state of Nigeria.

It is a discrete level tone language with two level tones: high (H) and low (L) as stated by Mafeni [1969] and Elugbe [1977]. Both previous writers, without any discussion, mention the possible use of intonation in Isoko. Though Mafeni claims the existence of downstep and downdrift, Isoko has neither. Moreover, the phenomenon of final low tone raising discussed by Elugbe is a feature of the intonation system. In this paper we give a preliminary analysis of intonation in the language.

For an adequate understanding of the subject of this paper, it is necessary to make some basic observations about the tone system in general. In addition to the H and L tones, contour tones also exist at the phonetic level. They are derived from underlying sequences of level tones: a segmental and a floating. The floating tone that forms one end point of the contour may arise from segmental processes such as Elision, Glide Formation, etc. or they could be underlying floating tones which have purely intonational function. Amongst the justifications for the latter claim is the fact that contour tones are generally simplified, but where one end point of the contour is a feature of intonation, the simplification of that contour is blocked since it is likely to obscure the intonation signal. Consider for instance

\[ \begin{align*}
&\text{yam'} \\
&\text{yam?'}
\end{align*} \]

A simplification of the contour tone on this final \( \varepsilon \) would yield a positive statement since, where it applies, all contours are simplified to (H).

The lexical function of tone is only prominent in the noun. The tone of the verb stem (Vs) is continually changing from one construction to another; our phonological analysis assigns no underlying tone to the Vs since there is evidence that it acquires its tone by polarisation.

We also postulate a process of tonal replacement. In its citation form the subject concord marker (a linker of the subject and verb phrase (VP)) bears a L tone. This L tone of the subject concord marker (Scm) does not undergo changes generally. However, in conditional and relative clauses at the beginning of a sentence, the L tone of the Scm is always replaced by a H. Compare examples (2a, b) with (2c).

\[ \begin{align*}
&\text{he Scm} \\
&\text{he will go}
\end{align*} \]

\[ \begin{align*}
&\text{he Scm} \\
&\text{he will go'}
\end{align*} \]

\[ \begin{align*}
&\text{he Scm} \\
&\text{he will go'}
\end{align*} \]
c. ṭọ ọ tẹ nà → ṭọ ọ tẹ nà 'if he goes...'

The L tone of the Scm in the imperative is also replaced by a H, after which the Vs is assigned a tone opposite to that of the Scm.

(3) ṭẹ ṣẹ dẹ → ṣẹ ṣẹ dẹ → ṣẹ ṣẹ dẹ 'buy!' (singular)

The process of tonal replacement is seen as a tonal manifestation of intonation in Isoko.

The relationship between intonation and other tonological processes is such that the former is superimposed on the latter. Hence, contour tones that are manifestations of intonation are not subject to simplification. The process of replacing the basic tone of a syllable to give a particular intonation contour usually takes precedence over other tonological processes. Consider, for instance, the replacement of the L tone of the Scm by a H after which the Vs is assigned an opposite L-tone by the process of Polarisation, as in (3) above. If the order is reversed, i.e. polarisation before replacement, the result will be an unacceptable tonal sequence. Compare (3) with (4).

(4) ṣẹ ṣẹ dẹ → ṣẹ ṣẹ dẹ → ṣẹ ṣẹ dẹ

2. Data and Analysis

2.1. Methodology. In addition to our auditory perception and the writer's knowledge as a native speaker, the data for this analysis were documented by test recordings from seven subjects which were analysed on a pitch metre. The pitch level of items with different tone patterns were first observed in isolation. These were then recorded in different positions in various grammatical constructions such as simple statements, yes/no questions, and a few complex sentences. The speed of the utterances is that of a normal conversational register, and we also extracted utterances from ongoing dialogues.

The average intonation curve for each utterance was taken by comparing pitch curves as produced by each of the subjects (from the pitch meter onto the mingo-graph). On the basis of this, we were able to identify five distinct pitch levels which are represented schematically below.

(5)

Pitch level 1 is a downglide, a low-falling pitch. It is unmarked on the vowel. Pitch level 2 (PL.2) is the level of a low tone and marked ' on the vowel. Pitch level 3 (PL.3) is that of a mid tone and it is marked -. This is the raised low. Pitch level 4 (PL.4) is the level of a high tone and marked'
on the vowel. PL.5 is slightly higher than the level of a high tone. This is referred to as the superhigh level and marked '. What generally obtains here is an upglide from high to the superhigh level.

It is necessary to add that sometimes we do get intermediate levels. For instance, a H may be raised above PL.4 but below PL.5 as in the case of the Hs in the High tune of the positive declarative statement (2.2.2). PL.5 is set up on the basis of what obtains in the negative interrogative sentence. The sentence itself is marked by a final floating H tone. Floating tones H or L functioning as features of intonation have remarkable effect of raising or lowering the pitch of the syllable (or vowel) on which they are segmentalised.

L tones are also observed to be raised without getting to PL.3. PL.3 is set up on the basis of what obtains in the case of final raised L tones in positive statements. The overall pitch contour of each sentence is labelled a tune, and this is considered a unit of intonation.

2.2. The intonation tunes. From our examination of the pitch contours of different constructions, the following tunes are identified: Low, High, Rising, and Falling. Each of these tunes can mark more than one construction type.

2.2.1. The Low Tune. This marks negative declarative and positive interrogative sentences. In negative declarative sentences, every syllable before the negative particle maintains either a PL.2 or PL.4 depending on the basic tone of the vowel. A distinctively lowered pitch is realised on the syllable immediately before the negative particle as well as the negative particle itself, which is marked by an initial floating low tone that has a lowering effect on preceding syllables.

(6) a. òbò ònà ọ dc òbl 'hVV
doctor the Scm buy book neg.

(b.) òbò ònà ọ dc òbl 'hVV
the native doctor didn't buy a book'

b. òmù òrà ọ to ụyọ 'hVV
child yours Scm steal money neg.

'your child didn't steal (any) money'

Note that the final H tone of the negative particle is below PL.4 while the preceding L tone is below PL.2.

In the case of the positive interrogative sentence, all tones remain at PL.2 or 4 as the case may be up to the final vowel of the sentence, where a distinctly lowered pitch is heard. This sentence is distinguished from its negative counterpart by the presence of a final floating L tone.
2.2.2. The High Tune. This tune marks positive declarative and imperative sentences. Whereas in the declarative, the high pitch is identified towards the end of the verb phrase, in the imperative, it is identified from the beginning to the end of the VP.

All final L tones in these two sentence types are raised to PL.3. There is no limit to the number of L tones that can be raised in this position provided there is no intervening H [Elugbe 1977]. In the case of positive declarative sentences it happens, however, that there is always an intervening H tone before the last syllable of the VP. This is because the Vs is either preceded by a L tone Scm, to which that of the Vs polarises to become H, or a H tone Am.

All high tones in these sentences remain at PL.4 except those within the VP which are raised slightly above PL.4 but below PL.5.

As for the imperative sentence, the L tone of the Scm is replaced by a H tone while the rest of the sentence ends like the positive declarative sentence.
There may be a few unanswered questions with the high tune we have set up. A high tone is not raised before the VP. Within the VP all Hs are raised to the same level. Within the VP a L tone which is followed by a H tone is not raised. If the key point of identifying this pitch contour is within the VP, why is the raising of the L tones restricted to only final position while that of the H is not?

2.2.3. The Rising Tune. This tune marks the negative interrogative sentence. As opposed to the positive interrogative sentence, it has earlier been said that the negative interrogative is marked by a final floating H tone.

There is a slight lowering of the pitch of the voice at the beginning of this sentence, followed by a gradual rise of the pitch of every syllable until one gets to the last syllable where a distinct rise is heard. All Ls and Hs before the last syllable are heard above PL.2 and PL.4 respectively. One gets the impression that there is a gradual updrift in the pitch of the whole sentence.

The final floating H tone, as usual, contracts with a preceding H while it forms a LH rising contour tone with the preceding L. The contracted H tone gives an upglide while the LH contour tone rises from a level slightly above PL.2 to PL.5.

(10) a. ṣbọ̀ ṣọ́nà ọ̀ de ọ̀gbì 'doctor the Scm buy book Nq
'didn't the native doctor buy a book?'

b. ṣmó ọ àọ to ògò 'child the Scm steal money Nq
'didn't the child steal any money?'
2.2.4. The Falling Tune. This tune marks all non-final pause groups such as conditional clauses, relative clauses at the beginning of a sentence, and coordinate constructions of the listing type.

The falling pitch contour of the conditional and relative clauses is identified at the beginning and at the end of the VP. In this construction type the L tone of the Scm is replaced by a H which is slightly raised above PL.4 while the construction itself is marked by a final floating L tone which occurs after the last vowel of the VP.

This floating L tone causes a preceding L tone (by contraction) to fall with a downglide from PL.2 to PL.1 and a preceding H tone to fall from PL.4 to a level slightly below PL.2 (by contour tone formation). All other tones in the sentence remain at PL.2 or PL.4 as the case may be.

(11) a. ọmọ ọ té to Ṽu ọ́ / ọ́ ọ́ tf de Ṽu ọ́
    child Scm if steal money he Scm will buy cup
    ọmọ ọ́ té Ṽu ọ́
    ọ́ ọ́ tf Ṽu ọ́

    'if the child steals money he will buy (a) cup'

b. ọbl ọnọ ọbà ọ́ de ọ́ ọ́ vrù nọ
    book which doctor Scm buy it Scm lose already
    ọblọnọ ọbà ọ́ de ọ́ ọ́ vrù nọ

    'the book that the native doctor bought, is lost'

When a sequence of pause groups occur in a list each group is marked by a final floating L tone; the pitch of the final syllable of each group thus ends with a fall. The fall is either from PL.4 to 1 depending on the basic tone of the syllable.

(12) ọnl ọmè́ / te ọsè ọmè́ / gbè ọbọ́́ /
    mother mine, with father mine, as well as doctor,
    ọvọ́ ọmọ́ ọnà́́ / à́ à kpe èkl
    and child the they Scm go market
The Intonation System of Isoko

Unlike the relative and conditional clauses, all the high tones of the listed items of the co-ordinate construction are heard above PL.4 before the fall on the last syllable. This construction type differs from those marked by a Low tune in the sense that in the latter, all tones, H or L of every syllable occurring before the end of the VP (which is the key point of identifying the Low tune) remain at PL.4 or PL.2 as the case may be.

3. Summary

Isoko is a tone language with intonation. Since both tone and intonation are pitch features, it is difficult to understand how intonation operates in the language without a preliminary knowledge of the overall tone system.

The two basic tones H and L, apart from their lexical function, are often used as manifestations of intonation. Whereas the H and L tones function as floating tones distinguishing various pitch contours, there is no floating mid tone. The mid tone is only a raised form of the L tone, just as the upglide is a raised form of the H tone marking the pitch of certain utterances. It is not a downstep High as Mafeni [1969] claims.

All Hs after Ls maintain their constant pitch levels and can only be raised or lowered when they occur at the key place for identifying particular intonation contours. These changes are certainly not characteristic of downdrift which Mafeni also claims to be present in Isoko. The error in his analysis notwithstanding, his pioneer work in this area has served as an incentive for the study of Isoko intonation.

This is only a preliminary analysis of intonation in Isoko, since in the investigation of more complex sentences, more tunes may be identified. However, it is hoped that this analysis will inspire further investigation of the presence and use of intonation in Isoko and in languages of the Edoid family.

REFERENCES


1. Background

Autosegmental phonology, depending on how it is constrained, offers a variety of related phonologies. Clearly then, one level of discussion within this approach will have to do with finding the most appropriate version of autosegmental phonology for each language. This paper explores the potential for a "segmental" version for Mende, a Southwestern Mande language spoken primarily in Sierra Leone.

2. Formalization

Briefly the formal properties of the "segmental" version are defined as follows:

(1) All autosegmental formal properties, e.g. tonal tier, segmental base, lines of association, no crossing of lines, obtain.

(2) In Mende lexical representations, all tonal associations are fully specified and obey the autosegmental conditions of association.

(3) The melodic tier conforms to the Obligatory Contour Principles (hereafter OCP).

(4) Mende does not permit three associations per tone-bearing segment. In fact two associations per segment are termed "tense" and are under pressure to resolve themselves to expel one tonal association.

While other formulations of Mende tone may provide workable treatments of Mende tonal phenomena, I offer this one because it permits an interesting treatment of (1) lexical and derived grammatical downstep and (2) an apparent tone rule conspiracy.

3. The Facts of Mende

3.1. Downdrift. The approach taken here is to treat both downdrift and downstep as a register phenomenon rather than one involving the gradual slippage of tone as is the usual practice, an approach which necessitates placing an iterative N-ary rule of downdrift within the domain of binary rules. With the imposition of the OCP, a Mende tonal register can be defined as an H L melodic sequence, regardless of the number of syllables involved. Thus instead of downdrifted high tones, there are downdrifted registers as shown in (1).

(1) H L CVCVCV
(2) H L CVCVCV
(3) H L CVCVCV
3.2. Downstep. A common historical derivation of downstepped high tone is through the loss of a low note between two highs, whether through a deletion or a raising of the low. In a register approach, this results in the adjacency of the high notes of two registers as shown in (2) below.

\[
\begin{array}{c}
\text{HL} \\
\text{CV} \\
\text{H} \\
\text{CV} \\
\end{array}
\rightarrow
\begin{array}{c}
\text{H} \\
\text{CV} \\
\text{H} \\
\text{CV} \\
\end{array}
\]

The representation in (2) appears to violate the OCP because it permits two adjacent high notes. However, the original purpose of the OCP was to prevent a problem of over-distinctiveness (3) where the same phonetic form could have two distinct representations.

\[
\begin{align*}
\text{(3) a. } & \text{H} \\
\text{CVCV} \\
\text{b. } & \text{H H} \\
\text{CVCV}
\end{align*}
\]

However with the concept of register, these forms can be seen to be distinct, for the second H in (3b) marks the onset of a new register which normally downshifts.

Not only does the OCP make possible the identification of registers as H L sequences, it allows for the lexical representation of downstep without the need for special symbols. In (4) the second syllables of (b) and (d) are downstepped.

\[
\begin{align*}
\text{(4) a. } & \text{ngulu } \text{to} \text{to} \\
\text{ngongongo } \text{gone} \\
\text{b. } & \text{H H LH H L} \\
\text{ndoi x lo x mbei x yeyama} \\
\text{c. } & \text{ngi-bi-gava-ni } \text{x lo} \\
\text{ngongongo } \text{gone} \\
\end{align*}
\]

As long as registers can be characterized in this way, there is no need to insert an n-ary downdrift rule in among the binary rules. Consequently the register marking of downstep avoids the problem raised above of interspersing binary and n-ary rules.

This use of register raises the question of whether register boundaries correspond to grammatical boundaries. To this end, I note that Mende has an automatic register change (shown in the following examples by x ) between subject NP and VP (5a) and between object NP and VP (5b) (but not between clitic object pronouns and the verb (5c)).

\[
\begin{align*}
\text{(5) a. } & \text{mahi x lo x gilima} \\
\text{mahciji x lo x gilima} & \text{'the chief is thinking'} \\
\text{b. } & \text{ndoi x lo x mbei x yeyama} \\
\text{H H L H L} & \text{'the child is buying rice'} \\
\text{c. } & \text{ngi-bi-gava-ni x lo} \\
\text{ngi-bi-gava-ni x lo} & \text{'I-you-cheat-ed'}
\end{align*}
\]
From the available evidence, Mende appears to change registers automatically across word boundaries but not within where, when two H's are juxtaposed as a result of suffixation, no downstep is observed.

(6) \( \overline{\text{H} \ H} \) OCP \( \overline{\text{H}} \)
\( \text{ngulu} + i \rightarrow \text{nguli-i} \)

The process illustrated in (6) could be handled by proposing that the OCP operate across morpheme boundaries but not word boundaries. This proposal would eliminate the need to state that register shifts occur at word boundaries since this is an automatic consequence of the proposal. This would thus eliminate the need to mark such boundaries through the grammatical insertion of floating low tones.

With reference to register, I also note that Progressive Reduction (see next section) does not apply across register boundaries (7).

(7) \( \overline{\text{H} \ H \ L} \rightarrow \overline{\text{H} \ H \ L} \)
\( \text{maha} \text{ngulu} \rightarrow \text{maha} \text{ngulu} \rightarrow \text{not maha} \text{ngulu} \)

The task of defining register shifts in terms of word boundaries will necessitate a redefinition of what constitutes a Mende word, something which I feel on other grounds has been warranted for some time.

3.3. Tonological tension. In section two, I introduced the concept of tonological tension. Interestingly, the tension reduction requires the conspiratorial action of three distinct rules.

3.3.1. Progressive Reduction. The first choice for the resolution of tension will apply if the rightmost association of a tense segment connects with a tone which has at least one other association (to the right). Here the melody is preserved through the dissolution of the rightmost tense association, as illustrated in (8).

(8) a. \( \overline{\text{L} \ H \ L} \)
\( \text{pele-i pele-i pele-i} 'the path' \)
b. \( \overline{\text{L} \ H \ L} \)
\( \text{nyaha-ngaa nyaha-ngaa nyaha-ngaa} 'women' \)

3.3.2. Low Expulsion. The second choice of tension reduction differs from the first in two ways: (1) it applies to the left-most tone association of a segment, and (2) it applies only if that association involves a low tone. The process involves the dissolution of that low tone association. If, as in example (9a), that low tone is also associated with the preceding syllable, the dissolution of the low tone association results in a simple L H tone pattern in a way not unlike Progressive Reduction.

(9) a. \( \overline{\text{L} \ H \ L} \)
\( \text{pele pele pele} 'a path' \)
But if the tense low tone is preceded by a high note (9b), the dissolution of the low tone results in the loss of the disassociated low note. At first glance, this appears to represent a case where the melody has been lost. However, given the redefinition of the OCP, the sequence of two high notes remains distinct marking a change of registers and with the second high tone showing up as a "downstepped" phonetically.

3.3.3. Utterance finally. The two reductions described above will reduce all tonologically "complex" segments except for falling tones occurring at the end of an utterance. In this situation neither the high nor the low tone are expelled and the segment remains a falling tone. However my data suggests that there is an optional rule which can apply in this situation to reduce the tension by adding another tone-bearing segment to the string. This segment, being tonally unmarked, acquires a low tone association. Then progressive reduction would apply to reduce the tension of the falling tone. Consequently the net result of this process, shown in (10), is a long falling tone.

(10)  
\[
\begin{array}{l}
\text{L HL V Addition} \\
\text{nyaha}\hat{\#} \\
\text{L HL L Association} \\
\text{nyaha}\hat{\#} \\
\text{L HL Prog. Red.} \\
\text{nyaha}\hat{\#} \\
\text{L HL} \\
\text{nyaha}\hat{\#} \\
\text{L HL} \\
\text{nyaha}\hat{\#} \\
\text{L HL} \\
\end{array}
\]

3.3.4. Underlying Representations. It should be borne in mind that because of the reduction operations described above, underlying tonally tense segments will only appear where the possibility for alternative reductions are possible, that is, at morpheme boundaries. Furthermore, given the priority of the progressive reductions, these alternations are only possible morpheme finally. Morpheme internal and morpheme initial tone bearing segments will therefore lack tonal alternations and can only appear lexically as H or L.

3.4. Tone Extension. In Southwestern Mande, when tone copying rules are characterized as the extension of a line of association to the next tone bearing unit, the result will be a tonologically tense tone as in the case of the Mende compound rule (11).

(11)  
\[
\begin{array}{l}
\text{Mende Compound Tone Extension} \\
\left[ \begin{array}{c}
<\text{H}> \\
...\text{CV} \\
\text{nom}
\end{array} \right] \\
\text{comp} \\
\rightarrow \\
\left[ \begin{array}{c}
<\text{H}> \\
...\text{CV} \\
\text{nom}
\end{array} \right] \\
\text{comp}
\end{array}
\]

These derived tense tones are also reduced by the above-mentioned tone reduction rules as shown in (12).

(12) a.  
\[
\begin{array}{l}
\text{bi-b\&\varepsilon\#-i} \\
\text{bi-b\&\varepsilon\#-i} \\
\text{bi-b\&\varepsilon\#-i} \\
\end{array}
\]

'your trousers'
b. \[H \quad L \quad H \quad \Rightarrow \quad H \quad L \quad H \quad \Rightarrow \quad H \quad H \]
   \[
   \begin{array}{c}
   | \\
   bi-we-i \quad bi-we-i \quad bi-we-i
   \end{array}
   \]
   'your rice (def)'

In Bandi, which is variously considered a dialect of Mende (Welmers) or a closely related language (Dwyer), we find in addition to two high tone extension rules a low tone extension rule. When seen as a rule which extends the domain of a L one (half) syllable to the right, the observed results of the Low Tone Extension rule can be obtained by using these same reduction rules (8), (9), and (10) as shown in (13).

\[
(13) \quad a. \quad \begin{array}{c}
\text{CVCVCV} \quad L \quad H \quad \Rightarrow \quad L \quad H \quad \Rightarrow \quad L \quad H
\end{array}
\]

\[
(13) \quad b. \quad \begin{array}{c}
\text{CVCVCV} \quad L \quad H \quad L \quad \Rightarrow \quad L \quad H \quad L \quad \Rightarrow \quad L \quad H \quad L
\end{array}
\]

\[
(13) \quad c. \quad \begin{array}{c}
\text{CVCVCV} \quad CVCVCV \quad CVCVCV \quad CVCVCV
\end{array}
\]

Case (13a) conforms to Progressive Reduction while (13b) and (13c) show the effects of Low Expulsion, which returns the string to its initial configuration.

4. Conclusion

This paper has explored the possibilities inherent in an analysis of Mende tone using an Autosegmental version which proposes (1) that Melodic-Tonal associations are fully specified and (2) that the Obligatory Contour Principle is in force. This formalization, when coupled with a register view of downdrift and downstep, permitted the identification of register changes thus eliminating the need for special symbols for marking either register or downstep. The paper then went on to point out that register changes also occur at word boundaries although at this point some details are still unclear. The active use of the OCP within word boundaries but not across them permits the prediction of word boundary register shifts without having to resort to floating tones or the like.

Principle (1) above also enabled the establishment of a concept of phonological tension (occurring whenever two lines of association were drawn to the same segment). Phonological tension then served as the principle which lay behind a Mende tone rule conspiracy; that is the different rules conspired, each in their own way, to resolve (reduce) the tension of both lexically tense segments as well as those which arise through tone extending rules.

[The revised version of this paper has profited from comments made by Larry Hyman and John Goldsmith at the 15th Conference on African Linguistics where this paper was read.]
SWAHILI GESTURES:
COMMENTS (vielezi) AND EXCLAMATIONS (vilingizi)

Carol M. Eastman, University of Washington
Sh. Yahya Ali Omar, University of London (SOAS)

While putting together an instructional video program entitled "Swahili Gestures",¹ we found we were variously naming some gestures or using whole sentences to represent others. As a gesture was to be done we would, for example, make reference to it by the verbal expression used to accompany it, e.g. ng'o ng'o, or use a verb referring to the gesture's behavior, e.g. kuramba kishogo 'to lick the back of one's neck', or use a phrase indicating what the gesture means, e.g. hichi changu 'this is mine'. In this paper we present some of the specific gestures independent of speech, requiring accompanying speech, or substituting for speech. We also describe the function of these pragmatic context-sensitive "Parts of Communication" making some attempt to relate this gestural aspect of Swahili to the study of language use in context. We assert that whereas exclamations (vilingizi) are verbal/audible gestures, vielezi as comments (lit. explanations) linked to gesture are gestural/visible speech.

Forms labeled kl. (kielezi, vielezi pl.) in the Kamusi ya KiSwahili Sanifu cover many usual language functions including adverb, exclamation, and preposition. None are listed in the dictionary as requiring gestures when they are used even though we "know" we should, e.g. wiggle the hands with the adverbial aihasili to illustrate the idea of many things mixed together. Where certain vielezi may be accompanied by gestures to illustrate the comment they make, others go with gestures in such a way that the two form a unitary expression meaningless apart.

Here we also focus on some culturally specific gestures used by Swahili speakers as an extension of the class of exclamation (vilingizi) to the nonverbal realm. In this category fall a number of gestures that may be referred to by verbs. That is, the activity they represent may be referred to in conversation though the gesture itself is nonverbal. One such verb "meaning" gesture is ku-suta. To -suta, one points the middle finger of the right hand down repeatedly over the head of another person accusing the other of having done the thing under discussion. Usually women do this to each other, e.g. if two women are talking about a third who had an affair, the first may -suta the second meaning, "Who are you to talk about affairs?" much in the sense of English, "It takes one to know one."

In Swahili the verb phrase -rampa kishogo 'to lick the back of one's neck' like -suta refers to a gesture but is never said in conjunction with it. The phrase refers to the act of sticking out your tongue as a sign of contempt when someone's back is turned from you to nonverbally tell one person you don't like another.

There is a gesture said to be used only by women who hold up the right thumb

¹"Swahili Gestures" as a video program is available from the Instructional Media Center, University of Washington, Seattle, WA 98195 (attn. Jack Armstrong). It is produced by Dr. Karen Morell as part of the University of Washington's African Encounter film series.
and index finger to form a circle while keeping the mouth open to match. This
is used, for example, when a woman finds that another has done something she too
has done, and both are embarrassed about it. The gesture is a comment of shared
complicity. The gesturer might say something like ela mena wewe! 'I say—me
AND you'.

In general it seems that the less associated with a verbal form a verbal indepen­
dent gesture is, the more exclamatory it is (opposed to being explanatory, com­
mentative) in function. Thus, we assert that emblematic gestures used in a
Swahili-speaking community which are not referred to by verbal forms have an ex­
clamatory function while verbally independent gestures that are associated with
speech, e.g. -suta, -ramba kishogo, serve as comments within the context of
discourse. Both gestural exclamations and comments are distinct from strictly
verbal exclamations and strictly verbal comments, i.e. other "words" categorized as
vielezi. The three categories of communication (excluding non-gestural
vielezi) are interpretable only in a discourse context functioning either to
 evoke an expected behavior or comment on one that has just taken place. That is,
gestural exclamations and comments and verbal exclamations are pragmatic aspects
of communication.

In addition to such gestural exclamations, gestural comments, and spoken ex­
clamations, there is a fourth category of communicational forms in Swahili that
are pragmatic in nature. These require both the auditory and visual communica­
tional channels working together in order to be interpreted. They involve the
union of the spoken with the seen, the heard with the performed. Where many spo­
ken exclamations (viingizi) are evoked by the discourse context, and all in­
tend to evoke behavior in context, all gestural/speech units are evoked by con­
text and are a reaction to it. They are emphatic responses to behavior with em­
phasis indicated by the combination of the two communication channels.

Gestures without necessary verbal accompaniment such as -suta, -ramba
kishogo, and the thumb plus index finger circle in front of the mouth, are
gearied to the visual channel only. We now turn to the verbal dependent gestures
that emerged during our gesture video-taping project. All may be seen to func­
tion as vielezi, i.e. to express views on the cultural behavior taking place
in the discourse context, and are evaluative, i.e. used when either speech or
gesture alone would be inappropriate. The verbal dependent gestures starting
with the most arbitrary language and gesture combinations as the most opaque to
the most transparent (visually and linguistically rather obvious) include:

salala - performed by placing both hands palm down on top of your head while ut­
tering the syllables as a commentative response to a question such as itakuwaje
mtu mmoja anywe pipa zima la maji? 'How would it be if a person were to drink a
whole cask of water?' The comment explains that this is "impossible", and one
should not even say such a thing.

tongo - said while placing the extended forefinger of one hand below the center
of the eye (left hand/left eye/and vice versa) and tugging the skin downward ex­
pressing the idea that what you, the gesturer/speaker, have just reported is ab­
solutely accurate. The eye-pull has the effect of showing there's no tongo
'sleep' in your eye to have interfered with the accuracy of your perception. If
you were inaccurate, perhaps you deserve tongo to the extent you'd be blind.
Unlike the European eye-pull (gesture-only form), tongo has an alertness or
awareness after the fact dimension, emphasizing the accuracy of a previous comment made in the discourse.

**hng'?ng'** - involves a breathy syllabic velar nasal followed by a glottal catch and another syllabic velar nasal. Clearly aberrant in syllable structure and usual sound combination for Swahili, the vocalization itself is constrained by the gesture that accompanies it. Both hands push the nose inward on each side starting from the top of the nose and moving down while the lips are closed, and the sound is emitted with a high pitch. When the hands reach the tip of the nose they release the nose, the lips are opened, and the vocalizing ceases in an audible release. This spoken gesture indicates to others that the gesturer detects a bad smell. Without the vocalization, this gesture might be thought of as mime. The specificity of the vocalization, however, lends the gesture combined with it enough of the requisite abstraction and convention to analyze it as a symbolic form. The hands on the nose might mean "smell" imitatively but the value "bad" assigned to the verbal/gestural unit clearly involves an abstraction from and understanding of the syllables emitted via the hands moving down the nose while the lips are closed and then released.

**us (ush)** - like **hng'?ng'**, is tied to and constrained by its gestural aspect. It requires that the index finger be placed to the parted lips at the time of vocalization. At first glance, this, too, might appear more imitative than symbolic. But, **us** as a verbal/gestural unit means more than the obvious exclamatory warning "shut up!" conveyed by **us!** alone without the associated gesture. It is used after certain behavior has led you to believe someone is about to make a scene, create noise, or start to cry in earnest. When used by an adult to a child it means, "I sense something about to start." It is a comment on behavior that has led the gesturer to expect that a scene is imminent.

**fu (bu)** - like **salaala** and **us** is an exclamation when used without an accompanying gesture and a comment on behavior when matched with its specific gestural form. It stands midway between **us** and **salaala** in terms of syllable structure being appropriately CV but somewhat aberrant as a single "word" in being monosyllabic. **fu** as a comment is uttered while a person wipes the right hand (palm inward) across the mouth from left to right. With this gesture **fu** is uttered with a high pitch (falsetto). It is used commonly after the verbal expression of an unfortunate deed, e.g.

ankila    fu
'she ate it'  "all of it" (...and, it was supposed to be saved and shared (with you))

**ng'o** - (also **ng'o ng'o**) is an expression, often perceived and intended as rude, accompanied by placing the right thumb under the front teeth (while the other four finders are loosely held in a fist form). While saying **ng'o**, wrist motion may be used to flick the thumb out from behind the teeth. **ng'o**, like **fu** often follows a verbal statement to reinforce what the speaker/gesturer has just said and to add more emphasis to it, e.g.

sikupi    ng'o
'I won't give you any'  "you're not anyone I'd ever give something to, you so and so"
Some verbal dependent gestures take phrasal verbal forms rather than single words or syllables. One of these is hichi changu (or hiki changu) which has the literal meaning, "This is mine." hichi 'this' refers to 'chin' (kidevu), and the gesture accompanying the phrase involves the thumb and index finger of the right hand encircling the chin (index finger on top, thumb below) and then pulling the chin sideways to the right. This spoken gesture asserts that a person, who appears not to, "should have respect for other people's property" [Claessen 1982:24]. One context for using this (pronounced hiichángugu when gesturally accompanied) would be if your child had bad companions, you might perform hichi changu to mean "It's my chin", i.e. if the child screws up, it's going to reflect on you, and what's more, you feel the child will screw up given the company s/he is keeping.

We may look at these Swahili verbal forms that are associated with gesture as speech plus gestural combinations ranging from units of referential vocalization plus imitative gesture to units of arbitrary vocalization plus arbitrary gesture. The forms described here represent five different types: language-like as exemplified by hichi changu, nonverbal or gesture-like as seen with ng'o, units of referential vocalization plus arbitrary gesture such as salala, cases of onomatopoetic vocalization plus imitative gesture as with hng'ng' and us, and one instance of onomatopoetic vocalization combined with an arbitrary gesture fu.

When we look at gesture as communication, and, as here, communication closely linked to language, we begin to blur any language use/communication distinction as well as any distinction between verbal and nonverbal forms of communication. In fact, we begin to see communication as a multi-channeled form of expressive behavior. Here we saw hand and facial movements particularly related to the concomitant production of sound with a communicative function. In [Eastman 1983] I described exclamations in Swahili as a "special" part of speech in which every member is associated with particular cultural behavior and understood only with reference to its context of use. In essence, exclamations in Swahili were seen to be verbal gestures, tied together by the fact that they call for behavior to take place in that context.

Here we've seen another class of items in Swahili that are context sensitive. We've seen further that not all such items are actually "words" in any real speech sense per se. Some are expressions gesturally independent of speech, some may be referred to in speech (with verbs expressing the act they represent). Thus, one may -suta a person, openly "fingering" them commenting that they did a certain deed; one may -ramba kishogo behind someone's back to express contempt for that person. These gestures without verbal accompaniment that are "named" function as vielezl, i.e. comments on and elaborations of certain behavior.

Other gestural forms stand for whole utterances that remain unsaid. They are looked at here, too, as functioning to elaborate the verbal discourse or cultural scene. These gestures (such as holding up the thumb and index finger in shared complicity), we suggest, like those with verbs referring to them, also fall into the kielezi 'comment' (explanation) category.

Described in most detail here were gestures always associated with specific speech forms, i.e. gestures perceived as unique combinations of visual and audi-
tory stimuli functioning to comment on context-based behavior. We feel that the more verbally associated a gesture is, whether united to a word (salala), expressible by one (-suta), imitative of a phrase (hichi changu), or onomato-poetic (ng'o), the more likely it is to function more as a comment than as an exclamation. We are intrigued by the crossover of linguistic function seen here between verbal and visual speech. Particularly noteworthy to us is the observation that linguistic exclamations and imitative gestures function primarily to call for behavior while speech + gesture units and gestures that substitute for utterances are commentative.

REFERENCES


EGYPTIAN ARABIC AND ENGLISH:  
NATIVIZATION PROCESS

Atteya Yussif El-Noory  
University of Illinois-Urbana

"Anyone who speaks English is a citizen of the world." These were the words Bruce Pattison [1956] used in describing the internationally widespread use of English. Kachru [1980a], on the other hand, distinguishes between two different varieties or uses of (non-native) English: institutionalized varieties, e.g. English in India, and performance varieties, e.g. English in Egypt. In both types of use, English has undergone what came to be known as nativization (Kachru, among others), namely, it has been, consciously and/or unconsciously, to some extent adapted to suit the norms of the host languages. In this paper I will attempt to examine some of the nativization processes that speakers of Egyptian Arabic have imposed on borrowings from English as well as a few other related issues.

1. English in Egypt

English came to Egypt with the British occupation (1882-1952). Habry and El-9azzawy [1960] reported that the occupation authorities worked for the replacement of Arabic by English as a medium of instruction in primary and secondary schools. At the same time, the British administration encouraged foreign missionaries to set up preaching schools on entirely foreign lines. After independence (1952), however, the introduction of English was delayed until the seventh grade where it continues to be taught as a compulsory subject all the way through college. Conrad and Fishman [1977] reported that 1,554,809 Egyptian students studied English in secondary schools in 1971, thereby ranking Egypt, at least statistically speaking, first in Africa in this respect in that year.

Certain social, political and economical factors have led to the predominance of English as the major foreign language taught and used in Egypt. The so-called "open-door" economical policy of the government, claims Bowers [1983], has increased the employment projects of the fluent user of English. On the other hand, tourism from countries where English is spoken either as a first or a second language has stressed the need for the use of English within the country.

At least two English language newspapers are currently published in Egypt: the Egyptian Gazette whose 1970 circulation was, according to Smith [1970b], 45,000, and the Egyptian Mail. Besides, both radio and television feature varied programs in English. Such programs are, however, listened to by the highly educated elite as well as by youths belonging to affluent families.

2. Lexical borrowings of English in Egyptian Arabic

English borrowings can be heard in Egypt in radio and television programs and commercials, naturally in lectures, and in everyday conversation. Newspapers, popular magazines and books, though to a lesser extent, all use borrowings from English. There is a cline of foreignness for English loans [Stanlaw 1982]. Some words seem to be regarded as completely Arabic by Egyptians, e.g. /radyu/ 'radio', /sigârah/ 'cigarette', film, etc. Others, especially very recent
borrowings, are short-lived or have limited circulation, e.g. /kündfšan/ 'air condition', /kásit/ 'cassette', /vºfi.dyu/ 'video', etc.

English loanwords may be brought in to fill certain "lexical gaps" as in the case of "video" above, but many other examples show that the filling of lexical gaps is only one reason for the English borrowing process. Kachru [1978a] uses the term "contextual units" to explain the "lexical sets" of borrowed items. The borrowing may depend on register or style; or there may be a mutual dependency between the lexical borrowing from a particular language and a specific register.

English loanwords are used by Egyptians regardless of age, sex or education. The range and complexity of English borrowings will, however, increase as one goes up the educational ladder, culminating in the use of technical and scholarly terms by specialists.

Following is a partial list of English loanwords in Egyptian Arabic (modified from Zughoul [1978]), divided wherever convenient by area.

2.1. Items for modern inventions and machinery. Most of these words are internationally used: /kündfšan/ 'air condition', /kumbýtľ/ 'computer', /lºýuri/ 'lorry', /mºvatľ/ 'motor', /tjalýafon/ 'telephone', etc.

2.2. Food and related items. /hambu'irgär/ 'hamburger', /bíbs(i)/ 'pepsi', /sérár/ 'lorry', /wúr/ 'motor', /išß/Cálfal¥l/ 'telephone', etc.

2.3. Professions. /duºactuºur/ 'doctor (male)', /miºakaníki/ 'mechanic', /sikírºrah/ 'secretary', etc.


2.5. Sports, recreation, and related items. /kásit/ 'cassette', /gºñ/ 'goal', /bãši/ 'pass', /bilmáº/ 'penalty', /tínº/ 'tennis', etc.

2.6. Clothes and related items. /b(i)lºzhah/ 'blouse', /faniliºh/ 'flannel', /shult/ 'shorts', etc.

2.7. Miscellaneous. /awºísbirň/ 'aspirin', /kárt/ 'card', /shºk/ 'check' (n.), /bank/ 'bank', /kirºsmas/ 'Christmas', etc.

3. Processes of Nativization

3.1. Phonology. Most of the borrowings are phonologically treated as if they were Arabic words. First, no consonant clusters exist initially in Egyptian Arabic, and a maximum of two consonants in a cluster is allowed medially and finally. Therefore vowels are inserted between two consonants in most borrowed English words, e.g. /kírºmás/ 'check', /asbirň/ 'aspirin'.

Second, some English sounds are usually altered to meet the fundamental phonetic habits of Egyptian Arabic: /p/ → /b/ as in /asbirň/ and /bíbs(i)/ (this rule is blocked if /p/ is followed by a voiceless stop, e.g. /kæptín/ 'captain'); /v/ → /f/ as in /šifúrlº/ 'Chevrolet' (the sound /v/ will tend to remain unchanged in the speech of the "educated"); /ch/ → /ʃ/ as in /šik/ 'check' (n.); /s/ → /ç/ next to /a/ as in /bãši/ 'pass' and /šalüºn/ 'saloon'; /e/ → /i/ as in /kásit/ and /bíbs(i)/; /v_i/ → /v,1/ as
in \textit{kartūnəh}/ 'carton' and \textit{galūn}/ 'gallon'. This change takes place mostly in the second syllable of a two-syllable word.

3.2. Morphology. All the borrowings are morphologically treated as if they were Arabic words. They are pluralized either by adding the plural inflection of Arabic or by changing the form to fit any of the Arabic paradigms. The following nouns take the plural feminine inflection \textit{\textipa{a}}t/:

\begin{itemize}
  \item /\textit{sinama}/ /\textit{sinam-\textipa{a}}t/ 'cinemas'
  \item /\textit{f-vult}/ /\textit{f-vult-\textipa{a}}t/ 'volts'
\end{itemize}

Examples of plural paradigms that are fitted into are:

<table>
<thead>
<tr>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td></td>
</tr>
<tr>
<td>/masruf/</td>
<td>/masruf\textit{Tf}/</td>
</tr>
<tr>
<td>/kartunah/</td>
<td>/karat\textit{Tn}/</td>
</tr>
<tr>
<td>Model 2</td>
<td></td>
</tr>
<tr>
<td>/xadd/</td>
<td>/xadd\textit{d}/</td>
</tr>
<tr>
<td>/bank/</td>
<td>/bun\textit{u\textipa{k}}/</td>
</tr>
<tr>
<td>/kart/</td>
<td>/kur\textit{u\textipa{t}}/</td>
</tr>
</tbody>
</table>

Arabic does not have a neuter gender. Accordingly, the borrowings are given either masculine or feminine genders depending on how they sound. For example, bank, doctor and film are treated in the singular as masculine whereas words like cinema and lamp \textit{\textipa{lambəh}/ are treated as feminine. English words that apply for both the masculine and feminine such as "doctor" are considered masculine and inflected to indicate femininity as in the following examples:

<table>
<thead>
<tr>
<th>masc.</th>
<th>fem.</th>
</tr>
</thead>
<tbody>
<tr>
<td>/\textit{dva\textipa{k}tu\textipa{vər}/}</td>
<td>/\textit{dva\textipa{k}tu\textipa{vər}\textipa{wəh}/}</td>
</tr>
<tr>
<td>/\textit{sikir\textipa{t}er}/</td>
<td>/\textit{sikir\textipa{t}er-\textipa{a}h}/</td>
</tr>
</tbody>
</table>

The largest percentage of the borrowings are nouns. However, some verbs were either borrowed or derived from these nouns. The verbs are treated as other verbs in Egyptian Arabic and inflected for present and order to fit the three forms of verb, e.g.

<table>
<thead>
<tr>
<th>past(root)</th>
<th>present</th>
<th>order</th>
</tr>
</thead>
<tbody>
<tr>
<td>safir</td>
<td>yisafir</td>
<td>safir</td>
</tr>
<tr>
<td>kansi</td>
<td>yikansi</td>
<td>kansi</td>
</tr>
</tbody>
</table>

Most of the borrowed items have maintained only one of their original meanings when used in Egyptian Arabic, depending on the area in which the word was borrowed. For example, the word \textit{\textipa{fi\textipa{b}sh}/ is only used in photography, the verb \textit{\textipa{du\textipa{b}lər}/ (from "to double") is used to mean 'to repeat a course or a year in school', whereas the lexical items \textit{\textipa{\textipa{f}w(i)}t}/ 'out', \textit{\textipa{gən\textipa{v}i}/ 'goal', \textit{\textipa{bə\textipa{g}ə\textipa{z}i}/
'pass' occur only in meanings related to soccer. Very few words have expanded in meaning, e.g. /yikansil/ 'to cancel' is used to mean both that and "to ignore".

4. Attitudes towards English mixing in Egypt

Two studies by El-Dash [1973] and El-Dash and Tucker [1975] showed that the Classical Arabic speakers, and, in general, the speakers of American English, British English, and Egyptian English were considered more highly educated than the vernacular speakers. However, opinions vary when it comes to the question of code mixing with English, or any other language for that matter. Al-Bakry [1976] bitterly criticizes such "lamentable innovation which some call franco-arab" and wished that they had elected (if they had to) to either use the foreign language all the way or, better still, use their own mother tongue, not excluding the vernaculars in this respect. At the public level, there seems to be a cline of acceptability for code mixing and borrowing in general. For example, the instances of code mixing cited by El-Bakry [1976] which includes /iftah ilwindu piTz/ 'open the window, please' and /?innahardah ilgaww viri nays/ 'today the weather (is) very nice' will certainly call for sarcasm. On the other hand, utterances such as /?ammatt ilfilm fi ilstudyu illi ganb ilbank/ '(I) developed the film at the studio which (is) next to the bank' will go unnoticed.

5. Conclusion

The possibility of institutionalizing English in Egypt is a farfetched one regardless of how highly some Egyptians may think of English. A number of factors distinguish the Egyptian context from, say, the Indian or Nigerian context. First, Egypt enjoys an ethnic and a linguistic homogeneity. Second, Egyptians think highly of Arabic; in fact most of them would think in terms of Arabic, at least at the lexical level, influencing English rather than the opposite. Third, the belief in the richness and flexibility of Arabic, among other reasons, has been behind the enthusiastic attempts for Arabicization. For this purpose, language academies have been established in Cairo, Damascus, and Baghdad. Sharaf [1976] believes that Arabicization has started to pay off, giving illustrations of a considerable number of previously popular (foreign) lexical items, e.g. refereee, team, etc. that have actually given way to their Arabicized counterparts.

Taking into consideration all the attempts at Arabicization as well as the official and public feelings about it, what would the Arabic of the future look like? Ferguson [1959] attempts an answer to this question: "...There is almost full unanimity (in the Arab world) that Arabic of the future will not be a form of colloquial Arabic. It will be a 'modern' slightly streamlined form of Classical Arabic, purified of all regionalism or of excessive foreign vocabulary ... Some believe it will take ten years, others go as high as fifty...." Twenty-five years have already passed since Ferguson first wrote this. Do we already have that modern, purified Arabic? Maybe not. Yet things are surely a lot better now than they were twenty-five years ago.
REFERENCES


QUESTION FORMATION IN KUSAAL

Nora C. England and William A. Ladusaw
University of Iowa

1. Introduction

Kusaal is the language of the Kusasi, who live in the northeast part of Ghana and adjacent Upper Volta. It is a member of the Gur (Voltaic) branch of Niger-Congo, sister to Dagbani and Moore (among others). Our analysis is based on data from Robert Agunga, a native speaker of the Bawku (Ghana) dialect.

In this paper we present a brief sketch of its structure and discuss the formation of Yes/No Questions (hereafter, merely "questions"), a process which seems to fall midway between inflection and cliticization.

Questions are formed from declaratives by the addition of a final Low tone, which is associated with the last word of the sentence. The addition of the tone triggers segmental operations which are in some cases phonologically predictable, while in others dependent on the morphological features of the last word in the sentence.

2. Grammatical Sketch

2.1. Syntax and Morphology. Kusaal is a rigidly SVOX language. Its NPs have prenominal possessor NPs, but postnominal adjectives, relative clauses, and determiners (in that order). There is no agreement marking on the verb. Nouns can be divided into inflectional classes according to which pair of singular and plural suffixes they take. Words which translate into English as adjectives may occur with inflectional suffixes drawn from a subset of the noun suffixes or as true verbs (or as both). The former type of adjective we refer to as N-form adjectives and the latter class as V-form adjectives. In general, number is marked on an NP only on the rightmost inflecting element. Hence Nouns in construction with N-form adjectives and Nouns and N-form adjectives in construction with certain determiners appear in their base forms. Thus, there is no real agreement between nouns and adjectives within the noun phrase. These features of Kusaal are illustrated in (1)-(4):

(1) awín gót té bɔːs: 'Awin is looking at a dog'
(name) look at-Impf dog
(2) ɔ yɛt puː ꓂ vɛ ꓂ Tŋ 'he sees a beautiful woman'
he sees woman beautiful
(3) m yɛt du ꓂ wɔ́ wɔ́ kia 'I see the tall man'
I see man tall-the
(4) wɔ́m ꓂ vɛ ꓂ kɔŋŋa ꓂ mʊyɛŋa 'this beautiful calabash is red'
calabash beautiful this is-red
2.2. Phonology. The inventory of underlying segments is given below.

<table>
<thead>
<tr>
<th>p</th>
<th>t</th>
<th>k</th>
<th>kp</th>
<th>i</th>
<th>u</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>d</td>
<td>g</td>
<td>gb</td>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td>m</td>
<td>n</td>
<td>°</td>
<td>(m?)</td>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td>f</td>
<td>s</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>v</td>
<td>z</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>y</td>
<td>l</td>
<td>r</td>
<td>h</td>
<td></td>
</tr>
</tbody>
</table>

(plus weak harmonizing vowel [i, o, a])

There are a labial, alveolar, velar, and labiovelar series of stops. The parenthesized [m?] occurs only word-finally, but its distribution is so far unpredictable. The units [kp] and [gb] occur only morpheme-initially. There are no word-final occurrences of [v], [z], [w], or [y]. The velar nasal [ŋ] occurs only finally or as an intervocalic geminate. There are seven distinctive vowels: [i], [e], [ɛ], [u], [ɔ], [ɔ], and [a], which may occur as either [±long] or [±nasal]. Additionally there seems to be a weakly harmonizing vowel, realized variously as [i], [o], or [a]. There are three distinct level tones: H ('), M ('), and L ('). Mid tone is comparatively rare.

Final oral stops are devoiced and nasally released. Vowels following nasal consonants are nasalized, though the nasal/oral contrast is maintained before nasal consonants. The glides [w] and [y] are nasalized (almost to the point of nasal consonant articulation preceding nasal vowels). The morphologically sensitive rule in (5) is of relevance here.

(5) Stem Truncation: \[
\begin{align*}
CV & \\
CV: & \\
CV_{I?V} & \rightarrow CV? / \_\_#
\end{align*}
\]

When a verb or adjective occurs in its base form, the contrast between short vowel, long vowel, and V-?-copy V is neutralized to a short vowel, checked by a glottal stop. Noun stems similarly neutralize a long vowel/short vowel contrast. The singular noun suffixes [-og], [-on], and [-ok] are realized as rounded [ŋ], [ŋ], and [ŋ] respectively when attached to stems which end in nonround vowels (cf. wawk in (3)).

3. The Formation of Questions

From this short sketch, let us highlight two points crucial to our analysis of questions. A question is marked by the addition of a low tone to the end of the sentence. The tone is associated as the last tone of the last word of the sentence. Given the syntax of Kusaal, the last word of the sentence can be of any category, as illustrated in (6-7). The last word may end in any vowel or any of the consonants [p], [b], [m], [t], [d], [n], [k], [g], [ŋ], [m?], [f], [s], [ʃ], or [r]. Questions corresponding to the sentences in (6) are given in (7).

(6) a. awín wá?ád 'Awin is dancing'

(7) a. awín wá?ádá: 'is Awin dancing?'
Question Formation in Kusaal

(6) b. awín yɛt bɔ:ɡ
   'Awin sees a goat'
   (7) b. awín yɛt bɔ:ɡâ:
   'does Awin see a goat?'

c. awín yɛt bɔ kódɛ
   'Awin sees an old goat'
   c. awín yɛt bɔ kódɛ:
   'does Awin see an old goat?'

d. awín yɛt bɔ:ɡwə
   'Awin sees this goat'
   d. awín yɛt bɔ:ɡwə:
   'does Awin see this goat?'

e. ḏ jà yǒ:m bɛ:ɡə
   's/he will sing tomorrow'
   e. ḏ jà yǒ:m bɛ:ɡə:
   'will s/he sing tomorrow?'

Though the formation of questions by the addition of a formative at the margin of a sentence is quite common for verb-final languages, we do not know how common it is for SVO languages. One's initial expectation for Kusaal would be either that the formative marking the question would be invariant or that any variation in its realization would be phonologically predictable. In short, we would expect a full formative or a clitic.

This expectation arises from the same principles which stand behind Zwicky and Pullum's [1982] criteria for distinguishing between cliticization and inflectional affixation. They offer three criteria to which we would add D below:

A. Clitics exhibit a low degree of selection with respect to their hosts, while inflectional affixes exhibit a high degree of selection.
B. Arbitrary gaps in the set of combinations are more characteristic of inflectional affixes than of clitics.
C. Morphological idiosyncrasies are more characteristic of inflectional affixes than of clitic groups.
D. Inflectional affixes are restricted to attaching to a limited range of syntactic categories, though clitics may attach to a wide range of hosts.

Question formation in Kusaal shows a mixed set of properties with respect to these criteria. The formative attached to the last word regardless of its syntactic category, and for every declarative sentence there is a well-formed question. In these respects it is like cliticization. Yet the segmental bearer of the added tone cannot be phonologically predicted. The generalizations must in some cases mention the syntactic category of the host word and its morphological features.

4. Analysis

The low tone which is added to a declarative to form a question must be incorporated into the last word of the sentence. Where the last element in the sentence ends in a long vowel, this tone is simply associated with the final vowel, creating a falling tone, as in (8).

(8) a. fɔ gɔt bɔ:
   'you look at a dog'
   b. fɔ gɔt bɔ:
   'are you looking at a dog?'
One might expect then that the form of a question would follow from independently motivated rules governing the association of tones. This is only partly true.

We assume that only one tone may be associated with a short vowel, but that a long vowel may (but need not) bear two tones. Tone patterns in monosyllabic lexical items support this assumption; there are no stems with moving tone on a short vowel, but there are stems with moving tones on long vowels, such as those in (9):

(9) ūn: 'horn', nītq0 'bird', nāf 'cattle', yo:m 'song'

There are then two possibilities for adding a sentence-final low tone to a word which ends in a short vowel which already bears a tone: the new tone could replace the original tone, or the vowel could lengthen in order to accommodate the association of two tones, the original (lexical) tone and the question-forming tone. The second possibility, that of lengthening the final short vowel, is what occurs in Kusaal, as shown in (10) and (11).

(10) a. fō bōd zā? 'you plant late millet'  b. fō bōd zā! 'do you plant late millet?'
(11) a. bā wōmyā? 'they sing-Per'  b. bā wōmyā: 'have they sung?'

The addition of the low tone blocks the application of stem truncation (5). Example (12) shows that a short vowel which arises by Stem Truncation from a V.?V. stem bears the low tone on the vowel after the glottal. Therefore the form of all vowel-final elements to which the question tone is added seems predictable from the rule in (13).

(12) a. m nā wā? 'I will dance'  b. m nā wā?ā 'will I dance?'
(13) T T T T
    T
  V V:

It is with consonant-final words that the phonological predictability breaks down. In these cases (excepting a class of forms which end in [m]), a vowel is added which bears the tone. The added vowel is one of the low vowels [e], [a], or [o]. Which vowel is added is not completely predictable from the phonology of the word. One must make reference to the morphological features of the word to which the tone and its vowel are incorporated.

To state the patterns of vowel association with consonant-final words, it is necessary first to distinguish verbs (and V-form adjectives) from nouns (and N-form adjectives). Every constituent which is not a noun, verb, or adjective either patterns like a noun or ends in a vowel. It is also necessary to exclude and treat separately three noun classes: the class which forms its plural with the suffix -nam, that which forms its plural with the suffix -b (the "People" class), and mass nouns terminating in [m]. Finally, at least one class
of derived nouns must be excluded: those formed with the diminutive -bil.

Verbs pattern differently from nouns. Verbs in the aorist and future can be consonant final, while other tense/aspect suffixes end in vowels and follow the lengthening rule. Consonant-final verbs in the Aorist may terminate in [d], [t], [m], or [n] and all add [a]. Verbs in the future that end in consonants other than [m] add [e]. Verbs ending in [m] lengthen the [m] as if it were a vowel.

<table>
<thead>
<tr>
<th>Declarative</th>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. ḏūn</td>
<td>ḏūnnaː; (Aorist) ḏūn</td>
</tr>
<tr>
<td>a. ḏūm</td>
<td>ḏūmmə; (Aorist) ḏūm</td>
</tr>
<tr>
<td>a. ḏūgfd</td>
<td>ḏūgfdəː; (Aorist) ḏūg</td>
</tr>
</tbody>
</table>

Excluding the four exceptional classes mentioned for nouns, then, the rules may be stated in terms of the final segment of the word. These rules seem, however, far from natural.

Nouns ending in the suffixes [-b] or [-f] take [a]; those ending in [m] or a coronal take [e], and those ending in velars take either [a] or [o]. The distribution of the question vowel with velars depends partly on whether the velar is in the stem or is a noun-class suffix, and (if a suffix) on the form of the suffix. The noun-class suffixes (singular) [-og], [-on], and [-ok] take [o] as the question vowel, while the suffix [-g] or a stem final [ŋ] or [k] take [a].

The four exceptional noun-classes mentioned above are exception in that the vowel which is added in the question form is not the one expected from the pattern laid out above. Their exceptionality is unified by the morphological properties of the words.

Though all other [l]-terminal words add [e], if the noun ends with the diminutive suffix [-bil], [a] is added. Hence the question form of [nɔłuː] 'little toe' (cf. [nɔbrɛːw] 'toe') is [nɔɓiː]. Similarly for [wɔɗbólica] 'star' (cf. [wɔɗgilə] 'moon').

The noun class whose plural is formed with the suffix [-nam] and the class whose plurals are formed by adding [-p] or [-b] all take [a] in their question forms. The majority of nouns in both classes are words for people, e.g. 'man', 'woman', 'witch', 'father', 'chief', 'midwife'. Some forms appear in both classes. The [nam] class also contains emphatic pronouns and some borrowings, e.g. [lɔːr] 'truck'.

Finally mass nouns generally lengthen their final [m] just as [m]-final verb stems do in the future, but unlike count nouns and aorist verbs. Hence the question form of [zɛːm] 'blood' is [zɛːmː], that of [yɛːzam] 'wisdom' is [yɛːzamː]. Yet even here we find two apparent mass nouns (dɛːzɛm, 'algae' and tɔːzɛm 'rope fibre') whose question forms suffix [e].
The schizophrenic nature of [m] in this process (acting now as a vowel, now as a consonant) is intriguing. We have not yet found any way of predicting which final [m]'s act in which way, nor have we found any independent reason to posit two distinct segments.

5. Conclusion

It is obvious that a fair amount of memorization is required to master this process. Our informant has occasionally been unsure of which vowel to add for some nouns, in exactly the forms that seem to admit of more than one morphological analysis. For instance, he gave both [a] and [o] as question vowels associated with the form mumuŋ 'small calabash', and eventually decided that he preferred the [a] terminal. This uncertainty is expected, since the form mumuŋ could be interpreted as a member of the [-g] suffixing class (in which case it should take [a]) or the [-og] suffixing class (in which case it should take [o]). In most cases, however, there was no uncertainty, since the generalizations are sensitive to information that is readily accessible to the native speaker.

In spite of a certain amount of lexical idiosyncrasy, most of the forms are entirely predictable. Taken bit by bit, each part of the question formation process involves reasonable rules, but as a whole, it is a generalization that cuts across the usual domains for rules: the phonological, the lexical, and the syntactic.

REFERENCE

RIVERS PIDGIN (CREOLE) ENGLISH:
TONE, STRESS, OR PITCH-ACCENT LANGUAGE?

Nicholas Faraclas
University of California, Berkeley

1. Possible Analyses of Pitch Phenomena in Rivers Pidgin English (RPE)

In Rivers Pidgin English (the dialect of West African Pidgin English spoken in the eastern part of the Niger Delta and throughout the southeast of Nigeria) we find such minimal pairs of words as (an acute accent indicates high pitch while a grave accent indicates low pitch):

/gó/ 'go' Á gó máket. [á gó 'máket] 'I went to market'
/gó/ future marker Á gó gó máket. [á gó gó 'máket] 'I will go to market'
/móda/ 'mother (biological)' [mómá]
/móda/ 'school marm' [mómá]

Is the pitch distinction between the two members of each pair cited above best analyzed as tonal (as in Gokana bá 'hand' vs. bà 'eat') stress-related (as in English import vs. impórt) or part of a pitch accent system (as in Igbo aká 'tooth' vs. aká 'maize')? Some of the arguments for and against each analysis are now summarized.

1.1. Arguments for a stress analysis. The argument for a stress analysis of pitch phenomena in RPE is supported by the existence of a phrase-level system of accentuation which is clearly stress-related, the limited scope and primarily grammatical nature of almost all critical pitch distinctions in the language, and the clear correlation of stress placement over English words to the assignment of high pitch over most RPE words derived from English. An apostrophe precedes a syllable which receives phrasal stress-accent:

/jám/ [já:m] 'yam' [ásí máket 'já:m] 'I found the market yams'
/máket/ ['máket] 'market' [ásí jám 'máket] 'I found the yam market'

1.2. Arguments for a tonal analysis. The argument for a tonal analysis of RPE involves the correspondence of pitch patterns over words borrowed into Pidgin from other Nigerian languages with the tonal patterns found over those words in the source languages, the failure of such words to be affected by phrase-level stress-accent, and the fact that fundamental frequency (pitch) level is the only reliable cue for accentuation phonetically in RPE. Another tone-like aspect of the behavior of units bearing high pitch in RPE is the fact that they do not cause reduction (lowering of pitch) of surrounding pitch-bearing units:

Igbo: ákóm 'fever' Pidgin: ákóm shótóp [ákóm jô'tó:p] 'ákóm shut up'
(malaria medicine)
1.3. Arguments for a pitch-accent analysis. The pitch-accent argument revolves around the way in which strings of connected speech are parsed for phrase-level accent assignment, which corresponds rather closely to that found in Japanese [McCawley 1965] and [Williamson, personal communication]. Another accent-like characteristic of the RPE system is that for items derived from English (the great majority of words), only a very restricted number of pitch patterns may occur in relation to the potential number of patterns, given the existence of two distinctive levels of pitch. An equal sign = symbolizes an accent group boundary:

/no=im båj jåm=fò mák[èt] ['nò: im båj 'jå:m fò 'mákèt']

'no, he bought yams in the market'

2. RPE: A Mixed System

The only satisfactory solution to the problem outlined above, that is, the only analysis that can predict pitch patterns over utterances in a unified way involves the interaction of tonal, pitch-accent, and stress units. Tone is assigned to words lexically, words from tonal languages being fully specified (one tone per syllable) and words from English being underspecified in most cases (often one tone per word). Underspecified items would then be assigned additional pitches on the basis of their position within a stress-accent group or phrase as well as in relation to the type of tone assigned to them lexically. The existence of a mixed system of pitch assignment and realization in RPE reflects the mixed origins of the language which include stress languages (English, Portuguese), tonal languages (Igbo, Yoruba, etc.), and pitch-accent languages (Ijọ). At the surface, one is tempted to apply a stress analysis but the actual system is in many ways a reinterpretation of stress in terms of tone and pitch-accent.

2.1. Implications for theories of pidginization/creolization. The RPE data re-affirms the importance of substrate-basilectal speech patterns on the underlying structures and processes in Pidgins and Creoles. It also indicates that speakers of Pidgins and Creoles operate from unified systems rather than from a loosely bound set of parasystems, each corresponding to the system as it exists in one or another of the input languages. In other words, a Pidgin or a Creole may behave at the surface in a way which is very much like the acrolect and perhaps very much unlike the basilect. A more careful analysis usually results in the discovery that the strategies used by speakers in the production and processing of surface forms are strikingly similar to those typical of basilect, rather than acrolect speakers.

2.2. Implications for phonology. Linguists are only now coming to recognize the importance of processes such as pidginization/creolization in the development of all languages. One of the implications of this realization will have to be the recognition of the possibility that mixed systems like the one outlined for RPE above are not restricted to languages generally classified as pidgins or creoles but may in fact be quite widespread and that such mixed systems where they do exist, function as unified systems rather than as sets of parasystems. This may explain the recent successes of the autosegmental model for
the analysis of suprasegmentals, which in effect is supple enough to at least begin to accommodate itself to the analysis of systems which have elements of both stress and tone, tone and accent, etc.

2.3. **Implications for the study of African languages.** There seems to have been a certain reluctance on the part of many Africanists to even consider the existence of any non-tonal pitch phenomena in the languages that they study. It is obvious that such attitudes can do nothing to advance, but much to hinder the scientific description and analysis of African languages. In many parts of Africa intermarriage, bi- and multilingualism, and other forms of interethnic contact and exchange are the rule rather than the exception. The traditional bias in linguistics against pidginization/creolization as a model for language change has had for this reason even more of a negative impact on the study of African languages than it has had in other areas. Africanists must learn to recognize, describe, and analyze mixed systems wherever they occur. Mixed systems can in fact shed much light on some of the most important unresolved questions facing the linguist in Africa. For example, the analysis of RPE pitch patterns as the result of a mixed system has allowed the division of lexical items into distinct classes each of which corresponds to a particular historical stage or source of borrowing of vocabulary items into the language. This in turn has provided critical evidence concerning such problems as the nature and extent of the influence of Sierra Leonian/West Indian Krio on Nigerian Pidgin English during the 19th century.
The aim of the present paper is to propose, for the first time, a comparative study of the form and function of logophoric systems in Chadic languages and to formulate a hypothesis regarding logophoricity in Proto-Chadic.

1. The Two Systems

1.1. The three sets of Mapun. The logophoric system of Mapun consists of the following sets of pronouns. (Tones are not relevant to the problems of this paper and hence are left unmarked; 'b and 'd indicate labial and alveolar glottalized stops respectively):

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 p. m. sg.</td>
<td>wur</td>
<td>'di</td>
<td>gwar</td>
</tr>
<tr>
<td>f. sg.</td>
<td>war</td>
<td>'de</td>
<td>paa</td>
</tr>
<tr>
<td>pl.</td>
<td>mo</td>
<td>'du</td>
<td>nuwa</td>
</tr>
</tbody>
</table>

The deciding factor for the use of the three sets is whether the pronoun of the embedded clause refers to the speaker, i.e. the subject, or to the addressee of the main clause (cf. Stanley [1982]).

1.2. Reference to the subject. In sentential complements following the verb sat 'say', set A indicates disjoint reference and set B indicates coreference with the speaker of the main clause, i.e. the subject of the verb sat. Compare (1a) and (1b).

(1) a. wur/war/mo sat ni wur/war/mo ta dee n jos  
     'he/she/they say Compl he/she/they stop stay Prep Jos  
     'he(1)/she(1)/they(1) said that he(2)/she(2)/they(2) stopped over in Jos'

    b. wur/war/mo sat ni 'di/'de/'du ta dee n jos  
     'he(1)/she(1)/they(1) said that he(1)/she(1)/they(1) stopped over in Jos'

Set B above has also its object counterparts, viz. 'din (m), 'de (f) and 'dun (pl). The object pronouns of the sentential complements behave in the same way as subject pronouns. The system further makes explicit whether or not 3rd p. sg. is a part of the set referred to by the 3rd pl., e.g.

(2) a. wur/war sat ni n nas mo  
     'he(1)/she(1) said that I beat them(2)',  
     i.e. "he/she" is not part of "them"
1.3. Reference to the addressee. If the third person pronouns in the embedded clause refer to the addressee rather than to the speaker of the main clause, then the pronouns must be drawn from set C but never from set B. Compare the following examples with pronouns drawn from set C, followed by an ungrammatical example with pronouns drawn from set B:

(7) a. n sat n war ni paa naa ki n k' es makaranta
1sg say Ben 3f Compl 3f look Compl 1sg finish school
'I told her, look, I have finished school'

b. n sat mo ni nuwa naa ki n k' es makaranta
3pl 3pl
'I told them, look, I have finished school'

c. n sat n wur an ni gwar tan me mbi n an
3sg 1sg Compl 3sg find some thing Ben 1sg
'I told him that he should find something for me'
1.4. The function of sets A and C. If the verb of the main clause is sat, then pronouns of either set A or set C may be used in the subordinate clause.

Pronouns of set C are used when the third person represents the addressee of the reported speech. Pronouns of set A will be used when the third person is not the addressee of the reported speech.

(9) a. n sat n wur taji wur dim n kaano
   Prohib. go
   'I told him(1) that he(2) may not go to Kano'

b. n sat n wur taji gwar dim n kaano
   'I told him(1) that he(1) may not go to Kano'

One can use the set C pronoun in the subordinate clause without having any third person pronoun in the main clause, e.g.,

(10) an ni ko a iri 'dak 'di gwar sin an 'be an mbi cin ni
    lsg Compl any Copula kind work Rel 3sg give lsg consec 1sg Fut do it
    'I said that any kind of work which he gives me, I will do it'

Thus the function of set C is fundamentally different from the function of set B. The pronouns of set B indicate coreferentiality with the subject of the main clause, are syntactically determined, and have an anaphoric function. The pronouns of set C do not indicate coreferentiality and are not anaphoric. They have a deictic function and are not syntactically, but rather pragmatically, determined.

1.5. Origin of the three sets. Set A pronouns are identical to the set of subject or object pronouns of the main clause. Pronouns of set B all have the same initial consonant '/d/' and they differ only in the vowel that follows it. It is possible that the pronouns of the Set B derive from demonstrative pronouns, which also have 'd' as initial consonant.

There is no clear indication about the origin of set C. The word /gwar/ occurs as the independent lexeme meaning 'man'. I do expect that the remaining elements of the set, viz. /paa/ and /nuwa/ also derive from independent words.

2. Comparative Data

2.1. Angas Group. At least two other languages, Sura and Angas, have logophoric pronouns. The pronouns in Angas are cognate with the pronouns of Mapun:

Set B: dyii (m.) 'da (f.) 'du (pl.)
Set C: gwa (m.) pe (f.) nywe (pl.) [Burquest 1973:195]

In Angas the two sets have functions similar to those of sets of Mapun. The data for Sura [Jungraithmayr 1963] indicate only the anaphoric function.

On the basis of the three languages belonging to the Angas group one can certainly postulate that logophoricity as described for Mapun was also a feature of Proto-Angas.

2.2. West Chadic. The only West Chadic language outside of the Angas group for which logophoric pronouns have been observed and described is Pero (cf. Frajzyngier [in press]).

In Pero there are two sets of pronouns, which differ only in the form of the second person markers. In one system these pronouns are ka (m) and ci (f) and in the other, peemu (m) and peeje (f).

The "peemu" set in Pero may, but does not have to, occur in sentential complements of the verb 'say'. The data clearly indicate that the "peemu" set refers to the second participant, i.e. to the addressee of the reported speech. Thus Pero has only the equivalent of set C of Mapun, i.e. in the logophoric context it has only the deictic and not the anaphoric function.

It is important to note that the second element of the logophoric pronouns in Pero also seems to be derived from demonstrative sets, for the form -mu is similar to the demonstrative suffix -mo.

2.3. Other branches of Chadic. In East Chadic a logophoric system has been reported for Kera [Ebert 1979:260 ff.]. The logophoric pronouns, identical with independent pronouns, indicate coreferentiality with the subject of the main clause. They correspond thus to set B of Mapun, i.e. they have an anaphoric function.

3. What Can be Reconstructed for Proto-Chadic?

3.1. Anaphoric set. The anaphoric set has been noted in two branches, West and East Chadic. In West Chadic the anaphoric set was certainly present in Proto-Angas. There are no claims that any other language of this branch has the logophoric system of set B. This set cannot be reconstructed as a retention from Proto-West Chadic. Although a logophoric system has been reported in Kera, such a system is not universal to the East branch. There is therefore no evidence to show that Proto-Chadic had a set B logophoric system. Unless more Chadic languages are shown to have such a system one would have to assume that set B logophoric systems in Angas languages and in Kera result from borrowing of an areal feature from one or more of the non-Chadic languages. In the same general geographical area, one may note studies of logophoricity in the Ubangi languages by Cloarec-Heiss [1969], Tuburi by Hagege [1974], Gokana by Hyman and Comrie [1981], Ewe by Clements [1975], and Tikar by Stanley [1982].

3.2. Deictic set. Since no evidence for the existence of set C has been found outside of the West Chadic branch one cannot postulate the existence of the deictic logophoric pronouns in Proto-Chadic. Within West Chadic, however, the markers of set C are cognate, mainly because of the forms involving /p/ in the
Pero "peemu" set and "paa" in the Angas group. The fact that outside of the Angas group only Pero has the set C mitigates against reconstruction of the set C for Proto-West Chadic. It is possible that this set developed within the West Chadic branch only as a result of borrowing of an areal feature.

[The work on this paper was conducted while I was a Fellow of the Center for Applied Humanities, University of Colorado. The data on Mapun were gathered with the help of an NEH grant, and also with the help of the Council on Research and Creative Work, University of Colorado.]

REFERENCES


0. Introduction

The initial goal of this paper is a brief exposition of Babanki noun tonology, in the form of an autosegmental reanalysis of Hyman [1979a]. There follows a sketch of the immediate ramifications of this analysis, which range from a reconsideration of the well-formedness conditions on autosegmental association to a series of arguments concerning the proper representation of downglide.

1. Tonal Alternations

Babanki, a Western Grassfields Bantu language of Cameroon, exhibits four surface phonetic levels: H, M, L, and X (a lowered low tone). In the ensuing discussion, it will be argued that there are actually six different tones, or tone reflexes, and that these are derivable from only two underlying tones, given an abstract enough underlying form.

Nearly sixty percent of all Babanki nouns are L-H in isolation, where the L tone falls on a prefix and the H on the monosyllabic noun stem. The rest of the nouns, in isolation, are either L-L or L-X. When the L-H nouns appear in different environments, however, they break down into four groups, which will be exemplified by kakam 'crab', kafo 'thing', kefo 'medicine', and koŋki? 'belt'.

The varied environments used in this paper will be the two noun slots in the associative construction. The construction is similar to the postnominal genitive in English, such that the first noun is the head and the second the possessor. In addition, both constructions have a particle separating the two nouns, although in Babanki this particle is segmentally identical to the prefix of the head noun.

When head of the associative construction, all four nouns are followed by a H associative prefix. In this position, kakam remains L-H. The other three are also L-H, but the associative particle and all subsequent H tones are lowered to a M level, which suggests downstep. Following a suggestion by Clements and Ford [1979] that downstep is a "floating" L tone between two "anchored" H tones, it would appear that the underlying representation of the latter three nouns includes a L tone, while the former does not.

When the possessor noun and preceded by a L associative particle, kakam and kofo appear as L-L, while kofo and koŋki? are L-H. Since kakam is argued to be without a L tone above, then some rule must be responsible for "bumping" the H tone. Apparently the latter two are resistant to this rule.

When the possessor noun and preceded by a H associative particle, kofo is H-L, which suggests that some rule has spread the associative H to the prefix. Kofo and kakam surface as H-M, in which the M-level tone is actually downstepped, given a rule of H-spread, the L tone which is displaced by that rule, and the preceding discussion of downstep representation. Finally, koŋki? is
M-H, which is the only true M tone found in this paper, since it is followed by other H tones.

Since the above nouns, when in head position, occasion H associative particles, it is tempting to posit an underlying H tone prefix. Given this hypothesis, a rule which subsequently lowers the initial syllable of a string and rules which spread H and L tones and create M tones, it is possible to reconstruct underlying forms with only H and L tones, although it is then necessary to posit underlying floating tones in some cases. Kakam is H-H, kọfo H- HL, kẹfo H-LH, and kẹŋkwi L-HL. The unusual L tone prefix for the latter noun turns out to be characteristic of all nouns which begin with nasal clusters, and it is also related to the appearance of M tones. Note that there are also classes of nouns which are H-L, L-L, L-LH, and L-H, although they have been ignored for the sake of brevity. X tones will be introduced and explained in a later section.

2. Autosegmental Analysis

The logic of the present analysis thus far parallels Hyman's original account. The introduction of the autosegmental framework, however, makes possible a massive clarification and simplification over Hyman's segmental analysis. Abstractions and redundancies which are inseparable from the previous work, e.g. contour tones as intermediate forms and two sources for M tones, disappear completely. In addition, the number of rules necessary is cut in half. Finally, a transparent rule ordering, not at all obvious before, is discovered.

The first step in any autosegmental analysis is the rule by which tones are associated with their tone-bearing units (TBU's). The rule adopted here differs only minimally from the one in Goldsmith [1976].

Association rule: Associate each TBU with a single tone, one-to-one, beginning with the leftmost TBU:

\[
\begin{align*}
\text{kọfo} & \quad \text{tọ-kọwọtọ} \\
\text{H- HL} & \quad \text{H- L}
\end{align*}
\]

Note that the level of association is the morpheme, i.e. that prefixes are associated separate from the roots.

Every autosegmental analysis incorporates well-formedness conditions (WFC's) to further regulate underlying association. Some are uncontroversial, others essential.

WFC 1: All TBU's must be associated with some tone:

\[
\begin{align*}
\text{tọ-kọwọtọ} \\
\text{H- L}
\end{align*}
\]

WFC 2: Association lines do not cross.

Since there are no contour tones in Babanki, save for those created by vowel elision, it is unnecessary to adopt a third (language-specific?) WFC.
WFC 3: All tones must be associated with some TBU:

\[ \text{\textbf{*ka-f0}} \]

\[ \text{H \, HL} \]

The rules which mediate between the underlying associations and the surface forms are of two distinct types. The first type, which includes H-bump and L-bump, alter associations. H-bump "floats" an anchored L tone when it is preceded by an anchored H tone and followed by any tone (whether floating or anchored); L-bump "floats" a bracket-final anchored H tone when preceded by an anchored L tone. The second type, which includes M-formation, X-formation, and initial TBU lowering, changes the value of anchored tones. The two types can be distinguished in many ways: autosegmental rules, the first type, are ordered first and sensitive to floating tones, tone-depressors, and tones attached to segmentally-null morphemes; segmental rules are blind to information on the autosegmental tier and only refer to adjacent information.

3. Representation of Downglide

Given the breakdown of the phonological rules into autosegmental and segmental, it is possible that downglide may be stated at either of the two levels, or, alternatively, at neither. As it turns out, attempts have been made to claim all three positions.

Some investigators of African tones (cf. Stewart [1971]) have attempted to group downstep and downglide together, because they both lower tones. As downstep (in this paper) is a L tone floating between two anchored H tones, one might suppose that downglide is an anchored L tone followed (or preceded) by a floating L tone. There are two reasons why this is incorrect: it necessitates a redundant LL melody, which contradicts the Obligatory Contour Principle [Leben 1976], and it leads to contradictions in the H-bump rule which lead into a paradox.

Alternatively, downglide could be a result of segmental rules. In fact, X-tones, hitherto ignored, are phonetically identical, even if X-tones are restricted to very few positions outside of string-final position, where they often become indistinguishable from downglide. If the X-formation rule is rewritten from lowering any anchored, bracket-final L tone preceded by an anchored L tone to any anchored, bracket-final L tone, the two tones can be collapsed, although it overgenerates madly. Another alternative is to reorder the rule before H-spread, which bleeds it. However, this step introduces what is at best an opaque rule-ordering, if not an outright violation of levels, and in addition, fails to provide the correct results.

Therefore, given the failure to provide a representation of downglide at either the autosegmental or segmental levels of phonological representation, it is concluded that downglide is instead to be represented as a final anchored L tone at the phonological level, which is subsequently lowered at the phonetic level.
REFERENCES


1. Summary of the Hypothesis

The familiar structure of the Bantu verb takes the form

Subject Tense Object Radical Extensions Final
Marker Marker Marker Marker Vowel

In many Bantu languages, there are one or two past tense markers of the shape -a-. When there are two, they differ in the degree of remoteness of the past time referred to, and they also differ in their tonal properties. In this latter respect they may differ, in the first place, in their own inherent tonal marking, one being High, the other Low; in addition, frequently it is found that one (usually the Far Past) places a High tone on the next mora or else can be analyzed as having descended from an earlier period in which that behavior could be found.

Thus at one stage, -a- was a "post-High" morpheme in the sense that it placed a High on the following mora. It seems likely that this unusual behavior was linked to its unusual phonological shape, being of the form -V-, not -CV-. At an even earlier stage, we may hypothesize that the tone shift was due to syllable merger with consequent simplification of the resultant contour tone syllable:

(1) CV - a - CV
    \__________________________\ \______________________
    T H (T) \__________________________\ \______________________
    C G a - C V \__________________________\ \______________________
    T H (T) \__________________________\ \______________________

2. General Comments and Terminology

Correspondences to early Bantu High tone will be marked with either a High tone or an asterisk (*) on a vowel, since it is often no longer easily identifiable as a High tone synchronically. In a number of cases, such as Tonga or Shi, early Bantu High tone most clearly corresponds to Low tone.

I will use the term "Meeussen's Rule" to refer to the rule lowering High after High [Goldsmith 1983, 1984], as in (2), and the term "plateauing", after Hazel Carter, to refer to the raising of Lows between Highs, as in (3). In some languages this is restricted to a single Low found between neighboring Highs; in others, it applies to any sequence of Highs.

(2) Meeussen's Rule: H H \rightarrow H L

(3) Plateau: Bounded: H L H \rightarrow H H H
    Unbounded: H L H \rightarrow H H H
If there is a post-High tense marker, in the simplest case we will find the following effects:

1. No contrast between the tone patterns of High and Low toned radicals when no object markers are present, with the neutralization in favor of the High toned pattern.

2. If there are Low toned object markers, the contrast between High and Low toned radicals will reappear after these object markers; furthermore, what would otherwise be Low toned object markers behave High toned.

3. **Luganda**

   In the Far Past in Luganda, the tense marker -a- has the "post-High" behavior noted above. This High placement is followed by Meeussen's Rule and then by a rule of Unbounded Plateauning. In the example below, we see a High tone on the first syllable of the radical for both Low tone verbs and High tone verbs in this tense; the tone patterns of the two kinds of verbs are neutralized in favor of the High tone verb pattern (examples from Hyman and Comrie [n.d.]):

   (4) Low tone verb \( \text{'he economized'} \) | High tone verb \( \text{'he looked after'} \)
       \[ \text{y a b\text{\'a}l\text{\'i}l\text{\'i}l a} \quad \text{y a l\text{\'a}b\text{\'i}l\text{\'i}l a} \]
       \( \text{H} \quad \text{H} \)

   This tense actually places a High tone in addition on the extensions, associating with the 2nd mora of the stem. In the examples above, it has been suppressed, since it is lowered by Meeussen's rule. It is included in the examples below, taken from Stevick [1969]:

   (5) High        Low
       \[ \text{y a l\text{\'e}e\text{\'e}t a} \quad \text{y a g\text{\'u} l\text{\'e}e\text{\'e}t a} \]
       \( \text{H} \}(\text{H}) \quad \text{H} \}(\text{H}) \)

       \[ \text{y a k\text{\'w\text{\'e}k a} \quad \text{y a g\text{\'u} k\text{\text{\'w\text{\'e}k a} \quad (Plateauning here)\]
       \( \text{H} \}{(\text{H}) \quad \text{H} \}{(\text{H}) \)

   NB: Object marker as noted, and /gu/ is underlying High; /a/ places a High on what follows; the final H is a past tense tonal suffix.

4. **Tonga**

   It is argued in Goldsmith [1983] that the anomalous tonal behavior of the Recent Past in Tonga should be analyzed as a shifting of accent induced by the Tense Marker -a- in the strong form of the Recent Past when it is accented. This is schematized in (5), where the -a- is the tense marker in question:

   (6) * * * * a v v + a v v
The accent, symbolized as "*", is itself realized as Low in tone generally in Tonga. Vowels between accents are High. This is illustrated in (7), using data from H. Carter:

(7) example of tone realization from -!a- tense:

```
  ba !a mʊ bon a
  L   H  L
```

The unusual behavior of the Recent Past (focus on verb form) is illustrated in (8):

(8) Recent Past (focus)

```
  *   *   *
  nd(i) a silik a > nd(i) a silik a > [nd a silik a]
  L     L
```

But if the verb stem is unaccented, then -a- is not followed by a High tone. That is, it can no longer be analyzed as simply placing a High on what follows, though this was undoubtedly the original condition (according to our hypothesis):

(9) 

```
  *   *
  nd(i) a tobēl a [nd a tobēl a]
  L
```

5. **Ruri**

The material discussed here on Ruri (CiRuri, closely related to Jita and Kwaya) comes from Massamba [1982, 1983]. An underlying High is realized on the following syllable except when it is on the penultimate syllable of the word; in that case, it is realized on its own syllable. This is schematized in (10). A phrase-final High tone is realized phonetically as a Rise-fall pattern spread over the last two syllables; a High on other word-final syllables shifts to the first syllable of the following word.

(10) 

```
  *   *
  -a- CV → a - CV
```

For example:

```
  o ku sumĩk a 'to tie'
  o ku Birim a 'to run'
```

In the Far Past, the Final Vowel is assigned a High tone, and both High and Low stems act like High tone stems.
We cannot test further behavior in this tense with object markers because all
object markers are underlyingly High toned and remove a High tone (by Meeussen's
Rule) from an immediately following verb stem.

6. **Digo**

The data and much of the analysis assumed here comes from Kisseberth [1983],
although I will make a crucial assumption that is significantly at odds with
Kisseberth's proposal, as I will note below.

In Digo, a single High tone (indicated here by *) always migrates to the Fi­
nal Vowel (I will call this the End Run Rule). In addition, Digo manifests
Meeussen's Rule and Unbounded Plateauing (the latter blocked by depressor conso­
nants).

What if there are two High tones? There are two possible analyses: 1.
that of Kisseberth [1983], where the rightmost High shifts to the end and the
leftmost shifts rightward to the object marker or radical; 2. that the leftmost
High undergoes End Run (if Final Vowel has no High tone).

If the latter course is chosen—and this is still open to dispute, although
a good case can be made for it—then the -a- Past is another relevant case for
our hypothesis. The -a- places a High tone on the stem that follows; the Fi­
nal Vowel carries a High tone, much as in Rundi.

7. **Rundi**

In Rundi (Ki Rundi), the recent past is marked by a Low tone -a- tense
marker and the Far Past by a High tone -a- tense marker. The Far Past marker
is followed by a -ra- Focus Marker (when focus is on the verb or the entire
verb phrase) and then by object marker(s) and the verb stem. The verb stem is
always on a High tone regardless of the lexical tone of the stem in the Far
Past, e.g.:

```
(13) ba-å-ra-söm-a
     H   H
     placed there by the tense
     they-past-Focus-read-Final Vowel
     'they read (far past)'
```
This, too, would look like a post-High placed by the Tense Marker except for the -ra- Focus marker that is in the way. However, the Focus marker is itself a recent innovation shared with Kinyarwanda but no other languages (a reanalysis of what was formerly the present tense marker, presumably). Thus the post-High behavior of the tense marker was reanalyzed as a morphological rule (rather than being triggered by linear position following the -a- Tense Marker) when the Focus Marker arose as a slot in between the Tense Marker and the Stem or Object Markers.

8. Conclusions

I have presented several sets of facts from Bantu tone systems which have been studied in some depth over the last few years. These detailed studies permit us to isolate the more puzzling synchronic characteristics, such as those we have looked at here. I hope that these observations may call attention to the rich prospects inherent in these materials, so that a thorough comparative study of this and other tenses can be undertaken, a study which will provide us with a deeper understanding of the evolution of tonal systems and, equally, with a deeper understanding of the nature of morphemes that place tone and accent on the following syllable ("post-accenting", "post-High" morphemes).

REFERENCES


1. Purpose of the Study and its Theoretical Framework

The impetus for a detailed and deeper analysis of cleft constructions came from the revolution in linguistics initiated by Chomsky [1957]. Since then a considerable number of studies on various aspects of cleft constructions have been done, e.g. Akmajian [1970], Bach [1970], Schachter [1973], Hawkins and Hyman [1974], Givon [1976], Simon [1982].

The purpose of this study is the analysis of the structure of cleft constructions in present day Ethiosemitic (ES). The approach is generative-transformational. The illustrative examples will be drawn from Amharic (hereafter Amh), primarily; citations from other ES languages may also be used for illustrative purposes of when the particular language shows some divergence from Amharic.

2. Description of Cleft Clauses

Cleft and relative clauses have very similar structure. In both, the absence of a major constituent (except for the main verb) is a characteristic of their surface structure. Both clauses are marked by a particle which we designate as rm (= relative marker); rm is prefixed to the main verb of the clause if there is no overt auxiliary. If there is an overt auxiliary rm is prefixed to it. In (1) and (2) below, the respective clauses are shown within labeled brackets. The abbreviations df and dm stand for definite article and direct object marker, respectively.

Cleft:

(1) a. \[\text{NP}_S \quad \text{Almaz} \quad \phi, \text{v}^\phi - \text{sd}^d \text{db} - \text{oc}^c \text{c} \text{c} (-w)] \quad [\text{VP} \quad \text{tamari-occ-u-n} \quad \text{naw}] \\
\quad \text{Almaz} \quad \text{rm-insulted-she(-him)} \quad \text{students-pl-df-dm is-he} \\
\quad 'it is the students that Almaz insulted'

b. \[\text{NP}_S \quad \text{Almaz} \quad \phi, \text{v}^\phi - \text{sd}^d \text{db} \text{y}^\phi - \text{n}^\phi \text{bbar}(-\text{aw})] \quad [\text{VP} \quad \text{tamari-occ-u-n} \quad \text{naw}] \\
\quad \text{Almaz} \quad \text{he-insulting-Ehem} \quad \text{rm-was(-him)} \quad \text{students-pl-df-dm is-he} \\
\quad 'it was the students that Almaz was insulting'

Relative:

(2) a. \[\text{NP}_S \quad \text{Almaz} \quad \phi, \text{v}^\phi - \text{sd}^d \text{db} - \text{oc}^c \text{c} \text{c} \text{c} \text{c} \text{c} \text{w}] \quad \text{tamari-occ} \text{c}_1 \text{c} \\
\quad \text{Almaz} \quad \text{rm-insulted-she-them students}_{\text{i}}_{\text{1}} \\
\quad '(the) students whom Almaz insulted'

b. \[\text{NP}_S \quad \text{Almaz} \quad \phi, \text{v}^\phi - \text{sd}^d \text{db} - \text{oc}^c \text{c} \text{c} \text{c} \text{c} \text{w}, \text{v}^\phi - \text{n}^\phi \text{bbbar}(-\text{u}-\text{t})] \quad \text{tamari-occ} \\
\quad \text{Almaz} \quad \text{she-insulting-them rm-was(-they-df) student-pl} \\
\quad '(the) students whom Almaz was insulting'
The structural similarities between the two clause types is clearly evident in the examples cited. Nonetheless, there are already differences in the grammar of the two clause types. For example, in (1) ṭāmāri-očč is not in the subject NP; in (2) it is the head of NP. In the section on derivation below an explanation will be given which correctly accounts for the differences.

In ES, the basic word order is

(3) S C V

If V is transitive, C is an O; if V is the copula, i.e. the equivalent of be, then C stands for its complement ranging over NP, AP, PP, and other major grammatical functions.

(4) Yonas sah-n sābbār-y plate broke-he

(5) Yonas ṭāmāri nāw student is-he

(6) Yonas ḏ-bet-nāw at-home is-he

In constructions where the main verb is doubly transitive and VP includes PP, word order is normally as in (7), as sentence (8) shows:

(7) S IO DO PP V

(8) Almaz la-ṭāmāri-očč mig-b bā-mākīna lak-očč to-student-pl provisions by-car sent-she

'Almaz sent provisions to students by car'

In matrix/simplex (hereinafter matlex) sentences in which the equivalent of the verb be is main verb, a right dislocation rule, which moves the subject to the end of the clause, may be operative. The triggering factor is TOPIC(alization) when it is applied to (one or more combination of the major constituents of) the subject clause or to the complement. Schematically, the dislocation rule works as follows:

(9) \[ \text{TOPIC} \]

\[
\begin{array}{ccc}
S & C & V \\
1 & 2 & 3 \\
\emptyset & 2 & 3 & 1 \\
\end{array}
\]

Application of rule (9) to sentences (5) and (6), results, respectively, in

(10) ṭāmāri nāw Yonas student is-he Yonas

(11) ḏ-bet nāw Yonas at house is-he Yonas

'Yonas is a STUDENT'

'Yonas is AT HOME'
Cleft constructions are basically sentences in which the matlex verb is the copula. The subject is a headless clause. The targeting of the complement by TOPIC triggers the operation of rule (9). Non-topicalized constituents, except for the complement, are moved to the end of the sentence.

(12) a. \[ NP [\text{Yonas yS-ayyS(-w)}] [VP Mary-+n nSw] \\
\hspace{1cm} \text{rm-saw(-he) (-him)} -dm is-he \]

b. Mary-+n nSw Yonas yS-ayyS(-w) [from topicalizing the complement]

c. Yonas Mary-+n nSw ya-ayyS(-w) 'it is MARY that Yonas saw' [from topicalizing the subject of cleft clause]

As is indicated in (12), the dislocation rule (9) has various strategies for moving (one or more constituents of) the subject clause. To incorporate this fact, rule (9) should perhaps be reformulated as in (13) below when it is to be applied to cleft constructions.

(13) \[ NP [X V_2] C V_1 \]

a. 1 0 3 4 2

b. 0 0 3 4 1 2

where \( V_1 \) is the copula, \( V_2 \) is the main verb of the cleft clause (hereinafter cleft verb) and \( X \) ranges over the other major constituents of the cleft clause, either singly or in combination. Thus, (12a) is a result of TOPIC ranging over \( X \); (12b) is a result of TOPIC ranging over the C, the focus constituent. Thus in ES the focus and the topic constituents need not be the same. Whichever strategy is used, the cleft verb is moved and placed at the end of the resulting structure.

3. Control of Concord in Cleft Constructions

In all other constructions in which the verb is finite, the subject controls agreement on its clause-mate verb. In cleft constructions, the situation is a little more complex. The subject of a cleft sentence is a headless, non-adjectival clause (see Palmer [1962], Abraham and Hailu [1966], Gragg [1972], and Hailu [1972] for other views of the structure of cleft clause in ES). As a result we may find the cleft clause controlling agreement on the copula. In such cases, irrespective of the person, gender, and number of the major constituents of the cleft sentence, and provided the missing constituent is not the subject of the cleft verb, the copula may be marked for agreement with 3msg (in Tny), and has to be so marked in other ES languages, e.g. Kistanińya (Kist) and Amh.

(14) Tny: \( n+hnə \emptyset z+i-r+i? na (-yo,-ya) n+i-Abrəhāt iyyu we \text{rm-saw-we} (-him,-her) dm- is-he \)

(15) Amh: \( +nəna \emptyset yS-ayyS-n(-Sw,-a+) Abrəhāt-+n nSw we \text{rm-saw-we} (-him,-her) -dm is-he \)
(16) Kist: Ḫānā yā-aṣṣā-nū (-inn, -inna) yā-Abraḥāt-in
we rm-saw-we (-him, -her) dm -is-he

All three sentences mean 'it is Abrahat that we saw'. If, however, the comple­
ment is nominative (Nom), the missing constituent in the cleft clause is the
subject of the cleft clause. In all such constructions there is person, gen­
der, and number agreement among the cleft verb, the complement and the copula.

(17) Amh: Ḫū Almaz-in yā-ayya-n(-at, -*bw) Ḫānā nūn (*nūw)
Almaz-dm rm-saw-we(-her, -*him) we are-we (*is-he)

(18) Kist: Ḫū yā-Almaz yā-aṭṭū-nū (-inna, -*nn) Ḫānā nūn (*-n)
dm-Almaz rm-saw-we(-her, -*him) we are-we (*is-he)

(19) Tny: Ḫū n+-Almaz z+-r?+n(-lya, -*lyya) n+hna ina (*-iyyu)
dm-Almaz rm-saw-we(-her, -*him) we are-we (*is-he)

'it is us who saw Almaz'

Again, rule (13) with its strategies may apply. The dislocation rule has no ef­
effect on agreement. The rules of agreement may be summarized as follows:

A. If complement is non-Nom
   (i) in all ES copula is 3 msg (see (14-16))
   (ii) in Tny, alternatively, subject of cleft clause may control agreement
        on copula. Compare (20) below with (14).

(20) n+hna z+-r?+n(-yla, -ya) n+-Abraḥat ina
we rm-saw-we (-him, -her) dm-Abraḥat are-we

'it is Abrahat that we saw'

B. With respect to the recalling of a missing DO or IO in the cleft clause,
   the pragmatics of the discourse context plays a decisive role. Consider
   the following:

(21) a. Ḫānā yā-matta-n-at Almaz-in nūw 'it is Almaz that we hit'
we rm-hit-we-her Almaz-dm is-he
b. Ḫānā yā-matta-n-∅w Almaz+n nūw

- him

In sentence (21a) the presupposition is that some female was hit. In (21b) the
object suffix -∅w 'him' attached to yā-matta-∅ 'that we hit' reflects
the presupposition that some hitting took place, and in ES an unspecified NP is
marked as 3 msg.

The same kind of consideration accounts for the two modes of formulation of
cleft sentences in which the complement is in an IO case. Thus, for instance,
in the cleft clause (22a), the presupposition is that the book was given to
some female; in (22b) the presupposition is simply that someone was given a
book.
(22) a. +ne mas' hashaf-u n y8-s8t'+8-hu-at la Almaz n8w
   I book-the-dm rm-gave-I -her to-Almaz is-he

b. +ne mas' hashaf-u n y8-s8t'+8-hu-tal la Almaz n8w
   -him

'it is to Almaz that I gave the book'

In contrast to the cleft clause, the relative clause, being a modifier or adjectival, forms a unitary major grammatical category with its head. As such, the grammar does not allow the removal of any material out of the NP of which it is a part.

The derivation of cleft and relative clauses in ES that we suggest is roughly represented by (23) and (24) respectively (cf. Schachter [1973], Gragg [1972], Hailu [1972]).

(23) VP
   S
   
   NP
   
   S
   
   Comp
   (rm)

(24) NP
   S
   
   VP
   
   S
   
   C1
   V
   Copula

   ...∅...V

In (23) ∅ represents a gap (absence of information) which is interpreted as being filled by the complement both in terms of grammatical category/relation and in terms of reference. S1, in structures like (23), represents an underlying structure for cleft clauses and cannot have a modifier function. Additionally, no material is extracted from it, for the candidate for extraction is phonologically null. In (24), on the other hand, S1 is a modifier, hence a relative clause. The inner NP, upon deletion leaves, except under specifiable conditions, a resumptive suffix pronoun reflecting its grammatical function in S1. To illustrate, we represent (25) and (26) as in (27) and (28) respectively.

(25) Kist: Almaz ∅ y8-allaft-tt b8-makina-in
   rm-left -she by-car -is-he

(26) Amh: Yonas ∅ y8-agbba-at seyyo
   rm-married-he-her woman married

'it is by car that Almaz left'

'the woman who(m) Yonas married'
4. **Conclusion**

To summarize, a cleft clause is characterized by a configuration which is much like that of a sentential complement governed (by the equivalent of) the complement "that" with no head; the only formal difference is that there is an obligatory absence of a major category/grammatical relation in the underlying source of a cleft clause. The presence of a gap or an empty category gives it the appearance of a relative clause. The non-existence of a head relates it to a sentential subject complement.
REFERENCES


SOME SYLLABLE STRUCTURE BASED RULES
OF TUNISIAN ARABIC

Hafedh Halila
University of Southern California

The title of this presentation refers to my ongoing investigation of the morphemic variations exhibited by surface forms of Tunisian Arabic (TA). For the purpose of this presentation I will focus on variations in the feminine possessive forms and propose an analysis of these forms using the notion of extrametricality [Hayes 1982]. I will first present some general properties of a model for the morphology and phonology of TA. Given the space and time limitations on this presentation, these properties will be treated as assumptions. Then I will present three different possible analyses of the feminine possessive forms and argue for the one based on the above assumptions and using the notion of extrametricality.

1. Morphology and Phonology of Tunisian Arabic

TA differs from other dialects of Arabic in that it allows a wide range of surface syllable structures and surface clusters of up to three consonants. Syllabification and syllable structure in TA can be captured straightforwardly through a multiple template approach à la Clements and Keyser [1983], subject to certain Negative Syllable Structure Constraints. Some of these constraints are given in (1).

(1) a. *[CV..]
   b. *[VCCC]
   c. *[VVCC]
   d. *[C]
   e. *[CV]

Furthermore, I assume that morphological processes in TA are of at least two types. The first type is lexical and involves the formation of lexical items and syntactic categories of the X°_type (all references to syntactic categories will be given in the framework of X theory). This includes categories such as verbs, adjectives, adverbs, masculine and feminine nouns, their dual and plural forms, etc. These formation processes proceed by prosodic derivational templates arranging the distribution of consonants and vowels, in pretty much the same way proposed by McCarthy [1981] for Classical Arabic. These templates represent the underlying representations and may differ from the actual surface forms.

The second type of morphological processes is a different sort of word formation which takes place at the phrase level. It involves suffixation, prefixation, compounding and cliticization. Crucially, these processes depend on information from the syntax. They operate on X° categories within larger categories, namely X and X categories. These morphosyntactic processes take place post-
lexically through the operation of morphological and phonological rules. The overall picture of the model that is being proposed here for the morphology and phonology of TA is shown in (2).

(2)

Lexical Underlying Representations

\[
\begin{align*}
\text{non-derived lexical items} & \quad + \quad \text{syntactic morphemes} \\
\downarrow & \quad + \quad \text{derivational templates} \\
\text{syntax = cliticization} & \quad \downarrow \\
\text{lexical insertion} & \quad \downarrow \\
\text{phonological and morphonemic rules} & \quad \downarrow \\
\text{output} &
\end{align*}
\]

The morphonemic alternations in TA argue in favor of such a model. In particular, tests of stress assignment and the internal structure of cliticized forms show that it is less redundant to have phonological rules and morphonemic rules apply postlexically. Since these rules are sensitive to syllable structure and also affect syllable structures, I will assume that syllabification applies before and after the application of each rule.

Before we move on to the analysis of the feminine possessive forms, we need to point out two major morphonemic rules in TA. The first striking characteristic of surface forms in TA is the presence of a vowel alternation between "neutral" forms and suffixed forms. Wise [1983] accounts for these alternations by a rule of posttonic vowel truncation which accounts for the alternations in (3), and a rule of metathesis which accounts for the alternations in (4).

(3) \( \gamma'\alpha\lambda'i \) 'wrong (masc.)' \( \gamma'\alpha\lambda\iota\epsilon \) 'wrong (fem.)'
\[
\begin{align*}
\text{måxzi} & \quad 'storage room' \\
\text{måxznii} & \quad 'my storage room'
\end{align*}
\]

(4) \( \text{šhān} \) 'a plate' \( \text{šhānna} \) 'our plate' \( \text{sānnii} \) 'my plate'
\[
\begin{align*}
\text{šhār} & \quad 'a month' \\
\text{šhārni} & \quad 'two months'
\end{align*}
\]

In Wise's analysis, the feminine form and the possessive form in (3) are derived from the masculine and "neutral" forms, respectively, through the operation of suffixation and rule (5).

(5) Short Vowel Deletion
\[
V \rightarrow \emptyset / \ C \_CV
\]

Rule (5) is in fact a reflex of a general surface constraint in TA which disal-
lows sequences of open light syllables. The analysis of the feminine possessive forms below will argue for a different statement of the same rule.

As for the forms in (4), Wise proposes the following metathesis rule:

$$\{C \} \text{ CCVC}^\#V \rightarrow \{C \} \text{ CVCCV}$$

We will show subsequently that rule (6) is also a reflex of the same rule of short vowel deletion.

The last rule to be mentioned is that of i-epenthesis which is quite common in other dialects of Arabic. In TA, i-epenthesis is a consequence of syllabification rules and the Negative Syllable Structure Constraints. An epenthetic i is inserted before a segment left stranded by syllabification.

$$\emptyset \rightarrow i / \{\_C \} \quad \text{where } C \text{ does not belong to a } \sigma .$$

(7) must follow all the other rules discussed so far but must apply before stress assignment.

2. An Analysis of the Possessive of Feminine Nouns

The possessive forms of feminine nouns in TA present an interesting problem. In these forms, a [t] which does not appear in the citation forms of these nouns surfaces when possessive suffixes are added to feminine nouns or in the genitive constructions with feminine nouns, as shown in (8) and (9).

(8) Possessives of feminine nouns

a. bâgra 'cow' bâgrit 'my cow' bâgritna 'our cow'
b. krûuma 'neck' kruumontik 'your neck' krumitkum 'your (pl) neck'
c. kîlma 'word' kîlmtu 'his word' kîlmtbhum 'their word'
d. mânṣfa 'towel' manṣâfți 'my towel' manṣbîtña 'our towel'

(N.B. The vowel final possessive suffixes in (8 a,c,d) have underlying long vowels which surface as short ones).

(9) Genitive constructions with feminine nouns

a. bâgrít màmmûud 'Maâmmûud's cow'
b. manṣâfți hàmmàam 'a bath towel'

Notice also the vowel alternations in the above forms. The final vowel of the unsuffixed forms drops when a suffix is added and an i appears in its position in the plural forms. In (8d), another vowel alternation involving a is present.

By contrast, masculine nouns do not have [t] in their possessive forms but exhibit the same type of vowel alternation.
Studies in African Linguistics, Supplement 9

(10) Possessives of masculine nouns

a. kālib 'dog' kālibi 'my dog' kālibna 'our dog'
b. tmār 'dates' tmāru 'his dates' tmārhum 'their dates'

The question now is how to analyse this [t] in the feminine possessive forms and how to account for the vowel alternations in the stems and possessive suffixes.

One possible analysis is to assume that the ending of feminine nouns takes three different surface forms:

- a before a pause
- incinn before vowel initial suffixes and in genitive constructions
- incinn before consonant initial suffixes.

In fact, this analysis was proposed by Wise [1983]. But it can easily be shown that even within the framework proposed by Wise, this analysis gives the wrong results. Consider the derivation of [manṣāfti] 'my towel' and [māríti] 'my wife', as given in (11), using the metathesis rule proposed by Wise.

(11) metathesis

\[ /\text{manšf} + \text{it} + \text{ii}/ \] and \[ /mr + \text{it} + \text{ii}/ \]

output \[ *\text{manšífti} \] \[ *\text{mírtí} \]

The above analysis is missing a generalization about endings of feminine nouns in TA and particularly fails to recognize that the i in the feminine ending [-it] is the result of i-epenthesis.

The second possible analysis of the possessives of feminine nouns is based on diachronic considerations. In Classical Arabic, feminine nouns end in [-at]. This final t appears in all of the regular morphological derivations involving feminines in Classical Arabic, such as the dual forms and the regular plurals. If we assume that TA, in its diachronic changes, has kept basically the same derivational templates of Classical Arabic, then we should expect the final [t] to appear in the dual forms and plurals of feminine nouns in TA.

This is the case as shown in (12). The dual is derived by adding the suffix [-ii] to the singular and the plural by lengthening the vowel of the final syllable.

(12) a. Dual of feminine nouns

\[ \text{bagrtii} \] '2 cows'
\[ \text{manšaftii} \] '2 towels'

b. Plural of feminine nouns

\[ \text{bagraat} \] 'cows'
\[ \text{mnaašif} \] 'towels'
\[ \text{manšfaat} \] 'towels' (marginally acceptable)

The evidence in (12) can be taken in favor of an analysis which assumes that a reanalysis of the possessive forms of Classical Arabic has happened which resulted in the [-t] of the feminine being analyzed as part of the possessive
suffixes and dual suffixes. Consequently, TA ends up having two classes of possessive suffixes. One class is added to consonant final nouns and another one, having [t] as initial consonant, is added to the feminine nouns ending in [a].

To test this hypothesis, we will consider the derivation of [manšäftī] 'my towel' and of [manšfatna] 'our towel'.

\[
\begin{array}{c|c|c}
\text{metathesis} & \text{N} & \text{N} \\
\text{rule (5)} & \text{N} & \text{N} \\
\text{i-epenthesis} & \text{N} & \text{N} \\
\text{output} & \text{manšäftī} & \ast \text{manšfatna} \\
\end{array}
\]

The possible analyses proposed above, besides leading to wrong surface forms, fail to capture the generalizations that we proposed for underlying and surface forms of TA. In what follows, I would like to propose an analysis which considers the alternations exhibited by the possessive of feminine nouns as part of the general morphonemic variations associated with suffixation and cliticization in TA. More specifically, I propose to characterize this funny final [t] as extrametrical, where by extrametrical I mean peripheral. For the purpose of this analysis, I keep the assumption that the lexicon in TA contains the core templates of masculine and feminine nouns. The templates of feminine nouns ending in [a] end in an extrametrical [t]. This [t] is peripheral and lies outside the core template and therefore can join any clitic suffix subject to syllabification postlexically. It ceases to be peripheral when a vowel initial suffix is added. Because of (1d), it must syllabify with it. But crucially it cannot join the coda of the preceding syllable in the core template.

Furthermore, if we assume that TA has kept the same derivational templates as Classical Arabic and that these templates are entered in the lexicon as underlying representations, vowel alternation can be accounted for by reconstructing the alternating vowel in the underlying representations of the forms which exhibit morphonemic vowel alternation. Consequently, we may find sequences of light open syllables in underlying representations and surface consonant clusters are created through the operation of the short vowel deletion rule. However, we need to reformulate rule (5) as rule (14).

(14) Short Vowel Deletion

\[
\begin{array}{c|c}
V \rightarrow \emptyset / (C)CV(C) \\
\end{array}
\]

Rule (14) scans the relevant domains from left to right and applies iteratively whenever its environment is met. Being a postlexical rule, it cannot have access to the internal structure of the lexical word unless an element of the lexical word is included in its environment. The crucial examples for this analysis are words with surface clusters of three consonants. Again, the derivation of 'my towel' is given in the left hand column of (15) and that of 'our towel' in the right hand column. The square brackets labeled LW mark the boundaries of
the lexical word and those labeled PLW delimit the boundaries of the postlexical word. The horizontal lines mark the domain of application of rule (14).

\[(15) \quad \text{PLW}_1 [\text{LW} \text{manşafa} \text{til}] \quad \text{PLW}_2 [\text{LW} \text{manşafa} \text{tnaa}] \]

\[
\begin{array}{c}
\text{Rule (14)} \\
\text{NA} \\
\text{mansafati Output mansaf}tna
\end{array}
\]

Output 'my towel' 'our towel'

Notice that in the derivation of 'my towel' rule (14) cannot delete the a of /-şa-/ because it does not have access to the lexical word. It must scan elements from the derived environment so that its application is triggered.

3. Conclusion.

In this presentation then, I have presented an analysis of the possessive of feminine nouns which fits into the overall model proposed for the phonology and morphology of TA. I have shown that by making certain minimal assumptions about the nature and properties of the underlying representations in TA, which are independently justified elsewhere in the grammar, we are able to account for the morphonemic alternations of the possessives of feminine nouns in a simplified and unified manner. Moreover, I have presented evidence against the existence of a metathesis rule [Wise 1983] and shown how the notion of extrametricality can be used to characterize the [t] of the feminine forms.

REFERENCES


1. Introduction

This paper is an argument for the true generalization condition [Hooper 1976:13]. It deals with the basic facts of Arabic so-called "doubled verbs" (roots of the pattern 122), comparing an "autosegmental" grammar with one governed by the condition to show the superiority of the latter with regard to these facts. Another purpose is to argue against unprincipled grammatical theories in general.

Principled grammatical theories are those governed by objectively applicable, empirical principles. Not among these are informal, subjectively interpreted principles such as Postal's [1968:56] "naturalness condition": "the relation between phonological and phonetic structure is a natural one." A preference for "naturalness", so defined, is as helpful as a similarly defined preference for "abstractness", or "concreteness". Such preferences for grammars is not helpful because it is obvious that languages have both natural and unnatural and abstract and concrete aspects. As linguists we would agree that language is so important to knowledge, and understanding the structure of language so important to human understanding in general, that the meanings of words like "natural", "concrete", and "abstract" are bound to be profoundly affected by whatever understanding we achieve of language, both as a social institution and a cognitive structure.

Nor are would-be formal principles of languages helpful if they lack empirical applicability, for example, those which merely declare something to be "costly" rather than to exclude it, such as Kiparsky's 1974 [1968] "alternation condition". There is inflation in linguistics as in economics. It does no good to have a constraint on rules unless you have a constraint on lexical entries, nor to have a constraint on lexical entries unless you have one on rules. It doesn't do any good to claim, for example, that the feature SYLLABIC is autosegmental and then allow ad hoc exceptions to the claim.

My claim, which I believe is the essence if not the totality of the true generalization condition, is that rules don't change, move, or delete features or segments. This is not entirely consistent with Vennemann's or Hooper's practice in their work on "natural generative phonology".

2. Arabic Doubled Verbs

The Arabic doubled verbs illustrate a little conspiracy against certain sequences of like consonants separated by a short vowel. Two alternations take place in fulfillment of the conspiracy:

(1) a. samm-a 'he poisoned' (*samam-a; cf. katab-a 'he wrote')
    b. ya-summ-u 'he poisons' (*ya-smum-u; cf. ya-ktub-u 'he writes')
b. jarr-a 'he pulled' (*jarar-a ; cf. katab-a )
c. ya-jurr-u 'he pulls' (*ya-jur-r-u ; cf. ya-ktub-u )

The first alternation is the presence and absence of a second vowel in the stem of form I perfects. The 3m.sg. form I perfect of smm 'poison' is samm-a , not samam-a , which is the pattern of the regular verb roots such as ktb 'write', i.e. katab-a . The second alternation is in the left-to-right arrangement of the second consonant of the root and the vowel of the form I imperfect stem. The 3m.sg. form I imperfect of smm is ya-smum-u , not ya-smum-u , which is the pattern of the regular verbs, e.g. ya-ktub-u . These distinguishing characteristics of the doubled verbs appear in other forms which, like the 3m.sg., have vowel-initial suffixes.

The doubled verb alternations are not unusual. They, like the pronunciation prab'lly for probably and the inspiration for tongue twisters like rubber baby buggy bumpers, are owed to the physiological complexity in repeating like consonants separated by only a short or unstressed vowel.

3. Two Analyses of the Doubled Verb Conspiracy

The tradition of generative grammar is to think of these alternations of form in the perfect and imperfect stems as owed to transformations of the stem-pattern of the regular verbs. This is the approach of the autosegmental analysis of Arabic verb morphology by John McCarthy [1982]. In this analysis, the stem of a verb is the tier of consonants and vowels or of the feature SYLLABIC. Input to the autosegmental derivation of the form I perfect and imperfect stems katab/ktub are the following: the root tier ktb (said to be all the features of ktb less SYLLABIC); the vowel tiers a perfect/ u imperfect; and the "template" tiers for CVCVC/CCVC, consisting of the feature SYLLABIC only. This is shown in (2).

(2) Perfect I Imperfect I

\[
\begin{array}{c|c|c}
\text{ktb} & \text{ktb} & \text{ktb} \\
\text{CVCVC} & \text{CVCVC} & \text{CCVC} \\
\text{a} & \text{a} & \text{u} \\
\end{array}
\]

The root tier ktb associates appropriately with the Cs of the template tier according to the autosegmental convention of grammar of "left-to-right" association. The single vowel of the vowel tier replicates by another convention.

The doubled verb conspiracy is expressed by McCarthy [1982:399] as rule (3). This is effectively a collapsing of a metathesis and a deletion rule. This is not obvious since rule (3) is unprecedented in form and effect.

(3) 1 2 3 4 5

\[
<\text{V}> \text{CVCV} \rightarrow 1 <3>_b 2 4 5
\]

Condition: a \(\supset\) ~ b
Interpretation of the angle brackets requires the subscripts $a$ and $b$ and the footnote to the rule. In the perfect, where $a$ occurs, $b$ is deleted. In the imperfect, where $a$ doesn't occur, $b$ does (contradicting the usual interpretation of angle brackets), but then $b$ (that is, 3) inverts with 2. The alpha, with its lines, is said to provide that the affected consonants are like consonants, and the $x$ that the left-most consonant is unlike these—that is, the rule affects a doubled verb.

I suggest that this rule doesn't provide a very revealing account of the doubled verb conspiracy. It seems strictly ad hoc. It's unusualness, considering that the doubled verb alternations are quite reasonable from the point of view of articulatory phonetics, may reasonably be an argument against the autosegmental analysis of which it is a part. An alternative understanding, which makes reference to the explanation mentioned, is the negative condition (4). Presumably at first purely phonological, this condition is now (I think) morphologized by the boundary.

$$(4) \quad * \left\{ V \right\}_{a}^{c} C_{b} V C_{b}^{a} V$$

In (5) are shown the non-transformational observationally valid rules of Arabic for expanding roots as stems of the perfect and imperfect. Note the straightforward relation of association between input and output consonants in (5).

$$(5) \quad$$

a. $CCC \rightarrow CaC(a)C$ / perfect I (i.e. $smm \rightarrow sam(a)m$ )
b. $CCC \rightarrow C-C,u-C$ / imperfect I (i.e. $smm \rightarrow s-m,u-m$ )

The left-to-right order of input consonants is necessarily preserved in the output, a necessity which follows from the true generalization condition prohibition of movement rules.

Perhaps the main argument given by McCarthy [1982:375] for the autosegmental analysis of Arabic root and pattern morphology is that in the usual transformational grammar all the input and output consonants in rules like (5) must be indexed, so that, assuming always the possibility of movement, these can be matched up correctly. The autosegmental grammar solves this problem with its convention of left-to-right association. According to the true generalization condition this is a principle of analysis: movement rules are prohibited as violations of the condition.

The second thing to notice about the rules (5) is the parentheses in the rule for the perfect stem. This is traditional in interpretation: the parenthesized vowel may or may not appear—this is the vowel absent in doubled verbs when a vowel-initial suffix follows. Finally notice the hyphens and comma in the rule for the imperfect stem. This notation abstracts invariant linearity from this stem. It expresses the possibility that the segments between the hyphens may or may not occur in the left-to-right order given. These are the segments which vary linearity in doubled verbs.

The parenthesis and comma-hyphen notations interact with the negative condition (4), the explanatory generalization over the doubled-verb alternations, to provide the correct outputs of doubled verbs. The necessary implication of (4),
given the variable outputs from (5) provided by these notations, is the expression of the doubled verb conspiracy. This is shown in (6):

(6) a. sam(a)m]a + sam]a by (4), otherwise, e.g. kat]a
   b. ya-s-m,u-m]u + yasumm]u by (4), otherwise, e.g. ya-k]u

Since each of the rules of (5) provide two possibilities, and according to rule (4) one of these is impossible for the doubled verbs, it follows deductively that the doubled verbs take their allowed form. By a general convention of the grammar—effectively proper inclusion precedence [Sanders 1974]—the other, regular, verbs are considered to take the other form of the alternatives provided by (5).

The statement of the facts of Arabic (4) and (5), with the prohibition of movement and deletion rules following from the true generalization condition, are a principled and explanatory understanding of the doubled verb conspiracy. The analysis (3), by contrast, is just a description, and not a very clear or interesting one at that.

4. Other Evidence: a Language Game

After the oral presentation of this paper Bruce Hayes reminded me of an argument given by McCarthy in a footnote (p. 396). In the autosegmental grammar the doubled verbs may be treated as biliterals, e.g. sm rather than smm, since by the left-to-right convention CaCC would be completed as s]mC, and by the convention for associating a remaining C or V on the CV tier (p. 382), this would become, correctly, samm. This approach is said to follow from the "obligatory contour principle" (p. 383) according to which identical adjacent segments of a tier are disfavored (not disallowed). McCarthy suggests (p. 396) that this prejudice against identical adjacent segments in a tier explains the limitations of Arabic root structure discovered by Greenberg [1950], e.g. *smm, but in fact these are diverse, extending well beyond any simple prohibition of identical adjacent segments on an autosegmental tier (for some discussion see Lightner [1973]).

The argument for biliteral doubled verbs is found in an Arabic language game, in which ktb can be realized variously as tkb, btk, etc., but the doubled verb Ỉonly as 칙. This follows from lexical ǔl, permuted as ǔh, which forms the correct stem ǔh. But these same facts follow very reasonably in the principled grammar of (5) given 闪过, since this could be permuted by the game rule as 闪过 and 闪过, but neither of these patterns (C,C,C, and C,C,C) are generally found as Arabic roots; the latter never, the former only rarely [Greenberg 1950]. They may be considered systematically disfavored. But if we assume that, in the game, doubled verbs are retained as doubled verbs, the impass is avoided by the game pronunciation 칙.

Notice that on this analysis movement DOES occur (in the conversion of ktb to tkb, etc.), but in the game lexicon, not in derivations. This is what is significant about language GAMES—they do things not possible in grammars but still in linguistically principled ways. Here the game violates the true generalization condition, but only as in a sound change (metathesis), not as synchronic movement. In the autosegmental-transformational theory, on the other
hand, movement is possible anywhere (cf. McCarthy's metathesis-derivation of form VIII, p. 390), so the fact that the language game allows this only in the lexicon would not be predicted. It happens that the autosegmental grammar handles the language game facts, but so does the non-transformational grammar, and only the latter does it in a fully principled way.

REFERENCES


WH-MOVEMENT AND RELATIVIZATION IN TWO CREOLE LANGUAGES:
CAPE VERDEAN (CV) AND HAITIAN (H)

John P. Hutchison
Boston University

1. Introduction

The structural similarities between WH-question formation and relative clause formation (and the often related cleft, pseudo-cleft, and embedded question constructions) have been observed with regard to many languages. In cleft constructions, a left focussed constituent is followed by a clause bearing a syntactic resemblance to the relative clause construction in a significant number of languages. Given [1979:217-218] explains this by suggesting that the cleft structure arose diachronically as a result of tacking the presupposed relative clause portion onto the focussed constituent as an afterthought, paratactic construction.

The languages here under consideration attest these paratactic RC-like structures, for many NP positions, under questioning and relativization. Thus the RC-like embedded S may be characteristic of the WH-question, the RC construction, the embedded WH-question, and the cleft and pseudo-cleft constructions. The similarities are indeed apparent from the following data set. Here the CV COMP KL is shown in capital letters, as is the H DET LA, the latter being shown without its usual phonological assimilation. CV precedes H.

(1) WH-QUESTION
   (ki) kuz(-e) KL bu kre?  'what do you want?’
      WH thing COP COMP 2S want
   (ki) sa u te ašte LA?   'what did you buy?’
      WH that 2S PST buy DET

(2) RESTRICTIVE RELATIVE CLAUSE
   kel kumida KL bu kre e kaçu pa
      the food COMP 2S want COP
   m we li v yo u te ašte LA
      1S see book PL 2S PST buy DET
   'the food that you want is kacupa'

(3) EMBEDDED WH-QUESTION
   n ka sabe (ki) kuz-e KL bu kre
      1S NEG know WH thing COP COMP 2S want
   m pa kon ki sa u te ašte LA
      1S NEG know WH that 2S PST buy DET
   'I don’t know what you want'

(4) CLEFT
   e kumida KL bu kre
      'it’s food that you want’
Cape Verdean and Haitian Creole

se liv u te ašte LA

'it's a book that you bought'

(5) PSEUDO-CLEFT/FREE RELATIVE

kuza Kl n kre e kumida

'what I want is food'

sa u te ašte (LA) bo

'what you bought is good'

In addition to marking the RC-like structures of (1-5), both the CV COMP Kl and the H DET LA play important roles in their respective systems of complementation, as shown in (6-7).

(6) NP COMPLEMENTATION

vini I vini (LA) neve m

'the fact that he arrived upset me'

(7) VP COMPLEMENTATION (that-S)

n fra-I Kl-el ten ki bai

'I told him that he must go'

1S tell him COMP he must go

m te di li te vini (LA)

'I said that he had come'

1S PST say he PST come DET

H LA may be optional in embedded S-final position both in WH-questions and in examples like (6-7). Lefebvre [1982:46] is the only author to describe the presuppositional role played by LA in the S-final position. Indeed there is an expectation, or the presupposition, that the COMP-S will be or has been realized when the LA occurs.

2. WH-Questions in CV and H.

In both languages, the questioning of subject, direct object, and also indirect object NPs involve a gapping strategy. Both have obligatorily dative-shifted indirect objects which take no adpositional marking either in declaratives or when questioned or relativized. In H, when the subject is either questioned or relativized, kl occurs obligatorily in subject position in the stead of the questioned or relativized subject NP constituent. In its relative use, Hall [1979:49] referred to this kl as a relative pronoun which is the subject of the embedded clause.

In (9), in the presence of the full analytic CV WH form kl kuza , with or without the copula, the COMP Kl is obligatory. When the reanalyzed kuze with apparently accreted copula is used, the COMP Kl is then optional. The obligatory use of the COMP Kl in the presence of the WH-word Kl suggests that a RC is involved since in all RCs the COMP Kl is obligatory. However, whether or not kuza 'thing' is the head of the RC or whether the RC structure is a paratactic free appositive RC of the type proposed by Givon is unclear.

This is due in part to the freedom to delete or omit the copula and the freedom of movement of the copula in these questions, as in clefts and pseudo-clefts, and also due to the ongoing processes of reanalysis and natural change in creole languages. The following are WH-questions on major NP positions.
(8) ken(-e) KI ben?  
who COP COMP come  
'who came?'  
*ken (e) ben?  
ki mun ki te vini LA?  
WH-person SUB PST come DET  
'who came?'  
*ki mun te vini (LA)?

(9) ki kuz -e KI bu ta faze?  
WH thing COP COMP 2S PRC do  
'what are you doing?'  
*ki kuze bu kre?  
ki sa u ap fe LA?  
WH-thing 2S PRG do DET  
'what are you doing?'  
*ki sa ki u ap fe LA?

It is possible that the structure involved in (9) is comparable to that of Igbo WH-questions as analyzed by Goldsmith [1981]. He treats such structures as the verbless juxtaposition of a WH-word and a RC. Thus for (9), 'what is the thing that you are doing?' rather than 'which thing is it that you are doing?' If this is the underlying structure of the CV example then it does not represent a case of WH-movement. In the alternative structures without the COMP KI in CV, WH-movement to initial position is obligatory, and the in situ WH-question is unacceptable in all cases. This suggests that WH-movement is becoming obligatory for the set of reanalyzed WH-words in CV.

In oblique WH-questions, a preposition is required if one is also required in its unquestioned form. For many temporals and locatives prepositions are not necessary in either case.

(10) ku kuza(-e KI) n ta korta pon?/kuza(-e) KI n ta korta pon ku el?  
'with what am I going to cut the bread?'

(11) ak ki sa m pra kupe pe LA?/*ki sa m pra kupe ak (li) LA?  
'with what am I going to cut the bread?'

3. Relative Clause Formation Strategies in CV and H

In all CV relative clauses, the COMP KI is obligatorily preposed to the RC. In H, LA is obligatorily postposed to the RC. Similarly the formation strategies for the two languages match within the RC, differing only with regard to subject relativization, Haitian having the relative pronoun-like KI in subject NP position within the RC, CV having a gap since CV KI is here analyzed as a non-WH COMP element. Indirect object relativization also differs among the two in that in CV a resumptive pronoun may occur obligatorily in data-shifted position, whereas in Haitian only the gapping strategy is used. Both languages use resumptive pronoun strategies for all prepositionally marked oblique NPs and for genitives.

(12) MAJOR NP RELATIVIZATION

kel omi KI bai  
'the man who went'

kel omi KI n oja  
'the man that I saw'

mun LA ki te vini LA  
'the person who had come'
The word order of the Haitian noun phrase constitutes a typological anomaly. This anomaly is at the root of noun phrase complementation and the structure of the complex NP of the language, in all of the forms here under consideration. In Haitian, all of the definitizing and deictic/anaphoric elements of the determiner system are postposed. Thus the genitive determiners, demonstratives, the definite article LA, and RCs all follow the head noun. Modifying prepositional phrases also follow the head. All quantifiers and adjectives however precede the head N. It is the DET LA which has been observed in its role in the NP and VP complementation systems. Hall [1964:32] describes LA as the "Determinant, or 'definite article', which serves to make a nominal phrase out of any phrase to which it is added, and makes it refer to a specific thing or things, acting as a 'nominalizer'." Hall thus captured the role of LA in relation to the embedded S.

The distribution of LA in embedded S-final position indicates that it has been reanalyzed as more than a simple DET. In the literature there has been significant disagreement on how the determiner LA should be treated. Different understandings of the role of DET LA in relation to the relative and non-relative embedded S are the reason. Lefebvre [1982:174] states, "We should remember that LA is obligatory in the environment of a restrictive RC, and that the presence of LA in this environment is essentially deictic." She completes her analysis of LA by ranking it together with the S-initial COMP position, as a "DET of S" [1982:46].

If COMP can evolve from a DET, then any language with postposed DETs is a candidate for having a postposed COMP, H would certainly not be unique in having a COMP which has evolved from a DET. Indeed, other creole and non-creole languages demonstrate comparable characteristics. Sankoff [1981:32] reports on the evolving Tok Pisin relativizer ia, whose development she has traced from a place adverb (from the English 'here'), through demonstrative or deictic, to relative marker. She argues that Tok Pisin uses ia on either side of the RC to syntactically bracket the RC as a step toward the syntacticization of embedding. I would instead interpret the Tok Pisin structures as instances of a RC construction with a definite head noun marked by the postposed DET ia, followed by the RC also marked by ia in its COMP function.
5. Conclusions

It has been observed that in CV and H, WH questions typically have more than one strategy: the clefted strategy in which the COMP element KI and the H DET cum COMP LA are used, and the non-clefted strategy in which these COMP-elements are absent. Further, for many positions, the WH-question strategies do not match those of their relative counterparts. The H WH-questions on positional NP positions do not allow the resumptive PRO strategy of their in situ relative counterparts. Instead they require fronting of the entire WH phrase with preposition and no possibility of stranding a preposition. Questioning the same positions in CV, we observe either the same WH-fronting strategy as in Haitian or the in situ PP, just as in the CV relative counterpart. I would suggest that WH-movement is not characteristic of the basilectal forms of either of these creoles. WH-words and movement are certainly absent from the RC-formation strategies where deletion and resumptive pronouns are used. In questions, I would argue that the cleft construction is basilectal and that WH-movement characteristics are coming into these languages as the result of the reanalysis of Wh-words and of contact with superstrate languages.

The present paper assumes an analysis of the H determiner LA as a complementizer and thus posits two COMP nodes in H, one preceding and one postposed to an embedded S. Therefore the sentence-LA structure represents an embedding, whether it be a nominalization or a clefted RC-like structure.

REFERENCES


DEPENDENCY RELATIONS IN SYNTAX: THE MYSTERIOUS CASE OF THE EMPTY DETERMINER IN AGHEM

Larry M. Hyman
University of Southern California

[This paper is a summary of the invited plenary address—Ed.]

In Hyman [1979a] I presented a list of environments in which the determiner -5 occurs in Aghem, a Grassfields Bantu language spoken in Cameroon. In this paper I would like to replace this list with an outline of a formal account of this problem. To begin, consider the examples in (1).

(1) a. ó mò tān tɪ̣-bvé 'nè 'he counted dogs today' [P₁ = today]
   he P₁ count dogs today past tense
   b. ó kà tān bvé 'tò nè 'he didn't count dogs today'
   he NEG-P₁ count dogs D today
   c. tān bvé 'tò nè 'count dogs today!'
   Ø count dogs D today

In the earlier study it was demonstrated that -5, which occurs in (1) with class 10 tɪ̣- agreement, is a determiner (D), since it (a) occurs in the same "slot" as the three demonstratives in the language, e.g. following adjectives but preceding numerals, and (b) is mutually exclusive with these demonstratives. In (la) we observe that a noun immediately following an affirmative verb (in a main clause) must occur without -5, while the corresponding object noun must occur with -5 after a negative verb in (lb) or an imperative (in (lc)). Thus, we are dealing with a problem in complementary distribution: environments where -5 must appear vs. those where it must not. (In the case where the D "slot" is filled by a demonstrative, the Ø/-5 complementarity is of course not observed.) Finally, the tɪ̣- prefix of tɪ̣-bvé 'dogs' not found in (lb,c) is deleted by a general rule of prefix deletion which applies whenever a head noun is followed by an agreeing element other than a numeral (cf. bvé 'tān̄̄ 'my dogs', bvé tɪ̣̄n̄ 'these dogs', etc.).

I propose that the D node is obligatory in Aghem, either being filled lexically by one of the three demonstratives or remaining as an empty element eD. If the eD is "syntactically well-formed" it surfaces as Ø; if it is not well-formed, it must be spelled out post-lexically via the -5. The well-formedness conditions on eD are of three types: (a) internal conditions on the NP; (b) external conditions on the NP vis-à-vis its governing head; and (c) external conditions on the NP or its governing head with respect to modality and clause-type. I illustrate each briefly below.

In order to capture the conditions on eD internal to the NP, consider the Aghem noun phrase structure rules in (2).

(2) a. \[ \tilde N \rightarrow \tilde N \text{ SPEC} \]
   b. SPEC \rightarrow \{ E \ (\text{Num}) \}

In Hyman [1979a] I presented a list of environments in which the determiner -5 occurs in Aghem, a Grassfields Bantu language spoken in Cameroon. In this paper I would like to replace this list with an outline of a formal account of this problem. To begin, consider the examples in (1).

(1) a. ó mò tān tɪ̣̄-bvé 'nè 'he counted dogs today' [P₁ = today]
   he P₁ count dogs today past tense
   b. ó kà tān bvé 'tò nè 'he didn't count dogs today'
   he NEG-P₁ count dogs D today
   c. tān bvé 'tò nè 'count dogs today!'
   Ø count dogs D today

In the earlier study it was demonstrated that -5, which occurs in (1) with class 10 tɪ̣̄- agreement, is a determiner (D), since it (a) occurs in the same "slot" as the three demonstratives in the language, e.g. following adjectives but preceding numerals, and (b) is mutually exclusive with these demonstratives. In (la) we observe that a noun immediately following an affirmative verb (in a main clause) must occur without -5, while the corresponding object noun must occur with -5 after a negative verb in (lb) or an imperative (in (lc)). Thus, we are dealing with a problem in complementary distribution: environments where -5 must appear vs. those where it must not. (In the case where the D "slot" is filled by a demonstrative, the Ø/-5 complementarity is of course not observed.) Finally, the tɪ̣̄- prefix of tɪ̣̄-bvé 'dogs' not found in (lb,c) is deleted by a general rule of prefix deletion which applies whenever a head noun is followed by an agreeing element other than a numeral (cf. bvé 'tān̄̄̄ 'my dogs', bvé tɪ̣̄n̄ 'these dogs', etc.).

I propose that the D node is obligatory in Aghem, either being filled lexically by one of the three demonstratives or remaining as an empty element eD. If the eD is "syntactically well-formed" it surfaces as Ø; if it is not well-formed, it must be spelled out post-lexically via the -5. The well-formedness conditions on eD are of three types: (a) internal conditions on the NP; (b) external conditions on the NP vis-à-vis its governing head; and (c) external conditions on the NP or its governing head with respect to modality and clause-type. I illustrate each briefly below.

In order to capture the conditions on eD internal to the NP, consider the Aghem noun phrase structure rules in (2).

(2) a. \[ \tilde N \rightarrow \tilde N \text{ SPEC} \]
   b. SPEC \rightarrow \{ E \ (\text{Num}) \}
The rule in (2b) accounts for the fact that there can be at most only one demonstrative (or \(\bar{\gamma}\)) and only one numeral per maximal NP. Thus, while one can say 'books of child these' or 'books of child that', one cannot say *'books these of child that' or 'books of child these that', etc., and similarly for numerals. Turning to the problem at hand, the eD is well-formed internally if it is preceded only by lexical nouns. Thus, in (3a) we see that an adjective occurring anywhere in the NP will make the eD ill-formed, and in (3b) we see that an empty head of the NP will do likewise (these phrases without \(\bar{\gamma}\) are ungrammatical):

\[
\begin{align*}
\text{(3) } a. & \quad \text{rat big of dogs} \\
\text{b. } & \quad \text{that of the dogs}, \text{ e.g. the rat of the dogs of dogs} \\
\text{c. } & \quad \text{rat of dogs}
\end{align*}
\]

Had there been no adjective, and had the head noun been lexical (rather than empty), then eD would have been well-formed, as it is in (3c).

The tree structures generated by the rules in (2) for (3a,b), which are given in (4a,b), demonstrate that the eD is ill-formed when it fails to be properly governed by every N within the NP, where proper government requires a lexical governor which c-commands the empty element (c-command meaning here that the first branching node dominating the governor must also dominate the empty element). Structures (4a,b) are ungrammatical, therefore, by a chain version of the empty category principle of Chomsky [1981] and others (cf. below).

\[
\begin{align*}
\text{(4) a. } & \quad \text{rat big (of) dogs} \\
\text{b. } & \quad \text{cl.7 (of) dogs}
\end{align*}
\]

Note that an adjective occurring anywhere in the NP will render eD ill-formed, while the offending empty noun head is grammatical only in the highest position of the maximal NP.

The sentences in (5) demonstrate that when the direct object NP is separated via focus-movement by either other post-verbal element(s) or by the subject, the eD is ill-formed, even if the NP-internal conditions are met.
The Empty Determiner in Aghem

(5) a. 'bév tó 'hevnt counted dogs today' (*dogs e)
   he P1 count today dogs D
b. 'nwó qá bgá 'bév tó 'children counted dogs today'
   DS P1 count children dogs D today [DS=dummy subject; *dogs e]

Noting that the immediate post-verbal position is the focus position [Watters 1979], and accepting the basic sentence structure in (6),

(6) 
   \[ S - N - V - Y - Z \text{ etc.} \]

a post-verbal NP will be governed by the verb only if it is in X (focus) position, assuming the same definition of c-command given above. From (1a) we observe that the eD is externally well-formed in X position, but not, as in (5a, b), if its NP is in Y (or Z, etc.) position. Thus, a necessary external condition on a post-verbal eD is that its NP be governed by the verb.

This does not explain, however, why the negative and imperative sentences in (1b,c) and the focused auxiliary sentence in (7) disallow eD in the immediate post-verbal NP object.

(7) 'bév tó 'he did count dogs today' (*dogs e)
   he P1-FOC count dogs D today

For (1b) we note the often observed "looser bond" between a negative verb and its object which may result in the failure of the verb to assign accusative case to its object and/or provide a prosodic break between the (negative) verb and its object (cf. Hopper and Thompson [1980]; Hyman and Watters [1984]. The auxiliary features NEG, IMPER, and FOC all have this effect in Aghem, i.e. the effect of removing the verb’s ability to govern its object. Though I shall attempt a different solution below, one might try to account for this formally by either requiring that the X position of (6) be empty just in case one of these features is present or, equivalently, that these features "copy" into the X position itself, thereby assuring that the NP object will not be governed by the verb. In (8) it is noted that only the object of the first verb is affected by negation in a serial verb construction:

(8) 'kó ñl fó kó ñl kó-fú 'he didn't take the knife and cut the
   he NEG-P1 take knife D cut rat rat' (*knife e; *fú kó 'rat D')

Turning to clause-type considerations, the sentences in (9) show that eD is well-formed in subject position in a main clause (9a) but ill-formed in subject position in a relative clause (9b).

(9) a. tó-bév e tó-bév ñl ñl 'the) two dogs fell' (*bév tó )
dogs D two P1 fall FOC
b. bé 'k¶ á bvé 'tó t¶-bhà mó z¶
fufu DEM REL dogs D two P₁ eat
'(the) fufu that (the) two dogs ate' (*dogs e)

Two additional facts are that (a) if and temporal clauses have exactly the same properties as relative clauses, and (b) eD is not well-formed in an immediate post-verbal NP in any of these subordinate clauses, as seen in (10).

(10) a. fú k¶'ò k¶ mó tàñ bvé 'tó '(the) rat that counted (the) dogs'
rat DEM REL SM P₁ count dogs D (*t¶-bvé e)
b. bughó ó mó tàñ bvé 'tó 'if he counted (the) dogs'
if he P₁ count dogs D (*t¶-bvé e)
c. gh¶'á ó mó tàñ bvé 'tó 'as he counted (the) dogs'
as he P₁ count dogs D (*t¶-bvé e)

However, as seen in (11),

(11) wizá wli á ó mó dzé nh¶'á wô mó tàñ t¶-bvé
woman DEM REL SM P₁ say that you P₁ count dogs
'(the) woman who said that you counted (the) dogs'

a that-complement clause does not cause the eD to be ill-formed in object (or subject) position, even if embedded in a relative clause as it is in (11). We therefore need to group together main and complement clauses vs. relative, if, and temporal clauses (among others), a distinction which I identify with Emonds' [1976] root/non-root S dichotomy. The table in (12) summarizes the differences:

<table>
<thead>
<tr>
<th></th>
<th>Root S</th>
<th>Non-root S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-verbal subject:</td>
<td>eD = OK</td>
<td>eD = *</td>
</tr>
<tr>
<td>Post-verbal NP :</td>
<td>eD = OK/*</td>
<td>eD = *</td>
</tr>
<tr>
<td>INFL (&quot;aux&quot;) :</td>
<td>full range of tenses (P₁ ≠ P₂) &amp; [+/-FOC]</td>
<td>P₁ = P₂, no [-FOC] distinction</td>
</tr>
<tr>
<td>FOC marker no :</td>
<td>OK</td>
<td>*</td>
</tr>
</tbody>
</table>

Not only is there a difference in the well-formedness of eD, but there are important differences concerning the tense morphology and the distribution of the focus marker no [Watters 1979] according to the root/non-root status of an S.

To follow through with the above discussion of government and its relation to the eD, it is necessary to say that for these purposes the subject of a root S is governed, while both the subject and the object of a non-root S are not. The analogy with tensed/non-tensed S's is intriguing. To account for these facts a combination of feature specification plus tree configuration will be invoked. I propose the following: (a) each S has a focus feature [+F]; (b) this [+F] is assigned either to the COMP, in which case we have a non-root S, or to
the S, in which case we have a root S; (c) if [+F] is assigned to COMP, it is copied as well onto INFL; and (d) the auxiliary features NEG, IMPER, and FOC have a second "intrinsic" [+F] ("auxiliary focus"; cf. Hyman and Watters [1984]). With these features assigned, a head attraction rule roughly of the form in (13) joins the head (X') of a maximal projection to a preceding [+F] element.

(13) \[ [+F] \leftarrow X \rightarrow Y \] (informal approximation)

The result, as seen in (14), is that the N of a subject NP will be attracted to the [+F] of a non-root S COMP and the V of the VP will be attracted to a [+F] INFL.

(14)

With such potential lexical heads attracted out of their respective phrases, we now can generalize on what was said concerning proper government and the \( \text{eD} \): an \( \text{eD} \) will be well-formed if it is the endpoint of a chain of proper governors up to the S node. In (14), the subject \( \text{eD} \) is not properly governed because the head of this NP has been attracted to the [+F] COMP; similarly, the \( \text{eD} \) of the object is ungrammatical because, while properly governed by its N head, the NP itself is not properly governed by the V which has been attracted to the [+F] INFL. Thus, when the chain of proper governors is broken, \( \text{eD} \) is ungrammatical.

Further evidence for this view is seen from (15), where it is observed that the two prepositions \( \text{a 'to/for'} \) and \( \text{a 'with/and'} \) are not proper governors.

(15) a. \( \text{a b\text{vè 'tò 'to/for dogs'} \ b. \text{a b\text{vè 'tò 'with dogs'}} \)

As a result, \( \text{dogs e} \) is ill-formed after both prepositions. (The third preposition in the language, locative \( \text{á}'n \), has peculiar properties of its own and will not be discussed here.) Finally, note in (16) that while \( \text{eD} \) is well-formed in citation form NP's, it is ill-formed in corresponding vocatives:

(16) a. \( \text{t'-b\text{vè e 'dogs'}} \ b. \text{b\text{vè 'tò 'dogs'}} \)

I assume that (16a) is a defective S with a [-F] COMP, while (16b) is a defective S with a [+F] COMP, capturing the fact that vocatives are "out of focus"
like the backgrounded clauses I have identified as non-root.

REFERENCES


RAISING IN FUR

Janice L. Jake
University of Illinois

1. Introduction and Discussion of the Problem

In this paper I examine the syntax of sentences like those in (1) in Fur. The data are from the Zalingi dialect, from the speech of Al Amedin Abdel Moniem. I show that the underlined nominals in (1b) bear the Subject, 1, and Indirect Object, 3, relations in the matrix clause, although in (1a) they are constituents of the subordinate clause. The synonomy of the sentences in (1) means that, at some syntactic level, the underlined nominals in (1b) bear grammatical relations in the lower clause. (In the examples below I have not marked tone, as it is irrelevant to this paper. Although Fur has borrowed extensively from Arabic, its syntax remains unaffected.)

(1) a. mumkin i [alaŋ Eli dogola-si kitaba inni] possible be/3SG [COMP Ali children-OBJ books give/3SG] 'it is possible for Ali to give books to the children'

b. Eli-si kitaba mumkin ge [alaŋ dogola-si inni] Ali-OBJ books possible be/3PL [COMP children-OBJ give/3SG] = (a)

The analysis of sentences like (1b) is a problem for Relational Grammar, although there are a number of possible analyses in which more than one nominal can bear a relation in a lower clause and, in a later stratum, a higher clause. I show that a multiple ascension analysis of (1b) involves a violation of the Host Limitation Law, the Noninitial Term Demotion Ban, and the Relational Succession Law (for clauses). On the other hand, a Clause Union analysis of (1b) is empirically unjustifiable, as well as theoretically problematic.

In this paper I present evidence for the rules of Raising to 1 and Inversion to 3 and argue that, of the analyses noted above, only the multi-ascension analysis can be empirically motivated. This analysis, illustrated in (2), is one in which a nominal ascends to 1 and, in a later stratum, inverts to 3, and a second nominal ascends to 1. In stratum 4, a 1 serves as host for an ascension to 1, violating the Host Limitation Law, since the ascende fails to assume the relation of its host. Furthermore, a non-initial 1 inverts to 3 in stratum 3, a violation of the Noninitial Term Demotion Ban [Perlmutter 1984].

(2) Partial representation of (1b):

```
Raising
Inversion
Raising

Eli kitabe dogola inni mumkin ge
Ali books children give possible be
```
Although I present arguments for the analysis in (2), such an analysis entails weakening the theory of Relational Grammar because it necessitates the modification or loss of a number of relational laws. However, as noted above, there appears to be more support for this analysis than others. I now present evidence for (2) based on Raising and Inversion. Where relevant, I note how other possible analyses fail to adequately account for the data.

2. Evidence for the Analysis in (2)

2.1. Evidence for sentential 1's. Sentential complements in Fur occur as tensed clauses with overt complementizers, as in (1), and as nominalized clauses. A nominalized sentential 2 can be optionally suffixed with -si in normal pre-verbal position. This is parallel to the case marking of inanimate non-sentential 2's. Compare (3), with a non-sentential 2, and (4), with a sentential 2. An extraposed 2 is obligatorily suffixed with -si, as is illustrated in (5) and (6).

(3) ka kitab(-si) 'agal
I book(-OBJ) see/1SG
'I see the book'

(4) ka [Eli-ŋ joo(-si)] wEEl
I [Ali-GEN going(-OBJ)] want/1SG
'I want Ali to go'

(5) ka 'agal kitab-si/*Ø
I see/1SG book-OBJ/*Ø

(6) ka wEEl [Eli-ŋ joo-si/*Ø]
I want/1SG [Ali-GEN going-OBJ/*Ø]

= (3) = (4)

The examples in (3) through (6) show that sentential 2's are case marked like other inanimate 2's, optionally suffixed with -si preverbally and obligatorily when extraposed. The fact that nominalized sentential complements of predicates analogous to A-Raising and Tough-Movement predicates in English cannot be suffixed with -si provides evidence against their 2-hood. This is illustrated by the examples in (7) and (8).

(7) [Fanne-ŋ jam(*-si)] daruugi i
Fatma-GEN eating(*-OBJ) necessary be/3SG

'it is necessary for Fatma to eat'

(8) [Eli-ŋ nuuŋ-in jiro(*-si)] muškila i
Ali-GEN food-GEN cooking(*-OBJ) problem be/3SG

'it is a problem for Ali to cook food'

The examples in (7) and (8) show that these sentential complements cannot be suffixed with -si. In this respect they are like non-sentential 1's which are not suffixed with -si, as is illustrated in (9) and (10).

(9) Fanne (*-si) nyanja
Fatma (*-OBJ) come/3SG

Fatma is coming

(10) k-illa toŋa(*-si)
PL-DEM houses(*-OBJ) big be/3PL

'those houses are big'
To sum up, the facts about case marking of sentential and non-sentential complements show that the sentential complements under investigation are 1's rather than 2's. Note that unaccusative advancement is obligatory for these predicates, if it is assumed that they occur with initial unaccusative strata.

2.2. Evidence for Raising to 1. I now show that predicates occurring with sentential l's govern ascension of l's, 2's, 3's, Instruments, and Locatives. The second in each pair of examples in (11) through (15) illustrates the ascension of a nominal bearing one of these relations. In the (a) sentences, main clause verb agreement is third person singular, or neutral. In the (b) sentences, the raised nominal controls main clause verb agreement and precedes the verb, in subject position.

(11) a. mumkin ialal [ałąŋ dogolak ali-si kəgilEl] possible be/3SG [COMP children ali-OBJ see/3PL] 'it is possible for the children to see Ali'

b. dogola mumkin ge [ałąŋ ali-si kəgilEl] children possible be/3PL [COMP ali-OBJ see/3PL] = (a)

(12) a. sayl i [ałąŋ ali ustas-si jəgil] easy be/3SG [COMP ali teacher-OBJ see/3SG] 'it is easy for ali to see the teacher'

b. ustas sayl i [ałąŋ ali jəgil] teacher easy be/3SG [COMP ali see/3SG] = (a)

(13) a. terra i [ałąŋ Fanne ali-si ʃay inni] tough be/3SG [COMP Fatma ali-OBJ tea give/3SG] 'it is difficult for Fatma to give tea to Ali'

b. Ali terra i [ałąŋ-si Fanne ʃay inni] ali tough be/3SG [COMP OBJ Fatma tea give/3SG] = (a)

(14) a. sayl i [ałąŋ dogola ki murtəna kee] easy be/3SG [COMP children INST horses go/3PL] 'it is easy for the children to go by horse(s)'

b. murtəna sayl ge [ałąŋ dogola kee kee] horses easy be/3PL [COMP children INST go/3PL] = (a)
(15) a. mumkin i [alaŋ Eli madrasa-le jee]  
   possible be/3SG [COMP Ali school-LOC go/3SG]  
   'it is possible for Ali to go to school'
   
b. madrasa mumkin i [alaŋ Eli dele jee]  
   school possible be/3SG [COMP Ali LOC go/3SG]  
   = (a)

   In (13a) a 3 is obligatorily suffixed with -si. Because -si is suffixed to both 2's and 3's I refer to it as an objective case marker. There is, however, a difference between 2's and 3's, even though both are suffixed with -si. In a host out of which a 3 ascends, -si cliticizes to another constituent of the host clause, usually the complementizer, as in (13b). Ascension of a 3 without cliticization is ungrammatical. In contrast, ascension of a 2 with -si cliticization is ungrammatical. The ascension of an Instrument or a Locative is similar to the ascension of a 3 in that there is cliticization, kee for Instruments and dele for Locatives. The grammatical relation of an ascendee is always recoverable because 3 of the relations leave clitics in the host clause and 1 continues to control verb agreement in the host.

2.3. Evidence for inversion. The example in (16b) illustrates inversion with a member of the class of cognitive predicates; the inverted nominal occurs in objective case, obligatorily suffixed with -si. Verb agreement in the inversion clause is neutral.

(16) a. ka 'aI0n-si [alaŋ Fanne jee]  
   I think/1SG-SI [COMP Fatma go/3SG]  
   'I think that Fatma is going'
   
b. ka-si en-qi-s-10n  [alaŋ Fanne jee]  
   I-OBJ in-1/20-SI-think/3SG [COMP Fatma go/3SG]  
   = (a)

   The example in (16) shows that the cognitive predicate is also marked with -si. A raising predicate governing Inversion optionally occurs with -si, but only if there is Inversion, as is shown in (17).

(17) a. dogola-si mumkin(-si) i [alaŋ Eli-si kagi|El]  
   children-OBJ possible(-SI) be/3SG [COMP Ali-OBJ see/3PL]  
   'it is possible for the children to see Ali'
   
b. dogola-si terna(-s) i [alaŋ ustas jəgi]  
   children-OBJ tough(-SI) be/3SG [COMP teacher see/3SG]  
   'it is possible for the teacher to see the children'

   The clitic which occurs on predicates governing Inversion has an optional phonological alternate, -s, after vowel final morphemes. It can be distinguished
from the objective case marker on these and other grounds. However, the fact
that a raising predicate exhibits a clitic similar to that of other predicates
governing Inversion suggests that the -si suffixed nominals in (1b) and (17)
are 3's. This, together with the fact that any relation which can occur as an
ascended 1 can also occur as a 3, provides evidence for an Inversion analysis.
Empirical evidence for the 3-hood of these nominals is that a -si cliticizes
in a host out of which they ascend, as is illustrated in (18).

\[(18) \quad \text{kwa mumkin ge} \quad \text{[alaŋ-si sayl(-si) i [alaŋ Eli jagil]]}
\]
\[
\text{men possible be/3PL [COMP-OBJ easy(-SI) be/3SG [COMP Ali see/3SG]]}
\]
\['it can be easy for Ali to see the men'\]

3. Arguments Against other Analyses

The data presented thus far provides evidence against a clause union analy­
sis because both a 1 and a 3 need not occur in clauses with tough and possible
predicates and because it is not predictable which relation in the lower clause
will bear the 1 or 3 relation after clause union. That is, a lower 1 could be­
come a matrix 1 or 3, a lower 2 could become a matrix 1 or 3, a lower 3 could
become a matrix 1 or 3, etc. A clause union analysis consistent with these
facts would be unlike any other clause union in two ways: there would be no
predictable relation borne by a nominal in the lower clause and its relation in
the matrix, and either of the two matrix relations can occur independently of
the other.

Now I briefly outline why a different kind of multi-ascension analysis must
be rejected. The evidence is based on an animacy constraint on 3's. The ex­
amples in (19) illustrate that an inanimate nominal can ascend to 1 but it can­
not occur as a 3. This is in contrast to the examples in (1b) and (17) in
which an animate nominal occurs as a 3. Furthermore, there is no animacy con­
straint 2's, as is illustrated in (20).

\[(19) \quad \text{a.} \quad \text{ton} \quad \text{terra i} \quad \text{[alaŋ ka dele ſaw]}
\]
\[
\text{house tough be/3SG [COMP I LOC go/1SG]}
\]
\['the house is difficult for me to go to'\]
\[
\text{b.} \quad *\text{ton-si} \quad \text{terra i} \quad \text{[alaŋ ka dele ſaw]}
\]
\[
\text{house-OBJ tough be/3SG [COMP I LOC go/1SG]}
\]
\[= (a)\]

\[(20) \quad \text{ka ton-si} \quad \text{weEl} \quad \text{[alaŋ apa I]}
\]
\[
\text{I house-OBJ want/1SG [COMP big be/3SG]}
\]
\['I want the house to be big'\]

The facts illustrated in (19) and (20) mean that, if it is assumed that the
sentential complement of predicates like tough are 2's and a nominal can ascend
to 2 and retreat to 3, then an ad hoc filter must disallow all sentences with
inanimate 3's resulting from 2 to 3 retreat. Note that this retreat is obliga­
tory, since tough and possible predicates do not occur with 3's and that this
rule is otherwise unmotivated. Moreover, this analysis would also violate the Chômeur Advancement Law because the sentential $\hat{2}$ can advance to 1 and act as host for a second ascension to 1.

In conclusion, I have shown that a multi-ascension analysis involving a violation of the Host Limitation Law and the Noninitial Term Demotion Ban is the analysis of sentences like (1b) for which there is the most evidence. I have also shown that such an analysis is theoretically problematic for Relational Grammar. Perhaps another, less problematic, analysis of these sentences will become apparent upon further examination.

REFERENCE

**EPENTHESIS: A POSITIONAL TREATMENT OF SWAHILI PRONOMINAL CLITICS**

Camillia N. Keach  
Temple University

1. Introduction

In this paper, I will present an analysis which accounts for the phonetic distinction between two sets of pronouns in Swahili. The discussion is carried out in a framework which characterizes sentence grammar as a set of subcomponents: a base, a syntax, a phonology, and a semantics.

Swahili exemplifies a rich network of agreement morphemes whereby a pronominal clitic (PRO) associated with a particular noun class agrees with some NP position in the clause. The subject agreement morpheme (hereafter, SA) generally occurs verb initially. In the amba relative clauses in (1), there are two PROs: the SA agrees with the relativized subject, and the relative pronoun, suffixed to amba, agrees with the head of the relative.

(1) a. kisu amba + cho ki + ta + faa 'the knife which will do'  
   knife REL SA FUT suffice  
   b. yai ambalo litafaa 'the egg which will do'  

The SAs and their corresponding relative pronominal forms are given in (2).

(2)  

<table>
<thead>
<tr>
<th></th>
<th>SA</th>
<th>RELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>u</td>
<td>o</td>
</tr>
<tr>
<td>b</td>
<td>i</td>
<td>yo</td>
</tr>
<tr>
<td>c</td>
<td>zi</td>
<td>zo</td>
</tr>
<tr>
<td>d</td>
<td>ki</td>
<td>cho</td>
</tr>
<tr>
<td>e</td>
<td>vi</td>
<td>vyo</td>
</tr>
<tr>
<td>f</td>
<td>li</td>
<td>lo</td>
</tr>
<tr>
<td>g</td>
<td>ya</td>
<td>yo</td>
</tr>
<tr>
<td>h</td>
<td>pa</td>
<td>po</td>
</tr>
<tr>
<td>i</td>
<td>ku</td>
<td>ko</td>
</tr>
<tr>
<td>j</td>
<td>mu</td>
<td>mo</td>
</tr>
</tbody>
</table>

Here I propose that there's a rule, o Epenthesis, which systematically relates the SA and REL forms in (2) by attaching /o/ to any word which ends with a PRO.

(3) o EPENTHESIS  

\[ \emptyset \rightarrow o / [X \ [PRO] \ ____ ] X \]

Assuming that X is a variable ranging over preterminal nodes, and, for simplicity, that amba is a complementizer, (4) will apply to a form like (4a), yielding (4b).
2. The Case for Epenthesis.

In this section I show five cases of PRO final words. For an extensive discussion of the range of meanings associated with each form, I refer the reader to the references.

2.1. kwa. The preposition kwa may be followed by a lexical NP. When no lexical NP follows, a PRO is suffixed to kwa, an /o/ final PRO.

(5) a. kata nyama kwa kisu 'cut the meat with a knife'
   b. kata nyama kwacho 'cut the meat with it'
   c. *kata nyama kwaki

2.2. na. In its various semantic roles, the preposition na may be syntactically followed by a lexical NP. When no such NP follows, an /o/ final PRO is attached to na.

(6) a. karatasi zile, angalia ufike nazo papers those take care that you arrive with them 'see that you arrive with those papers'
   b. *karatasi zile, angalia ufike nazi

2.3. kuwa and locative noun phrases. An /o/ final PRO may be suffixed to the copula kuwa when a locative phrase is present or absent.

(7) a. nitakuwa hapa 'I will be here'
   b. nitakuwapo hapa 'I will be here'
   c. *nitakuwapa hapa

2.4. -inge-. The word -inge-, translated as "some NP of the same kind", agrees pre- and post-stem with the noun it modifies. Its suffix is always an /o/ final PRO.

(3) a. nione watu wengineo 'let me see some people of the same sort'
   b. *nione watu wenginewa

(9) Ah walimu wa Nairobi wa macho kwa jambo hili na jingelo. teachers of Nairobi have eyes for matter this and others of the same kind
   'Ah, the teachers of Nairobi are awake as regards this matter and any other like it.'
2.5. The general relative. In the general relative, there is no overt tense marker. The relative pronoun, an /o/ final PRO, is suffixed directly onto the verb stem.

(10) a. kazi i + tu + faa + yo work SA us suffice REL
    b. *kazi itufaai

We have seen five cases where an /o/ final PRO occurs word finally. Importantly, it does not occur exclusively in relative constructions.

3. The Traditional Position: + [already mentioned]

According to the traditionalists, the phonetic distinction between the two PROs reflects a semantic feature, [+already mentioned] (hereafter, AM). Haddon [1955] gives a succinct and fair summary of the position.

(11) "Referential Particle: This formative uses the 0 of reference completed with a pronominal concord to construct a particle, usually a suffix, which refers to something already mentioned or is understood." (p.9)

Evidence that AM obfuscates Swahili sentence grammar comes from word order. Because the subject NP normally precedes the verbal unit, it is, in fact, already mentioned. Yet the SA is never an /o/ final PRO.

(12) kisu kitafaa 'the knife will do'
    *kisu chotafaa

Even though the subject precedes the SA that agrees with it and hence is positionally already mentioned, the SA is never an /o/ final PRO.

It is not difficult to imagine what it means for a PRO to refer to something already mentioned or understood. But if AM is to have any content at all, we would expect PROs lacking the feature to be non-referential or to refer to something not already mentioned or understood. Since statements like (13) are completely acceptable in well formed Swahili discourse, PROs like SA would seem to be referential in exactly the same way as /o/ final PROs.

(13) killanguka 'it fell'

The appeal of AM may derive from its handy description of the near demonstratives, cases where /o/ final and non-/o/ PROs occur side-by-side. The forms in (14a) are the near speaker demonstratives and those in (14b) are the near listener demonstratives.

(14) a. kitabu hiki yai hili 'this book (near me)'
    b. kitabu hicho yai hilo 'this book (near you)'

However, by making the following assumption about lexical entries, it is possible to account for the near demonstratives without the feature, AM.

(15) Two words are separate lexical entries just in case there is at least one feature not shared by both words.

The unshared feature in (14) is based on proximity, nearness to the speaker or listener (or as Leonard [1980] would have it, on "concentration of attention"). Hence in accordance with (15), both demonstratives are separate lexical entries, remaining unstructured at the level at which ə Epenthesis applies and consequently inaccessible to the rule. Thus the demonstratives can be accounted for without AM.

In the light of this discussion, we can assume that both pronominal forms are referential, more appropriately a feature of discourse grammar. And we have seen that the appearance of /o/ final PROs is strictly a sentence grammar phenomenon.

4. Implications

ə Epenthesis predicts the non-occurrence of word internal /o/ final PROs. But in tensed relatives such a PRO is cliticized onto one of four tense affixes: -li- 'past'; -taka- 'future'; -na- 'present'; or -si- 'past, present, or future negative'.

(16) a. kitabu ki + na + cho + somwa 'the book which is (being) read'
    book SA PRES REL be read

b. mii + taka + yo + faa 'trees which will be suitable'
    tree SA FUT REL suffice

c. ya! lilo! anguka 'eggs which fell'

Taking the ə Epenthesis proposal seriously, the tensed relative is analyzed as two words: one formed around the verb stem, and the other around the tense morpheme where the relative pronoun ends a word. ə Epenthesis applies to the word final PRO, suffixing /o/, yielding the relative pronoun.

(17) a. TENSE \ TENSE PROC \ anguka
    V \ V
      TENSE \ TENSE PROC \ anguka
    \ lili \ \ lili

An important aspect of this analysis is that it imposes structure on the verb. Evidence for a structured verb comes from secondary stress, widely reported in traditional literature for relatives and non-relatives as well.

(18) kama ni + ngali + jua kusoma ni + ngali + nunua kitabu if SA COND know to read SA COND buy book

'if I had known how to read, I would have bought a book'
These stress facts may be accounted for nicely if the primary stress rules apply to two words, Verb Stem and the material dominated by the Tense node.

(19) \[ V \quad V_{+[stress]} / \quad \text{(C) (C) V (C) #} \]

After Tense Cliticization, a stress reduction rule follows yielding the well formed stress contours in (18) and (20).

(20) a. wall\posim\’ama 'where they stood'  
    b. vitabu all\vyovit\’aka 'the book that he wanted'

In summary, the stress analysis supports a structured analysis of the verb, a conclusion also reached by Wald [1973, 1976] and Givon [1971a,b] based on historical evidence.

5. Conclusion

The distribution of pronominal clitics in Swahili can be accounted for positionally without invoking a meaning difference between the two forms. Such an analysis is available within a framework having a level of sentence grammar.

[I am grateful to Mark Baltin, Nick Clements, Jane Grimshaw and Ken Hale for their helpful comments, and to my primary informants: Sheikh Yahya Omar, Khalfan Kasu and Samira Fakih. All errors I acknowledge as my own.]

REFERENCES


ON CORE SYLLABLES IN MODERN ARABIC DIALECTS

Michael Kenstowicz
University of Illinois

The grammars of most Arabic dialects generate CVCC syllables in addition to the universal CV and CVC. The construction of CVCC syllables is constrained by the sonority properties of the individual consonants in familiar ways. In a study of the phenomenon for Lebanese Arabic Haddad [1984] shows that the first consonant must be more sonorous than the second in order for them to cluster together in the coda, as shown by the denominal verbs of (1).

(1) CVCC       CVCC
    dârâs      'he studied'  dârâs  'a lesson'
    hâlim      'he dreamed'  hâlim  'a dream'
    hâmal      'he carried'  hâmil  'a load'
    dâfan      'he buried'   dâfin  'a burial'
    ramây-na   'we threw'    râmi   'throwing'

When the first postvocalic consonant is more sonorous than the second the underlying CVCC shape may emerge phonetically, e.g. dârâs, hâlim. When the consonants do not exhibit a falling sonority profile a syllabic rime cannot be erected under them. Languages seem to differ in their response to such unincorporable consonants. Some grammars delete the extrasyllabic consonants while others such as Arabic take the tack of incorporating them into an anaptyctic syllable by either vocalizing the consonant, e.g. râmi, or epenthesizing a supporting vowel, e.g. dâfin, hâmil. This paper considers exactly where and how the doubly closed CVCC syllables are formed in the grammar.

The first observation to be made is that the creation of the anaptyctic syllable must occur after the word stress has been assigned. Stress is assigned by the familiar principle that the rightmost heavy syllable is accented and otherwise the first syllable (with an antepenult limitation).

(2) hâmal      'he carried'  hâlim  'dream'    hâmil  'load'
    hâmâl-na    'we carried'  hâmâl-na  'our dream'  hâmîl-na  'our load'
    hâmâl-u     'they carried' hâmâl-ak  'your dream'  hâmîl-ak  'your load'

The nominal hâmîl-na appears to be an exception to the stress principle since the surface heavy penult has been skipped in favor of initial accent. Starting with Brame [1973] virtually all students of Levantine Arabic have interpreted these data as requiring an ordering of stress assignment before epenthesis (/hâmîl-na/ → /hâmîl-na/). If we are to maintain this ordering of the rules along with the assumption that epenthesis is a phonological response to extrasyllabic consonants then it follows that the phonological representation cannot be parsed exhaustively into syllables at a single point in the derivation.
Rather syllabification must take place in stages with the intermediate application of phonological rules. In particular, we must assume an initial parse into onset and rime categories to provide the proper representations for the application of stress assignment. Only after the stress has been assigned may the unincorporated consonants be organized into an anaptyctic syllable.

In our analysis CV and CVC are core syllables constructed in the lexical phonology by ordered rules assigning a vowel to a rime, a prevocalic consonant to onset, and a single postvocalic consonant to rime. Postlexical phrasal rules construct CVCC and anaptyctic syllables by assigning a string of extra-syllabic consonants to alternating rime and onset positions starting from the right. Adjacent rimes that exhibit a falling sonority pattern are collapsed together optionally; otherwise, vowelless rimes are supplied with a vowel through vocalization or epenthesis.

(3) \[ \text{Lexical} \]

\[
\begin{align*}
\text{hilm-na, himl-na} & \rightarrow \text{hilm-na, himl-na} & \rightarrow \text{hilm-na, himl-na} \\
\text{OR OR OR OR OR OR OR OR} \\
\end{align*}
\]

\[ \text{Postlexical} \]

\[
\begin{align*}
\text{hilm-na, himl-na} & \rightarrow \text{hlm-na, inappl.} & \rightarrow \text{inappl., himl-na} \\
\text{OR R OR OR R OR OR} & \text{OR OR} \\
\end{align*}
\]

In the rest of the paper we try to justify some of the assumptions underlying this analysis. First, it's clear that anaptyxis is postlexical since it operates phrasally [Abu-Salim 1980]. In Levantine Arabic there are contrasts such as verbal /fi}$\text{him}$ il$\text{walad} 'he understood the boy' vs. nominal /fi}$\text{him} il$\text{walad} 'the understanding of the boy'. Nominal /fi}$\text{him} derives from CVCC (cf. /fi}$\text{him-na 'our understanding'}). If anaptyxis is lexical, nominal /fi}$\text{him}/ will have merged with verbal /fi}$\text{him}/ before they are embedded phrasally, preventing us from explaining why the verb fails to lose its vowel. In fact, the i of the definite prefix is epenthetic as well, as we can see by comparing fi}$\text{him} il$\text{walad} 'understanding of the boy' with fi}$\text{him} lim$\text{sallim} 'understanding of the teacher'. These phrases receive the postlexical derivations of (4).

(4) \[ \text{fi}$\text{hm} il$\text{walad, fi}$\text{hm} lim$\text{sallim} + \text{fi}$\text{hm} il$\text{walad, fi}$\text{hm} lim$\text{sallim} + \]

\[
\begin{align*}
\text{I V} & \text{I V} & \text{I V} & \text{I V} & \text{I V} \\
\text{OR OR OR OR OR} \\
\text{OR R OR OR R OR} \\
\text{OR R OR OR R OR} \\
\end{align*}
\]

\[ \text{fi}$\text{hm} il$\text{walad, fi}$\text{him} lim$\text{sallim} \]

\[
\begin{align*}
\text{I V} & \text{I V} & \text{I V} & \text{I V} & \text{I V} \\
\text{OR OR OR OR OR} \\
\text{OR R OR OR R OR} \\
\text{OR R OR OR R OR} \\
\end{align*}
\]

The analysis sketched in (3) also claims that the m of hilm and himl-na is in the rime of the initial syllable lexically but in the onset of the anaptyctic syllable postlexically. The first claim is forced by the existence of the syllable building rule assigning a postvocalic consonant to the rime. This rule must apply lexically in order to create the branching rime structure crucial to the correct application of the stress rule. The grammar would have to be complicated considerably in order to prevent the application.
of this syllable-building rule to the m of /ḥiml/. Since this assumption is strongly implicated by the overall analysis, one would want to confirm it by looking at other lexical rules that are sensitive to the distinction of closed versus open syllables. However, it turns out that all attempts to verify this assumption for the Levantine dialect are thwarted by the lack of a crucial test case. The stress rule is sensitive to heavy versus light syllables and our analysis claims that the initial syllables of ḥiml and ḥiml-na are heavy at the point where stress applies. The initial stress of these words is consistent with our analysis but does not confirm it since the initial stress might follow from a representation in which there is no heavy syllable and hence is initial by default.

One of the advantages of working on a language as dialectally diversified as Arabic is that sometimes questions that can be posed but not answered for one dialect can be answered by looking at another dialect. In the Bedouin dialect of the Bani-Hassan of Jordan the definite prefix is al-. The vowel of this prefix may receive the primary word stress, as seen by the data in (5).

(5) wálad 'boy' bint 'girl'
al-wálad 'the boy' al-bint 'the girl'

Since the definite prefix counts for stress we can now test for the structure of the initial syllable of ḥiml. If it is closed in the lexical phonology, as our analysis predicts, then primary stress should not fall on the prefix. Before revealing the answer, let's look at some additional data bearing on our question. BRA possesses the common Bedouin rule raising short vowels in an open syllable. This rule is inhibited by adjacent gutturals or a following dental sonorant. We illustrate its effect on stems with the underlying shape CaCaC. These data also evidence the Bedouin rule eliding short a from an open syllable when followed by short a in an open syllable.

(6) ḥámal 'he carried' ḥmál-u 'they carried'
kṭab 'he wrote' kṭǔb-u 'they wrote'
wálad 'boy' wāl-d-ak 'your boy'
fāras 'mare' frās-ak 'your mare'

We now have two rules (stress and elision) sensitive to branching rime structure that can be utilized to test for the underlying syllable structure of CVCC nominals. Recall that our analysis claims that the initial syllable of such nouns is closed in the lexical phonology. We thus predict that in the Bani-Hassan dialect the initial syllable of the root will take primary stress (instead of the prefix) and will fail to show vowel raising. The data in (7) confirm both predictions.

(7) CVCC CVCC
    ḥilm 'he dreamed' al-ḥilm 'the dream'
    ḥámal 'he carried' al-ḥiml 'the load'
On Core Syllables in Modern Arabic Dialects

'difan 'he buried'  
ad-dafin 'the burial'

fuṭan 'he remembered'  
al-fāṭin 'the intelligent one'

To illustrate, our analysis assigns the verb dffan and the noun ad-dafin the lexical syllable structure of (8).

(8) dafan  
\[ \begin{array}{ll}
\text{III} & \text{V} \\
\text{OROR} & \\
\end{array} \]

al-dafn

Primary stress is assigned to the root syllable of the noun instead of to the prefix al- since, by hypothesis, it is heavy. The branching rime of the nominal also blocks raising, which does apply to the nonbranching rime of the initial syllable of the verb. Postlexical anaptyxis derives the epenthetic vowel of the noun, as in Levantine Arabic.

Our analysis in (3) also claims that there is resyllabification to onset position in front of the epenthetic vowel. Once again BHA provides supporting evidence. Like all Arabic dialects, BHA has resyllabification to onset across word boundary: dārab ōlwāḥad is pronounced [dāra.b ōlw.ā.had]. Now unlike in LA, BHA has extended the syncope of high vowels in open syllables to the phrasal level. While ṣfrīb il-māyya 'he drank the water' and ṣfrīb il-māyya 'drinking of the water' contrast in LA, they fall together in BHA as ṣfrīb al-māyya. We thus assume the derivation of (9).

(9) ṣfrīb al-māyya  
\[ \begin{array}{ll}
\text{IV} & \text{V} \\
\text{OR 0} & \\
\end{array} \]

As in LA, epenthesis operates at the phrasal level in BHA. /katab ktaab/ 'he wrote a book' is realized as [ka.ta.b ik.taab] by the postlexical rules assigning the initial consonant of /ktaab/ to rime position, followed by epenthesis, followed (by hypothesis) by resyllabification. We would have reason to treat resyllabification in front of the epenthetic vowel as a phonological, as opposed to mere phonetic, process if some phonological rule were sensitive to the derived representation. The fact that BHA /fihim ktaabu/ 'he understood his book' is realized as [fihm ik.taa.bu] shows that the phrasal syncope rule is sensitive to the open syllable created by resyllabification in front of the epenthetic vowel. We assume the derivation in (10).

(10) fihim ktaabu  
\[ \begin{array}{ll}
\text{I} & \text{IV} \\
\text{OROR OR ROR OR OR ROR OR ROR ROR ROR OR ROR ROR ROR ROR ROR} \\
\end{array} \]

The BHA phrasal syncope rule also permits us to show that CVCC syllables are constructed postlexically. An underlying string of the form CVCCIC will resyllabify the final consonant before a vowel, opening the final syllable, leading to the syncopation of the i, creating, we assume, an extrasyllabic consonant
out of the erstwhile onset. This consonant is treated in exactly the same fashion as the final consonant of a CVCC nominal, as the data in (11) show.

(11) yifrṣd 'he ambushes'
yifrṣd ṣalwalad 'he ambushes the boy' (cf. /dars/ → ḏārs 'lesson')
yīlmis 'he touches'
yīlims ṣalwalad 'he touches the boy' (cf. /ḥillm/ → ḥīlm 'dream')
yīlimis 'he smoothes'
yīlims ṣalaraṣ 'he smoothes the land' (cf. /ḥiml/ → ḥīmil 'load')
yīfnid 'he confirms'
yīfinid ṣalxabar 'he confirms the news' (cf. /dāfn/ → ḏāfin 'burial')

REFERENCES


1. Introduction

One of the features of Tangale morphology is the suffixation of genitive pronominals to noun radicals on the one hand and objective pronominals and/or tense-aspect markers to verb radicals on the other hand. When a pronominal or tense-aspect marker is suffixed to a noun or verb radical ending in a vowel, that final vowel gets elided. Final vowel elision is also observed in a suffix that is followed by another suffix and in the first of two successive words, provided the words belong to the same phrasal category.

(1) yáró + nò → yár-nò 'my arm'
  wùdó + dò → wùd-dò 'her tooth'
  sèèrò # Lǎkû → sèèr Lǎkû 'Laku's friend'
  dòb + ʒi + gò → dòb-ʒ-gò 'called you(f)'
  dòbú + Lǎkû → dòb Lǎkû 'call Laku!'

Final vowel elision plays an important role in Tangale phonology because it is associated with some phonological alternations, namely: vowel ~ zero, consonant ~ zero, vowel ~ glide, and voiced~voiceless obstruents (as well as some tonal alternations). In this paper, I attempt to formulate the rules underlying the alternations vowel ~ zero, consonant ~ zero, and vowel ~ glide.

1.1. Theoretical framework. The theoretical framework assumed in this paper is that of generative phonology enriched with the autosegmental representations introduced in recent work by Halle and Vergnaud [1980]. Under this framework, phonological units are represented in terms of separate tiers and relations between the tiers are specified in terms of association lines by the Well-formedness Condition:

(a) associate a vocalic segment with a V-slot and a consonantal segment with a C-slot
(b) association lines may not cross.

In this paper, phonological units are represented in terms of at least two tiers: the melody and the skeleton. The melody tier consists of the segmental units; the skeleton consists of the elements C and V representing the overall syllabic profile of the segmental string. The skeleton serves as the terminal point for the organization of the segmental units into Onsets and Rimes, the immediate constituents of the syllable.

As for the feature specification for the segmental units, I propose the in-
corporation of the shared feature convention proposed in Steriade [1982]. According to this convention, the feature(s) shared by the segments are percolated on a separate tier, and the percolated feature(s) and segmental units are linked by association lines. In this paper, the "Place Features" (PF) and the "Manner Features" (MF) are represented on separate tiers. The feature tiers and the segmental tier are linked to each other by association lines. The association lines are also subject to the "no crossing of association lines" constraint. The proposed phonological representation is illustrated in (2a,b).

(2) a. PF [+ant][+hi][+ant] b. PF [+ant][−hi] [+ant] [−hi]
    MF [−son][+rd][+nas] MF [+son][−rd][+nas] [−nas][−rd]  
    s u m a n d a  
    a a a  
    C V C C C C

In (2b), the use of the convention allows segments that share the same PF to be linked to the same feature matrix. The linking of the segments to the features occurs in the lexicon. As will be seen, the convention plays a significant role in the formulation of one of the rules.

1.2. The goals. My goals in writing this paper are three. One, it is to show that the vowel ~ zero, consonant ~ zero, and vowel ~ glide alternations are to be accounted for by the rules of epenthesis, consonant elision, and vocalization respectively. Secondly, and most importantly, I wish to show that the formal statements of the rules in question are simplified and the motivations for the rules become more transparent when the rules are viewed as operating on representation in which segments have been organized into syllables. Thirdly, I will show the utility of the shared feature convention in providing explanation for the consonant elision rule.

2. Epenthesis

Tangale allows the following syllable structures: V, VV, VC, VVC, CV, CVV, CVC, and CVVC, e.g. ḍo 'eye', Ṽu 'run', ḡk 'body', ḍom 'blood', and Ṽom 'strong'. Tangale thus disallows initial and final clusters as well as medial three consonant sequences. When violations of these phonotactic constraints arise as a result of the concatenation of morphemes in underlying forms, an epenthetic u/조사 is inserted to break up the impermissible consonant cluster.

(3) wǒd-ꔼ-gó → wǒdúɡó 'bit you (f)'
    wǒd-g'lawọ → wǒdūɡ lawọ 'bit a child'
    kór-n-gó → kórʊŋgó 'protected'
    kór-n-ŋ-gó → kórʊŋŋgọ 'protected me'
Morpheme Alternation in Tangale

In (3) we observe that, for a CCC sequence, \( /u/\) is inserted between the first and second consonants while for a CCCC sequence, the vowel appears between the second and third consonants. For a CCCC cluster, the vowels occur in two positions: between the first and second consonants and between the third and fourth. As for the word-final cluster, the vowel is inserted between the ultimate and penultimate consonants.

It is possible to formulate an epenthesis rule within the linear framework. Under this framework, two types of epenthesis rules, final and medial epenthesis, are required.

(4) FINAL EPENTHESIS

\[
\emptyset \rightarrow \ V \ / \ V\\
\begin{array}{l}
+\text{high} \\
+\text{back} \\
\text{aATR}
\end{array}
\]

(5) MEDIAL EPENTHESIS

\[
\emptyset \rightarrow \ V \ / \ V \ C\_CC \\
\begin{array}{l}
+\text{high} \\
+\text{back} \\
\text{aATR}
\end{array}
\]

Condition: Apply iteratively from right.

The two epenthesis rules are ordered in such a way that final epenthesis must precede medial epenthesis since the reverse order will generate \(*k\text{g}r\text{yn}\text{ug}\) for \(k\text{g}r\text{yn}\text{ug}\).

Let us now consider the syllable approach adopted in this paper. Under this approach, the epenthesis rule is viewed as a reflex of the general scheme for imposing syllable structure on a successive string of consonants. This scheme is achieved via the syllable building rules (SBR) which are ordered among themselves and among other phonological rules in the grammar.

The two universal SBRs that assign (i) a V-slot to a Rime (R) and (ii) a prevocalic C-slot to an Onset (O) apply in Tangale. In addition, Tangale has the following SBRs: (iii) assign a string of one or more unassigned C-slots to alternating R and O starting from the right end of the string; (iv) if R fails to dominate a V-slot, adjoin a V-slot as a left daughter to the R; (v) assign and link, to a V-slot that is unlinked to a segment and feature matrix, the feature \([+\text{rd}]\) on the MF tier, and the features \([+\text{hi}]\) on the PF tier; then spell \([+\text{bk}]\)
out these features as U on the segmental tier. (Vowel harmony rule converts the U into u or u). These SBRs are followed by resyllabification. Using the SBRs we derive ָוָֹוָֹוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָוָو
3. Consonant Elision

A consonant elides if it is homorganic with a preceding consonant provided the two consonants are homomorphemic. The elision phenomenon is exemplified in (8).

(8) 
\begin{align*}
\text{d'nd} & \rightarrow \text{d'niy} & \text{'your(f) pot'} \\
\text{d'nd} \text{mala} & \rightarrow \text{d'ni mala} & \text{'Malai's pot'} \\
\text{k\textcircled{lo}d} & \rightarrow \text{k\textcircled{lo}m\textcircled{u}} & \text{'our sheep-pen'} \\
\text{s\textcircled{mb}} & \rightarrow \text{s\textcircled{mg}} & \text{'met'} \\
\text{k\textcircled{w\textcircled{a}}\text{\textcircled{l}}} & \rightarrow \text{k\textcircled{w\textcircled{a}}\text{\textcircled{g}}} & \text{'tied'} \\
\text{p\textcircled{n\textcircled{n}}} & \rightarrow \text{p\textcircled{n\textcircled{y}}} & \text{'knew me'} \\
\text{p\textcircled{n\textcircled{d}}} & \rightarrow \text{p\textcircled{n\textcircled{y}}\textcircled{d}} & \text{'knew thoroughly'} \\
\text{k\textcircled{e\textcircled{m\textcircled{d}}} & \rightarrow \text{k\textcircled{e\textcircled{m\textcircled{y\textcircled{d}}}} & \text{'filled'}
\end{align*}

However, these consonants do not elide in the forms in (9).

(9) 
\begin{align*}
\text{s\textcircled{mb}} & \rightarrow \text{s\textcircled{mb\textcircled{y}}} & \text{'met me'} \\
\text{s\textcircled{mb\textcircled{g}'k\textcircled{a}l}} & \rightarrow \text{s\textcircled{mb\textcircled{y}g} k\textcircled{a}l'} & \text{'met Kai'} \\
\text{m\textcircled{\textcircled{m\textcircled{\textcircled{d\textcircled{g}}}}} & \rightarrow \text{m\textcircled{\textcircled{d\textcircled{y\textcircled{m\textcircled{g}}}}} & \text{'forget us'} \\
\text{k\textcircled{\textcircled{w\textcirccircled{a}}\textcircled{l}}\text{\textcircled{g'}k\textcircled{d\textcircled{d}}k}} & \rightarrow \text{k\textcircled{\textcircled{w\textcirccircled{a}}\textcircled{l\textcircled{y\textcircled{g}} k\textcircled{d\textcircled{d}}k}} & \text{'tied a rope'}
\end{align*}

In (9), consonant elision fails to apply where epenthesis has applied. We could formulate a consonant elision rule within the linear approach as in (10).

(10) 
\[
\begin{array}{c}
\text{[+cons]} \\
\text{[+ant]} \\
\text{[\textcircled{Ecor}]} \\
\end{array} \rightarrow \emptyset / + C _{\text{ant}} + \text{C} _{\text{Ecor}}
\]

Considering the forms in (9), consonant elision must be ordered to apply after final epenthesis (4) and medial epenthesis (5).

Let us now consider consonant elision within the syllable approach adopted in this paper. Like epenthesis, consonant elision is viewed as another process by which syllable building in Tangale is achieved. At this point, the shared feature convention plays an important role in the syllable building in respect of the forms in (8). In (11a,b) we illustrate the significance of the effect of the shared feature convention (see example on following page).

In (11b), we successfully apply SBRs i-iv, but SBR v is blocked because linking the features \([+\text{hi}]\) on the PF tier to the melody and the unlinked V-slot results in violation of the well-formedness condition, "no association lines may cross". In Tangale the ill-formed structure from SBRv can be repaired. The repair includes the erasure of the ill-formed association line, which in effect wipes out the unlinked V-slot. We are thus left with the struc-
ture we started with in (11a). Resyllabification applies leaving a consonant unattached as shown in (11c).

(11) a. PF
\[
\begin{array}{l}
\text{MF} [+\text{nas}] [-\text{nas}] \\
\text{SBR} 1-v \\
\text{MF} [+\text{nas}] [+\text{rd}] [-\text{nas}] \\

\hline
\hline
\text{din} \quad \text{dzi} \\
\text{CV} \quad \text{CCV} \\
\text{ORO} \quad \text{ORO} \\
\hline
\end{array}
\]

b. PF
\[
\begin{array}{l}
\text{MF} [+\text{cor}] \\
\text{SBR} \quad \text{b} \quad \text{cor} \\
\text{MF} [+\text{nas}] [-\text{nas}] \\

\hline
\hline
\text{din} \quad \text{dzi} \\
\text{CV} \quad \text{CCV} \\
\text{ORO} \quad \text{ORO} \\
\hline
\end{array}
\]

c. PF
\[
\begin{array}{l}
\text{MF} [+\text{cor}] \\
\text{SBR} \quad \text{d} \quad \text{cor} \\
\text{MF} [+\text{nas}] [-\text{nas}] \\

\hline
\hline
\text{din} \quad \text{dzi} \\
\text{CV} \quad \text{CCV} \\
\text{ORO} \quad \text{ORO} \\
\hline
\end{array}
\]

In conformity with the syllable structure in Tangale, the extrasyllabic conso-
nant elides.

The use of the shared feature convention may appear cumbersome. However, it
has the following advantage: because the shared feature is assigned in the lex-
icon, it allows for the distinction between homomorphemic and heteromorphemic
homorganic consonants (cf. the difference between the homorganic consonants in
\text{diniz} \quad \text{and} \quad \text{pōŋydgo} ). The distinction between these consonants made by the
shared feature eliminates the use of morpheme boundary in the statement of con-
sonant elision rule.

4. Vocalization

In Tangale glides alternate with vowels in the forms in (12).

(12) indefinite \hspace{1cm} definite \hspace{1cm} possessive 'my'
\begin{align*}
\text{āṭā́y} & \quad \text{āṭáw-} & \quad \text{āṭáu-} & \quad \text{'mat'} \\
\text{pōŋgɔ́} & \quad \text{pōŋgɔyll-} & \quad \text{pōŋgɔoll-} & \quad \text{'dish'} \\
\text{yəyə} & \quad \text{yəy-} & \quad \text{yaT-} & \quad \text{'farm'}
\end{align*}
Morpheme Alternation in Tangale

<table>
<thead>
<tr>
<th>indefinite</th>
<th>definite</th>
<th>possessive 'my'</th>
</tr>
</thead>
<tbody>
<tr>
<td>ləwò</td>
<td>ləw- PureComponent</td>
<td>lə́y-nọ</td>
</tr>
<tr>
<td>təgyò</td>
<td>təgy- PureComponent</td>
<td>təgyi-nọ</td>
</tr>
<tr>
<td>kūrwā</td>
<td>kūrw- PureComponent</td>
<td>kūruu-nọ</td>
</tr>
<tr>
<td>yáúll</td>
<td>yáúl- PureComponent</td>
<td>yáwul-nọ</td>
</tr>
<tr>
<td>pājdo</td>
<td>pāj- PureComponent</td>
<td>pāyö̃-nọ</td>
</tr>
</tbody>
</table>

Using the linear approach, we could express vocalization as in (13).

\[
\begin{align*}
\text{[-cons]} & \rightarrow \left[\begin{array}{c}
\text{[+syll]} \\
\text{[ard]}
\end{array}\right] / \{C\} \\
\text{[-syll]} & \rightarrow \left[\begin{array}{c}
\text{[ard]}
\end{array}\right] / \{\emptyset\}
\end{align*}
\]

Considering the forms such as yáwul-nọ and pāyö̃-nọ, the vocalization rule must be ordered to apply after the final and medial epenthesis rules.

From the point of view of the syllable theory, the glide - vowel alternation in Tangale is simply a matter of the two distinct phonetic realizations of a non-syllabic and non-consonantal segment in Onset and Rime positions. In Onset position, this segment is realized as a glide; while in Rime position, it is realized as a vowel. Because the forms such as yáwul-nọ and pāyö̃-nọ favor the analysis with glide in UR, the forms in (12) are accounted for by vocalization. This rule may be stated as in (14).

\[
\begin{align*}
\text{[-cons]} & \rightarrow \left[\begin{array}{c}
\text{[+syll]} \\
\text{[ard]}
\end{array}\right] / \{\emptyset\} \\
\text{[-syll]} & \rightarrow \left[\begin{array}{c}
\text{[ard]}
\end{array}\right] / \{\emptyset\}
\end{align*}
\]

Utilizing the SBRs and vocalization rule, pōgōl lə́y-nọ, kūruu-nọ and yáwul-nọ receive the derivations in (15a-d), respectively.

(15) a. pōgōl lə́y-nọ c. kūruu-nọ d. yáwul-nọ

\[
\begin{align*}
\text{SBRs} & = \text{i-v}
\end{align*}
\]
REFERENCES


There are words in Kinyarwanda such as kugabura 'to feed' and kujabura 'to feed uncooked food' which differ in only one sound, i.e. g and j in this case. These pairs of words are also semantically related. The difference in meaning is very slight. This slight semantic difference is caused by the minor phonetic difference. This linguistic phenomenon is not to be confused with "phonetic variation" whose function is only "emotive". These are "doublets". The existence of doublets in language is a case of linguistic "iconicity". It is an isomorphism: a one-to-one relationship between form and meaning. Thus the more similarity there is between two referents, the more similarity on their respective signs is found. The existence of doublets supports Peirce's theory of semiotics which claims that signs are not arbitrary, but start as icons or indices and later on become symbols because of semiosis. This paper is not only interested in iconicity in language, but is also concerned with the process of doublet formation and how one can tell its directionality. I argue that the markedness theory has to be fully developed because in many cases whether it is doublet formation obtained through borrowing or derivation, it seems to help a lot.

1. Doublets in Kinyarwanda

Doublets in Kinyarwanda differ in only one sound. In the examples provided, these sounds are underlined:

(1) kugabura 'to feed'  kujabura 'to feed uncooked food'
kuzerera 'to wander'  kuyerera 'to wander aimlessly'
gukona 'to castrate'  gukuna 'to elongate the vagina's lips'
kuyoberwa 'not to know'  gushoberwa 'not to know what to do'
gushiguka 'to stand up suddenly'  gushibuka 'to release suddenly' (intransitive)
gukwegura 'to stretch resulting in lengthening'  gukwetura 'to stretch resulting in widening'
gukamura 'to squeeze wet or juicy things'  gukanyura 'to squeeze firmly'
gukokora 'to prune a tree'  gutotora 'to prune vegetables'
gucyeha 'to level horizontally'  gucyaha 'to level vertically'
gushinjagira 'to walk elegantly with dignity'  kujinjagira 'to walk awkwardly'
Studies in African Linguistics, Supplement 9

kibuta 'tall and strong'  kibata 'short and strong'
kuzunguruka 'to go around'  kuzenguruka 'to circle'
kureka 'to spring' (water)  kureta 'to foam'
gukaraga 'to rotate something in the hands'  gukoraga 'to stir milk with a stick'
gusenga 'to pray'  guseng(era) 'to court a powerful individual'
kurondora 'to tell in detail'  gutondora 'to arrange side by side'
kumanunya 'to break into pieces'  gutanyura 'to break off'
kuvika 'to put something in the water to prevent it from falling out'

umugaba 'military general'  umugabe 'monarch'
kununza 'to drink dregs with a straw'  kunyunyaza 'to suck/to exploit'
ubwenge 'knowledge'  ubugenge 'tricks'
kwonka 'to suck'  gukonka 'to suck into juicy fruits'

Triplets and quadruplets are also found as the examples in (2) show:

(2) a. gufubika 'to cover in order to cause warmth'
gufunika 'to put a lid'
gupfunyika 'to put in a bag'

b. kunyeganyega 'to shake (big object)' (intransitive)
kuregarega 'to shake from the handle'
kujegajega 'to shake (small object)'
kuyegayega 'to move very slowly'

Obviously, examples of doublets just given are not "phonetic variations" of words. Each has its own independent referent. As Coupez [1975] has correctly observed about Bantu languages, "lexical variability" has an "expressive function" or to use Jakobson's [1971] terminology an "emotive function". Thus kugengesera, kubungsera, gusengesera in Kinyarwanda are individual phonetic variations of the same referent "to carry with great care". The referent "very small" has 40 phonetic variations [Coupez 1975] which don't in any way affect the meaning:

(3) kaniya, kanzinya, kaninya, kanzizinya, kaniginya, kanigirinya, kanyigi, kanyigiri, kanyigiriri, kanuya, kanzunya, kanyuyu, kanuyu, kanzunyu, kanuru, kanunuya, kanyunyuya, kanunuyu, kanzuzunyu, kanunuru, kanyunyuyu, kanzuzurunya, kanzuzurunyu, kanuguya, kanzugunya, kanugunya,
Doublets in Kinyarwanda

The phonetic variation doesn't change meaning but gives information on the psychological state of the "interpretant": happy mood, anger, intimacy, formality, dialect, social class, gender, etc. At the formal level, phonetic variation and doublets look alike since they don't differ greatly from the basic word. The difference is that doublets refer to distinct objects and thus obey the "one meaning one form" principle of semiotics whereas in the case of phonetic variation, it is "one meaning with different forms" although each form carries its own "emotive" connotation.

2. Doublet Formation

Doublet formation resembles other types of sign-production: regular morphological derivation (prefixation, infixation, suffixation, zero-derivation), compound-word formation, blending (smog, brunch, motel), and tropicality (metaphors and metonymies). They can be coined consciously by language professionals as the French examples in (4) indicate where words have been taken from Latin to express concepts closely related to the ones conveyed by older words of Latin origin also:

(4) nager 'swim' naviguer 'sail'
écouter 'listen' ausculter 'sound with a stethoscope'
mâcher 'chew' mastiquer 'masticate'
frêle 'frail' fragile 'weak/fragile'
âcre 'sour' aigre 'tart'

Regular doublet formation starts like phonetic variation, but because of the "one meaning one form" principle of semiotics, the emotive connotation becomes "referential". This referentiality is also tropical. There is an association between the initial referent and the derived referent. More precisely it is a similarity. This similarity is not metaphoric but diagrammatic. In many cases referents of metaphors have other signs. In the case of doublets, the new referent and its sign don't have an independent existence. Both the signans (signifier, form) and the signatum (signified, content) of the new sign develop from the signans and the signatum of another sign. Thus in Kinyarwanda, the referents 'to prune a plant' and 'to drink dregs with a straw' created the referents 'to prune vegetables' and 'to suck blood' respectively. Signs which refer to the new concepts were also created from the signs of the initial referents. Doublet formation took place as schematized in (5).

(5) gukokora x gukokora
'to prune' 'to prune' gukotora
plants' vegetables' 'to prune' 'to prune'

kununuza x kununuza
'to drink' 'to suck' kunyunyuya
'dregs' 'blood' 'to suck' 'blood'
When a sign starts acquiring two signata which are semantically related, a new signans is created for it which is phonetically related to the signans of the initial referent. A lot of doublets are also obtained through borrowing from cognates of sister languages as English words shirt/skirt or Kinyarwanda impeta 'ring/traditional military decoration' / ipeta 'military medal' (from Swahili) show. The language can also borrow from itself (archaic language, jargons, poetic language, ritualistic language, slang, etc.) or from its different dialects.

3. Directionality of Doublet Formation

The task of historical linguistics and semiolinguistics in general since it is concerned with sign production, creativity, and dynamism is to establish criteria that help to tell which sign is the original one and which one is derived. The "markedness" theory greatly helps. The sign which is marked in usage, in meaning or on form or on both levels is usually the derived one.

2.1. Markedness on the form. Words which have extra markers (segments, syllables, or tones) are newer. Thus in Kinyarwanda kwonka 'to suck' and ubwenge 'knowledge' might be older than their respective counterparts gukonka 'to suck in skins of juicy fruits' and ubugenge 'tricks' because the latter have extra consonants k and g. In English history must be newer than story, although both of them come from Latin. The former is marked by hi-. It is the same thing with charm and charisma. The latter is more marked on the form. In some cases, the derived doublets carry suprasegmentals that the initial words don't have. In Kinyarwanda, when doublets are equally marked (same number of consonants and vowels), the new word carries a high tone in addition:

(6) kuyoberwa [kuyoberga] 'not to know'
gushoberwa [gûšoberga] 'not to know what to do'
kurondora [kuroon do ra] 'to tell in detail'
gutondora [gûtöondo ra] 'to arrange side by side'

3.2. Markedness on the content. When doublets don't have any distinguishing markers on the form, the meanings can help tell which one is older and which one is derived. The word which has a more general meaning is usually older. In other words, the derived sign has a specialized meaning and can therefore be defined by the former. For instance, both gukokora 'to prune a plant' and gutotora 'to prune vegetables' are equally marked, but gukokora must be older than gutotora because the latter can be defined by gukokora, which has a more general meaning. Kujabura is 'kugabura uncooked food', history is 'a documented or written story', etc.

3.3. Markedness in usage or frequency. Words which have a higher frequency or are used more often are usually older and those which have a specialized meaning or low frequency are newer. Since warranty is only used in legal terms, it is newer than guaranty; turtle is newer than tortoise since it refers to a specific type of tortoise. Words which have marked sounds are also newer. Since the pal-
atal nasal ny is more marked than the bilabial nasal m, one can tell that kunyunyuza 'to suck blood' derives from kununuza 'to drink dregs' and gukanyura 'to squeeze firmly' from gukamura 'to squeeze wet or juicy things'. j is more marked than g and y more marked than z. Therefore, kujabura 'to feed uncooked food' and kuyerera 'to wander aimlessly' are newer than kugabura 'to feed' and kuzerera 'to wander' respectively.

There are doublets which are equally marked on the form, meaning, and usage for which the markedness theory fails to give the right solution, however. Consider the following examples:

(7) gushiguka 'to stand up suddenly' gushibuka 'to release suddenly'
gukwegura 'to stretch resulting in lengthening' gukworera 'to stretch resulting in widening'
gucyeha 'to level vertically' gucyaha 'to level horizontally'

There are cases of "markedness reversal" also in which the form, the meaning, or the usage of the new sign has become more generalized than that of the initial sign. Sometimes, doublet formation can result by accident. For instance the Kinyarwanda doublets gucukura 'to dig' and gufukura 'to dig a well' were obtained by accident. Gufukura is a denominal verb from the word ifuku 'mole-skin'. A moleskin digs holes in which it hides its food and things it has stolen. The verb gufukura has thus been derived metonymically. "To dig" is to behave like a moleskin. In English also, ridiculous and ludicrous are considered as doublets, but they have a different historical origin. The former comes from Latin ridiculus 'laughable' and the latter from ludus 'game'. Without the knowledge of the history of the sign or the language, the markedness theory may fail to predict the directionality of derivation. But its failure might be due to the fact that it is not fully developed.

REFERENCES


THE ma- PREFIX IN AFROASIATIC

Karen Lahaie
Stanford University

Joseph Greenberg [1963:48] noted in Languages of Africa that "Hausa and certain other Chad languages have a m- prefix which forms nouns of place, instrument and agent. This prefix occurs very commonly in Semitic, Egyptian, Berber and Cushitic with the same general range of meaning." Since this m- prefix is found in all five sub-branches of Afroasiatic, we can assume its presence in Proto-Afroasiatic as well, just as similar case systems in various Indo-European languages indicate that Proto-Indo-European had cases. Whether or not some of the daughter languages have lost some or all of a construction present in the protolanguage is of little importance as long as the construction is found in a widely-distributed sample of daughter languages.

In the Berber branch of Afroasiatic, -ma- forms agents from verbs, for example a-ma-ssaal 'porter' from asa 'carry' and a-ma-rir 'singer' from urar 'sing' (a- is the masculine singular pre-prefix). In Egyptian we have the instrument mb't 'scales' from h 'measure', as well as possible agents and locatives in m-. In Coptic, locatives are quite productively formed with a ma- prefix. In the Cushitic branch of Afroasiatic, where the ma- prefix is rare, we do find it in Bedauye, for example, mignaf 'camping place' from giniif 'cause to barrack'. In Chadic, the ma- prefix is very productive in Hausa for agents, instruments, and locatives, with some evidence for it in other Chadic languages as well. In Semitic, we find locatives and instruments in ma-, and agents as active participles of derived verb forms in mu- have been reconstructed as Proto-Semitic [Moscatì et al 1964].

Several theories have attempted to explain the ma- prefix. Semiticists such as Bravmann have theorized that it arose from a relative pronoun *maa (and there is a limited relative pronoun ma/maa in Semitic and Berber). Another theory is that *ma- originally meant 'place' (and we do find a word ma meaning 'place' in Coptic [Crum 1939] and a Chadic language, Tera [Newman 1970]). It has also been suggested that *ma was once a pronoun (and we do have a pronoun ma in two Chadic languages, Angas and Margi). The present theory can account for all three of the above theories, while differing from them in one important respect: it does not assign the agentive, instrumental, and locative meaning to the ma- at all, but rather to the changes in the stem and suffix of the ma word itself.

In all Afroasiatic languages that have a ma- prefix for agent, instrument, and locative, there is a distinction among them in their stem as well as their suffix. In Hausa, for example, a tone difference on the stem differentiates agents from instruments and locatives, while a suffix difference sets locatives aside from both agents and instruments [Kraft and Kirk-Greene 1973, Newman and Newman 1977, Cowan and Schuh 1976]:
agent  ma'aunii  'one who sells grain by measure'
instrument  ma'aunii  'scales'
locative  ma'aunaa  'place where selling by measure is going on'

All from  'aunaa  'buy by measure'.

This pattern has become completely regularized in Hausa and is very productive. Margi, another Chadic language, has a more productive way of forming agents, but still has relics (Hoffmann [1963] wasn't sure how productive) of this formation in m'tlalabu 'pupil' from tlalabu 'learn'. Note the -i suffix and low tones on the stem (although a certain class of verb may take the opposite tones). Hoffmann also notes some locatives in Margi, such as m'cifubu 'corral, fence' from cibuy 'to fence in', m'ontu 'crack' from ntu 'to crack', m'ndalal 'escape hole' from ndalu 'to ooze'. In these we see suffix vowels other than -i. In other Chadic languages that have relics of agents, such as Angas (mwam 'caterpillar' from wam 'to spoil' [Foulkes 1915]), the final vowels dropped from nouns in general, and not enough information is available to determine what might have once differentiated the instrument, agent, and locative. Presumably only one of the three meanings remained in many languages due to phonological changes that eliminated the distinctions between their forms.

In Semitic, we find that Arabic [Wehr 1971, Abboud et al 1975] distinguishes instrument from locative much as Hausa does, replacing locative -a- with 'agentive' -i-: m'nalb 'cooking utensil' vs. m'alb 'kitchen' from jababa 'to cook'. Arabic active participles of the simple verb stem function as agents without the ma- prefix, but active participles of the derived verb forms (intensive, reflexive, causative, etc.) have a characteristic mu- prefix with -i- between the last two radicals, which goes back to Proto-Semitic [Moscati et al 1964]. This seems to correspond to the Chadic "agentive" -i found in Hausa agents and instruments. In addition, the agent seems to be based on a u-stem in Arabic (although this is probably based on the imperfect in Semitic; the imperfect, in turn, was perhaps made up of a -u- element). While no one knows for certain, there is evidence for a *w or *u copula, which became a passive marker in some language groups (Chadic -u passive in Hausa, Berber -t-w-passive in which -t- seems cognate with the reflexive Semitic -t-, and Semitic -u- passive; the Cushitic wuu copula (Central Somali) and Egyptian copula lw [Gardiner 1951:461-463]). In any case, it seems that Hausa agents are also based on the -u stem, which has all low tones, but here the -u is replaced by the "agentive" -ii suffix.

The instruments and locatives seem to share a stem type. In Arabic this is the causative (form IV) stem, which is usually in the perfect: tählab for jababa. -? seems to be a reflex of Proto-Afroasiatic *5- causative [J. H. Greenberg, p.c.], as is the -r suffix in Hausa. In Arabic, ma- plus -i- (instruments) or -aa- (infixed here?) for locatives replaces the causative marker -? but uses the causative stem. Surprisingly enough, Hausa also uses the causative stem, i.e. all high tones, but the "agentive" suffix for instruments, or the -aa suffix for locatives replaces the causative marker -r. Use of the perfect stem in Semitic does not surprise us, since the Arabic perfect was once a nominal suffix conjugation used to record a state or condition [Moscati et al 1964:133]. Some instruments in Arabic are formed with the verbal noun stem of
the causative verb: ?iftaab from fataha 'open', miftaab 'key'.

If the root and suffix alone can determine whether a deverbal noun is an agent, instrument, or locative, we would expect to find these types of words without a ma- prefix as well. As a matter of fact, they are not hard to find. In Semitic, the active participles, i.e. agents, of the simple verb form do not have the ma- prefix, but rather a distinctive aa-i vocalization: kaatib 'writer' from kataba 'write' (Arabic). In Egyptian, which usually didn't express vowels, we do find an -y suffix which later became -amy forming active participles which were either human or non-human agents: sāmi 'hearer, judge', later written sāmy 'one who listens'. In Berber, causative stems form instruments and locatives without a ma- prefix (the Berber s- causative prefix is used, another reflex of Proto-Afroasiatic *s- ). According to Prasse [1973,IV: 68], the locatives are plural formations of the instruments.

If ma- is unnecessary in these constructions, and does not carry the meaning as previously thought, what is its origin and function? It so happens that ma reflexes are prepositions throughout Afroasiatic. In Egyptian, a preposition expressed only by m meant 'into, in, from, on, at, with, out from, as, like, according to'. Prefixed to other particles, new prepositions were formed, such as m' ( m + ayn, the sign for 'hand') 'by the hand of, by means of, because'. The prepositions in Semitic, such as Hebrew ʾālm, Arabic maa, Ugaritic ʾam, seem cognate with this Egyptian compound form and mean 'with'. The simple *ma preposition seems to have been replaced by the h- and l- prepositions in Semitic, but some of its derivatives remain. In Berber ma means 'as for' [Basset 1948], which can be paraphrased by prepositions such as 'to' or 'about'. In Chadic, the Dghwede (Wandala group) suffix -ma means 'in, into'. Angas ma means 'of, into', and most Chadic languages have indirect object forms of the form *ma + dependent pronoun which have undergone vowel harmony and erosion: Angas mfa 'of me', munu 'of us'; Hausa mna 'to/for me', mana 'to/for us'. While ma is still found as the preposition 'to' dialectally in Hausa, it is being replaced, particularly in the standard dialect, by wa before nouns. In fact, the possessive marker mā in Hausa, which means 'owner of' before nouns and is an agentive marker before verb phrases, was originally ma + yā, that is to say 'to/for him'. There is even some possible connection between the -i: agentive suffix and ya (the third person, masculine singular verb prefix). This verb marker is found throughout Afroasiatic (cf. Semitic y-, Berber i-, Hausa ya- ). Further evidence of *ma as a preposition includes its ability to substitute for other prepositions, notably ke in Angas, kš and ma are both prepositions with overlapping meanings: kš 'of, with, by, from'; ma 'to, of'. Chadic languages often make use of prepositions as particles between the subject marker and the nominal form of the verb to show imperfective (cf. Hausa ʾān and ʾakš, where na also means 'of'). In Angas, ma is the particle used with plural subject markers, while kš is used with singular subject markers. Since the third person subject markers may drop out before these particles, it is not surprising that kš has become re-interpreted as the third person singular subject marker in non-past tenses, while ma is the third person plural marker. This explains the development of ma as a pronoun in Chadic. Again, in Tera, another Chadic language, ke has become generalized as a subject marker, yet it is dropped when ma 'in order that' is used. Thus they occur in complementary distribution: tem ke zur 'we (subject) fry'
The ma- Prefix in Afroasiatic becomes ma zur mi 'in-order-that fry we'.

Ma as a preposition would also explain the use of ma as a relative "pronoun" in some languages such as Berber. In Egyptian relative clauses were expressed by a preposition instead of a relative pronoun: sdmw-n-f meant 'the voice heard by him' and thus 'the voice that he heard' secondarily. (cf. also Gardiner [1957:196] non-verbal relative clauses with m-).

If ma developed from a preposition, we would expect to see infinitives formed with ma- as well. In languages of the world infinitives are often formed from a dative construction, as in English 'to do'. In Afroasiatic we find the same structuring of infinitives. In Chadic, Mofu-Gudur has all infinitives formed with ma- , e.g. mątěkóčey 'to hide', mąkóčey 'to hit' (the y is a Mofu-Gudur innovation according to Barreteau [1977]). In Mandara, another Chadic language, am is a preposition meaning 'in, on, from, among'. When this occurs before verbs it becomes an infinitive: ta vaha am palasa aa vacia 'they spend-day in suffering of sun'; ta higa am shansa aa tare 'they are-pleased to find them' [Fluckiger and Whaley 1981]. Bedaye, a Cushitic language, has "action nouns" from verbs, such as mądar 'murder' from dir 'kill' [Roper 1928]. In Amharic, a Semitic language, all infinitives are formed with ma- [Woldegabir 1972]. Arabic verbal nouns of some types of verbs are formed with ma-, for example musaa-a(åda(t) 'help(noun)' from saaåda '(to) help'.

One reason that we don't find the ma- prefix to be very productive in Cushitic may be due to the Cushitic SOV word order, which tends to have postpositions rather than prepositions. The common ma- negative in Cushitic may have also discouraged use of a ma- prefix in the same contexts.

We have seen strikingly similar constructions in distantly-related languages, and I think we can not only consider the ma- prefix as a Proto-Afroasiatic phenomenon, but also the use of causative stems to form instruments and locatives: *mą + causative nominal form of verb + i meaning 'by it causing to do' for instruments, and *mą + caus. nom. form + a meaning (?) 'at the (instrument)', for example 'at the stove' meaning 'kitchen'. The locative -a here may be related to the Afroasiatic oblique (dative or accusative) case, and the adpositions such as Chadic a 'to' as well as the Berber distance demonstrative -a suffixed to nouns. On the other hand, it may just be a feminine [Newman 1979] or a plural marker [Prasse 1974]. We have also seen a u-stem for the agent both in Chadic and in Semitic and, most importantly, an agentive -i that seems well-established as Proto-Afroasiatic. *mą + u-stem + i could perhaps be thought of as 'by him ( ma ya ) done' or 'owner of doing' (since we don't know if the u-stem was passive or not).

From this discussion I think it becomes clear that we have to examine the stems and suffixes of ma- words, rather than the ma- prefix alone, to determine the origins of these forms, and I hope this paper will stimulate further investigation of this construction.
REFERENCES


INTONATION IN CHADIC LANGUAGES

William R. Leben
Stanford University

In this paper I describe certain aspects of intonation in three Chadic languages: Hausa, Tera, and Angas. For a starting point, I assume a model of intonation based on some work done on Swedish at Lund (by Costa Bruce and by Eva Garding) and revamped along the lines of Pierrehumbert [1980].

The Lund model of Bruce [1977] factors a complex intonational gesture into three separate components, the word accent, the sentence accent, and the terminal juncture, all of which can be realized by varying the fundamental frequency, as shown in (1):

(1) WORD ACCENT, SENTENCE ACCENT, TERMINAL JUNCTURE

Pierrehumbert modifies this approach by resolving the dynamic tones represented by arrows in (1) into sequences of level tones. Pierrehumbert maintains the separation of intonational gesture into three components, which she designates the pitch accent, the phrase accent, and the boundary tone. She uses these different terms to distinguish, e.g. between the two utterances illustrated in (2):

(2) a. bulldozer drivers union?
   L* [H-] L%

b. bulldozer drivers union?
   L* [H-] H%

In both cases, the main stress of the phrase receives a pitch accent, which others might have called a nuclear tone. In (2a), the pitch pattern is (ideally) level after the rise of the pitch accent, while in the second example, there is an additional rise at the end. The rise from a nuclear L is described in both cases by positing a H phrase accent, indicated by H[-] in (2). The difference between the final level pattern and the final rising pattern is captured by positing a low boundary tone (L%) at the end of (2a), which is interpreted as not changing the value specified by the H[-] to its left, as compared with a high boundary tone (H%) at the end of (2a), which is interpreted as creating a rise from the value of H[-].

The question now arises as to whether we may also describe the behavior of intonation in tone languages using this same sort of mechanism, even though it was constructed for English, a stress language, and for Swedish, a pitch accent
language. I think that quite a number of different intonational effects in Chadic languages can be described quite well in Pierrehumbert's framework.

1. Hausa

Newman and Newman [1981] mention three distinct prosodic gestures that can be involved in Hausa questions: a Low tone, added to the rightmost High tone in the utterance; length, added to the final syllable of the utterance if it is not already long; and key raising, an upward shift in the pitch levels of H and L. Note that these three gestures are quite distinct. For one thing, length is added to the final syllable, while the Low tone is added to the final high-toned syllable. More importantly, key raising can take place in questions that don't receive a Low tone at all, as we will see directly.

In yes-no questions, the final H (and, sometimes, preceding H's) is raised, and any following L is raised as well, including the interrogative L that is obligatory with yes-no questions. But in wh-questions, the interrogative L is added only optionally. Furthermore, H raising applies only if interrogative L has been added. Finally, this interrogative L or any other L is not subject to raising by Newman and Newman's key raising rule. That is, only the types of intonation illustrated in (3a,b) are found in wh-questions. The types illustrated in (3c,d) are ill-formed:

(3) a. Wàà ya sàyi rîgaa? b. Wàà ya sàyi rîgaa?
  - - -

  - - -

(N.B. (3d) is acceptable as an echo question.)

By contrast, in yes-no questions only the intonation pattern parallel to (3d) is well-formed, because these require an interrogative L, which must be raised.

The Newman and Newman analysis does not properly distinguish the behavior of yes-no questions from wh-questions. And by tying the raising of H to the raising of L (through its notion of key raising), that analysis provides no way of accounting for (3b), where the H but not the L has been raised.

To capture the facts just noted, we may redefine the interrogative L as a phrase tone L[-] added after the rightmost H of an intonational phrase: this tone is added optionally in wh-questions and obligatorily in yes-no questions. H raising is conditioned by this interrogative L; thus, it fails to apply in wh-questions like (3a), explaining the ill-formedness of (3c).

The raising of interrogative L must have a different source: let us attribute it to a H boundary tone added to the extreme right of the utterance. To account for the fact that this L raising effect appears only in yes-no questions, we posit that this boundary tone is added only to these, not to wh-questions.
Thus, phrase tones and boundary tones both appear to function in Hausa intonation.

2. Tera

There is, to my knowledge, no literature on intonation in Tera. Newman's detailed study of the language does not offer any data on the subject. A study of Tera intonation might have some special interest, since Tera is a language with three phonemic tone levels, High, Mid, and Low. My elicitations from a single speaker in summer, 1983, give some initial idea of what Tera intonation is like. The overall picture is one of considerable variability in the intonational realization of tones. But this variability seems to correlate quite well with where the pauses are in an utterance. For example, compare what is encircled in the (a) and (b) tokens. (Mid tone is left unmarked. '!/' marks pauses.)

(4) bird climbed monkey yesterday
   di:-ka wà:-bì yi:-dì bì -dà   'the bird climbed the monkey yesterday'
   a. 295 x / x 220 280/267 267/300 x   285 x / x 210 260/270 267/307 x

It seems in general that pauses set the boundaries at which many tonal processes take place. For example, downdrift in Tera does not seem to apply across pauses:

(5) di:-ka wà:- bì vì:- dì:- bì -dà? 'did the bird climb the monkey yesterday?'
   a. 170 175 /125 167 175 x x x x
   b. 167 170 160 145 155 /165 165 / 180 120

The encircled numbers differ by 20 Hz, and the only relevant difference between them is the presence of a pause earlier in the string. (From this set of differences, we may derive the differences in the values of the L tones in the word wà:lbì.)

Tera also provides motivation for another use of the boundary tone. Note that in the following utterances, the first of the two initial L tones in the sentence is considerably higher than the second:

(6) a. elephant ate greens where?
   jù - wàn zam lò - yòs fì - nà?
   210 145/ 225/ 195 200/265 150
   165 130/ x 130 130 175 130
   -150
   in forest
   b. jù - wàn wà zam lò - yòs sne ndò - yù 'the elephant ate greens in the forest'
   250 150 / 200 233 / 215 213 / 215 215 200
   -150

This is perhaps a sign that utterances beginning with a sequence of L tones are assigned a H boundary tone, which would condition the raising of the first L.
3. Angas

In Angas, there is a rule which realizes a final L tone as rising. Burquest [1974] notes that this effect takes place at the end of the phonological sentence. An example from my data is:

(7) monkey bit nose my yesterday
jó - lí at gê - zên fê - na ndô - ndôn 'the monkey bit my nose
180 130 /152 152 151 158 151 148 140 yesterday'

(8) I beat hen with stick
nâ cêk ki shi ndân 'I beat the hen with a stick'
165 200/ 160 / 175 135
-173 -165

Assuming that the use of boundary tones is reserved for syntactically defined clause boundaries, example (8) calls for an intonationally motivated nonlow tone other than a boundary tone. Thus, (8) provides prima facie evidence for the most innovative of Pierrehumbert's intonational features, the phrase tone, which would take the form of H[-] (or possible M[-]) in (8).

Similarly, an interposed phrase tone would explain the fact that in the following examples, where joli is followed by a brief pause, it ends with a higher pitch than when not followed by pause. The final syllable of joli is underlined in each case:

(9) jó - lí at gê zên fê - na ndô - ndôn 'the monkey bit my nose
a. 190 154/165 165 165 165 158 159 148 yesterday'
    -130 -168
b. 180 130 152 152 151 158 151 148 140
    -155

(10) shi kwâdân do jó - lí at gê - zên fê - nê:?' why did the monkey
a. 176 165 163 176 190 160/167 176 173 175 185- bite my nose?'
    -180 150
b. 166 158 159 171 185 150 165 158 167 168 177-
    142

4. Conclusion

I have shown that in a few examples from different Chadic languages, phrasal tones and boundary tones help to describe the facts. Also, even the limited data presented here illustrate a wider variety of intonational devices in tone languages than is normally recognized.

It might seem reasonable to think that if a tone language permitted too wide an array of intonational effects, this would interfere with the functional load to be maintained by the lexical tone system. But such a functionally based belief must be seriously defective, given that there are documented cases of neutralization of phonemic tone contrasts by intonational effects. A fitting exam-
ple is Newman and Newman's demonstration that the final lengthening in Hausa interro-gatives neutralizes the contrast between short and long vowels.

The present paper provides some evidence that tone languages permit intona-tional modifications of pitch that are not dissimilar to those reported for non-tonal languages.

REFERENCES


RELATIVE CLAUSES IN Aŋlo Ewe

Marshall Lewis
Indiana University

1. Introduction

A detailed study of relative clause formation (RCF) in the Aŋlo dialect of Ewe yields results of both language-specific and typological interest. The investigation is pursued against a background of research in language universals as exemplified by Keenan and Comrie [1977] (henceforth, KC). The various NP positions defined by Ewe syntax are surveyed to determine what strategies (in the sense defined by KC) are utilized for relativizing NPs at each position. Among the conclusions which emerge are the following:

(a) The set of NP positions provided for in KC's universal inventory is insufficient. In particular, the facts of Ewe oblige us to recognize an additional position called "Concomitant Object" (CO), translatable as [with __]. CO is motivated as a distinct position in Ewe by various kinds of evidence, including behavior in RCF.

In addition, the position of Object of Comparison (OC) is justified as a distinct position, even though such an NP is the object of the (non-initial serial) verb wu 'exceed' (cf. KC's assertion that no separate OC position exists in Yoruba, but rather is just a sub-case of DO because the comparative construction is likewise serial verbal in nature).

(b) The inventory of RCF strategies provided for by KC (as well as that of Maxwell [1979]) is too gross. The characterization of all markings of the RC-internal site associated with the head of the RC as "resumptive pronouns" is challenged. In Ewe, it is necessary to distinguish between truly pronominal marking, involving the third person singular and plural pronouns (3% and 3%%, respectively) agreeing with the head NP for number, and an invariant marker called "Relative /e/" (R/E), having the form of 3% regardless of the number of the head NP.

(c) Some of the cross-linguistic generalizations based on KC's Noun Phrase Accessibility Hierarchy (NPAH) appear to be falsified by the Ewe RCF data. Claims about pronominal strategies in particular are called into question in the light of the two kinds of marking just mentioned, and the dubious pronominal status of the invariant one.

2. NP Positions in Ewe

Ewe is an SVO language with fixed constituent order, illustrated in (1):

(1) Kofi fia nu dëvi-a-wo 1ɛ sukuu kpiɛ dzidzo wu Ama
SU DO IO OP CO OC
K. show thing child-DEF-PL at school with joy exceed A.

'Kofi taught the children at school more joyfully than Ama'

These positions are defined by the following diagnostic frames:
In addition there is the genitive position (GEN), with the frame [__ POSS NP], which has no fixed position in the sentence, as it may occur within any NP.

CO is the only one of the above positions not provided for in KC's universal inventory. (The position they call "Oblique Object" is an umbrella term which in Ewe amounts to the more specific Object of Preposition.) Furthermore, the word kple which constitutes the defining environment for CO has been assigned by previous authors to the category Preposition (or "verbid"; see Ansre [1966a,b] and Clements [1973]). Evidence of various kinds, however, indicates that kple should not be so analyzed, and that CO should be recognized as a distinct NP position.

The word kple has two guises: (i) NP linker, as in avu kple dadi 'dog and cat'. We shall not be concerned with this use. (ii) Introducer of what I will call "Concomitant Object" (CO), a clause-final NP which has affinities with the subject, in that the entity it denotes contributes to bringing about the action expressed by the verb. A human CO is interpreted as an ancillary agent; a concrete CO as an instrument; an abstract CO as an adverbial of manner. These possibilities are illustrated in (2):

(2) Kofi tu kple Komla/ ba/ dzidzo
    K. build house-DEF with K./ mud/ joy
    'Kofi built the house with Komal/ mud/ joy'

OPs never have affinities with subject, but rather express loci or goals of action. A syntactic correlate of this semantic difference involves an alternative construction for expressing such subject-affiliated NPs by means of a serial verb construction (SVC) wherein the object of one VP participates in the subjecthood of a following VP, as in (3):

(3) a. Kofi kple Ama yi Togo (e) 'Kofi conducted Ama to Togo'
    K. lead A. go T. S/E
b. Aaron tsq sika wo nyivi (e) 'Aaron used gold to make a calf'
    A. take gold make calf S/E
c. Moses do dziku fo kple-a (e) 'Moses angrily struck the rock'
    M. kindle anger strike rock-DEF S/F

The final /e/ in (3a-c) is called "Serial /e/ " (S/E) by Lewis and Schwartz [1983], who investigate its syntax and semantics in detail. Just in case the DO of a VP in a SVC participates in the subjecthood of a following VP, the latter is (optionally) marked with S/E. (For a brief discussion of the relationship between CO and SVC, see Clements [1973:214-215].)

A number of arguments, some likewise based on S/F, also establish OC as an independent position, not a sub-case of DO, even though the comparative construction in Ewe is clearly a SVC, as illustrated in (4):
(4) Kofi a qu mòlu (a) wu Ama 'Kofi will eat more rice than Ama'
K. FUT eat rice FUT exceed A.

The diagnostic for verbhood is the ability to (optionally) take overt marking
for the tense/aspect shared by the VPs of a SVC, as pointed out in Ansre [1966a, b].

Some of the strongest evidence for CO and OC as distinct positions in Ewe is
provided by their behavior in RCF, to which we now turn.

3. RCF in Ewe

In the terms of KC's taxonomy of RCF strategies, all Ewe RCs are postnominal,
i.e. the restricting clause follows the head. The canonical form is:

\[ [N (\text{ADJ})^*(\text{QNT}) (\text{SPEC})] \text{ si } [(\text{LNKR}) \text{ NP}]^* \text{ [S] (}\text{\texttt{l}\text{a}}) \]

Head + Modifiers REL NP(s) in apposition with head Restricting clause

All NPs can be relativized, with one exception: OCs compared to NPs other than
SU. The final \texttt{\textbar{l}\text{a}} occurs if the RC is non-final.

KC also distinguish between so-called [+case] and [-case] strategies. The
former are characterized by the presence of some overt element which unambiguously identifies the syntactic role of the head NP within the RC while the latter have no such marker. In Ewe, the primary RCF strategy (that used to relativize SU) as well as the dominant strategy (that used to relativize the most positions) is [-case], as in (5):

(5) a. am\text{e} s\text{ i } [__ f\text{ i} e \text{ agba\text{i} e-a}] 'the person who bought the book'
  person REL buy book-DEF

b. agba\text{i} e \text{ si } [Kofi f\text{ i} e __] 'the book Kofi bought'
  book REL K. buy

c. am\text{e} s\text{ i } [Kofi f\text{ i} e agba\text{i} e-a na __] 'the person Kofi bought the book
  person REL K. buy book-DEF for

This is the orthodox strategy used to relativize SU, DO, and OP (5a, b, c respectively); it is a possible strategy for every position except CO.

As Dzameshie [1983] first observed, however, we sometimes find an RC-internal marker at the site corresponding to the head NP. He gives examples where a morpheme having the form of 3% instantiates a [+case] strategy for relativization of DO, OP, CO, and OC. He refers to this as "pronoun retention", but offers no argument for a pronominal analysis and gives no conditions for this "retention".

An exploration of RCF at the various NP positions reveals that two different phenomena need to be distinguished, but which were neutralized in Dzameshie's
small sample. Like virtually every other writer on RCF, Dzameshie fails to con­sider the effect of varying the grammatical number of the head of the RC. When we do so, we find that in some cases the RC-internal marker covaries for number with the head, taking the form of 3% and 3% accordingly; this is clearly pro­nominalization. In other cases, however, the marker is invariant, always hav­ing the form of 3%. I will refer to this invariant marker by the neutral term "Relative /e/ " (R/E), deferring discussion of its categorial analysis.

Table I shows the possible RCF strategies available for the NP positions in Ewe. Strategies of [+case] are either [+ #] (when the marker covaries for num­ber with the head) or [- #] (when it is invariant). The entry +/- means that both possibilities obtain.

<table>
<thead>
<tr>
<th>SU</th>
<th>DO</th>
<th>IO</th>
<th>IO</th>
<th>OP</th>
<th>CO</th>
<th>OC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>+/-</td>
<td>+/-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[case]</td>
<td>[ # ]</td>
<td></td>
<td>[+-(?)]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex.</td>
<td>5a</td>
<td>5b</td>
<td>7c</td>
<td>7a</td>
<td>7b</td>
<td>6a</td>
</tr>
</tbody>
</table>

Table I. RCF strategies in Ewe, by NP position

R/E regularly occurs when RCF applies to CO and OC, obligatorily in the for­mer case, optionally in the latter, as in (6a,b):

(6) a. amezdro si (wo) [Kofi yi Togo kpli -i/ *wo] 
stranger REL PL K. go T. with R/E 3%% 
'the stranger(s) Kofi went to Togo with'

b. amezdro si (wo) [Kofi du nu wu (-i/ *wo)] 
eat thing exceed R/E 3%% 
'the stranger(s) Kofi ate more than'

(R/E, like 3%, is realized as [-i] after high vowels.) As Table I shows, CO behaves uniquely in RCF and differently form OP in particular. Likewise, beha­vior in RCF distinguishes OC from DO.

At all other positions except for IO, the [-case] strategy is used (5a-c). The situation at IO is complex, depending on whether the IO is part of a SVC and, if so, whether it belongs to the first or a subsequent VP. A [+case] strategy is nearly always optional for RCF of IO. An IO from the first VP of a SVC will be marked with R/E (see 7a); the IO site in a non-initial VP, on the other hand, is marked with a pronoun (7b). In non-serial structures, there is considerable inconsistency in the selection and evaluation of variable vs. invariant marker (7c), but the pronoun seems basic.

(7) a. ntsu si (wo) [m-a bia ga (e/ *wo)][(a) fie atike] 
man REL PL 1%-FUT ask money R/E 3%% FUT buy medicine
'the man/men I'll beg money from to buy medicine'
b. nubiala si wo [m-a dø to fada] [(a) na ga (wo/ *e)]
  beggar REL PL 1%-FUT send ear priest FUT give money 3%% R/E
  'the beggars I'll obey the priest and give money to'

Thus, Dzameshie's examples of [+case] at CO and OC represent normal RCF strategy at those positions, while his [+case] DO and OP examples conflict with Table I. The occurrence of markers at these sites, however, is limited to cases where the relativized NP is possessor of another NP which precedes it in the RC, as in (8).

(8) a. amε si [ɛ fɛ avu ɡu-i]  'the person whose dog bit him'
  person REL 3% POSS dog eat-3%

b. amε si wo [wo fɛ avu du wo/ *i]  'the people whose dog bit them'
  PL 3%% 3%% R/E

As (8) shows, this marker covaries for number with its antecedent. This kind of pronominalization is independent of RCF and needn't be dealt with as a relativization strategy.

4. Status of R/E

The question remains whether R/E is to be analyzed as a pronoun. At stake is KC's treatment of all RC-internal [+case] markers as resumptive pronouns, as well as Dzameshie's language-specific analysis. The case for a pronominal status is three-pronged: (i) the marker has the form of 3%, including the same paradigm of alternations determined by the preceding vowel; (ii) it is found in a position normally occupied by an NP; (iii) it is interpretively associated with the head of the RC.

Against (i), it can be pointed out that there is at least one other morpheme homophonous with 3% which is clearly not a pronoun, viz. S/E. The force of (ii) and (iii) depends on what one counts as defining properties of pronouns. The problem with a pronominal analysis is that in all other cases, Ewe pronouns agree in number with their antecedents. If R/E is not a pronoun, its invariance follows; if it is, however, the failure of some pronouns (but not others) to agree in number would be an apparently arbitrary fact in need of explanation.

Perhaps rather than choose categorically between a pronominal and non-pronominal analysis, it would be more profitable to regard R/E as a focus of linguistic change, i.e. a middle stage between a clear-cut resumptive pronoun and an invariant, clearly non-pronominal marker like S/E.

In any event, the typological implications are that KC's taxonomy of RCF strategies is too gross, blurring distinctions amongst [+case] strategies involving RC-internal markers. These might be expressed by a feature [+Pron] or, if markers like R/E are treated as pronouns, by a feature [+ #]. The value of such features, of course, depends on the extent to which they afford a basis for further generalizations. As previous studies have almost uniformly neg-
lected to investigate plural heads of RC, the wider relevance of this aspect of the Ewe data remains to be assessed.

5. Hierarchy Constraints

KC's major generalization is the NPAH, SU > DO > IO > GEN > OBLO > OCOMP, which asserts a universal hierarchy of accessibility to RCF. On the basis of the NPAH, KC posit a number of additional "hierarchy constraints". Maxwell [1979] offers additional constraints of the same sort, based on an enriched taxonomy of RCF strategies (which does not, however, make provision for [+Pron] markers). The Ewe data as summarized in Table I appear to present counter-examples to one of KC's constraints and one of Maxwell's. According to KC, if a primary, i.e. subject-relativizing, strategy applies to a position on the NPAH, it applies to all higher positions. As the primary [-case] strategy applies to OC but not CO in Ewe, this claim is falsified. Maxwell states that if a pronoun strategy ([+case], RC-internal) applies to a position on the NPAH, it applies to all lower positions which are relativizable at all. As IO, CO, and OC use a pronoun strategy, but the intervening OP does not, Ewe fails to conform to this "universal". One way out of this is to treat OP as a sub-case of DO in Ewe, as suggested in Clements [1973](cf. Ansre [1966a,b]). In any case, Maxwell's claim is difficult to evaluate without knowing how the marker R/E should be regarded.

6. Conclusion

Previous typological work on RCF has proven very promising, but areas of serious neglect remain to be sedulously explored before we can feel quite comfortable in our generalizations. In particular, research needs to be pursued on (i) RCs with plural heads and possible effects on markers in [+case] strategies, (ii) RCF applied to verbal arguments in SVC, and (iii) behavior of the various NPs when subject to other kinds of extraction, e.g. focussing or WH-question formation.

[I am indebted to Dan Avorgbedor for most of the data and for many insightful remarks thereon.]
REFERENCES


ON CERTAIN NOMINAL PATTERNS IN TIGRINYA

Jean Lowenstamm
Université du Québec à Montréal

Jean-François Prunet
McGill University
and
Université du Québec à Montréal

1. Introduction

Major category items in Tigrinya are modelled according to morphological patterns specifying a given arrangement of vowels and consonants. We give below two examples, an adjectival pattern qatăl and a nominal pattern qatégay.

(1) qatégay:
  qatégay: bâraḥ 'bald'
  qatégay: nágadjay 'merchant'
  qatégay: ṣâmâm 'deaf'
  qatégay: ṣâmâm 'weaver'

According to Leslau [1941], nouns fall into two categories, those that can be readily assigned to a pattern and a residual class, for which pattern membership is unclear. Interestingly, the items of the latter category are vowel-final. While it is conceivable that not every noun is an instantiation of a given pattern, we feel that the kind of exceptionality just mentioned is hardly accidental and can be reduced a great deal when the relevant phonological processes are identified and understood.

Having illustrated the workings of nominal patterns in Semitic, we can now address the main question of this paper, viz. does every noun belong to a given pattern? The evidence adduced to support a positive answer will have several theoretical consequences. First, it will illustrate the role of nominal patterns in Semitic morphology. Second, it will shed light on the nature of lexical representations. Third, it will uncover a phonological entity, viz. the light diphthong, the existence of which is predicted by the formal model adopted here.

2. The Theoretical Framework

This model of phonological representation assumes an autosegmental account of non-concatenative languages based on the one developed by McCarthy [1981]. Vocalic and consonantal segments are represented on separate, quasi-autonomous, tiers and are viewed as autosegments which link to positions on the skeletal tier. Our model departs from McCarthy's in that it does not make use of a CV tier. Rather, it relies on the existence of a skeletal tier consisting of skeletal points as in Kaye & Lowenstamm [in press] and, in a somewhat different sense, Hyman [1983].

With respect to the general organization of the phonological component we adopt a parametric approach. This approach relies on a few basic and unordered operations comparable to move-α in syntax. The characterization of phonological processes must, consequently, be derived from principles of Universal Grammar within which language-specific grammars may select parameters. It follows from this approach that (ordered) phonological rules must be absent from grammars.
Before proceeding with our proposal, it may be necessary to provide the reader with a few facts of Tigrinya regarding syllable structure and phonological processes.

3. Background on Tigrinya

The syllable structure of Tigrinya is best described as falling under two distinct categories, word-internal structure and word-final structure. The analysis presented follows that of Charette [1985] for Wolof, a language which exhibits the same syllable structure.

The various types of word-internal syllables encountered are shown in (2a) whereas imaginable but non-occurring ones are given in (2b). They are given with the CV notation for the sake of clarity.

(2) a. CV   b. *CC...
   CVC   *CVCC
   CVV   *CVVC

From this distribution we may conclude that the possible word-internal structures are those of (3).

(3)

The impossible structures of (2b) are ruled out by the fact that Onsets and Codas may not branch and the Prosodic Government Principle mentioned in Lowenstamm and Kaye [in press].

Word-finally the types of sequences, exemplified in (4a), are slightly different. Non-occurring but imaginable sequences are given in (4b). (5c) will be explained later.

(4) a. CV   b. *CC...
   CVC   *CVCC
   CVV   *CVVC

Notice that (4a) exhibits a natural class, i.e. the presence of an optionally branching Nucleus. This generalization is only possible in a model recognizing syllable-internal constituents. We agree with Charette's interpretation of (4): (1) the optional, word-final consonant is dominated by the Appendix (a non-branching, word-final position), and (2) only open syllables appear...
Nominal Patterns in Tigrinya

4. 

4.1. Suffixation. Consider the following paradigms obtained by adding pronominal suffixes to two nouns, [gaza] 'house' and [faras] 'horse'.

(5) a. faras
   farasu
   farasa

   gaza
   gaz?u *gaza?u 'his...'
   gaz?a *gaza?a 'her...'

The nature of the syncope observed in (5b) need not concern us here. For a discussion of this phenomenon, see Lowenstamm and Prunet [in preparation].

Let us see how such facts would be accounted for in a more traditional approach.

4.2. Derivations. In a SPE type framework, the most likely characterization of the facts of (5) would be a derivation using ordered rules as in (6):

(6) Underlying representation
    /gaza + u/    /faras + u/
    Hiatus Breaking   gaz?u ------
    Syncope            gaz?u ------
    Output             [gaz?u]     [farasu]

As we underlined in 2, our framework does not recognize the existence of phonological rules. It follows that, even if the "rules" of Syncope and Hiatus Breaking postulated for (6) could be derived from principles and parameters, our framework does not allow for the crucial ordering of (6). We are thus forced to look for an alternative.

4.3. An alternative account. Let us postulate that [gaza] and [faras] have segmentally parallel underlying representations. More specifically, let us assume that the vocalic patterns of the two words are underlingly identical, viz. they both consist of a single vocalic autosegment /A/, and that the two words have triliteral roots as displayed in (7).

(7) IgAzA11.....[gaza]
    IfArAsI.....[faras]

Thus, the surface [a] of [gaza] must be interpreted as resulting from an underlying sequence /A+/ (A related, although distinct, process of vowel-lowering can be seen at work in the reverse sequence, viz. /? +A/ → [?a] as 'weaver' in (1.).)

The question now arises whether the parallelism between /gaza?/ , surface [gaza], and /faras/ , surface [faras], extends beyond the fact that both roots are triliteral. It should be noted here that there is no logical necessity for such an extended parallelism to obtain. In fact, it is a characteristic of non-concatenative morphology that root similarity, or indeed identity,
need not be matched by morphological pattern identity.

4.4. Two logical possibilities. Two logical possibilities suggest themselves: \([g\text{\textbf{a}za}]\) and \([f\text{\textbf{a}r\text{\textbf{a}s}}]\) are instantiations of the same nominal pattern or they are not. Let us first consider the latter possibility, represented in (8) below.

\[(8)\]

Under the analysis of (8), \([g\text{\textbf{a}za}]\) and \([f\text{\textbf{a}r\text{\textbf{a}s}}]\) are morphologically unrelated and nothing more is to be said. However, such undue reverence for phonetic dissimilarity entails the loss of an important generalization, as we will see.

Under the alternative represented in (9), both roots are associated to the same prosodic configuration, \(\sigma[xx]\sigma[xx]x\), i.e. two open syllables followed by an appendix.

\[(9)\]

4.5. Linking. When linking takes place, the final glottal stop of \([g\text{\textbf{a}za}]\) will be re-associated as shown in (10a), resulting in the formation of a complex segment (\(?\)), ultimately interpreted as \(\sigma\) (see example on next page). A study of word final configurations reveals that glides do not appear in Appendix position. Other consonants are not subject to this distributional restriction. Accordingly, they appear after long or short vowels word-finally. On the other hand, glides (and gutturals) do not appear after long vowels, clearly showing that they are barred from the Appendix. This distributional observation constitutes the first piece of evidence motivating the spreading (cum delinking) of \(?\) in (10a).

In Tigrinya, ungeminated /\(b, k, q/\spirantize to become \([b, k, q]\) in
Nominal Patterns in Tigrinya

(10) a. \( \sigma \) \( \sigma \) Ap. b. \( \sigma \) \( \sigma \) Ap.

post nuclear position. Interestingly, spirantization is optional in the following cases:

(11) \( \text{b\text{\textl}a\text{\textl}a\text{\textl}k\text{\textl}u} \) or \( \text{b\text{\textl}a\text{\textl}a\text{\textl}k\text{\textl}u} \) 'I ate'
\( \text{s\text{\textl}a\text{\textl}a\text{\textl}k\text{\textl}a} \) or \( \text{s\text{\textl}a\text{\textl}a\text{\textl}k\text{\textl}a} \) 'you worked'
\( \text{n\text{\textl}a\text{\textl}y b\text{\textl}i\text{\textl}r\text{\textl}a\text{\textl}y k\text{\textl}a\text{\textl}\text{\textl}s\text{\textl}a\text{\textl}d} \) or \( \text{n\text{\textl}a\text{\textl}y b\text{\textl}i\text{\textl}r\text{\textl}a\text{\textl}y k\text{\textl}a\text{\textl}\text{\textl}s\text{\textl}a\text{\textl}d} \) 'the ox's neck'
\( \text{k\text{\textl}i\text{\textl}d\text{\textl}a\text{\textl}\text{\textl}\text{\textl}n\text{\textl}i} \) or \( \text{k\text{\textl}i\text{\textl}d\text{\textl}a\text{\textl}\text{\textl}\text{\textl}n\text{\textl}i} \) 'clothes'

This optional character of spirantization is consistent with the ambiguous status of glides (and gutturals) with respect to which constituent of the Rime they link to. In the first set of cases, the glide (or guttural) is analyzed as forming a diphthong with the preceding vowel, i.e. both are under N. (For an important discussion of diphthongs, see Newman & Bello [1981]). Accordingly, \( k \) finds itself in post nuclear position and spirantizes. Alternatively, when the glide is analyzed as a coda, \( k \) no longer stands in post nuclear position and fails to undergo spirantization.

Additional evidence for this ambiguous analysis of glides can be gathered from the systematic option in pronouncing sequences of three consonants when the first is a glide. This option is illustrated in (12).

(12) \( \text{k\text{\textl}i\text{\textl}d\text{\textl}a\text{\textl}\text{\textl}n\text{\textl}i} \) or \( \text{k\text{\textl}i\text{\textl}d\text{\textl}a\text{\textl}\text{\textl}n\text{\textl}i} \) 'clothes'

In the first case, the glide is analyzed as forming part of a diphthong, while in the second case, it is analyzed as a coda. A sequence of a coda followed by two consonants is impermissible (see (2)), and an epenthetic vowel appears between the first and the second member of the offending sequence.

Of course, the possibility that a glide may appear in coda position is irrelevant in the case of the word-final glide of forms like \( [g\text{\textl}a\text{\textl}a\text{\textl}a] \), since all words of Tigrinya end in open syllables (possibly followed by an appendix). The important point for our argument is that they be legitimately analyzable as nuclear elements.

5. Conclusion

In conclusion, the arguments presented above allow us to achieve the following: if \( [g\text{\textl}a\text{\textl}a\text{\textl}a] \) and \( [f\text{\textl}a\text{\textl}a\text{\textl}a\text{\textl}a] \) exemplified different noun patterns, say \( g\text{\textl}a\text{\textl}a\text{\textl}a \) and \( g\text{\textl}a\text{\textl}a\text{\textl}a\text{\textl}a\text{\textl}a \), respectively, we could not explain (13).
(13) 1. every instance of $g\hat{a}ta$ is demonstrably glide-final
2. no instance of $g\hat{a}t\hat{a}l$ is glide-final.

On the other hand, we have shown that only one pattern, $g\hat{a}t\hat{a}l$ is involved in the cases discussed above and the facts of (13) are a mere consequence of our analysis. If this analysis is correct, we predict that for any surface pattern of type $C\hat{A}C\hat{A}C\hat{A}$, there will be a corresponding surface pattern such as $C\hat{A}C\hat{A}C\hat{a}$. Indeed, upon suffixation, $m\hat{a}l\hat{a}k\hat{a}t$ 'trumpet' and $k\hat{a}t\hat{a}m\hat{a}$ 'city' behave like $f\hat{a}r\hat{a}s$ and $g\hat{a}za$, respectively:

(14) $m\hat{a}l\hat{a}k\hat{a}t$ $k\hat{a}t\hat{a}m\hat{a}$
    $m\hat{a}l\hat{a}k\hat{a}tu$ $k\hat{a}t\hat{a}m\hat{a}u$ $k\hat{a}t\hat{a}m\hat{a}u$ 'his...'
    $m\hat{a}l\hat{a}k\hat{a}ta$ $k\hat{a}t\hat{a}m\hat{a}a$ $k\hat{a}t\hat{a}m\hat{a}a$ 'her...'

[We are grateful to our informant, Ms. Elisabeth Berhe, of Adwa, Tigray, for her patience and insights. Research for this paper was funded by grant # 410-81-0503, from the Social Sciences and Humanities Research Council of Canada to Jonathan D. Kaye, and grant # 410-82-0891X1, from the Social Sciences and Humanities Research Council of Canada to G. L. Piggott. J-F. Prunet thanks World University Service of Canada for an exchange scholarship and McGill University for a travel grant. For a fuller discussion of the material presented here, the reader is referred to Lowenstamm & Prunet [forthcoming].]
REFERENCES


Lowenstamm, J. and J. F. Prunet. in preparation. "gasa, gaza, gwasa, čirgiya: on projecting the phonology of a language from its vowel system."


INTRODUCTION

Doke [1935:118] defines the ideophone as follows: "The ideophone is in Bantu a special part of speech, resembling to a certain extent in function the adverb, together with which it is classified as a descriptive." The objective in this paper is to demonstrate that the ideophone in Bantu is predicative in function and should rather be classified with the verb as a non-verb verbal. Tsonga, one of the languages of Doke's [1954:23] South-eastern Zone will be used to demonstrate the syntactic predicative function of the ideophone.

THE IDEOPHONIC PHRASE AS A SYNTACTIC CATEGORY IN THE SOUTHERN BANTU LANGUAGES

C. T. D. Marivate
University of South Africa

0. Introduction

Doke [1935:118] defines the ideophone as follows: "The ideophone is in Bantu a special part of speech, resembling to a certain extent in function the adverb, together with which it is classified as a descriptive." The objective in this paper is to demonstrate that the ideophone in Bantu is predicative in function and should rather be classified with the verb as a non-verb verbal. Tsonga, one of the languages of Doke's [1954:23] South-eastern Zone will be used to demonstrate the syntactic predicative function of the ideophone.

1. The Syntax of the Ideophone as Compared with that of the Verb in the Deep Structure

Compare the following sentences (see end of paper for list of abbreviations):

(1) homu / yi nghena /ematini
   Subj / sc v / loc. n.
   'the ox enters into the water'
   VP

(2) homu (yi ri PHOMU), ematini
   Subj (sc aux v. id ) loc. n.
   ( IP )
   'the ox "does" PLUNGE into the water'

Note that the verb /yi nghena/ and the Ideophonic phrase (yi ri PHOMU) perform one and the same function, namely, predication. Semantically, they also convey the same idea, that of entering into the water. However, the ideophone expresses a picturesque idea of entering into the water by carrying a plus-factor, namely, that of PLUNGING. Another semantic plus-factor carried by the ideophone is its suggestiveness of the sound one hears when the PLUNGING takes place.

Both /yi nghena/ and (yi ri PHOMU) are syntactic categories belonging to the same class in that they are interchangeable. Culicover [1976:13] defines a syntactic category as "...a group of words or sequences of words in a given language that can replace one another in any sentence of the language whatsoever without affecting grammaticality." The Ideophonic Phrase does exactly that in this instance.

Note that the constituent elements of the verb predicate in sentence (1) above are /sc+v/, while those of the Ideophone predicate in sentence (2) are (sc+v+id). I will use Culicover's theory again to explain this phenomenon. Culicover [1976:13] says "... if a syntactic category contains single words only, it is a lexical category. When a syntactic category contains sequences of
The Ideophone in Southern Bantu

words, it is a phrase category." Using Culicover's theory to interpret this phenomenon, the predicate in sentence (1) comprises a lexical category, namely, the verb /-ngenya/, while the predicate in sentence (2) comprises a phrase category namely \(v + \text{id}\), i.e. \((-\text{ri} + \text{PHOMU})\).

It should be noted that the syntax of the ideophone is closely linked with the verb stem \(-\text{ri}\) (Tsonga/Venda), \(-\text{re}\) (Sotho), \(-\text{thi}\) (Nguni, i.e. Zulu, Xhosa, Swazi). The crux of the matter lies in defining the role of this verb stem in the context of the ideophone. Note that the verb stem \(-\text{ri}\) can occur as a normal verb as well as an auxiliary verb.

2. The Verb Stem \(-\text{ri}\) as Normal Verb and as Auxiliary Verb

Where the verb stem \(-\text{ri}\) is used semantically to mean 'say', it is a normal verb, as in the following example:

(3) Magezi u ri! 'tana!' lit. Magezi he says: 'come!'
    Subj sc v v 'Magezi says: "Come!"'

When it is used with ideophones, \(-\text{ri}\) is an auxiliary verb. In this context, it does not mean 'say' as is the case when used as a normal verb. In the context of the ideophone, it has an idiomatic meaning. It expresses 'to do', 'to act', signifying that the subject 'does' or 'acts' or 'is in a stative situation' implied by the ideophone.

3. The Concept of the Ideophonic Phrase

The verb stem \(-\text{ri}\) in the context of the ideophone in Bantu languages should be regarded as an auxiliary verb, and not as a normal verb. The auxiliary verb stem \(-\text{ri}\) plus the ideophone constitute a phrasal syntactic category capable of acting as a predicate, just like a normal verb, which is a lexical category. I have decided to call this syntactic construction the Ideophonic Phrase, abbreviated as IP. The IP is the underlying structure of the syntax of the ideophone.

4. The Internal Structure of the IP

The IP as a syntactic category consists of two constituents, namely, the auxiliary and the ideophone. Each of these constituents has its morphology and function in the IP.

4.1. The morphology and function of the auxiliary of the IP. The basic constituent of the auxiliary of the IP is the verb stem \(-\text{ri}\) with its semantic import 'do' or 'act' in the context of the IP. The basic function of the auxiliary is to put the IP in the different moods, tenses, and conjugations. Consequently, its morphology is grammatical.

Taking the main mood, viz. the Indicative Mood, as basis for illustrative purposes, the basic constituent morphemes of the auxiliary of the IP and their order is almost the same in all of the Bantu languages, namely \(sc + aux v\). Hence

\[
\text{IP} = (sc + aux v. + \text{id})
\]
Where the IP takes an object, the order of the constituent morphemes is $sc + oc + aux \, v.$, hence

$$IP = (sc + oc + aux \, v. \, + id)$$

For example:

(4) nguni (zu) ïnkômô (i - mú - thê BHUXE), ísísu, ngô - phôndo  
    Subj (sc - oc - aux \, v. \, ID ), obj, instr \, mo - NP  

lit. the ox (it him 'has done' BUMP), the stomach, with/by the horn  
'the ox BUMPED him on the stomach with the horn'

Note that the position of the oc is the same as in the case of the normal VP which takes an object, namely /sc + oc + v/. Note the following Tsonga example:

(5) gámô (ri ri \, xî BVUU), xîkûkwâną  
    Subj (sc aux \, v. \, oc \, id), obj \, ( \, IP \, )  

lit. the eagle (it 'does' it SNATCH), the chicken  
'the eagle SNATCHES the chicken'

However, where Tsonga predicates with a VP which takes an object, the order of morphemes is the same as the rest of the Bantu languages, namely, /sc + oc + v/.

4.2. The morphology and function of the ideophone in the IP. The ideophone root in most Bantu languages is, to a large extent, indeclinable. In whatever tense, mood or conjugation the IP may be, the ideophone generally retains its basic morphological shape. It is only the auxiliary which changes and/or accommodates other auxiliary morphemes in the process. The function of the ideophone is to supply the IP with its semantic import. It is the ideophone which determines the transitivity or otherwise of the IP. Some ideophones bear a transitive feature, in which case the IP may take an object. Others bear the intransitive feature and influence the IP accordingly. For example, in (5) above, the ideophone BVUU! (of SNATCHING) has a transitive feature. On the other hand, BULU! (of BURSTING open) is an example of an ideophone with an intransitive feature as in the following example sentence:

(6) òandžâ (rî rî \, BULU )  
    Subj (sc aux \, v. \, id) \, ( \, IP \, )  

lit. the egg (it 'acts' BURST)  
'the egg BURSTS OPEN'

5. Deletion of the Auxiliary in Surface Realisations

A very important and interesting factor about the auxiliary of the IP is that the whole auxiliary, together with all its constituent parts may be deleted in surface realisations, leaving the ideophone so to say to predicate alone in such instances. Hence, example sentence (2) above with the auxiliary of the IP deleted will take the following form:
The Ideophone in Southern Bantu

6. The IP in the Different Moods, Tense, etc.

Unfortunately the scope of this paper does not permit the demonstration of the IP in the different moods, tenses, passive relationships, etc. Suffice it to say that the IP is capable of performing all the predicative functions, just like the verb.

7. IP and VP Governed by the Same Grammatical Subject

The IP and the verb may be used simultaneously, being syntactically governed by the same grammatical subject. In such instances the IP generally follows the VP, but not necessarily so. Note the following examples:

(8) 1st Pred. 2nd Pred.
mpfundla / wâ țî¿iâ/, (wû ri TLUKWA) lit. the hare/it jumps/, (it
Subj / sc v /, (sc aux v. id ) 'does' JUMP)
/ VP /, ( IP ) 'the hare JUMPS'

(9) mpfundla (wû 1â (ri) TLUKWA), /wû țî¿iâ/
Subj (sc stv asp mo aux v. id ), /sc v /
( IP ) / VP /

lit. the hare (it has simply 'done' JUMP) /it jump/
'the hare simply JUMPED'

The first predicate in most cases is in the Indicative Mood, but it may also be in the Potential Mood. The second predicate, i.e. the one following, may either be in the Indicative or Dependent Mood, depending on the intention of the speaker. This syntactic phenomenon is one of the grammatical characteristics of Bantu languages. In such usages, the VP serves to state the action while the IP serves to "exemplify" or "dramatise" the action suggested by the VP. The simultaneous use of the VP and the IP lends emphasis to the utterance. In such cases, the deletion of the auxiliary of the IP in surface realisations leaves the two, namely the verb and the ideophone, to appear to occur in the same clause. Hence sentence (8) may appear in the following form:

(10) mpfundla /wâ țî¿iâ/, (- TLUKWA!)
/ sc v / ( id )
/ VP / ( IP )

Such surface realisations may have acted as the basis for regarding the ideophone as an adverb since with deletions, the ideophone normally occupies the position occupied by the adverb. However, the deep structure syntax proves the contrary.
8. Conclusion

In the underlying structure, the ideophone is incorporated in the syntax of the language by means of the IP whose basic function is predication. Consequently, the ideophone whose sole function is to appear in the IP, should be classified with the verb as a non-verb verbal.

Abbreviations

<table>
<thead>
<tr>
<th>asp</th>
<th>aux</th>
<th>IP</th>
<th>id</th>
<th>instr</th>
<th>loc</th>
<th>mo</th>
<th>n</th>
<th>NP</th>
</tr>
</thead>
<tbody>
<tr>
<td>aspect</td>
<td>auxiliary</td>
<td>Ideophonic Phrase</td>
<td>ideophone</td>
<td>instrumental</td>
<td>locative</td>
<td>morpheme</td>
<td>noun</td>
<td>Noun Phrase</td>
</tr>
<tr>
<td>obj</td>
<td>oc</td>
<td>sc</td>
<td>stv</td>
<td>Subj</td>
<td>v</td>
<td>VP</td>
<td>Zu</td>
<td></td>
</tr>
<tr>
<td>object</td>
<td>object concord</td>
<td>subject concord</td>
<td>stative</td>
<td>subject</td>
<td>verb</td>
<td>Verb Phrase</td>
<td>Zulu</td>
<td></td>
</tr>
</tbody>
</table>

REFERENCES


VOWEL HARMONY IN KIPSIGIS

Bruce Martin
University of Illinois

In this paper, I will examine the vowel harmony of Kipsigis, or Kalenjin, a Nilotic language of western Kenya. All of the data presented in this paper is from the speech of Erastus K. Cheruiyot, from the vicinity of Kericho, Kenya.

Kipsigis has a 20 vowel system consisting of 5 tense and 5 lax vowels, any of which may be either long or short, just as described by Hall et al. in their discussion of African vowel harmony systems [1974]. For ease in typing, I will use lower-case letters for tense vowels, and upper-case for lax:

<table>
<thead>
<tr>
<th>Tense:</th>
<th>I</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>lax:</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td>e</td>
<td>o</td>
</tr>
<tr>
<td></td>
<td>a</td>
<td>A</td>
</tr>
</tbody>
</table>

Although Kipsigis is tonal, I will not mark the tones, since they are irrelevant to the discussion.

Most of the examples to be given below contain mainly non-high vowels. This is not because of any skewing of the distribution of high and non-high vowels in Kipsigis, but because I found the tense/lax distinction far easier to hear for low vowels. Without going into a phonetic discussion of the distinction I have informally characterized above as tense/lax, I will refer to the relevant feature as +ATR, after Hall et al. [1974] and others.

Most words in Kipsigis have either all tense or all lax vowels. None of the many verb or noun stems collected contained both tense and lax vowels. The suffixes collected fall into one of four groups. The first two are shown in (1a-b).

(1) a. pAAñ-AAn 'that walk'
  aAñ-an 'that beast'
sAñ-O 'painting'
sab-o 'recovering'
pAAñ-dA 'the trip'
kwan-da 'the woman'
b. sal-uut 'a paint job' (cf. sAñ-O)
  maas-uut 'a hitting spot'
  sal-is-yo 'some act of painting'
  am-is-yo 'some act of eating'
  maas-e 'he is hitting' (cf. kA-mAAs
  'he hit')
  keer-e 'he is seeing'

The harmony exhibited by the suffixes in these two groups presents few problems for analysis. In (1a), the tenseness of the suffix vowel is determined by the tenseness of the stem. In (1b), the entire word is tense, whether or not the stem is underlyingly tense. There is no corresponding group of lax suffixes which cause the stems to become invariably lax. So this is what has been termed a "dominant-recessive" vowel harmony system, where the presence of a vowel with the dominant feature, here +ATR, or tense, causes other vowels in the word to assume that same dominant feature, here to become tense. The re-
cessive feature, -ATR or lax, surfaces only when no vowel in the word has the dominant feature underlyingly.

In the lexical phonology framework, as proposed by Kiparsky, Mohanan, Pulleyblank, and others, this sort of vowel harmony system may be analysed by the use of underspecification of certain features at certain lexical levels and default rules to fill in the unspecified features. The only feature which may be specified underlyingly in Kipsigis for tongue root advancement is +ATR, or tense. The vowels of the morphemes which show alternation in (1) are unspecified for tongue root advancement. The vowel harmony rule would then spread the feature +ATR to the unspecified vowels in the word. A default rule would then assign the feature -ATR to all vowels still unspecified for tongue root advancement after the application of vowel harmony.

For example, the two vowels in sAI-O 'painting' in (1a) would be unspecified as to tongue root advancement until they were specified as -ATR by the default rule. The suffix -wu† of sal-wu† 'a paint job' in (1b) would be specified lexically as +ATR. Vowel harmony (spread +ATR) would then spread the +ATR specification to the rest of the word, yielding sal-wu†. The default rule would of course not apply.

This analysis using underspecification has the advantage over a fully specified autosegmental analysis of not requiring a rule deleting the feature -ATR in the presence of a +ATR vowel in the same word, and the advantage over a metrical analysis a la Halle and Vergnaud of not requiring multiple specification of the same feature.

The data from the prefixes found in Kipsigis may be accounted for in the same fashion. All prefixes, with one exception to be noted below, harmonize with the rest of the word. Whether the stem is invariably tense or tense because of a tense suffix, the prefix will be tense. Elsewhere, the prefix will be lax.

(2) kEE-mAs 'to hit' klp-slr-I† 'the writing utensil'
kee-geer 'to see' kib-am-i† 'the eating utensil'
KEE-gEr 'to shut' am-muga 'which lid?'
a-geer-e 'I'm seeing' Ag-gOk 'which chicken?'
o-geer-e 'you (pl.) are seeing' a-ger-e 'I'm shutting'
o-gör-e 'you (pl.) are shutting'

All of the prefixes in (2) may then be analysed as unspecified for tongue root advancement. For example, the first person prefix A in kA-A-gEr 'I just shut' would be unspecified for tongue root advancement, as would the recent past prefix kA- and the stem gEr 'shut' since this verb surfaces with lax vowels as would be specified by the default rule. All of the vowels in a-ger-e 'I am shutting' are tense because the continuous suffix -e is underlyingly tense, and vowel harmony will spread this specification to the rest of the word.

The other suffixes alluded to above require further analysis. In these two groups of suffixes, the tense-vowelled suffixes do not cause the lax-vowelled stems to become tense, and the lax-vowelled suffixes do not become tense after
tense stems. These are the suffixes underlined in (3). I will refer to these underlined suffixes as stratum 2 suffixes and to the suffixes as seen in (1) as stratum 1 suffixes. As far as I know, the suffixes underlined in (3) are the only stratum 2 suffixes in Kipsigis.

(3)  

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ꝶAh-Eszę-aan</td>
<td>'those words'</td>
</tr>
<tr>
<td>kwEEs-ľj-aan</td>
<td>'those goats'</td>
</tr>
<tr>
<td>kwEEs-ľa-nun</td>
<td>'your goat'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńa-an</td>
<td>'their goat'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńwAA-n</td>
<td>'their bone'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńwAA-n</td>
<td>'your bone'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńwAA-n</td>
<td>'his bone'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńwAA-n</td>
<td>'our bone'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńwAA-n</td>
<td>'your (pl.) bone'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńwAA-n</td>
<td>'their bone'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńwAA-n</td>
<td>'this calabash'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńwAA-n</td>
<td>'this calabash of today'</td>
</tr>
<tr>
<td>kwEEs-ľa-ńwAA-n</td>
<td>'this calabash of yesterday'</td>
</tr>
</tbody>
</table>

Stratum 2 suffixes never cause a change in tongue root advancement of any stratum 1 suffixes in addition to not affecting the tongue root advancement of stem vowels. This can be seen in (4). Stratum 2 suffixes are underlined throughout.

(4)  

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>kwEEs-AaN-aan</td>
<td>'that goat of ours'</td>
</tr>
<tr>
<td>kwEEs-AaN-aan</td>
<td>'our painting'</td>
</tr>
<tr>
<td>kwEEs-AaN-aan</td>
<td>'that their beast'</td>
</tr>
<tr>
<td>kwEEs-AaN-aan</td>
<td>'their recovery'</td>
</tr>
</tbody>
</table>

As indicated by my reference to these apparently exceptional suffixes as stratum 2 suffixes, I will analyze these examples by making use of the stratification of morphological operations and rules of lexical phonology. The basic idea is that different morphological operations are assigned to different strata, and the phonological rules are marked as applying to the output of morphological operations in different strata or sets of strata. So in Kipsigis, those suffixes I have referred to as stratum 1 suffixes will be assigned to stratum 1, and the vowel harmony rule will apply in stratum 1. The stratum 2 suffixes will be assigned in stratum 2, where the vowel harmony rule will not apply. The fact that stratum 1 suffixes never occur after stratum 2 suffixes is also consistent with this analysis.

The prefix ṅa-, meaning 'bad' or 'too many' when prefixed to a noun, is almost invariably tense-vowelled, yet the vowels of the nouns to which it is prefixed may be either tense or lax.

(5)  

<table>
<thead>
<tr>
<th>Suffix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ṅAAbA</td>
<td>'bad dish'</td>
</tr>
<tr>
<td>ṅAAbA</td>
<td>'dish'</td>
</tr>
<tr>
<td>ṅAAbA</td>
<td>'too many women'</td>
</tr>
<tr>
<td>ṅAAbA</td>
<td>'women'</td>
</tr>
</tbody>
</table>

This apparently exceptional behavior is also easily explained by the assignment of this prefix to stratum 2, to which vowel harmony is not applicable. The prefixes in (2) above, which undergo vowel harmony, would then be assigned to stratum 1, where vowel harmony could apply.
If a stratum 2 suffix is followed by another stratum 2 suffix and is the penultimate suffix, it will assume the category of the following (ultimate) suffix. If a stratum 2 suffix is followed by two other stratum 2 suffixes, it will be unaffected. Examples can be seen in (6).

(6) sota-ni-gAA-nlwAAn  'their calabash of today'
sota-ni-غاا-n 'our calabash of today'
rOOb-l-gAA-nlwAAn  'their rain of today'
rOOb-l-غاا-n  'our rain of today'
OOrltnw-Ej-aan-gAA-jlwAAk 'those their roads of today'
OOrltnw-Ej-aan-غاا-jaak 'those our roads of today'

So here we have a second vowel harmony process which differs from the first in three significant ways. First, it applies only to vowels in stratum 2 suffixes followed by stratum 2 suffixes. Second, it is not a dominant-recessive vowel harmony, but spreads both tenseness and laxness to the affected vowel. Third, this kind of harmony does not operate bidirectionally, but leftward only. Given just the first two differences, the two harmony processes could easily be collapsed by positing the harmony as spreading of $\alpha$ATR, and adding a rule delinking the correct vowels from their ATR specification, allowing the feature to be spread by harmony. However, since I am not prepared here to argue for a general principle which would predict the difference in directionality of the application of the two harmony rules, the collapse of the two rules would have to be at too great a cost in other complications. For this reason, I am proposing an analysis with two separate harmony rules. I will refer to the vowel harmony seen in (1) as stratum 1 harmony and this other vowel harmony process as stratum 2 vowel harmony. For stratum 2 harmony, a possible rule would be as in (7).

(7) \[ \begin{array}{c}
\text{±ATR} \\
\text{±ATR} \\
\text{V} \\
\text{V} \\
\end{array} \]

This rule would delink a vowel in the penultimate stratum 2 suffix and assign it the specification of the vowel in the following suffix. Note that this rule assumes the full specification of the vowels for tongue root advancement at the point at which it applies, and note also that this rule must crucially refer to stratum 2 boundaries in order to prevent it from affecting stratum 1 suffix vowels or stem vowels. Thus, the default rule which assigns $-\text{ATR}$ to unspec- fied vowels must apply before stratum 2 vowel harmony in stratum 2 or after stratum 1 harmony in stratum 1.

One other problem concerning the level of rule application concerns the nature of the instances of stratum 2 vowel harmony. The output of this rule is gradient and variable. That is, not only do the effects of stratum 2 vowel harmony sometimes disappear in slow, careful speech, but also the affected vowels are often not entirely tense or entirely lax, but somewhere in between. This gradient, variable manner of application has been proposed by Kiparsky
Vowel Harmony in Kipsigis

[1983] to be a general property of the application of certain post-lexical rules. The problem with this in the case of Kipsigis stratum 2 vowel harmony is that, since this harmony process does not affect vowels in stratum 1 suffixes, and does not apply across word boundaries, the harmony rule must refer to stratum 2 boundaries, which are not available for reference at the postlexical level.

So gradience and variability as a property of rule application may not be claimed even to be a property only of a certain set of postlexical rules, since it may characterize at least one lexical rule as well, namely stratum 2 vowel harmony in Kipsigis. One property which these gradient rules do share, as far as I know, is that they all apply to features which were either unspecified or lexically marked as other than the output of the rule at an earlier level. I will not propose a characterization of which rules will have this gradient character, since I don't know one, but I simply want to point out that these rules may not be assumed to be a subset of postlexical rules.

REFERENCES


SOME ASPECTS OF CONSONANTAL SOUND CHANGE IN KIKAMBA

P. Mumo Maundu
Nairobi University

1. Introduction

This paper will discuss the following points:

(a) The loss of voiced velar fricative /γ/ and the lateral /l/ in Kikamba;
(b) The deletion of the alveolar consonant in the first syllable, if two or more syllables consisting of alveolar sounds follow one another (a case of dissimilation);
(c) Some analogical creations in Kikamba:

(i) generalization of /mb-/ as the concordial classmarker for classes 10 and 9 and
(ii) analogical extension of class 9/10 as adjective with /vv/ stems.

For the purpose of the paper Kikamba dialects will be taken to be as follows (minor variations will be ignored):

(1) Kitui North Dialect: It is spoken with slight variations in the northern part of Kitui District of Kenya.
(2) Kitui Central Dialect: It is spoken in the rest of the district with numerous variations increasing and becoming more marked eastwards.
(3) Kilungu Dialect: It is spoken with slight variations in the western Division of Machakos District.
(4) The 'Standard' and written dialect: It is spoken in other parts of Machakos but with greatest variation in the south almost resembling Kitui Central dialect.

Data used in this paper will strictly be based on the said dialects with only very occasional reference to Gikuyu to strengthen some of the claims the paper makes.

1. The Loss of the Voiced Velar Fricative /γ/ and the Lateral /l/

The loss of the two consonants is almost complete in most of the dialects except the Kitui dialects. The fricative is more prominently retained in the northern dialects while the lateral is retained in all the dialects but for various reasons. In most of Kitui dialects it is retained in the demonstrative "that" while in the Machakos dialect it has become a glide like in the following cases:
(1) **Kitui Dialect**                   **Machakos Dialect**

/\vando\ vala/                  /\vando\ vaya/              'that place'
/\kindo\ ke\i\la/             /\kendo\ ke\i\ya/        'that thing'

The other case of the appearance of /\i/ in Kikamba is in cases where homophony would result if the lateral was lost.

(2) **Words with /\i/**                  **Words without lateral**

[keto \i\] 'grasshopper' [keto \i\] 'that which emits'
[\i\eu\] 'food' [\i\eu\] 'knee'
[\i\ngi\i\] 'a thousand' [\i\ngi\i\] 'seed'
[\i\g\i\o\] 'monkey' [\i\g\i\o\] 'joke'

The other case where /\i/ seems to be present is in class 12/13:

(3) **Pre-form Basic forms**            **Class 12/13**             **Gikuyu**      **Gloss**

*ke\i\ema* [ke\i\ema]  [ka-\i\ema]  [\i\ema]   'hill'
*mb\i\oli* [mb\i\oli]  [ka-\i\oli]  [mb\i\oli]  'goat'
*mb\i\ale* [mb\i\ale]  [ka\i\ale]  [mb\i\ale]  'clan'
*may\i\olo* [ma\i\olo]  [tololo]  [may\i\olo]  'foot'
*\i\g\i\e\i\* [\i\g\i\e\i\]  [ka-\i\i\e\i\]  [\i\g\i\e\i\]  'tooth'

The appearance of /\i/ in this class is not a case where /\i/ has been retained in class 12/13. In such a case we can write a rule:

(4) \(^{ '+' \text{ consonant} } \) > \( \emptyset / V \_\_V \)

\( ^{ '+' \text{ lateral} } \) \( \longrightarrow \) **Class 12/13**

It is a case of an analogical extension in class 12/13 where /\i/ has been inserted between vowels regardless of the consonant that was lost. For instance Kitui north dialect has the following but all have /\i/ in class 12/13:

(5) **Kitui north**  **other dialects**  **class 12/13**  **gloss**

*[\i\g\i\e\i\]  *[\i\e\i\]  *[ka-\i\i\e\i\]  'tooth'
*[\i\g\i\o\o\]  *[ma\i\o\]  *[tu-\i\o\o\]  'feet'
*[\i\g\i\e\i\em\i\be\]  *[\i\e\i\em\i\be\]  *[k-\i\e\i\m\i\be\]  'hoe'

3. **The Case of Dissimilation at the Alveolar Region**

In one dialect of Kikamba, Kikilungu, it seems no allowance is made for alveolar segments to follow one another in more than two or three syllables.
(6) a. Other dialects
   [sisya]  [it'Ja]  'look'
   [tinda]  [unda]  'remain'
   [kanimi] [kaini]  'small'

b. [nene]  [nene]  'big'
   [motondo] [motondo]  'mud'
   [tanda]  [tanda]  'spread'

A comparison between (6a) and (6b) will show that it is only (6a) that undergoes this change. The only motivation visible is that of making the first sound unlike the following ones but instead of the whole syllable getting lost, only one consonant is lost. We therefore want to propose that this is a case of dissimilation. Lehmann [1962:169] suggests that haplogy sometimes can be regarded as a case of dissimilation if the loss makes two consecutive similar sounds more unlike one another. Rule 7 below forms part of the grammar of the speakers of Kilungu dialect:

(7) \[
\begin{array}{c}
C \\
+ \text{anterior} \\
+ \text{coronal}
\end{array} \rightarrow \emptyset \\
\begin{array}{c}
- V \\
\alpha \text{ant} \\
\alpha \text{cor}
\end{array} \rightarrow \text{ant} \\
\text{cor}
\]

We also want to claim that this rule is also diachronic in that wherever other dialects have the consecutive syllables with all the sounds present, Kilungu speakers delete the first sound.

4. The Generalization of mb- as the Concordial Class Marker for Classes 9 and 10

The two classes have one thing in common: the class markers for 9 and 10 are both represented as /N-/, but class 10 marks plural. There are classes other than 9 which take class 10 as their plural. Derived lexical items changed to nouns are placed in class 9/10 example, e.g. oka 'be bad or ugly' changes to noka 'an ugly person'. Class (9/10) has another semantic connotation: it implies maturity or a state of being grown-up, especially for animals, e.g.

(8) 12/13 9/10
   [ka-tolome]  'a small ram'  [ndo:mc]  'a ram'
   [ka-salo]  'a young or small bull'  [nzao]  'a mature bull'
   [ka-samba]  'a small cock'  [nzamba]  'a mature big cock'

/nd-/ or /nz-/ in the above cases appear as if they are the class 9/10 markers instead of /N-/. In Bantu, class markers are copied onto the adjectives that qualify them, and therefore the assumption above will be strengthened by the data below.
(9) **Noun**  **Adjective**

/mbɔsɔ N-iŋge/ [...mbinge] 'a lot of bean'

/mboi N-aʊ/ [...nzaʊ] 'a white goat'

/mbuə N-iŋge/ [...mbinge] 'a lot of rain'

/mbe N-iŋge/ [...mbinge] 'many goats'

So we can see that /mb-/ as a plural marker or concordial agreement is copied onto those stems with /VCV/ stems and /nz-/ to those with /V-V/ stems or roots, suggesting the following rule, where the rules refer to the noun prefix as a concordial marker for class 9/10 (it is also true for plural formation with singular words with the same structure):

(10) **a.** \( \emptyset \) \( \rightarrow \) /-b-/ /NVCV

**b.** \( \emptyset \) \( \rightarrow \) /-z-/ /NVV

8. **Conclusions**

Rules (7a,b) are historical rules which created the present (synchronic) forms of this class of adjectives. The rules offer the explanations regarding the lack of phonetic motivation for the surfacing of consonants (N + N = mb or nz).

The consonants which were lost in the environment before adjective stem vowels were not necessarily the ones surfacing synchronically. Mutahi [1977] showed that some dialects southern Mount Kenya have [β̃jo] as adjective for 'black' and [ŋ̃ẽro] for 'white'. We have shown that Kikamba has /nzio/ and /nzaʊ/ respectively for the same adjectives. It is clear, therefore, /-z/ and /-b-/ were not the original consonants in these lexical items. The evidence we have gathered from Kitamba, strengthened by Gikuyu, weakens the claim made by Hinnebusch [1974] regarding the resurfacing of /-b-/ and /-z-/.

It is therefore clear that restructuring most of the class 9/10 lexical forms are due to the analogical reinterpretation of the prefix in the cases of /VCV/ and /VV/ word stems. The new forms were taken as basic, and speakers are unaware of what were historical forms. The surfacing of the new consonants is a proof of it. The wide spread or occurrences of these new forms are due to analogous extension to forms which had other forms historically.

REFERENCES


For a learner of a foreign language the interference of the vernacular should never be overlooked. He unconsciously uses the deep-rooted patterns of his mother tongue which contradict those of the target language. That is quite apparent when Arab speakers try to learn English as a foreign language. The interference is encountered on different levels, phonologically, morphologically, and syntactically. Here we are concerned with some phonological problems that face the Arab learner of English.

One of the most important principles contributed by present day linguistics to the teaching of foreign languages is that the language to be learned must be presented in terms of an activity which depends on productive as well as receptive levels and that the language is, after all, a vivid activity which should be practised, as in real life, through speaking.

The most problematic hindrance for learning a foreign language is that the learner does not have the chance of using this foreign language as a vivid activity, by using it for communication and self-expression. In most cases the target language is introduced as written material. The native speaker is rarely found and the technical aids, if available, can never substitute for a teacher with whom there is occasion for lively conversation.

Some people have maintained that the written language is a visual representation of the spoken language. If that were true then learning to read would not involve learning a new communication system, but merely transferring a language skill already acquired from one medium which is the medium of speech to another medium which is the medium of writing, namely from sounds to visual symbols. Writing is not merely speech written down, it is in fact a separate activity, another channel for acquiring the target language. Naturally, speaking should come first and by speaking we do not mean reading aloud, but conversing or using the language as an everyday activity and not as it is preserved in books. I want to come quickly to the conclusion that the idea that the written language is a representation of the spoken and that it is quite sufficient to learn the foreign language by reading it as long as the chance of learning it through speaking is not easily available, is completely wrong. It is evident that reading aloud as a means of learning pronunciation has disastrous effects for it makes the learner speak as if he were writing, and it also hinders the main target of reading which is comprehension.

Needless to say that English spelling does not always accurately represent the sound of spoken words. English orthography does tally with English pronunciation. In addition to the chronological priority of spoken language linguists have put forward strong reasons for stating that speech has primacy over writing and the more frequent use of speech than writing. To cut this argument short and go back to our main line I quote Bloomfield [1933:21], who says,
"Writing is not language, but merely a way of recording language by means of visible marks."

That leads us to the conclusion that when we teach the foreign language through the written representation, we are doing it the wrong way. Suffice it to say that we have hundreds and hundreds of languages, all over the world, with no writing systems but not vice versa, any language that is not used as a spoken medium is a dead language such as Latin, old Egyptian, and Greek.

So, in our community, as well as in countries of the Third World where learning English has become a necessity, to be acquainted with progress in advanced countries, we are facing the lack of chances for our students to practise the language as spoken activity, and as learners are not given that chance of achieving fluency of pronunciation of the target language, there is an opportunity for the well established linguistic spoken habits of the vernacular to interfere and retard the process of learning. Even scientists and technologists who learn a foreign language solely for technical purposes find it necessary to speak the foreign language for they find themselves always in need of contacting foreign experts and scholars in the field of their specialisation.

There are many features of spoken language such as rhythm, tempo, intonation, voice quality, and so on, for which there are simply no corresponding writing conventions. The same is true of all the features of spoken conversational language which are interactional and concern verbal and non-verbal feedback between speaker and listener and therefore have no direct counterpart in written language, which is essentially one-directional.

When we make a comparison and contrast of the sound structures of English and Arabic we can go on from the analysis of English phonemes to the analysis of Arabic phonemes and find that certain phonemes of the English system do not occur in the Arabic system and vice versa. It is an elementary linguistic fact that the English segmental phonemes that do not occur in the Arabic sounds create a prospective phonetic difficulty for Arab learners of English. The bilabial voiceless stop /p/, the interdental voiceless spirant /θ/, the interdental voiced spirant /ð/, and the labiodental voiced spirant /v/ are good examples of sounds that are usually mispronounced by Arab learners and are changed into /b/, /s/, /s/, and /f/. Two things should be noticed in that respect: (a) Although /θ/ and /ð/ are found in classical written Arabic, they are not used in spoken Arabic. Therefore they create a pronunciation difficulty in spoken English and even in reading written English or Arabic. (b) This difficulty faces the Egyptian Arab speakers, but it is not problematic for some other Arab speakers such as the Iraqis because they are common in their vernacular. There are other minor changes such as the English alveolars /t/, /d/, /s/, /z/, /n/, /l/, /r/ which are pronounced by Arab speakers as dentals. The velar /ŋ/ is pronounced dental /n/ because Arabic has no phoneme corresponding to it.

Consonant combinations predict, as Hamlet says, "a sea of troubles" for Arabic learners of English. Speakers who are used to a language in which no more than two consonants are permitted in a cluster are going to face a complicated problem with a language which has four consonants without juncture /+/, e.g. /ekstra/ as in 'extra', and seven consonants with juncture, as in 'he glimpsed
streams of water' /glimpst + striymz/. The most conspicuous mispronunciation, to an English ear, of a consonant cluster occurs with the past tense. The past form -ed represents a full syllable only after /t/ and /d/. For the Arabic speaker any cluster of three or four consonants has to be broken by a "cluster breaker", a weak vowel that separates the consonants.

When two consonants end the simple form of the verb, as in 'ask' /ask/, then the Arabic speaker adds a vowel before the next consonant of the past tense ending, i.e. /askt/ is pronounced /askid/. For English verbs which occur after a verb-final consonant or consonant cluster, as in 'kept' /kept/ and 'begged' /begd/, when the following word begins with a consonant, the /t/ or /d/ is regularly dropped in normal speech due to elision as in 'kept quiet' /keptkwaiat/, 'begged me' /begmi/. While the native speaker omits the /t/ and /d/ unconsciously and the meaning is not changed, the Arabic learner insists on pronouncing them because of their grammatical role and orthographic shape. The poor production of consonant clusters would often impair the meaning or impede the comprehension of the listener.

The simple vowels of both languages have in common the characteristics of being relatively short and lax but a thorough study of the vowel charts of both languages indicates the great contrast between the structuring of these vowels. An Arabic vowel allophone may belong to a quite different phoneme in English. Arabic /i/ as an example covers such a wide range that Arabic speakers may use for this phoneme sounds which an English ear hears as belonging to either /i/ or /e/. This is the result of the fact that Arabic has only the three simple vowels /i/, /u/, /a/ while English has nine.

With respect to the complex vowels the contrast is strong. Arabic has five complex vowels which are long and tense. Phonemically these long vowels can be treated as a doubling of a simple vowel /ii/, /uu/, etc. and sometimes, for convenience, the sign of length /:/ is used to designate the complex sound. The corresponding English diphthongs are /iy/, /ey/, /ay/, /oy/, /aw/, /ow/, and /uw/. The wide gap between the two vowel systems is quite clear and the prospective difficulties need not be mentioned. A contrastive study of the two segmental systems illustrate the expected range of erroneous pronunciation.

Many theoretical contrastive phonological studies of English and Arabic have been mainly concerned with the problems involved in the teaching of English segmentals to native speakers of Arabic, but few studies have been basically devoted to the suprasegmental or prosodic features and the problems involved in practising them. Failure in the handling of some prosodic features such as pause, tone group and tonic may account for the fact that the English spoken by Arabic speakers is rather difficult for English native speakers to follow with ease.

Studies show that the high frequency of occurrence of pauses made by Arabic native speakers and their wrong distribution result in short incoherent information units, and together with a high frequency of occurrence of tonics and their wrong placing, may lead to difficulty in understanding the message or to confusion. By pause we mean silence intervals of varying lengths, overlapping rhythmicality, and intonation contours. It is a suprasegmental feature which can never be detected if learning the foreign language is attained by written
material. I remember when someone described the spoken language of a teacher by saying that he spoke as if he were reading a book. The intonation of a certain language is of the utmost importance in conveying the meaning. Difference of intonation might change the statement into a question or an exclamation and might alter the whole meaning as in the question, "Is your husband here?", with a change in the intonation of the utterance. In the hierarchical system, the tone group, which is the unit of intonation, consists of one or more rhythmic groups in the same way that the rhythmic group consists of one or more syllables.

As Halliday says, "It marks a unit of information in the message that the speaker is trying to convey." There is always a point within the tone group (one or more syllables) where the defining pitch (movement of the tone) occurs; this point is called the tonic. It is important to state that the place of the tonic in an English utterance is variable and that the variation is meaningful, creating an independant set of choices. The tonic syllable starts on a stressed syllable which is usually more prominent than the other stressed syllables in the tone group, because it carries, as Halliday says, the main burden of the pitch movement in the tone group.

To conclude, learning a foreign language should be mainly through the spoken medium. According to Allen [1959] the teacher of English should be a good model for his students by mastering English pronunciation and acquiring a good knowledge of the sounds of English so as to be able to help his students in their learning process. O'Connor [1976] mentions that listening carefully to a good model is the most important thing for ear training. He also stresses the point that learners can reach the stage of perfection in pronunciation through careful repetition.

When we bring the case to verdict all accusations point at the written form of the language which prevents the learner from following the natural channel of learning the language, namely, acquiring linguistic habits through hearing and imitating the model of pronunciation.
The issue of word order in Classical Arabic (henceforth Arabic) has been the subject of an as yet unresolved controversy. Some linguists contend that Arabic is underlyingly VSO; while others argue that Arabic is an SVO language (underlyingly). Both camps seem to agree that the other available word orders are derived by some kind of stylistic rules. My purpose in this paper is not to discuss the nature of these rules, rather it is an attempt to locate these rules in the grammar of Arabic.

Arabic allows a high level of word order variation in simple declarative clauses. Consider:

(1) a. kataba ?aḥmad-un kitaab-an 'Ahmed wrote a book'
wrote 3sg Past Ahmed-NOM book-ACC
b. kataba kitaab-an ?aḥmadun
c. ?aḥmadun kataba kitaab-an
d. ?aḥmadun kitaab-an kataba
e. kitaab-an kataba ?aḥmadun
f. kitaab-an ?aḥmadun kataba

Most Arabists agree that all the above sentences are grammatical and roughly synonymous. The choice of one order over another is largely dictated by pragmatic and functional factors such as topic and comment, old vs new information, etc.

The theoretical framework presupposed in this paper is the theory of government and binding as developed by N. Chomsky [1981, 1982] and his associates. According to this theory the organization of the grammar is as in (2):

(2) D-Structure (DS)
    Move α
    S-Structure (SS)
    Phonetic Form (PF)           Logical Form (LF)

Crucial to this paper is the assumption that there is no interaction between the PF and LF sides of the grammar. Furthermore, my understanding of the current literature is that the so-called stylistic rules apply in the PF. The central claim of this paper is that this is incorrect, in so far as Arabic is concerned. I will argue that given the organization of the grammar in (2), and given certain phenomena to be discussed below, these rules must apply at SS at the latest. The evidence will derive from restrictions imposed on word order...
when the interpretation of the possessive pronoun is involved.

For the purposes of this paper I will introduce the following definitions:

(i) when the possessive pronoun is interpreted as coreferential with the subject NP, I will call this the i-reading;

(ii) when the possessive pronoun is interpreted as coreferential with the object NP, I will call this reading the j-reading; and

(iii) when the possessive pronoun is interpreted as free in reference, I will call this the k-reading.

To exemplify, consider:

(3) wādā`a ?ahlad-un ?a`l-kitaaba fi ṣunduuq-i-hi
put 3sg Past Ahmed-NOM the-book-ACC in box-GEN-his

a. 'Ahmed, put the book in his box' = i-reading
b. 'Ahmed put the book in its box' = j-reading
c. 'Ahmed put the book in his/its box' = k-reading

When the possessive pronoun is given either the i- or the k-reading word order is irrelevant. That is to say, both readings are obtainable no matter which surface order is chosen. Consider:

(4) a. ramaa ?ahlad-un kurat-a-hu 'Ahmed threw his ball'
throw 3sg Past Ahmed-NOM ball-ACC-his

b. ramaa kurata-hu ,k ?ahladun
   i,k

c. ?ahladun ,rampa kurata-hu ,k
   i,k

d. ?ahladun ,kurata-hu ,k ramaa
   i,k

e. kurata-hu ,k ?ahladun ,rampa
   ,i,k

All the sentences in (4) are grammatical under either the i- or the k-reading.

The interpretation of the possessive pronoun under the j-reading is heavily restricted in terms of word order. Consider:

(5) a. *qaraba ᵢṣadiq-u-hu ᵢahlad-an 'his friend hit Ahmed'
hit 3sg Past friend-NOM-his Ahmed-ACC

b. qaraba ᵢahladan ᵢṣadiq-hu

j

c. ᵢṣadiq-hu ᵢgaraba ᵢahladan
   j

j

d. ᵢṣadiq-hu ᵢahladan ᵢgaraba

j

j

e. ᵢahladan ᵢgaraba ᵢṣadiq-hu

j

j

f. ᵢahladan ᵢṣadiq-hu ᵢgaraba
The sentences (5a-f) demonstrate that the j-reading can only be obtained if the antecedent precedes the possessive pronoun. Thus, the admissible surface word orders under the j-reading are VOS, OVS, and OSV. The remaining three orders are ungrammatical.

This restriction on word order with regard to the j-reading is neither restricted to matrix clauses, nor is it restricted to objects of verbs. Consider:

(6) 
\[
daxala fi \text{?al-}\text{daar-i } \text{\textashahib-u-haa}
\]
entered 3sg Past the-house-GEN owner-NOM-her
'into the house\_j entered its\_j owner'

In (6) the antecedent is an object of a preposition. Again the only permissible word orders are [V PP S], [PP V S], and [PP S V] (S = Subject).

As for embedded clauses, consider:

(7) 
\[
qarrarat \text{?al-}\text{\textahuukumat-u } ?\text{\textan yastarji\texti\texta}
\]
decided 3sg Past the-government-NOM that recover subjunctive
\[
\text{?al-}\text{\texta-arg-a zaari\textl-ha-haa}
\]
the-land-ACC cultivator-ACC-her
'the government decided that the land\_j be recovered by its\_j cultivator'

(7) is a very interesting sentence. The complementizer ?\textan requires as part of its subcategorization for a verb to follow it immediately. This requirement, coupled with the restrictions on the j-reading, makes VOS the only permissible word order in this and similar sentences.

Given the above evidence, in particular the j-reading, it follows that if the surface word orders are to be derived in the PF by stylistic rules, it seems that there is no straightforward way to account for these facts. I will not discuss whether there are stylistic rules applying at SS or whether these word orders are derived by the general rule Move \(a\). Rather, I will turn my attention in the paragraphs to follow to providing a constraint which will make the correct predictions.

In order to do this the notion of c-command will be crucially needed. More specifically the subject NP and the object NP must be somehow rendered as structurally unequivalent. Since the notion of c-command makes crucial reference to configurational structures, it is necessary to establish some kind of a constituent structure where the subject will be structurally in a position to c-command the object but not vice versa.

Several alternatives come to mind. Among them are the following: (i) Assume, following J. Aoun's proposal, that the verb is coindexed with those constituents it subcategorizes for. This coindexing mechanism can be built into our definition of c-command, thus giving us the required results. (ii) Assume, following Zubizarreta and Vergnaud [1981] concerning Japanese, that Arabic has virtual categories. More precisely, the only virtual category in Arabic is the VP. This virtual VP will serve as the barrier preventing the object NP from c-
commanding the subject NP. (iii) Assume, following Emonds [1980], that Arabic is underlyingly an SVO language. This will give the familiar tree, thus making the subject in a position to c-command the object NP but not vice versa.

The choice of any of these alternatives over the others goes beyond the scope of this paper. Insofar as I am able to determine, the conclusion to be arrived at here is compatible with any of these alternatives.

We are now in a position to provide an account for our problem. Let us assume that the notion of c-command is defined as in (8):

\[(8) \quad x \text{ c-commands } y \text{ if the first maximal projection dominating } x \text{ also dominates } y\]

Now, in all the sentences involving the i-reading the possessive pronoun is cliticized to an NP that the subject c-commands. Thus, the subject is in a position to c-command the possessive pronoun. This state of affairs does not obtain in the case of the j-reading, as the possessive pronoun is cliticized to the subject. That is to say the object NP is not in a structural position to c-command the possessive pronoun. I therefore propose that the following constraint be considered as part of the grammar of Arabic:

\[(9) \quad \text{(Informally) If an NP and a possessive pronoun are to bear the same index, and that NP does NOT c-command this pronominal, then the NP must (linearly) precede.}\]

(9) will now account for all the restrictions (or the absence of restrictions) imposed on word order in sentences involving the possessive pronoun. But there remain the problems posed by verbs of the "put" and "give" types. Consider:

\[(10) \quad \text{wağafa } \text{?aḥm-d-un kitaaba-hu } fi Şunduuq-i-hi\]

a. 'Ahmed put his book in his box'

b. 'Ahmed put his book in his (its) box'

The k-reading is irrelevant. The i-reading is always possible since the subject will always be in a position to c-command both possessive pronouns in (10). To obtain the j-reading in (10b) the object NP kitaabahu 'his book' must precede the possessive pronoun with which it is coindexed. For (9) to apply here it therefore follows that the object NP must not be in a structural position to c-command the possessive pronoun.

As for the "give" type sentences, consider:

\[(11) \quad \text{?aḥṭa } \text{?aḥmad-un xallil-an kitaab-a-hu}\]

a. 'Ahmed gave Khalil his book'

b. 'Ahmed gave Khalil his book'
Again the i-reading is always possible for the same reasons given above. And the same restriction still applies to the j-reading. This suggests that if (9) is correct the constituent structure of the VPs in (10) and (11) must be (12):

(12) (irrelevant details omitted)

Now a new question poses itself: what is the nature of \( V' \) in (12)? If it is going to block the lower NP from c-commanding the higher NP or PP, this \( V' \) must be a maximal projection, i.e. VP. (Recall the definition given in (8)). Whether to take this option or revise the definition in (8) in such a way as not to refer to maximal projections but to the first branching nodes I leave as an open question.

I hope that I have shown, given the phenomena discussed above, that stylistic rules in Arabic, if they exist at all, must take place in the syntax.

REFERENCES


Gwandara, spoken by some 50,000 people in the Middle Belt of Nigeria, is the Chadic language most closely related to Hausa. It is geographically totally separated from Hausa (and other Chadic languages), being surrounded by speakers of Kwa and Plateau languages. Gwandara and Hausa are not mutually intelligible, though many Gwandaras speak Hausa as a second language. The extent of difference between the two languages, as well as the degree of similarity, can be seen in the following examples. (All Gwandara data in the paper are drawn from Matsushita [1972, 1973, 1974]).

(1) R. nåa tşaayaa 'a kăn hāfāl səyā 'I stood on one foot'
   G. n cēcē a kəyə kāpə də

(2) R. wūkānnän baa tá dā kāifī 'this knife isn't sharp'
   G. Tnkaṃmū tā mə n kəpT bā

(3) R. bāi zōo bā tüküná 'he hasn't come yet'
   G. t tükū jō wō bā

(4) R. kūnāmāa tāa hārbēe nī 'a scorpion stung me'
   G. kūnāmā rībT nī

This paper addresses the following linguistic/ethnohistorical question: Is Gwandara an independent millennium-old sister language of Hausa, as implied by most Chadic language classifications, or is it, as suggested by oral history, a relatively recent offshoot of Hausa that has undergone massive, contact-induced change? In other words, is the relationship of Gwandara to Hausa comparable to that of Frisian and English or to that of Krio and English? A key to the solution of the problem lies in the comparative analysis of the tonal systems of the two languages.

Unlike Hausa, which has two distinctive tones inherited from Proto-West-Chadic, Gwandara has three tone levels: Hi ə, Lo ə, and Mid ə. The tonal correspondences are set out in the following examples, Hausa on the left, Gwandara on the right. (Long vowels in Hausa are indicated by double letters; Gwandara does not have distinctive vowel length.)

(5) H L > H M
    ƙāfāl kāpə 'foot'
    wākāl wənkə 'song'
    dūutsēe duncT 'stone'
    gāsəə gəshə 'roast'

(6) L H > L M
    ƙəhōo kəhō 'horn'
    yəəshī yənshT 'sand'
    kwādōo kwārō 'frog'
    ƙəərū kərū 'increase'
The above examples illustrate differences between Gwandara and Hausa, but do not explain the development of the additional tone level. Here are some of the sources for the three-way tonal contrast in Gwandara. First, there is truncation of disyllabic H H words resulting in a distinctive Mid tone on monosyllabic words.

(13) Loss of syllable: H H > M M > M

First, there is the effect of final vowel length. The H L > H M change presented in (5) actually only applies to words with an original long final vowel. Where the original final vowel was short, H L is realized as H L. Since Gwandara has lost distinctive vowel length, the former contrast between long and short final vowels has emerged as a tonal contrast between H M as opposed to H L.

(14) Effect of short final vowel: H L (short) > H L

Third, within Gwandara itself, a contrast between Hi and Mid has surfaced due to the suffixation of a floating Hi tone genitive/associative marker. Note the following examples within Gwandara of M M vs. M H, H M vs. H H, and L M vs. L H.
Tone Splitting and Gwandara Ethnohistory

(15) Suffixation of genitive/associative marker '.

<table>
<thead>
<tr>
<th>Tone</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>'sun/day'</td>
<td>nānā m'ja</td>
</tr>
<tr>
<td>'light'</td>
<td>ńskyē nānā</td>
</tr>
</tbody>
</table>

(cf. Hausa háskeē, hásken rāanāa)

<table>
<thead>
<tr>
<th>Tone</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>'hole'</td>
<td>rāmī kūntī</td>
</tr>
<tr>
<td>'month'</td>
<td>wōntā jīyā</td>
</tr>
<tr>
<td>'father'</td>
<td>ńbā m'jīt</td>
</tr>
<tr>
<td>'bone'</td>
<td>kēshī bāy</td>
</tr>
</tbody>
</table>

The mechanism by which Gwandara has increased its number of tones does not match the generally known processes of tone splitting [Maddieson 1974]. Rather, the Gwandara tonal system seems to have resulted from the reinterpretation of Hausa surface tones in terms of the tone spacing grid of a language or languages already having three tones. Note the similarity between the treatment of Hausa loanwords in Gwari and Nupe [Maddieson 1977], two Kwa languages with three tones, with the tonal correspondences between Hausa and Gwandara. The proposed explanation for the tone splitting in Gwandara is thus substratum, i.e. Gwandara arose from Hausa as pronounced by people whose first language had three tones like neighboring Kwa and Plateau languages. The substratum explanation fits in well with other major changes in Gwandara: loss of grammatical gender, loss of distinctive vowel length, and loss of glottalized consonants. These are all common features of non-native Hausa [Feyer 1947, Hodge 1960]. By contrast, glottalization is an extremely stable feature of Chadic languages [Wolff 1959], being resistant to loss even under extreme areal and contact pressures [Wolff and Gerhardt 1977].

This is not to say that Gwandara should be viewed simply as a deviant, pidginized dialect of modern Hausa [Gouffé 1973]. Since Gwandara exhibits older features of Hausa as well as certain Kano Hausa innovations, it cannot be identified with any present-day dialect. Rather, Gwandara has to be associated with a variant of Kano Hausa as it existed at a somewhat earlier period of time. Consider the following:

(16) Sokoto H. Kano H. Gwandara

<table>
<thead>
<tr>
<th>Tone</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>'daughter/children'</td>
<td>dīyāa m'īyāa</td>
</tr>
</tbody>
</table>

If Gwandara had taken its singular form from the Sokoto dialect (which preserves the original shape of the word) it would appear as dīyā or dīyā; if it had taken it from present day Kano Hausa it would appear as gāyā with Hi tone. The Gwandara form relates instead to a hypothetical Old Kano form *m'yā that existed after the *dīyā- to 'y innovation, but before the subsequent L H to H contour tone simplification rule characteristic of modern Hausa [Parsons 1955, Leinen 1971].

According to oral tradition [Hassan and Na'ibi 1962, Isichei 1982, Temple and Temple 1922], the Gwandara were Hausa people who fled from Kano rather than accept Islam. Unomah [1982] has proposed that the exodus most likely took place
during the reign of Yaji, 11th Emir of Kano (1349-1385), or Muhammad Rumfa, 20th Emir of Kano (1463-1499), while Adamu [1978] has pointed to the reign of Muhammadu Kutumbi, 29th Emir of Kano (1623-1648). The latter date, particularly, has been felt to be incompatible with the considerable linguistic differences between Gwandara and Hausa. However a time depth of as little as 300 years presents no problem once one recognizes that Gwandara is not an isolated sister language of Hausa that has diverged gradually through time, but rather is a sort of creolized Hausa that has undergone rapid change due to substratum influence. Linguistic, historical, and geographical factors suggest that there was never a mass migration of Gwandaras from Hausaland. Rather, a small group or groups of pagan Hausa found a refuge in the Middle Belt region where they became the nucleus for multi-tribal settlements containing various Kwa and Plateau speakers. Eventually the non-native Hausa of the majority of the community developed into the language that is now known as Gwandara.

REFERENCES


ACOUSTIC ANALYSIS OF VOWELS AND DIPHTHONGS 
IN CAIRO ARABIC

Kjell Norlin
Lund University, Sweden

1. Introduction

This paper reports an investigation on acoustic properties of vowels and diphthongs in Cairo Arabic. The two classes of plain and pharyngalized consonants in Cairo Arabic are also discussed, using the alveolar fricatives /s/ and /ʃ/ as an illustration [Norlin 1983]. The effects on vowels in the environment of these consonant classes are demonstrated.

Cairo Arabic has five contrasting long vowels /ii, ee, aa, oo, uu/ and three short vowels /i, a, u/. The long mid vowels are derived from the diphthongs /ai/ and /au/ . In Cairo Arabic the long mid vowels /ee/ and /oo/ are shortened under certain morphological conditions and merge with short /i/ and /u/ phonetically. In addition, Cairo Arabic does have three phonetic diphthongs [iu, au, ai] in a subset of the vocabulary. These are usually analyzed as /iu, aw, ay/ [Harrell 1957]. The following syllables occur: CV, CVC, CVV, CVVC, CVCC. The last two syllables can only occur in word final position and, of course, form monosyllabic words.

2. Procedure

Most of the data illustrating vowels are taken from real monosyllabic words of the types CVVC and CVCC, set in a sentence frame. All contrasting vowel phonemes and the three diphthongs occur in this position. To get examples of short [e] and [o] , some disyllabic words were included, since they cannot occur in monosyllables. Long /ii, aa, uu/ between pharyngalized consonants also occur in a stressed syllable on disyllabic words, since these consonants do not form monosyllables.

Six speakers of Cairo Arabic recorded the sentences. Mean formant values representing each vowel of each speaker were calculated. Using a lab computer the formant values in Hz were converted to mel and plotted on an acoustic chart with $F_1$ against $F_2$ in the usual way.

3. Results

3.1. Vowel duration. Vowel durations were first considered separately in plain and pharyngalized environment. The results show that vowel duration is not significantly different between plain and pharyngalized long vowels, nor between plain and pharyngalized short vowels in these environments. Therefore plain and pharyngalized vowels are considered together. The difference in length between long and short vowels is rather large. Short vowels are about half the duration of long ones.

3.2. Vowel quality. Figure 1 is a formant chart of plain, long vowels. Plain, long vowels are well separated with no overlapping, except for long /ee/ and
Vowels and Diphthongs in Cairo Arabic

Figure 1. Long plain vowels

Figure 2. Short plain vowels
Vowels and Diphthongs in Cairo Arabic

Figure 5: Short phonologized vowels

Figure 6: dispersion
/oo/ touching each other. Figure 2 shows a formant chart of the three plain, short vowels. These vowels are also well separated clusters. Figure 3 also shows plain, short vowels, but it includes the non-phonemic [e] and [o]. Here short [e] and [o] show nearly complete overlapping with short /i/ and /u/. Their non-phonemic status is stated in literature, but phonetic data supporting the linguistic analysis is generally not presented. In some textbooks there is even a claim that there is a phonetic difference between [i] and [e] on the one hand and [o] and [u] on the other [Abdel-Massih 1975]. Figures 4 and 5 show formant charts of the pharyngalized long and short vowels.

4. Discussion

4.1. Vowel length and vowel quality. Differences in vowel length influence vowel quality. For both plain and pharyngalized environments the long vowels are more peripheral, whereas the short vowels are inside the space of the long vowels, with the exception of long /aa/ and short /a/ where the quality difference is small. Both plain and pharyngalized long /i/ differ significantly (p<0.001) from short /i/ along both F₁ and F₂. The short /i/ vowels are lower and further back than the long /ii/ vowels. Short /u/ vowels are lower and more front than long /uu/ vowels. Both sets of long /aa/ differ to some extent from their short counterparts along F₂. The short vowels tend to be further back. The differences along F₁ between long and short /a/ vowels are non-significant; long and short vowels have the same vowel height. It seems as if vowels in Cairo Arabic are anchored on the low vowels and the short vowels /i/ and /u/ do not reach the vowel quality of the long ones.

4.2. Vowel quality in plain and pharyngalized environment. An earlier study [Norlin 1983] analyzed all the fricatives in Cairo Arabic. These include the pharyngalized fricatives /ʃ/ and /z/. In this study FFT spectra of fricatives were converted to critical band spectra. The center of gravity of the critical band spectra was plotted against dispersion, Figure 6. The results from this earlier study showed that the plain and pharyngalized consonants are different, even if the difference is small. Considering the effects of plain and pharyngalized environments on vowel quality, a comparison between long vowels shows that there is complete overlapping for both sets of long /uu/. Pharyngalized long /ii/ shows a small difference from its plain counterpart in that it is slightly lower and further back. Pharyngalized long /aa/, however, is greatly affected, showing a considerable difference in the F₂ dimension from its plain counterpart. The pharyngalized /aa/ is much further back than the plain /aa/ (see Figures 1 and 4).

In conclusion, it is evident that the pharyngalization process affects the whole syllable. On the one hand, plain and pharyngalized consonants differ consistently. On the other hand, vowels in plain and pharyngalized environment differ in more complex ways. Plain and pharyngalized low vowels always show a considerable difference in the F₂ dimension, regardless of length, the pharyngalized vowels being more back. High, long vowels show small or no difference, whereas short high vowels always are further back than plain ones.
Vowels and Diphthongs in Cairo Arabic

Figure 7

Figure 8

Figure 9
5. Diphthongs

Standard Arabic diphthongs /ai/ and /au/ have in Cairo Arabic developed into long /ee/ and /oo/. They also exist in Cairo Arabic, however, in a number of purely dialectal words and some Standard Arabic words commonly used in daily speech. In addition, there exists a third diphthong [iu] due to morphophonematic rules in verb conjugation. A comparison between the short vowels and the corresponding segments in the diphthongs shows some difference in vowel quality. In the diphthongs, F₁ frequencies are always identical with F₁ in short vowels, but F₂ is always lower, making diphthongs segments more back.

The rate of transition between the two components in diphthongs is fast, around 30-35 ms in all the diphthongs. It seems as if the diphthongs in Cairo Arabic are made by stringing the short vowels together with a fast transition (see Figures 7, 8, 9).

[the work reported here is supported by the Swedish council for Research in the Humanities and Social Sciences.]

REFERENCES


THE "INDIGENOUS VERSUS FOREIGN" CONTROVERSY ABOUT THE SOURCES OF SWAHILI VOCABULARY

Derek Nurse
University of British Columbia

Two different kinds of claims are often made about the historical sources of Swahili lexis: (1) Those based on Guthrie's claim that 44% of his PB lexical reconstructions are retained in Swahili today. This has sometimes been reversed to suggest that 44% of today's Swahili vocabulary is of PB origin. (2) Those proposing a heavy foreign, mainly Arabic, component for Swahili. Thus Krumm [1940:2] suggests that 20% of the items used in spoken Swahili are of foreign origin, 30% in written Swahili, and up to 50% in older poetry. Such proposals often imply such levels of loan material are out of line with other languages.

These claims raise several questions. One concerns Guthrie's corpus which contains slightly over 2,000 words, compared with over 400,000 words in the O.E.D., and the estimate of a passive vocabulary of 20,000 items for the average English speaker. Assuming the average Swahili speaker's vocabulary is similar in size to that of the average English speaker, then any estimate based on Guthrie's figures will assess only a small part of that vocabulary. A second concerns the basis of Krumm's methodology, which is unstated. The roundness of his figures indicates they may be rough estimates. A third is the lack of outside control. Even given a standard list of items for comparison, how would Swahili figures stand up against those for other languages, African or non-African? Until we have some idea about that, we cannot know about the degree of their "normality". Finally, the two claims above do not allow for the very real possibility that items may be neither of PB nor foreign origin. We now think that Swahili has existed as a discrete language on the coast for some 1200 years. Before that stretched at least another 2,000 years back to PB. During both periods Swahili must have absorbed material from other languages, Bantu and non-Bantu, yet presently available etymological dictionaries (except Sacleux) do not mention this possibility.

These claims about Swahili may usefully be compared with the types of statement current about English. Williams [1975:67], for example, distinguishes different kinds of vocabulary based on frequency of use (first 1,000 most frequently used words, second thousand, up to 10,000). Major etymological dictionaries differentiate inherited, loan, and other, vocabulary. Their main distinction is between inherited and borrowed elements, taken mainly from French, to a lesser extent from Danish, classical, and other languages. Among the most used English vocabulary is a high percentage of inherited items and a low proportion of loans. Moving into the lexis of lesser frequency, the level of loans from French and classical languages, not Danish, rises dramatically, to the point where after 4,000 words, inherited items form only a quarter of the total, while French and Latin provide 60%. The ratio of inherited to loan items varies along well-recognised parameters such as style, register, topic of discourse, educational level of speaker, etc.
Most of these loanwords can be ascribed to known major facts of the history of English, most obviously 300 years of its history, A.D. 1100-1400, when the ruling classes and their satellites spoke French almost exclusively while the rest of the population spoke English. There is also a long tradition of drawing on Latin and Greek for certain kinds of specialised lexis.

It seemed desirable to break Swahili down along the same general lines just sketched for English, but practical problems necessitated modification. To my knowledge there is no large frequency-based ranking of Swahili vocabulary. We know less of the external history of Swahili than of English. We have scanty knowledge of lexical change and innovation between PB and contemporary Swahili. Little has been done on differentiating foreign loans according to period and dialect of origin. Virtually nothing has been done on local African loans in Swahili. Finally, we do not know the degree of typicalness of Unguja in the overall historical process along the the Swahili dialect spectrum, so we need to compare it to other Swahili dialects, but we have even less data for them than for Unguja/Standard Swahili.

Given these obstacles, I adopted a modified approach and would stress that it represents only a beginning: (a) A slightly modified version of Swadesh's 100-word list was initially applied to 16 coastal Swahili dialects, from north to south: Mwiini, Bajuni, Siu, Pate, Amu, Jomvu, Mvita, Chifunzi, Vumba, Mtang'ata, N. and S. Pemba, Tumbatu, Hadimu-Makunduchi, Unguja, and Mwani. As a form of control, 11 other East African coastal languages (4 Comorian dialects, 4 Mijikenda, 2 Pokomo, plus Elwana) and English were also considered. (b) Each word on each list was carefully scrutinised, and the same basic distinctions drawn as are made for English: (i) Inherited (corresponding to PIE and Germanic), we distinguished PB, Eastern Bantu, North East Coastal Bantu, and PSabaki. Although the number of distinctions made is different, they correspond to all the pre-English, and pre-Swahili, stages respectively. (ii) Loans, divided into foreign (Arabic, Indian, Persian), local (Bantu and Cushitic), and material from other Swahili dialects, relevant when considering the southern Swahili dialects. (iii) Other, a minor category, comprising Swahili-specific items and those of unknown origin.

The results are set out in Table 1 (see p. 248), figures of 0.5 being the result of the presence of two words for some items. These results suggest the following tentative conclusions: (c) Percentages for inherited items for individual Swahili dialects range from just over 70% to 90%, average 80%. Thus, at this level, the bulk of Swahili basic vocabulary is inherited. The figures for the 11 control languages and for English also fall within this range.

(d) The second largest overall source of vocabulary is a loan set from other East African languages, mainly Bantu and Cushitic. The range is from 5% to 15%, average just over 8%. The (mainly Southern) Cushitic set is smaller and stable (older), the Bantu set more variable in size (i.e. more recent), ranging from zero to 10%. With one exception (Elwana), the East African control languages also fell within these limits.

(e) The third largest source is foreign loans, ranging from 2% to 8%, average just under 5%. The only relevant control languages, the Comorian dialects, have a comparable level of outside loans.
(f) With the exception of Mwiini, historical loans from other Swahili dialects are limited to dialects from Mombasa south, ranging from zero in the north to 10% in Unguja, average 3%. Two control languages also have high levels of Swahili loans.

(g) Total figures for loans (d-f above) range from 6% to 23%, average 16%. This compares to 12% for English and the other control languages. The latter figure would almost certainly be raised slightly if some of the "unknown" category were diagnosed.

(h) The remaining items are of uncertain/unknown origin or Swahili-specific (under 5%). The comparable figures for English (4%) and the other East African languages (8%) are similar or a little higher, the latter again reflecting our level of ignorance of these languages.

(i) The Swahili dialects of northern Kenya and Somalia, with the partial exception of Mwiini, have the highest level of inherited items (nearly 90%) and a correspondingly low level of loans from neighboring African languages. The dialects of Tanzania, particularly of Pemba and Zanzibar, have a significantly lower proportion of inherited items explained by their higher level of loans from other Swahili dialects and/or neighboring Bantu languages. The inter-dialect loans in the southern dialects, especially Unguja, derive from the northern dialects. Elsewhere [Nurse 1982:191] I have suggested they have a specific Amu/Pate/Siu shape. Loans from neighboring Bantu languages are most evident in the rural dialects of Zanzibar. The general level of penetration of the dialects south of Mombasa by loans from neighboring languages is almost identical to that of the mainland languages of the control group.

(j) The same northern dialects are also more internally consistent than those further south. This, and (i), apparently reflect the respective geographical position of the two groups: the northern communities are closely adjacent and relatively isolated, whereas the southern dialects cover a longer stretch of coast and islands and are more open to neighboring language communities.

The logical next step would be to apply a longer list, say of 1,000 items, to a selection of the foregoing dialects. Again, this ideal procedure was modified by the lack of available accurate data for many of the dialects for the 1,000 items, and four dialects were selected for which adequate data were available. The figures on the right are taken from Williams [1975] for a 1000-word list for English, but it should be emphasised that the lists used are different and thus not really comparable. William's list is frequency-based, whereas the list used on the four dialects was the more arbitrary one used in the Language Survey of Tanzania. (See Table 2 on page 250).

From Table 2 certain things are obvious. The proportion of inherited items drops, dramatically in the northern dialects, less so for the southern ones. The percentage of foreign loans rises, again dramatically for the northern dialects, less steep for the southern. Comparison with William's figures suggests that this tendency would be magnified if larger samples were taken. Such sampling would, however, encounter major problems in Swahili with many items of unidentified origin.

Other things in Table 2 are less clear, because the level of our ignorance
Table 1a: Sources of lexis for 100-word (Swahili dialects)

<table>
<thead>
<tr>
<th></th>
<th>Mwi</th>
<th>Baj</th>
<th>Siu</th>
<th>Pat</th>
<th>Amu</th>
<th>Jom</th>
<th>Mvi</th>
<th>Chi</th>
<th>Vum</th>
<th>Mta</th>
<th>NP</th>
<th>SP</th>
<th>Tum</th>
<th>Had</th>
<th>Ung</th>
<th>Mwa</th>
</tr>
</thead>
<tbody>
<tr>
<td>PBantu</td>
<td>44.5</td>
<td>44.0</td>
<td>42.5</td>
<td>44.5</td>
<td>44.5</td>
<td>39.5</td>
<td>39.0</td>
<td>42.0</td>
<td>38.5</td>
<td>38.5</td>
<td>34.0</td>
<td>36.0</td>
<td>34.0</td>
<td>33.5</td>
<td>36.0</td>
<td>40.5</td>
</tr>
<tr>
<td>EBantu</td>
<td>32.5</td>
<td>39.5</td>
<td>39.5</td>
<td>38.5</td>
<td>37.5</td>
<td>36.5</td>
<td>36.0</td>
<td>34.0</td>
<td>33.5</td>
<td>33.5</td>
<td>34.0</td>
<td>31.0</td>
<td>31.0</td>
<td>33.5</td>
<td>31.0</td>
<td>28.0</td>
</tr>
<tr>
<td>NECBantu</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>PSabaki</td>
<td>5.0</td>
<td>5.0</td>
<td>6.0</td>
<td>6.0</td>
<td>6.0</td>
<td>5.0</td>
<td>5.0</td>
<td>5.5</td>
<td>5.5</td>
<td>5.0</td>
<td>6.0</td>
<td>5.0</td>
<td>5.0</td>
<td>4.0</td>
<td>5.0</td>
<td>6.0</td>
</tr>
</tbody>
</table>

<p>| | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>total inherited</td>
<td>83.0</td>
<td>89.5</td>
<td>89.0</td>
<td>90.0</td>
<td>89.0</td>
<td>82.0</td>
<td>81.0</td>
<td>82.5</td>
<td>78.5</td>
<td>78.0</td>
<td>75.0</td>
<td>73.0</td>
<td>71.0</td>
<td>72.0</td>
<td>73.0</td>
<td>75.5</td>
</tr>
</tbody>
</table>

<p>| | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ECushitic</td>
<td>4.0</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
</tr>
<tr>
<td>SCushitic</td>
<td>1.0</td>
<td>2.0</td>
<td>2.0</td>
<td>1.5</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>3.0</td>
<td>2.0</td>
<td>1.5</td>
<td>2.0</td>
<td>3.0</td>
<td>4.0</td>
<td>3.0</td>
<td>1.0</td>
</tr>
<tr>
<td>other Sw.</td>
<td>4.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.0</td>
<td>3.0</td>
<td>1.0</td>
<td>2.5</td>
<td>2.0</td>
<td>3.0</td>
<td>5.5</td>
<td>5.5</td>
</tr>
<tr>
<td>Bantu</td>
<td></td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>2.0</td>
<td>3.0</td>
<td>6.0</td>
<td>7.0</td>
<td>8.5</td>
<td>9.0</td>
<td>7.5</td>
<td>9.0</td>
</tr>
</tbody>
</table>

<p>| | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>total local loans</td>
<td>9.0</td>
<td>4.5</td>
<td>5.0</td>
<td>4.0</td>
<td>4.5</td>
<td>8.5</td>
<td>8.5</td>
<td>9.5</td>
<td>13.0</td>
<td>13.0</td>
<td>14.0</td>
<td>15.5</td>
<td>18.0</td>
<td>18.5</td>
<td>17.5</td>
<td>14.0</td>
</tr>
</tbody>
</table>

<p>| | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>4.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>4.0</td>
<td>5.0</td>
<td>2.5</td>
<td>3.5</td>
<td>3.5</td>
<td>4.0</td>
<td>5.0</td>
<td>4.0</td>
<td>2.0</td>
<td>4.0</td>
<td>3.0</td>
</tr>
<tr>
<td>India</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>2.0</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>unclear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
<td>1.0</td>
<td>1.5</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

<p>| | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>total foreign loans</td>
<td>4.0</td>
<td>2.0</td>
<td>3.0</td>
<td>3.0</td>
<td>3.5</td>
<td>6.0</td>
<td>7.0</td>
<td>4.5</td>
<td>5.5</td>
<td>5.5</td>
<td>6.5</td>
<td>8.0</td>
<td>5.5</td>
<td>3.0</td>
<td>6.0</td>
<td>3.0</td>
</tr>
</tbody>
</table>

<p>| | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Swahili-specific</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>?</td>
<td>3.0</td>
<td>3.0</td>
<td>2.0</td>
<td>2.0</td>
<td>2.0</td>
<td>3.5</td>
<td>3.5</td>
<td>3.5</td>
<td>3.0</td>
<td>3.5</td>
<td>4.5</td>
<td>3.5</td>
<td>5.5</td>
<td>6.5</td>
<td>3.5</td>
<td>6.5</td>
</tr>
</tbody>
</table>
Table 1b: Sources of lexis for 100-word list (Sabaki languages and English)

<table>
<thead>
<tr>
<th>Source Type</th>
<th>Comorian</th>
<th>Mijikenda</th>
<th>Pokomo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ngaz</td>
<td>Mwal</td>
<td>Ndzw</td>
</tr>
<tr>
<td>PBantu</td>
<td>43.5</td>
<td>41.5</td>
<td>43.0</td>
</tr>
<tr>
<td>EBantu</td>
<td>36.5</td>
<td>35.5</td>
<td>35.0</td>
</tr>
<tr>
<td>NECB</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>PSabaki</td>
<td>3.0</td>
<td>4.0</td>
<td>4.0</td>
</tr>
</tbody>
</table>

| Inherited    | 84.0     | 81.5      | 83.5    | 82.0   | 84.5 | 83.5 | 81.5 | 74.0 | 82.0 | 77.0 | 73.0   | 84.0     |

| ECushitic    | -        | -         | -       | -      | -    | -    | -    | 4.0  | 1.0  | 9.5  |          |          |
| SCushitic    | 1.0      | 1.0       | 1.0     | 1.0    | 1.5  | 2.0  | 2.0  | 4.5  | 3.0  | 4.5  | 1.0    |          |
| Cushitic     | -        | -         | -       | -      | -    | -    | -    | 1.0  | -    | -    | 0.5    |          |
| Thagicu      | -        | -         | -       | -      | 1.0  | 1.0  | 3.5  | 2.0  | 1.5  | 2.0  | 6.0    |          |
| Swahili      | 1.5      | 1.5       | 2.5     | 1.5    | -    | 1.0  | -    | 9.5  | 2.0  | 9.0  | -      |          |
| Bantu        | 2.5      | 1.5       | 1.5     | 2.0    | 4.0  | 4.0  | 3.0  | 3.0  | 2.5  | 2.0  | -      |          |
| Maasai       | -        | -         | -       | -      | -    | -    | -    | 1.0  | 1.0  | -    | -      |          |

| Local loans  | 5.0      | 4.0       | 5.0     | 4.5    | 6.5  | 8.0  | 9.5  | 21.0 | 13.0 | 18.0 | 17.0   | 1.0      |

| Arabic       | 2.5      | 2.5       | 2.5     | 3.0    | -    | -    | -    | -    | -    | -    | -      |          |
| unclear      | 1.0      | 2.0       | 2.0     | 3.0    | -    | -    | -    | -    | -    | -    | -      |          |
| French       | -        | 1.0       | 1.0     | 1.0    | -    | -    | -    | -    | -    | -    | -      |          |

| Foreign loans| 3.5      | 5.5       | 5.5     | 7.0    | -    | -    | -    | -    | -    | -    | -      | 11.0     |

| ?            | 7.5      | 9.0       | 6.0     | 6.5    | 9.0  | 8.5  | 9.5  | 5.0  | 5.0  | 4.0  | 9.5    | 3.0      |

| Dialect      | 1.0      |          |         |        |      |      |      |      |      |      |        |          |
Table 2: Sources of lexis for 1000-word list (4 Swahili dialects)

<table>
<thead>
<tr>
<th></th>
<th>Bajuni</th>
<th>Pate</th>
<th>N. Pemba</th>
<th>Unguja</th>
<th>(English)</th>
</tr>
</thead>
<tbody>
<tr>
<td>inherited</td>
<td>68.0</td>
<td>69.0</td>
<td>66.0</td>
<td>65.5</td>
<td>(83.0)</td>
</tr>
<tr>
<td>African loans</td>
<td>9.0</td>
<td>9.0</td>
<td>8.0 ?</td>
<td>7.5 ?</td>
<td></td>
</tr>
<tr>
<td>Swahili dialect loans</td>
<td>-</td>
<td>0.5</td>
<td>2.0 ?</td>
<td>4.5 ?</td>
<td></td>
</tr>
<tr>
<td>foreign loans</td>
<td>7.5</td>
<td>7.5</td>
<td>10.0</td>
<td>10.0</td>
<td></td>
</tr>
<tr>
<td>total loans</td>
<td>16.5</td>
<td>17.0</td>
<td>20.0 ?</td>
<td>22.0 ?</td>
<td>(15.0)</td>
</tr>
<tr>
<td>Swahili-specific</td>
<td>2.0</td>
<td>2.0</td>
<td>1.5</td>
<td>2.0</td>
<td></td>
</tr>
<tr>
<td>unknown</td>
<td>13.5</td>
<td>12.0</td>
<td>12.5</td>
<td>10.5</td>
<td>(2.0)</td>
</tr>
</tbody>
</table>

rises. The proportion of items of unknown origin triples, making the figures for all local loans combined, and thus the total loans, less reliable. Since it is likely most of the unknown items are of local provenance, the figures for loans from local languages and other Swahili dialects would probably also rise if we could diagnose the unknown set.

The general conclusion from this limited investigation is that the proportion of inherited to loan material in Swahili is not significantly different statistically from that in English or the other East African control languages.

REFERENCES


A NEW ANALYSIS OF THE KRIO CLEFT PREDICATE

Dudley K. Nylander
University of Manitoba

The cleft predicate is a construction found in a number of African and creole languages. In this paper, I propose a new (transformational) analysis of the cleft predicate, using data from Krio (an English-based creole language spoken in Sierra Leone and elsewhere in West Africa). The basic characteristic of the cleft predicate is that the cleft element is a copy of the matrix verb, as shown in (1). The cleft element must be identical to the matrix verb, as shown by the contrast between (1) and (2). The matrix verb cannot be cleft, as shown by the contrast between (1) and (3).

(1) [\textit{S}' \textit{nà brúk} \textit{i} [\textit{S}' [\textit{S} m' pád' dè brúk \textit{in bùbà}]]] it is-wash- my-friend-PROG-\textit{wash} her-dress (\textit{PROG}=progressive aspect)

'what my friend is doing is washing her dress' (lit. it is wash that my friend is washing her dress)

(2) *[\textit{S}' \textit{nà brúk} \textit{i} [\textit{S}' [\textit{S} m' pád' dè kíñ ñ \textit{in bùbà}]])

(3) *[\textit{S}' \textit{nà brúk} \textit{i} [\textit{S}' [\textit{S} m' pád' dè e ñ \textit{in bùbà}]]]

The ungrammaticality of (3) can be attributed to the fact that the trace of the cleft verb is not properly governed. This implies that the auxiliary element dè is not a governor. Nor is \textit{in bùbà}, since it is an NP, i.e. a maximal projection. Two other features of the cleft predicate are (i) the cleft element is incompatible with an auxiliary element, as shown by the contrast between (4) and (1); (ii) the cleft element is incompatible with a direct object, as shown by the contrast between (5) and (1).

(4) *[\textit{S}' \textit{nà dè ñ brúk} i [\textit{S}' [\textit{S} m' pád' dè ñ brúk i \textit{in bùbà}]])

(5) *[\textit{S}' \textit{nà brúk} i [\textit{NP} \textit{in bùbà}] j [\textit{S}' [\textit{S} m' pád' dè brúk i [\textit{NP} \textit{in bùbà}] j]]]

The phenomena illustrated in (4) and (5) have led a number of linguists (including Välman [1978:46]) to resort to the notion of "deverbal", to account for the fact that the cleft element has undoubtedly lost its verbal status. However, the notion of "deverbal" raises more questions than it answers. In particular, how would a "deverbal" be represented in the lexicon? More precisely, if we adopt the notion, we would be forced to list the cleft \textit{brúk} in (1) as a "deverbal" and the matrix \textit{brúk} as a verb in the lexicon. Now, if a verb can be functionally defined as an element that can combine with auxiliary markers, e.g. dè, no such definition is available for a "deverbal". Furthermore, as shown by the contrast between (1) and (3), there can be no "deverbal" without the presence of the corresponding matrix verb.
In this paper, I would like to propose a new analysis of the cleft predicate. The analysis is based on the observation that Krio has a base rule which allows reduplication. The rule is given in (6) and its use is illustrated in (7). Krio also has the S'' expansion rule given in (8). The rule is independently required to account for the clefting of NPs, PPs, and APs.

(6) $VP \rightarrow V_1 V_i (NP)$

(7) $m' \, \text{padd} \, d' \, \text{bruk}, \text{bruk}, \in \text{bubâ} \quad 'my \, friend \, is \, constantly \, washing her \, dress'$

(8) $S'' \rightarrow \text{TOP} \, \alpha \, S'$

Suppose, now, we adopt the essentials of Chomsky's [1977] analysis of cleft sentences, i.e. (i) the cleft element is directly generated under S'' and (ii) a copy of the cleft element is generated in S and then moved into COMP and there deleted. Then, combining (7), (8) and the transformational rule "Move a ", we have four independent possibilities, illustrated in (9). (9a) and (9b) have the same meaning as (1).

(9a) $[S'' [\text{TOP na}] [V \, \text{bruk}], renamed \, V \, [S', [S m' \, \text{padd} \, d' \, \text{bruk} \, e_i \, \in \, \text{bubâ}]]]$

(9b) $[S'' [\text{TOP na}] [V \, \text{bruk}], renamed \, V \, [S', [S m' \, \text{padd} \, d' \, \text{bruk} \, e_i \, \in \, \text{bubâ}]]]$

(9c) $[S'' [\text{TOP na}] [NP \, \in \, \text{bubâ}], renamed \, V \, [S', [S m' \, \text{padd} \, d' \, \text{bruk} \, \text{bruk} \, e_i]]]$

(9d) $[S'' [\text{TOP na}] [NP \, \in \, \text{bubâ}], renamed \, V \, [S', [S e_i \, \text{bruk} \, \text{bruk} \, \text{bruk} \, \in \, \text{bubâ}]]]$

In (9a), $e_i$ (the empty category) is governed by $\text{bruk}$ (and not by $\in \text{bubâ}$, which, as we have seen, is a maximal projection). In (9b), $e_i$ is also governed by $\text{bruk}$ (and not by $\in$ —recall (3)). In (9c), $e_i$ is also governed by $\text{bruk}$. If we adopt Kayne's [1981] version of the Empty Category Principle (ECP) in which government by a lexical category does not remove the need for an antecedent, then, to account for (9a), (9b), and (9c), we are forced to assume that S'' (which contains the antecedent of $e_i$ in each case) is a projection of V, the lexical governor. That S'' is a projection of V is suggested by Chomsky [1981:140, footnote 27], but needs further investigation.

Sentence (9d) is more problematic. But, suppose we assume, following Rou-Vert & Vergnaud [1980:115] that S' deletes, since COMP is null. Suppose, further, that S is Chomsky-joined to the cleft NP. The result is (10).

(10) $[S'' [\text{TOP na}] [NP_j [NP_i m' \, \text{padd}]]_1 [S e_i \, \text{bruk} \, \text{bruk} \, \text{bruk} \, \in \, \text{bubâ}]]$
Suppose we assume, further, that (i) S is not a maximal projection and (ii) a maximal projection cannot be dominated by a category of the same type. (ii) allows NP\_j (but not NP\_j) to be a maximal projection. Then the noun pàd\_i governs e\_i, if we adopt the definition of government given by Sportiche and Aoun [1981], see (11). The relevant maximal projections in (10) are S\_j and NP\_j, both of which dominate pàd\_i and e\_i.

(11) Government
\[ \alpha \text{ governs } \beta \text{ if } \alpha = X^0 \text{ and } \alpha \text{ and } \beta \text{ share all maximal projections} \]

The preceding analysis implies that N is a governor, contrary to Kayne's [1982] claim. But it can be shown that Krio has features which Kayne claims not to exist, e.g. genitival constructions in which a non-maximal projection of N governs a maximal projection of N (see (12a)) and tensed clauses introduced by prepositions (see (12b)). As such, Krio shows the need for a parametrised version of Kayne's theory.

(12) a. [\[NP [N John] \[N" \{n ð5]\]]]
John\_i his-house

b. [\[g\ld\d\] f\{ w\{ n\{ d\{ k\{ à \]
I-bè happy-for-that-you-PERF-come
(PERF = perfective aspect)

Returning now to (4) and (5), the ungrammaticality of these sentences follows from a very simple fact, namely, that the verb brùk does not form a constituent with the auxiliary element or the direct object.

REFERENCES


PROBLEMS IN THE METRICAL REPRESENTATION OF TONE

David Odden
Yale University

Although there has been considerable research into deep tonal alternations in nonlinear phonology, less attention has been paid to relating abstract tonal entities to actual pitch values. A typical example of this problem is downdrift, whereby H and L tones do not receive an invariant translation into specific pitch levels, but are modified by a rule which gives the first H and L tones relatively higher pitch, and gives following tones relatively lower pitch: this iterates through the string so that every span of H and L tones is lower than the preceding. Clements [1981] has argued that such phenomena can be given a metrical treatment. I argue that the machinery of metrical tone register predicts that a wider range of tonal processes could be found and then show that much of the predicted range of variation in tone register systems is actually attested.

In the metrical treatment of tone terracing, different tonal registers are described in terms of a metrical tree built on H and L tones with the labels h and l, such as in (1).

(1) Shona

\[
\begin{array}{c}
\text{va na va no tI} \\
\text{va \textit{na} va no tI} \\
\text{\text{'children who say'}}
\end{array}
\]

This tree is generated by three rules:

(2) a. Every tonal matrix containing H preceded by L forms the left branch of a maximal n-ary branching tree.

b. Any remaining tonal matrices are gathered into an n-ary branching tree.

c. Trees constructed by (a) and (b) are gathered into right-branching binary trees labeled by the labeling principle h,l.

A change in one of these parameters for tree construction can yield a different tone register system. Consider first a process much like downdrift, namely automatic downstepping, found in Kishambaa, inter alia, which has predictable downdrift. Unlike the usual case of downdrift, every H is downstepped after a H as well. (See example 3 on the next page). The trees for this type of downdrift are constructed by rules like those for classical downdrift, except in Kishambaa, we specify that a H forms a left branch everywhere. The crucial difference is that in Kishambaa, when H is preceded by H, both H's form the left branch of a foot, so the first H forms the only member of its foot, hence adjacent H tones are "downstepped".
The metrical representation of tone register also handles the problem of Kipare downstep. In this language, there is no downdrift, which alone is handled by gathering all tones into a single foot. Although there is no downdrift, there is downstep, which arises under two conditions. It arises when a L tone is set afloat by a rule spreading H leftward in *niwakome makui* 'I should kill the dogs'. The tree construction rules of Kipare are very much like those found in Kikuyu [Clements 1981]. In both languages, only a floating L tone defines the edge of a tonal foot. Kipare differs from Kikuyu in that the drop in pitch between registers is less than the drop between different tones in Kipare, while in Kikuyu, the drop between registers is actually the same as the drop between different tones ("absolute downstep").

Kipare also has a rule assigning an utterance-final H to a lower register, so that underlying *ipängä* becomes *ipängä*. This can be described by another rule which makes a final H be the sole member of a degenerate foot. This process differs from related processes of downdrift in that the environment for defining the foot is the end of an utterance, rather than some particular neighboring tone. The rule labeling sequences of feet h,l which accounts for downsteps conditioned by floating L tones also applies to the foot created by the "marginal downstep" rule.

A similar process in final position is found in Kenyang, which makes an utterance-final L the sole member of a degenerate foot. Kenyang also has downdrift, so we independently need the labeling principle h,l for the feet constructed by the downdrift rule. Kenyang thus differs from Kipare in its utterance final register rules only by having a L tone rather than an H tone form the foot. At the utterance level, sequences of feet are grouped together into a tree labeled h,l just as they are in Kipare, Shona, Kikuyu, and Kishambaa.

The difference between these languages lies in what tones form left-branches, what environmental factors condition foot formation, and what the physical interpretation of register shift is. They have in common the fact that all groups of feet are gathered into a tree labeled h,l. But trees can be labeled l,h as well, and Clements [1981] shows that Zulu has a subsystem of downdrift, whereby a downstep becomes an upstep in a certain environment. Although upstep in Zulu is a modification of the basic labeling principle h,l, there are other register phenomena which require l,h labeling independently. Thus, the selection of the labeling principle h,l versus l,h is an independent variable, just as the conditioning for foot formation or the tones which form branches are in-
dependent variables.

A case of automatic upstep is found in Kimatuumbi. In this language, every H in a sequence of H's is automatically lower than the following H's. Kimatuumbi is then in one sense like Kishambaa; it has automatic stepping of H's after H's, so we will describe it by making every H the sole member of a degenerate foot. Kimatuumbi differs from Kishambaa mainly in the labeling principle, which is 1,h in Kimatuumbi, but h,1 in Kishambaa.

(4)

```
1 h
  1 h
   1 h
    L H H H H LH L
    baat'lyé k'ndyé l|lsó 'they ate the birds yesterday'
```

A similar upsweep rule is found in Bukusu, where a string of L tones is realised with ascending pitch, as described in Austen [1974]. Bukusu is thus described by the same mechanism used to describe Kimatuumbi, except that in Bukusu, L tones, and not H tones, form a degenerate foot.

There are similar processes which affect only a single tone standing at some phonologically defined margin. In Engenni, a H tone before a L stands in a higher register than preceding H's [Thomas 1974].

(5) mi móní wó 'I saw you' mi móní 'wó bhee 'I did see you'

```
[ H H H ] [ H H ]
[ L ] [ L LL ]
```

Engenni thus employs the labeling principle 1,h just like Kimatuumbi and Bukusu. The foot formation rule is different in Engenni; here, a H tone forms a degenerate foot only before a L tone, not everywhere. This is similar to the rule in Kipare where a H tone forms a degenerate foot before pause; the surface difference between these two languages results from the different labeling principles for the languages. A similar "marginal upstep" process is found in Imithupi Makua, where a H tone before a string of H's is lower than following H's. To describe this we merely specify that a H tone after a L or at the beginning of an utterance forms a degenerate foot, and that the following string of free H tones forms a foot. The labeling principle for Makua is 1,h; it is the position of the conditioning material that distinguishes Makua from Engenni.

There are a considerable number of tone-register problems in languages beyond simple downdrift and downstep. Yet these all can be handled in a metrical theory by variations in a few parameters, including class of input tones, the conditioning environment, the physical interpretation of register differences, the labeling principle, and the inclusion of free tones into feet. A metrical
The Metrical Representation of Tone

theory of tone register handles not just classical downdrift, but also seemingly aberrant forms of downdrift as are found in Kishambaa and Kipare, as well as a whole range of upsweep phenomena and a number of processes not discussed here. The machinery underlying the metrical theory of register predicts that such phenomena should be possible. The fact that the theory generalises to handle other phenomena besides those discussed in Clements [1981] provides quite strong support for the theory.

REFERENCES


LE FINI/L’INFINI OU
L’AFFIRMATION/L’INTERROGATION EN MOBA
(LANGUE VOLTAIQUE PARLÉE AU NORD-TOGO)

Annie Rialland
C.N.R.S., E.R.A. 433

1. Les Marques Prosodiques: le Fini/l’Infini

Nous avons montré que l’opposition entre affirmation et interrogation n’utilisait pas de marques intonatives au sens propre, c’est-à-dire de marques mélo-diques (tons de frontière ou suppression de downstep par exemple). La mélodie reste réservée aux réalisations des tons (4 tons et downstep; voir pour l’étude du système tonal, Rialland [1983]). Les marques prosodiques de l’affirmation/interrogation sont d’une autre nature: elles consistent dans le mode d’arrêt de la voix, brutal pour l’affirmation (le fini), doux et progressif pour l’interrogation (l’infini).

Une étude phonétique nous a permis de préciser la nature de cette opposition, tant du point de vue physiologique qu’acoustique. L’arrêt brusque de la voix est obtenu, physiologiquement, par un coup de glotte, l’extinction lente et douce par une ouverture progressive de la glotte à la fin de la voyelle finale qui s’était dans un souffle. Du point de vue acoustique, l’extinction lente se traduit par un décrescendo et une disparition progressive des harmoniques élevés au cours de la voyelle. Le mode d’arrêt de la voix a également pour corrélat la longueur vocalique: l’arrêt brusque a pour effet de raccourcir la voyelle finale, l’extinction lente, au contraire, de l’‘étirer’.

À titre d’exemple, nous nous en tiendrons, pour ce résumé, aux tracés correspondant à l’opposition /tâňä'/ [tâňä] 'des pierres', /tâňä.../[tâňä...] 'des pierres?' (les trois points indiquant l’étirement des voyelles finales) (Figure 1).

La voix s’arrête rapidement. La Fo correspond à la réalisation des tons, en l’occurrence des deux tons hauts (Figure 2).

On voit que la voyelle finale est "étirée" et que son énergie décroît progressivement, ce que traduit la forme en fuseau de l’oscillogramme. La Fo dont on peut constater qu’elle reste inchangée par rapport à la Figure 1, correspond toujours à la réalisation des deux tons hauts (Figure 3). N.B. Le zéro (0) du débit d’air est en haut de la figure. La remontée de la ligne traduit donc une diminution.

On notera que le débit d’air décroît rapidement à la fin de la voyelle et qu’il devient vite nul, ce qui est dû à la fermeture glottale (Figure 4, p. 260).

On note ici, au contraire, que le débit d’air augmente: la glotte s’ouvre et la voyelle se termine dans un souffle. Sur l’oscillogramme, la simplification de la forme d’onde indique que les harmoniques élevés s’affaiblissent et disparaissent progressivement.
Figure 1. Le fini: l'affirmation

Figure 3. L'affirmation

Figure 2. L'infini: l'interrogation
2. Les Marques Segmentales: les Formes de Base dans l’Interrogation

Dans un deuxième temps de notre communication, nous avons montré que les formes apparaissant dans l’interrogation étaient les formes de base tandis que, dans l’affirmation, ces formes avaient subi troncation ou métathèse. Ceci est évidemment contraire au sens habituel des dérivations ou transformations entre affirmation et interrogation. Le tableau suivant permet de comparer les formes des mêmes noms dans l’affirmation (ou à l’intérieur d’enoncé) et dans l’interrogation. Dans la dernière colonne, nous avons indiqué les noms équivalents dans une langue apparentée, très proche mais plus conservatrice, le gurma.

<table>
<thead>
<tr>
<th>Affirmation (ou intérieur d’enoncé)</th>
<th>Interrogation</th>
<th>Gurma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noms avec suffixe gu</td>
<td>faàk</td>
<td>faàgù</td>
</tr>
<tr>
<td></td>
<td>kôbùk</td>
<td>kôbgù</td>
</tr>
<tr>
<td>Noms avec suffixes plènù, plèn+ù</td>
<td>plènù...</td>
<td>plèmù</td>
</tr>
<tr>
<td>ū', ū', à tânà</td>
<td>tânà...</td>
<td>tânà</td>
</tr>
<tr>
<td>Noms avec suffixes bìk'</td>
<td>bì+k'</td>
<td>bìg'</td>
</tr>
<tr>
<td>kl, l', m'</td>
<td>tûòi'</td>
<td>tûòi</td>
</tr>
<tr>
<td>tîl, bl', l'</td>
<td>pûm'</td>
<td>pîmû</td>
</tr>
</tbody>
</table>

Note: Les formes tronquées ou métathèses sont indiquées. Les noms en Gurma sont donnés en fonction des contextes phonétiques fort ou faible.
La Prosodie en Moba

<table>
<thead>
<tr>
<th>Affirmation</th>
<th>Interrogation</th>
<th>Gurma</th>
</tr>
</thead>
<tbody>
<tr>
<td>kúť</td>
<td>kúď'</td>
<td>kúďT</td>
</tr>
<tr>
<td>nìb'</td>
<td>nìbî...</td>
<td>nìbā/nìbT</td>
</tr>
<tr>
<td>bát'</td>
<td>báď'...</td>
<td>bádo/bádT</td>
</tr>
</tbody>
</table>

(Pour une étude des formes "faibles" et "fortes" en gurma et de leurs emplois, voir Riallant [1981]).

On voit qu’à l'affirmation, la métathèse affecte la suffixe ġù et que la troncation affecte les suffixes de la forme ġ(1). On aura pu également remarquer que, segmentalement, les formes des suffixes moba dans l'interrogation sont identiques aux formes faibles des suffixes gurma. Ainsi, dans l'interrogation, se trouvent conservées des formes plus archaïques.

3. Relations entre Marques Prosodiques et Marques Segmentales ou le Rôle du "Poids" Prosodique

L'existence de la marque prosodique "infinie" explique la présence des formes de base ou formes "pleines" dans l'interrogation ainsi que le maintien de formes plus archaïques. En conférant plus de "poids" (pour reprendre ici le terme de L. Hyman [1983]) à la voyelle finale, elle l'a préservée.

4. Conclusion

L'opposition affirmation/interrogation, en moba, est caractérisée par des marques prosodiques qui ne mettent pas en jeu la mélodie: le fini (ou arrêt brutal de la voix)/l'infini (extinction lente de la voix). La marque prosodique "infinie", par l'accroissement de poids qu'elle confère à la voyelle finale, a eu pour effet de sauvegarder les formes pleines des mots, lesquelles, à l'affirmation, subissent métathèse et troncation.

[Les tracés ont été obtenus à l'Institut de Phonétique de Paris III. Les enregistrements, pour cette étude, ont été faits, en partie, à Paris, en partie, à Bombouaka (Nord-togo).]

REFERENCES


0. **Introduction**

The purpose of this paper is to give an account of the properties of one type of serial-verb construction (henceforth, SVC) in terms of Government Binding (GB) theory, as outlined in Chomsky [1981]. The language which forms the object of study is Gokana, an Ogoni language of Eastern Nigeria.

1. **Data**

   The structure of the SVC we are concerned with is given schematically in (1):

   $$(1) \text{NP} \quad \text{Tns(Neg)} \quad V_1 \quad \text{NP}_1 \quad V_2 \quad \text{ma} \quad \text{NP}_2$$

   In all the examples we will discuss, $V_1$ is $\text{tú} \ 'take'$, so we will refer to the construction illustrated in (1) as the $\text{tú} \ldots \text{ma}$ construction. Some examples are shown in (2):

   $$
   (2) \\
   \begin{align*}
   a. & \quad \text{aè tú fíra kùùmà nùtc} \quad \text{'he opened the door with the key'} \\
   b. & \quad \text{aè tú ny1manu bùmà nom} \quad \text{'he cooked the meat intelligence cook-ma animal ligently'} \\
   c. & \quad \text{mì tú kpègè zamá glorò} \quad \text{I-Past take money buy-ma slave} \\
   d. & \quad \text{aè tú gìma kroolì nòm} \quad \text{'he cut the meat with the knife'}
   \end{align*}
   $$

   We can see that $\text{tú}$ is necessary for grammaticality from (3):

   $$
   (3) \quad *\text{aè fíra kùùmà nùtc} \\
   \text{he-Past key open-with door}
   $$

   Similarly, (4) demonstrates the necessity of the presence of $\text{ma}$:

   $$
   (4) \quad *\text{aè tú fíra kùùrà nùtc} \\
   \text{he-Past take key open door}
   $$

   At this point, one might be tempted to question the categorial status of $\text{tú}$ . Is $\text{tú}$ a Verb? Morphological tests indicate clearly that it is.

   First, $\text{tú}$ obligatorily takes 1Sg and 2Sg present-tense agreement, which takes the form of an extra low-tone copy of the final vowel. This is shown in (5):
Serial Verbs and Government Binding Theory

Second, tů obligatorily shows 2Pl agreement in -i and the logophoric marker, a harmonically conditioned e/ε. Both of these affixes attach outside of other affixes, e.g. -ma, and to both Verbs of an SVC:

(6) a.  oò tui  gīma kpoomā nom
    you-pl-past take-2pl knife cut-ma-2pl animal
    'you cut the meat with the knife'

   b. aè ko aè tūe gīma kpoomāe nom
    he-past say he-past take-log knife cut-ma-log animal
    'he said he cut the meat with a knife'

Third, tů can itself appear with -ma attached, occupying the place of V₂ in (1):

(7) aè tū gīma tuumā nom lɔ̃ m’m bā
    he-past take knife take-ma animal remove in pot
   'he removed the meat from the pot with a knife'

So it seems that tů is a Verb. That V₂ is a Verb is seen from the attachment of ma, as this affix attaches only to Verbs.

The question we address in this paper is, then, "How can we derive the properties of the tů...ma construction from general principles of Universal Grammar (UG)?" To answer this question, we need to look at some of those principles.

2. Theory

Our account of the tů...ma construction centres on the Case theory and θ-theory modules of UG. The central structural relation of each of these modules is government, defined as follows:

(8) Government: A governs C in a configuration like

   \(( \_B \ldots C \ldots A \ldots C \ldots )\)

   where: (i) \(A = X^0\) (= a lexical element);
   (ii) where P is a maximal projection, if P dominates C, then either P dominates A, or P is the maximal projection of C;
   (iii) A c-commands C.

   [Belletti & Rizzi 1981]
The instances of government that will concern us are primarily the following:

(9) a. \((\text{VP} \ V \ NP)\)
    b. \((\text{PP} \ P \ NP)\)

In (9a), V governs NP and the head of NP; in (9b), P governs NP and the head of NP.

The principal condition of Case Theory is the Case Filter:

(10) **Case Filter:** *NP if NP is both phonologically realised and lacks (abstract) Case.

In order to not lack Case, and so pass the Case Filter, we assume that an NP's inherent abstract Case feature must be matched under government with the abstract Case feature of some other item. In (9a), for example, if V has an Accusative feature, NP must have an Accusative feature also. The features will match under government and so NP will pass the Case Filter. Note that the Case Filter does not hold of empty categories, as we have stated it here.

What happens to an NP in the environments shown in (11)?

(11) a. \((\text{NP} \ N \ NP), \text{e.g.} \ *\text{destruction the city}\)
    b. \((\text{AP} \ A \ NP), \text{e.g.} \ *\text{proud John}\)

Adjectives and Nouns lack Case features which can match in the way just described. That is why the examples given in (11) are bad. However, languages possess various ways of escaping the effects of the Case Filter in such examples. In English, the Preposition of can be inserted. This Preposition has a Case feature, and so the NPs can now match their inherent features with the feature of of. As a result, the Case Filter is satisfied, and we find NPs and APs of the type of destruction of the city, proud of John. We will see that Gokana also makes use of "dummy Case-assigners", analogous to English of.

If a Verb is subcategorised for an argument, e.g. a direct object NP or complement S/\overline{S}, it assigns a θ-role to that argument. θ-roles are semantic notions of the type Theme, Patient, Agent, etc. They thus express the semantic relations between Verbs and their arguments. θ-roles are assigned subject to the following well-formedness condition:

(12) Every argument must be assigned one and only one θ-role, and every θ-role must be assigned to one and only one argument.

(12) is the θ-criterion. θ-roles are lexically specified properties of Verbs, linked to subcategorisation. Lexical properties are guaranteed in syntax by the following requirement, known as the Projection Principle:

(13) Lexical properties must hold at all syntactic levels, i.e. D-Structure, S-Structure, and Logical Form.
Serial Verbs and Government Binding Theory

(12) and (13) combine to ensure that all θ-roles will be uniquely and exhaustively assigned at all syntactic levels. θ-role assignment takes place under government.

The Case Filter has been linked to θ-role assignment in the following way:

(14) NPs must be Case-matched in order to be assigned θ-roles.

(14) is called the Visibility Condition. Given (12) and (13), NPs (at least, those which are subcategorised arguments) must be Case matched in the manner described above.

Having presented the necessary theoretical notions, we now return to the Gokana tū...ma construction

3. Gokana

The properties of the tū...ma construction can now be deduced from the following statements:

(15) a. -ma is an affix which adds a θ-role to the Verb to which it attaches, but which lacks a Case feature;

b. tū is a Verb which transmits, rather than assigns, a θ-role, but which has a Case feature.

We will first support these statements and then show how they derive the properties of the tū...ma construction.

Given the θ-criterion, no argument NP is allowed unless it is assigned a θ-role by some Verb. What we find in Gokana is a consistent parallel between the number of arguments a Verb allows and the number allowed by that Verb + ma. In each case, the Verb + ma combination allows one "extra" argument NP in the clause. This is illustrated in (16):

(16) a. Intransitive Verb:

\[
\begin{align*}
\text{aè} & \quad \text{kpe} & \quad \text{he-Past pay} & \quad \text{'he paid'} \\
\text{aè} & \quad \text{tū kpe\textg kpe\textmā} & \quad \text{he-Past take money pay-mā} & \quad \text{'he paid cash'}
\end{align*}
\]

b. aè \quad kpo nom \quad \text{he-Past cut animal} \quad \text{'he cut the meat'}

\[
\begin{align*}
aè & \quad \text{tū gī\textma kpo\textmā nom} & \quad \text{he-Past take knife cut-mā animal} & \quad \text{'he cut the meat with the knife'}
\end{align*}
\]

c. Ditransitive Verb:

\[
\begin{align*}
aè & \quad \text{nè nüü nwīn} & \quad \text{he-Past give thing-assoc child} & \quad \text{'he gave the thing to the child'}
\end{align*}
\]
c. Ditransitive Verb:

\[
\begin{align*}
\text{aè tù nyìmànù neàma gìmà nwìn} \\
\text{he-Past take intelligence give-ma knife-assoc child}
\end{align*}
\]

'he intelligently gave the knife to the child'

The simplest way to account for this is by saying that \( \text{ma} \) adds a \( \Theta \)-role to the Verb to which it attaches.

Further support for this view comes from evidence of another affix, \( \text{-a} \), which attaches to Verbs and affects their \( \Theta \)-roles. In this case, however, the affix in question takes away one of the Verb's arguments. As a result, transitive Verbs receive an ergative or middle interpretation.

Examples of the productive affixation of \( \text{-a} \) and \( \text{-ma} \) to Verbs, with the concomitant effects on the Verb's \( \Theta \)-roles, are given below:

(17) a. \( \text{aè lìg fìra} \) 'he twisted the key'  
\( \text{he-past twist key} \)

b. \( \text{aè tù bà lìmà fìra} \) 'he twisted the key with his hand'  
\( \text{he-past take hand twist-ma key} \)

c. \( \text{fìra è lìgà} \) 'the key twisted'  
\( \text{key past twist-a} \)

(18) a. \( \text{aè zuge bà} \) 'he shook the pot'  
\( \text{he-past shake pot} \)

b. \( \text{aè tù bà zuma bà} \) 'he shook the pot with his hand'  
\( \text{he-past take hand shake-ma pot} \)

c. \( \text{bà è zuga} \) 'the pot shook'  
\( \text{pot past shake-a} \)

So the fact that Verb + \( \text{ma} \) combinations consistently have one more \( \Theta \)-role to assign than the corresponding Verb is evidence that \( \text{ma} \) has the property given in (15a).

All the examples we have given with \( \text{ma} \) so far also involve \( \text{tù} \). So it is reasonable to ask how we know that it is \( \text{ma} \) and not \( \text{tù} \), or the combination of \( \text{tù} \) and \( \text{ma} \), that adds a \( \Theta \)-role? Two pieces of evidence tell us that it is \( \text{ma} \) alone which has this property.

First, there is an alternative to the \( \text{tù...ma} \) construction available. In certain circumstances, \( \text{-ma} \) can take a low-tone and \( \text{tù} \) is impossible (the exact circumstances are phonologically determined and need not concern us here). This possibility is illustrated in (19):

(19) \( \text{aè kpoèmà gìmà nòm} \) 'he cut the meat with the knife'  
\( \text{he-past cut-ma (low) knife animal} \)

Here, \( \text{kpoèmà} \) has three arguments, even though \( \text{tù} \) is not present. On the
other hand, in an example like (4), tú is present, ma is absent, there are three arguments and the sentence is bad. Given the Θ-criterion, this contrast is sufficient to show that ma, and not tú, adds the extra Θ-role.

A second piece of evidence comes from the form of ditransitives, without ma. There are two possibilities for ditransitive clauses. These are as follows:

(20) a. òè nè nuwé nwin 'he gave the thing to the child'
    he-past give thing-assoc child (= 16c)

b. òè tú nu nè nwin 'he gave the thing to the child'
    he-past take thing give child

In (20b), tú appears but there is no extra Θ-role associated with nè. If tú added a Θ-role to the structure it appeared in, (20b) would be impossible, given the Θ-criterion. The grammaticality of this example shows then that tú does not add a Θ-role.

tú is in complementary distribution with the associative marker in (20). The associative marker, exemplified on nuwé in (20a), is a low tone that appears obligatorily following a mid-tone vowel and optionally following a low tone. Where the stem is CV, V is copied; otherwise the low tone appears on the final segment. The associative marker appears in the following environments:

(21) a. N^n NP:
    nom nwin
    animal-assoc child
    kroò nom
    cut-assoc animal
    mìn to
    inside-assoc house

b. A^n NP:
    miimìl nom
    red-assoc animal
    tÈè nwin
    three-assoc children
    dÈÈ mùù
    all-assoc water
    ñdaà gè
    my-assoc knife

(21a,b) are basically the environments where of-insertion takes place in English (cf. 11). So we propose that the associative marker is a dummy Case marker in Gokana. Now the fact that the associative marker is in complementary distribution with tú in (20) suggests that tú is a dummy Case marker also. More specifically, our suggestion is that in the tú...ma construction ma
contributes a θ-role but no Case feature, and ṭû a Case feature but no θ-role.

In addition to ditransitives, ṭû is in complementary distribution with the associative marker in transitives. Alongside examples like (2d), we find (22):

\[(22) \quad \text{aè kpɔomà nuû nom }\quad \text{'he cut the meat with the thing'}\]

In (22), -mà is low tone and nuû appears with the associative marker. In the discussion of low-tone -mà earlier (cf. 19), we showed that ṭû did not contribute an extra θ-role. Instead, it is a dummy Case marker. In (19) and (22), where ṭû is absent, low tone -mà acts as a Case marker. We may identify the low tone on -mà with the low-tone associative marker. Thus the two instances of complementary distribution between ṭû and tonal morphemes just discussed collapse to the same phenomenon; both are ways of Case marking the extra argument added by -mà.

We have now seen evidence to support the statements in (15). How do these statements derive the properties of the ṭû...ma construction?

Ma attaches freely to Verb stems in the lexicon, and adds a θ-role to the Verb's set of θ-roles. Accordingly, given the θ-criterion, an "extra" NP must appear in the government domain of V + ma. However, since ma lacks a Case feature, this NP will not meet the Case Filter (or, equivalently, the Visibility Condition). Thus some form of dummy Case marker must appear whenever ma appears. We have seen that there are two dummy Case markers in Gokana: the associative marker and ṭû. One of these items always appears with -mà. These are the features of the ṭû...ma construction that we can derive from (15) given the theoretical considerations in Section 2.

Certain questions remain, in particular the question of why the ṭû VP precedes the other VP in SVCs. A full discussion of this and other questions appears in Roberts [forthcoming].

REFERENCES


THE PHONOLOGY OF BANDA-TANGBAGO

Douglas L. Sampson
University of Texas
and
Summer Institute of Linguistics

0. Introduction

The purpose of this paper is to outline the phonology of the Tangbago dialect of the Banda language. Greenberg [1963] classifies Banda in the Eastern branch of the Adamawa-Eastern language family. There are numerous dialects, some of which are mutually unintelligible.

Most of the Banda people live in the Central African Republic (CAR), but about 10,000 from eleven dialect groups live in Sudan. Members of these dialect groups are in CAR as well. Tangbago is a major dialect in both countries.

The Tangbago described in this paper was studied in the Sudanese town of Sopo in 1982 and 1983. This description may serve as a comparison to the work done on the Linda dialect by Cloarec-Heiss [1969, 1972, 1978], which is the most definitive work done on any of the dialects in recent times.

1. Consonants

1.1. Inventory. Banda-Tangbago has 33 consonants as charted in Table 1, p. 270. At one point in the study, a consonant count from the lexical file showed the following order of frequency, from most common to most rare: /r, k, l, t, ng, d, g, y, m, b, w, nd, s, n, gb, mb, p, ngl,  בכלל, j, c, z, v, ่, ่, nj, f, mv, ่, nz, h, ่/.

1.2. Distribution in the word. In a CVCV-type stem, there are no restrictions for consonants co-occurring within the word. For both consonants to be identical, however, is relatively rare, and when they are, the vowels are usually identical, e.g. /jaj/ 'to confuse', /§§§/ (plant type). The only exception to this is /kākā/ 'leaf'.

In stems of three syllables and more, if two contiguous syllables contain identical consonants, the vowels of those syllables are always identical.

2. Vowels

2.1. Inventory. Banda-Tangbago has ten vowels as charted in Table 2, p. 270. This is two less than Banda-Linda as described by Cloarec-Heiss, which has two more diphthongs /æe, ia/.

2.2. Distribution.

2.2.1. In the syllable. /i/ and /e/ can never occur following /s, z, nz/; /t/ and /a/ can never occur following /§, ่, c, j, nj/; only /e/ and /a/ can occur in a V syllable.


<table>
<thead>
<tr>
<th>Phonoemes</th>
<th>Labial</th>
<th>Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Labio-velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>stopped</td>
<td>p</td>
<td>t</td>
<td>c</td>
<td>k</td>
<td>kp</td>
<td>?</td>
</tr>
<tr>
<td>voiceless</td>
<td>b</td>
<td>d</td>
<td>j</td>
<td>g</td>
<td>gb</td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>mb</td>
<td>nd</td>
<td>nj</td>
<td>ng</td>
<td>ngb</td>
<td></td>
</tr>
<tr>
<td>prenasalized</td>
<td>mv</td>
<td>nz</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>f</td>
<td>s</td>
<td>ž</td>
<td>h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiceless</td>
<td>v</td>
<td>z</td>
<td>Ž</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>voiced</td>
<td>mv</td>
<td>nz</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>n</td>
<td>ŋ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>liquid</td>
<td></td>
<td></td>
<td>l</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>retroflexed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lateral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flapped</td>
<td>ř</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approximated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>γ</td>
<td>w</td>
</tr>
</tbody>
</table>

Table 1. Consonant Phonemes

<table>
<thead>
<tr>
<th>Phonoemes</th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>ũ</td>
<td>ū</td>
<td>ū</td>
</tr>
<tr>
<td>mid</td>
<td>e</td>
<td>æ</td>
<td>o</td>
</tr>
<tr>
<td>low simple</td>
<td>a</td>
<td>æ</td>
<td>o</td>
</tr>
<tr>
<td>glided</td>
<td>ea</td>
<td>œ</td>
<td>œ</td>
</tr>
</tbody>
</table>

Table 2. Vowel Phonemes

2.2.2. In the word, /a/ can never fill a V syllable (word-initial) if the word-final vowel is low.

In word stems, there is a tendency away from the occurrence of front and back vowels in contiguous syllables. At one point in the study, front and back vowels co-occurred in only 6.8% of the CV.CV words filed in the lexicon.

CV.CV words tend to contain identical vowels. In the same set of CV.CV words, 47% contained identical vowels. In 37% of these, the vowels were /a/.

Also in word stems, only /e/, /a/, and /o/ can occur in weak syllables.
(see 4.2) if the vowel of the strong syllable is /a/ or /ea/. If that vowel is /o/, however, only /e/ and /o/ can occur. My data don't reveal /oa/ in strong syllables.

The high and mid vowels are also neutralized with free variation before nasals and prenasalized stops:

\[
\begin{align*}
&/k\text{on}j\text{a}/ \quad [k\text{un}j\text{a}] \quad [k\text{on}j\text{a}] \quad \text{'mat'} \\
&m\text{e}n\text{o}/ \quad [m\text{en}o] \quad [m\text{en}o] \quad \text{turtle type}
\end{align*}
\]

3. **Tone**

3.1. **Inventory.** There are five tones in Banda-Tangbago: High (\textasciitilde{}), Mid (\textendash{}), Low (\textendash{}), Falling (\textendash{}), and Rising (\textasciitilde{}). An allotone of Low occurring utterance final glides from low pitch to extra low pitch.

3.2. **Distribution**

3.2.1. **In the syllable.** All five tones occur in CV syllables. V syllables, however, can carry only H, M, or L.

3.2.2. **In the word.** Nouns usually contain only level tones, but there are some exceptions. All nine combinations of level tones are found in two-syllable words and all 27 combinations in three-syllable words. Verbs contain all tones but R. Adjectives contain only level tones, and usually the tones are identical. Adverbs and functors can carry all five tones.

3.3. **Functional load.** Tone has lexical function in all word classes. In verbs and pronouns, it has grammatical function as well:

\[
\begin{align*}
&/s\text{e} w\text{o}/ \quad \text{'he saw' (perfective)} \\
&3s \text{ see} \\
&/s\text{e} w\text{u}/ \quad \text{'let him see' (imperative)} \\
&/s\text{e} w\text{u}/ \quad \text{'he will see' (imperfective)}
\end{align*}
\]

4. **Syllables**

4.1. **Configurations.** There are two syllable configurations in Banda-Tangbago, V and CV. CV syllables can occur in any syllable in the word, but V syllables can occur only word-initial. Every syllable carries a tone.

4.2. **Strong and weak.** For the sake of convenience, CV syllables composed of a liquid and a low vowel are termed strong. Preceding CV syllables composed of a non-liquid and a simple vowel are termed weak. Occurrence of vowels in weak syllables is restricted to /e, a, o, o/, the latter three of which are shortened and altered in quality:

\[
/\text{ak\text{\textbar}}r\text{\textbar}/ \quad [\text{\textbar}k\text{\textbar}r\text{\textbar}] \quad \text{'peanuts'}
\]
5. Words

The following word shapes are found in Tangbago:

- **V /á/** 'not'
- **CV /na/** 'to go'
- **VCV /ãbã/** 'food'
- **CVCV /gūsū/** 'grass'
- **VCVCV /ãgbõlõ/** 'child'
- **CVCVCV /bëmărã/** 'lion'
- **VCVCVCV /õmbõrõpõ/** 'star'
- **VCVCVCVCV /måz+rågbã/** 'rhinoceros'
- **VCVCVCVCV /õfõrõkõngõ/** 'bushbabies'
- **VCVCVCVCV /gõgõnõkõmõvõ/** 'marabou stork'
- **VCVVCVVCVCV /õwõgõbõgõrõ/** 'December'

Noun stems take all shapes except V. CV is very rare. Adjective stems take VCV, CVCV, VCVCV, CVCVCV, and VCVCVCV. VCVCVCV does not occur in my data. Verb stems and adverbs take CV, CVCV, and VCVCVCV. Functors take V, CV, VC, and CVCV.

6. Word Boundary

6.1. Prose. Although there is no vowel elision at word boundaries in prose, vowel assimilation does occur. Obligatory assimilation is limited, and optional assimilation occurs depending on the rapidity of speech and the individual speaker. In any case, the tones remain the same.

The obligatory rules are:

1. \(/æ/ \rightarrow [a] \parallel /æ/
2. \(/ç/ \rightarrow [c]

The optional rules are:

3. \(/æ/ \rightarrow [ç] \parallel /æ/ 
4. \(/ø/ \rightarrow [Ø] \parallel /ø/ 
5. \(/V/ \rightarrow [æ] \parallel /æ/ 
   [-low]

6.2. **Song.** Although song has not been studied in detail, it has been noted in passing that vowel elision may occur in songs. In this case, it seems that the tone is elided with the vowel. In all examples so far, however, the elided element has always been a vestigial noun prefix /a-/ or /a-/`, so it is difficult to tell whether the phenomenon is phonological or grammatical. In the following example, the bracketed portion is what gets elided:

```
  tètèrà kpé [à]wà [à]bè è só lò [à]wù
  Etere fled fear riki-grass and was in earth
  'Etere avoided the riki grass and was in the earth'
```

7. **Pause Group**

A pause group beginning with a vowel is often initiated with an etic glotal stop, which is lost medially in the pause group, in contrast with an emic one, which is retained. Note the morphemes /?à/ 'we, our' and /àndà/ 'house' in the following examples:

```
/àndà nò ?à kò ̀lè / [?àndà nò ?à kò ̀lè] 'this is our house'
/?à wù àndà nò yè/ [/?à wù àndà nò yè] 'we've seen his house'
```

A pause group final /a/ is elided regularly along with its tone:

```
/mà só gù tè/ [mà só gù t] 'I'm coming back'
/àndà nò mè/ [?àndà nò m] 'my house'
```

Other pause group final vowels are elided in the same way following a voiceless fricative preceded by an identical vowel:

```
/mòròfà/ [mòròf] 'ten'
/bò$array$]/ [bò$array$] 'two'
```

Otherwise, a pause group may be terminated with an optional etic voiceless glottal fricative [h]:

```
/àmà bàndà/ [?àmà bàndàh] ~ [?àmà bàndà] 'the Banda language'
```

8. **Intonation**

There is not much in the way of contrastive intonation in Tangbago, but intonation is used to mark paragraph boundaries. That is, the first few phrases of a paragraph will exhibit a higher intonation level than the body of the paragraph. Also, the last phrase of the paragraph will drift down to a lower level. Thus, there are three levels involved.

The high level is also used to signal queries for information, astonishment, reprimands, and such like, and is maintained throughout the domain of the communicator's attitude or intent so being signaled. This level is usually one pitch above the mid intonation level, so that a (syllable level) mid tone in
normal statement intonation (paragraph medial) is on the same pitch as a (syllable level) low tone in question intonation (paragraph medial).

REFERENCES


1. Introduction

From the beginning, Generative Grammar has required from Linguistic Theory to be "simple" and "natural" in accounting for the actual relation between data and grammar, in order to justify languages' learnability. Recent improvements in phonology have increased the explanatory power of Theory, accordingly reducing idiosyncrasies, by adopting parametric analyses implying language-particular well-formedness conditions constraining choices among universal principles. This paper will show how an integrated view of three of these recent phonological models allows one to capture a varied set of phenomena in Kinyarwanda nouns.

The first sub-theory that will be assumed here, roughly schematized under (1), is that of "Three Dimensional Phonology", forwarded mainly by McCarthy [1979] and by Halle and Vergnaud [1980].

```
(1)                              + tonal/accentual tier
                                          + syllabic (prosodic) tier
                                           + skeleton
                                          + melodic (segmental) tier
                                           + harmonic (or else...) tier

H=L-H-L-L  
\--------\--------/  
|   \   |   /  
\-------\-------/  
V-C-C-V-C-V
\-------/  
z-z-z-z-z-z
```

According to this approach, elements of various phonological tiers are linked to the slots of a skeleton under conditions stipulating that segments and slots are linked one-to-one from left to right, that each slot is linked to at least one segment, and, conversely, that linking lines do not cross and that unlinked elements ultimately delete.

Second, we will adopt a "Syllable Geometry" that has been developed by Kaye and Lowenstamm [1979, 1980] and others. This theory, illustrated in (2), claims that the minimal representation of a syllable is that of a binary branching structure, composed of at least two main constituents, onset and rime, the latter being allowed itself to separate into two subconstituents, nucleus and coda. Each of these constituents may branch in its turn, the nucleus being the only universally obligatory constituent.

```
(2)   
    \  
   (O)    R
    /  
(x)  (x)  
   /  
U   V   W  
   /  
N   X   Y   Z
  /  
X   (x) (x)  
/   
Y   (x)
```

In such a model, the representation of a complex segment such as a prenasalized or palatalized consonant will be that of several segments linked to a single
slot of the skeleton, as UV in (2), and that of a long vowel will correspond to
a single segment linked to two slots of the skeleton, Y in (2); a given segment
will receive its traditional [+syllabic] interpretation from the nature of the
syllabic constituent that dominates the skeletal slot it is attached to (thus
accounting for the behavior of ambivalent segments such as high vowels and nasals, which are marked as being both [+vocalic] and [+consonantal]).

Furthermore, three universal principles will be assumed with Kaye and Low-
ennstamm. First, a constraint on syllabic inventory according to which the sym-
labic index of a given language is an ordered pair \((m,n)\), where \(m\) is the value
of the most complex onset and \(n\) that of the most complex rime, the former hav-
ing a complexity that is lower than or equal to that of the latter. This means
that a language having branching onsets will also have branching rimes (but not conversely). Second, a resolution principle, designed to eliminate from the
syllabic structure the null elements, which will first try to reanalyze an err-
atic structure by resyllabifying it without violating syllabic constraints of
the language; if impossible, it will proceed to the insertion of an epenthetic
segment. Third, a Nuclear Fusion Principle which will cause two adjacent nu-
clei to combine into a single one, thus resulting in a long vowel if the skele-
otal slots dominated by the nuclei are linked to a single segment.

The third model we will refer to is Hayes' [1980] "Metrical Stress-Assign-
ment Theory", according to which stress is assigned by constructing feet and
word-trees based on the adjustment of parameters. Thus, for example, a given
language will have the choice between binary or illimited and right or left dom-
inant feet, and between right-to-left or left-to-right and iterative or non-it-
ervative feet construction. Moreover, it will be assumed that a degenerate (non-
branching) foot is constructed at the edge of a domain when there is only one
foot terminal element left in a binary feet construction and that such a foot
can be deleted or not, depending on a further parametric choice.

2. Kinyarwanda: the Data

Let us now observe the following Kinyarwanda nouns:

(3) a. /u-mu-gabo/ \rightarrow [\text{umugabô}] \quad 'man'
b. /a-bo-gabo/ \rightarrow [\text{abegabô}] \quad 'men'
c. /i-ki-gabo/ \rightarrow [\text{ikigabô}] \quad 'big man (pej.)'
d. /u-tu-saaaz:/ \rightarrow [\text{udâsàaz}] \quad 'small flies'

Most of the nouns are composed of a radical preceded by a classificatory prefix
of the CV type and by a pre-prefix composed of the same vowel as that found in
the prefix. However, there are two classes, illustrated in (4), that are excep-
tional with respect to that pattern.

(4) a. /i-ri-óya/ \rightarrow [\text{irgóyà}] \quad 'feather'
b. /i-(ri)-no/ \rightarrow [\text{inò}] \quad 'toe'
In one of these classes, (4a-b), the ri prefix drops in all but two nouns. In the other, (4c), the prefix consists of an n, which drops too in many nouns, as in (4d).

Long vowels are found, which may have an independent morphological status since they can appear in radicals, as in (3d), or be derived in two ways, as a result of glide formation or of prenasalization. The first case is illustrated in (4a), where we see a high vowel change to a glide (and then get palatalized), the following vowel thus lengthening. The second case is exemplified in (5), where we see a short vowel becoming long when preceding a nasal followed by another consonant.

Moreover, it can be seen in (6) that a place particle can "replace" the pre-prefix in nouns.

(5) /i-ki-ntu/ → [lk'\text{n}h\text{u}] 'thing'

When this occurs with a noun having an n prefix and a radical beginning with a consonant as in (6c), the vowel of the place particle lengthens, but when it occurs with a noun whose pre-prefix drops as in (6d-e), a vowel appears before the radical, which lengthens when the place particle vowel changes to a glide.

Finally, Kinyarwanda has two tones, high and low. If a noun has a high tone on its last or penultimate syllable, this tone will be "lowered" (or will appear as a "falling tone") if the word occurs before a pause or in isolation, as in (7a). Similarly, the first of a sequence of high tones will be brought to a mid-tone (or "rising tone"), as in (7b).

(7a) a. [lg\text{ft}f] → [lg\text{ft}f] 'tree'

This leads to viewing these processes as resulting from the application of late phonetic rules. Rialland and Furere [1983] have called attention to the fact that, notwithstanding these phonetic rules, nouns appeared to obey a parity constraint, illustrated in (8).

(8) a. /i-(ri)-b\text{ba}b\text{a}/ → [l\text{b}b\text{ba}] 'wing'

b. /i-(n)-f\text{f}/ → [f\text{f}] 'fish'
c. /ka-bátá/ → [akábába] 'small wing'
d. /ka-fɛ/ → [ágáfɛ] 'small fish'
e. /mu-áana/ → [umwáaná] 'child'
f. /ki-sórógbá/ → [lgísrógbá] 'worm (sp.)'
g. /ka-tééká/ → [ágátéeáká] 'little law'
h. /i-(n)-tééká/ → [ítéeáká] 'law'
i. /mu-n-ká/ → [múùñhá] 'in the cow'
j. /mu-(n)-fɛ/ → [múùñfɛ] 'in the fish'

More precisely, all nouns exhibit an even number of high tones. When the number of these in the radical is even as in (8a,c), syllables preceding it will bear low tones. When it is odd as in (8b,d-j), the first syllable preceding it will bear a high tone, as if to maintain the parity. Notice that under such an analysis, long vowels must be considered as being part of two distinct syllables.

3. The Metrico-Syllabic Analysis

All the facts already mentioned will ensue straightforwardly from the universal principles exposed in the introduction without any "rule" having to be stipulated once a few (parametric) well-formedness conditions or choices will have been made clear.

First, we will claim that Kinyarwanda's most complex syllable is a CVV one. This means that we will not find any branching onset nor anyoda in that language and that its only permissible sequence of vowels will be that of two identical vowels. If two vowels of different quality meet over a word boundary, the first of these will drop. If they meet over a morpheme boundary, the first will either drop if it is low or change to a glide if it is high. The last case is what is found in words such as those of (4a) and (8e). In the theoretical framework that is assumed here, that process would be the following for a word like [iróoya].

That is, the high vowel i, preceding another vowel, detaches itself from its point of the skeleton and reattaches under that of the preceding consonant, dominated by an onset, which will make it "consonantize", i.e. change to a glide, and even palatalize under the influence of the r with which it shares a same point. Subsequently, the slot of the skeleton that has been vacated by the i will "attract" the only segment that can attach to it without crossing lines, the o. The Nuclear Fusion Principle will then integrate the two adjacent nuclei in a single one, thus melting the two syllables they are part of (both ill-
formed) in a single one.

Similarly, if we posit that the nasal consonant in the words of (5), (6c), and (8i) is not underlingly part of any syllable (maybe a remnant of a diachronic change), the long vowel preceding a NC cluster obtains along the same lines. For example, the derivation for \([\text{m}u\text{u}n\text{z}u]\) would be:

\[10\]

\[\begin{array}{c}
\sigma \sigma \\
\text{m} \text{u} \text{n} \text{z} \text{u} \\
\end{array}\]

The point to which the \(n\) is linked cannot be interpreted as being part of a nucleus (since Kinyarwanda does not allow syllabic nasals) nor as part of an onset (which would result in an ill-formed syllable missing the obligatory nucleus). It cannot either be attached to the coda of the preceding syllable (Kinyarwanda has none) nor to the onset of the following one (the language does not allow branching onsets, since these would be more complex than the most complex rime). The \(n\) thus detaches from its point, reattaches to that of the following onset (which will yield a prenasalized consonant), and the preceding \(u\) spreads to fill the orphaned slot, which will then be unified in a single nucleus.

Turning now to the status of the pre-prefix, we saw that there were nouns such as that of (4c), where the prefix was an \(n\). If we assume that the pre-prefix is usually constructed by some kind of a copy process that duplicates the prefix vowel in a preceding syllable, such words will not have any, since there is not any vowel that can be copied in the prefix. Moreover, when the \(n\) will attach to the following onset, an epenthetic \(i\) will be inserted to replace it, since there is not any vowel that can spread to do so. This accounts for the fact that we do not find any word beginning with a long vowel before a NC cluster. Similarly, in these words where the prefix drops, those of (4b,d) and (8a,b,h), no pre-prefix will be constructed either, and an epenthetic \(i\) will fill the nuclear slot deserted after the drop. If a place particle is then inserted before that \(i\), i.e. under the syllable where we usually find the pre-prefix as in (6d-e) and (8j), this will trigger the same process as that illustrated in (9), glidization of the place particle vowel and lengthening of the \(i\).

Finally, tones will be assigned to nouns before pre-prefix construction in the following way. First, beginning with the lexically accented point of the skeleton (the one that actually bears phonetically the last high tone in the radical), the Feet Construction (FC) will erect from right to left maximally-binary right-dominant feet on the points of the skeleton dominated by nuclear nodes. If a single syllable remains at the left edge of the domain, a degenerate foot will be constructed and immediately deleted. The Word-Tree Construction (WTC) will then build a right-dominant structure over the feet (if there are two or more) and the High Tone Assignment (HTA) will attach a high tone to every skeletal point that is a terminal element of a foot. Ultimately, a few tonal adjustments will occur at the phonetic level, such as the assignment of a low (unmarked) tone to all elements that are not already bearing a high tone,
the facultative lowering of the tone whose tone-bearing unit is dominated by weak nodes only in themetrical structure (provided that it is not preceded itself by another high tone), etc. Here are a few examples of complete derivations.

(11) **Lexicon**

\[
\begin{array}{cccc}
\hline
\text{gakana} & \text{nkana} & \text{nkana} & \text{nfif} \\
\hline
\text{w} & \text{s} & \text{w} & \text{s} \\
\text{w} & \text{s} & \text{w} & \text{s} \\
\text{x} & \text{x} & \text{x} & \text{x} \\
\text{m} \text{u} & \text{n} \text{k} & \text{a} & \text{m} \text{u} \ \text{i} \ \text{f} \\
\hline
\end{array}
\]

**FC**

\[
\begin{array}{c}
\text{x} \text{x} \text{x} \\
\text{w} \text{s} \\
\text{x} \text{x} \\
\text{w} \text{s} \\
\text{x} \text{x} \\
\text{w} \text{s}
\end{array}
\]

**HTA**

\[
\begin{array}{c}
\text{x} \text{x} \text{x} \\
\text{x} \text{x} \\
\text{x} \text{x} \\
\text{x} \text{x} \\
\text{x} \text{x} \\
\text{x} \text{x} \\
\end{array}
\]

**Drop and Insertion**

\[
\begin{array}{c}
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\end{array}
\]

**Place particle**

\[
\begin{array}{c}
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\end{array}
\]

**Pre-prefix**

\[
\begin{array}{c}
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\end{array}
\]

**Reanalysis**

\[
\begin{array}{c}
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\end{array}
\]

**NFP**

\[
\begin{array}{c}
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\text{N.A.} \\
\end{array}
\]

\[
\begin{array}{c}
[\text{agakâ}] \\
[\text{f} \text{nâ}] \\
[\text{mu'nâ}] \\
[\text{mu'rif}]
\end{array}
\]

'small cow' 'cow' 'in the cow' 'in the fish'

4. **Conclusion**

The analysis presented here thus succeeds in giving an integrated account of a whole series of morphological and phonological phenomena in a rather restricted area of a particular language, namely pre-prefix formation, prefix-vowel alternations, tone assignment, contour tones creation, ephenthesis, prenasalization, glide formation, and compensatory vowel lengthening in nouns of Kinyarwanda. Since the theories we have been referring to have been elaborated independently, in order to deal with completely different phenomena in completely
different languages, it can be said that the present analysis allows a deeper insight into their relatedness and brings further evidence for their relevance and their membership in Universal Grammar.

REFERENCES


THE DEVELOPMENT OF CASE AND FOCUS IN UMBUNDU

Thilo C. Schadeberg
State University at Leiden, The Netherlands

1. UMBundu

UMBundu is the single most important (Bantu) language of Angola. It is spoken by about two million people. I collected the data on which this study is based in Angola during 1981-82.

2. Noun Structure

Nouns have the structure: Augment - Nominal Prefix - Stem.

Stems occur in various lengths and tone types; the example used here has a disyllabic stem of the tone type Lo-Hi.

Nominal prefixes may have one of the shapes CV, N, or V, e.g. -ci- (class 7), -n- (class 9), -g- (class 5).

The augment is -g- in all classes. It is regularly absent (deleted) before a nominal prefix consisting of just one vowel. Thus we find, for example, -g-yò 'tooth', -g-vé-yò 'teeth' (classes 5/6). Syntactically, the augment is stable, i.e. there are no syntactic environments in which the augment is deleted. Some nouns, however, never have an augment; this is typically the case with proper names, often derived from common nouns, and kinship terms.

3. Nominal Tones

Each noun appears in two tonal shapes. Shape A (Port. alto) is characterized by an initial sequence of high tones, shape B (baixo) is characterized by an initial sequence of low tones.

'shippo\text{pota}t\text{mos}' 'N\text{geve}' (n.pr.)
shape A: g-n-\text{ge\text{"e}}vé n-\text{ge\text{"e}}vé
shape B: g-n-\text{ge\text{"e}vé} n-\text{ge\text{"e}vé}

4. Distribution of Tonal Shapes

The distribution on these two shapes A and B is syntactically defined. Interestingly, the distribution is not the same for nouns with and without an augment. Three classes of syntactic environments have to be distinguished, of which at least two have much in common with what is traditionally regarded as case.

<table>
<thead>
<tr>
<th>label</th>
<th>synt. env.</th>
<th>with augment</th>
<th>w/o augment</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRED</td>
<td>predicative, citation form</td>
<td>A: \text{ng\text{&quot;e}vé}</td>
<td>A: \text{ng\text{&quot;e}vé}</td>
</tr>
<tr>
<td>OC</td>
<td>first complement of affirmative main verb</td>
<td>A: \text{ng\text{&quot;e}vé}</td>
<td>B: \text{ng\text{&quot;e}vé}</td>
</tr>
<tr>
<td>CC</td>
<td>elsewhere</td>
<td>B: \text{ng\text{&quot;e}vé}</td>
<td>B: \text{ng\text{&quot;e}vé}</td>
</tr>
</tbody>
</table>
The function of Direct Object provides the name for the Object Case (OC). The most salient function of the Common Case is Subject; however, this case also marks the second complement, the first complement of a negative verb or of a subordinated clause, as well as nouns in several more specific syntactic environments.

5. Origin of the Distribution between OC and CC

Phonologically, A-forms can be derived from B-forms by positing a word-initial Hi tone. The further derivation by Hi-Spreading is regular.

I suggest that the augment originally only occurred in OC environments and that it had a high tone (see Meeussen [1967], De Blois [1970]). Later, the alternation (presence versus absence of augment) was eliminated by spreading the augment to all other environments. However, the original tonal distinction was preserved. This hypothesis is supported by the fact that those nouns that lack the augment in all syntactic environments also lack the tonal distinction between OC and CC; such nouns attest formerly augment-less forms in both environments. This development appears to provide a nice contrast with better documented cases where a tonal distinction originates from the loss of all phonic material except tone.

6. Source of A-forms for PRED

It is this better-known type of development which I propose to regard as the source of the use of A-forms for PRED. I assume that a vocalic predicative index such as is found in Kongo and Bemba existed in pre-UMBundu. When the augment became stabilized throughout the paradigm it occurred predominantly in pre-vocalic position where only its high tone has survived.

It should be clear that the augment and its high tone cannot be the source of the PRED A-forms. Firstly, an initial high tone is present even in words which have no augment and never had one, even in pre-UMBundu. Secondly, there is good comparative evidence that the predicative noun should be reconstructed as having no augment. Compare

Kinga: sinu 'it is food' (isinu c1.8 'food')
Ganda: bitabo 'they are books' (ebitabo c1.8 'books')

7. Focus Marking

UMBundu has not only developed a rudimentary case system, it also suggests a possible further development of focus marking. Consider relative clauses. The most conspicuous differences between relative and non-relative verb forms are again tonal.

ngèvó wàfá kòvítá 'Ngeve died in the war'
B:CC ABS A:OC

ngèvé wàfá kòvítá 'it is Ngeve who died in the war'
A:PRED REL B:CC

An alternative translation of the second of these two sentences would be
'NGEVE died in the war', with contrastive emphasis on "Ngeve". I do not think that this would be the proper analysis for UMbundu, but I think it shows how focus markings in nouns and in verbs could develop.

8. Typology

Greenberg [1978] has shown how a weak demonstrative or "definite article" may develop into a gender marker. UMbundu provides a typologically interesting alternative in which the same element, i.e. the augment, develops into a case marker. This is consistent with the important role of definiteness in conditioning alternative case markings of objects as described by Moravcsik [1978].

[Note: A full version of this paper is to appear in Africana Linguistica 10, Tervuren: Musée royal de l'Afrique Centrale.]

REFERENCES


DOWNSTEP IN THE KIPARE VERB COMPLEX

Deborah Schlindwein
University of Southern California

This paper deals with some selected aspects of the tone system of one dialect of Kipare, an Eastern Bantu language spoken in Tanzania.

Kipare downstep is the phonetic realization of a floating low tone sandwiched between two high tones. Downstep, in this language, often results from a rule of High Tone Anticipation (HTA), which operates as described in (1) and exemplified in (2):

\[(1) \text{ High Tone Anticipation (HTA)}
\begin{align*}
\text{HTA} & : \quad \begin{array}{c}
\uparrow \\
H
\rightarrow
\end{array} \quad \uparrow \\
L
\rightarrow
\begin{array}{c}
H
\end{array}
\end{align*}
\]

\[(2) \quad \text{magàré 'vehicles' + matàtu '3' } \rightarrow \quad \text{magàré 'matàtu '3 vehicles'}
\begin{align*}
\text{HTA} & : \\
\text{HTA} & : \\
H & \rightarrow L \\
& \rightarrow H \\
& \rightarrow H \\
& \rightarrow H
\end{align*}
\]

We will examine the operation of HTA in the Kipare verb complex and draw some conclusions about the nature and interaction of lexical and post-lexical rules based on what we find in the verb tonology.

The Kipare verb radical generally consists of either one or two syllables, the first of which is lexically specified for either a high (H) or a low (L) tone. Two examples are given in (3):

\[(3) \quad \text{kom 'kill' + som 'stab'}
\begin{align*}
\text{HTA} & : \\
H & \rightarrow L
\end{align*}
\]

The composite verb structure, which is similar or identical to that of the verb complexes in most Eastern Bantu languages, is shown in (4):

\[(4) \quad \text{[SM (neg) (tns/asp) - (OM) RAD (ext's) FV] verb}
\begin{align*}
\text{pre-base} & \quad \text{base}
\end{align*}
\]

The optional object marker (OM), as well as the radical, also carries an inherent lexical tone, H or L, with two examples given in (5):

\[(5) \quad \text{ni 'me' + tu 'us'}
\begin{align*}
\text{OM} & : \\
L & \rightarrow H
\end{align*}
\]
High-toned *tu*, the first person plural OM, will appear in subsequent examples. Example verbs are given in (6):

(6) a. *vénetúkóma* [ve + ne + tu + kom + a] 'they'll kill us'

\[
\begin{array}{cccc}
  & H & L & H & L \\
3p & \text{fut} & 1p & \text{kill} & FV \\
\end{array}
\]

b. *vénetúsoma* [ve + ne + tu + som + a] 'they'll stab us'

\[
\begin{array}{cccc}
  & H & L & H & L \\
3p & \text{fut} & 1p & \text{stab} & FV \\
\end{array}
\]

c. *vesikatúsómé* [ve + si + ka + tu + som + e] 'they won't stab us'

\[
\begin{array}{cccc}
  & L & H & L & H & L \\
3p & \text{neg} & \text{fut} & 1p & \text{stab} & FV \\
\end{array}
\]

(HTA effect)

We can see the effects of HTA in example (6c), where the low-toned radical *som*, because it occurs between a high-toned OM and a high-toned FV, is realized with a downstepped high tone. Its lexical L has been floated off of it. Surprisingly, however, we observe in the examples in (6) that there seems to be perhaps some strong "boundary" between the verb pre-base and the verb base across which HTA does not apply. In each of the examples (6a,b,c) we would expect the *ne* and *ka* of the pre-base to be realized with a downstepped high tone, rather than with a low tone.

I introduce this idea of a "boundary" existing between the pre-base and the base as a descriptive device, not as the answer to why we don't find HTA effects. I will try to derive the non-application of HTA shortly. In every tns/polarity combination that I have looked at in this language, it is precisely and invariably at the juncture between the pre-base and the base that the application of HTA is blocked.

The data presented in (7) lend further support to the idea of a "boundary" existing between the pre-base and the base:

(7) (repeated from (6) relative clause forms

<table>
<thead>
<tr>
<th>a. <em>vénetúkóma</em></th>
<th><em>vénetúkóma</em></th>
</tr>
</thead>
</table>
| \[
  \begin{array}{cccc}
    & H & L & H & L \\
  \end{array}
\] | \[
  \begin{array}{cccc}
    & L & H & L & H \\
  \end{array}
\] |

<table>
<thead>
<tr>
<th>b. <em>vénetúsoma</em></th>
<th><em>vénetúsoma</em></th>
</tr>
</thead>
</table>
| \[
  \begin{array}{cccc}
    & H & L & H & L \\
  \end{array}
\] | \[
  \begin{array}{cccc}
    & L & H & L & L \\
  \end{array}
\] |

<table>
<thead>
<tr>
<th>c. <em>vesikatúsómé</em></th>
<th><em>vesikatúsómé</em></th>
</tr>
</thead>
</table>
| \[
  \begin{array}{cccc}
    & L & H & L & H & L \\
  \end{array}
\] | \[
  \begin{array}{cccc}
    & L & H & L & L & H \\
  \end{array}
\] |

In the second column of (7) are shown the verbs originally presented in (6)
in their corresponding relative clause forms. The relative form marker is the L that appears as the leftmost tone of the verb complex. In (7a,b) it pushes the tones of the pre-base, but not those of the base, to the right. In (7c), the L which marks the relative has been deleted, or in another way of looking at it, it has been absorbed by the first L of the pre-base. In other words, the melody of the pre-base seems to obey the Obligatory Contour Principle.

This suggests the possibility that the melody of the pre-base is a floating melody, which I think is correct. That is, it is an unattached melody which is specific to each particular verb tense. We have seen HL for the future tense pre-base and LHL for the future negative. LH, which is absent from the presented data, is the melody of the past tense negative pre-base.

Taking as given the assumption that the melody of the verb pre-base is a floating melody, let us see if we can derive the effects of the "boundary" which blocks HTA between the verb pre-base and the verb base. We shall adopt a model of word building essentially like that of Kiparsky [1982] and exploit the resultant cyclicity and Elsewhere effects.

(8) Word building

(morphology) ==> \([tu + kom]_\text{base} \ 'us' + 'kill'\)

\(\begin{array}{c}
H \\
H \\
\end{array}\)

(phonology) ==> (none)

(morphology) ==> \([vene + tukom = a]_\text{verb} \ 'they will' + 'kill us'\)

\(\begin{array}{c}
H \\
L \\
H \\
H \\
L \\
\end{array}\)

(HTA) ==> N/A

(tone association) ==> \([vene + tukom + a]_\text{verb} \ 'they will' + 'kill us'\)

\(\begin{array}{c}
H \\
L \\
H \\
H \\
L \\
\end{array}\)

(identity) ==> venetukoma = vénetúkóma

\(\begin{array}{c}
H \\
L \\
H \\
H \\
L \\
\end{array}\)

\(\begin{array}{c}
H \\
L \\
H \\
H \\
L \\
\end{array}\)

In (8) the rule of HTA is inapplicable because HTA floats only an attached L. It would be meaningless to imagine it floating a L that is already floating. Tone association ensures that untoned vowels are matched up with unassociated tones from left to right.
Studies in African Linguistics, Supplement 9

(9) **Word building**

(morphology) \[\text{base} \quad [\text{tu} + \text{som}] \quad \text{'us'} + \text{'stab'}\]  
(phonoogy) \[\text{(none)}\]  
(morphology) \[\text{verb} \quad [\text{vesika} + \text{tusom} + \text{e}] \quad \text{'they won't'} + \text{'stab us'}\]  
(HTA) \[\text{verb} \quad [\text{vesika} + \text{tusom} + \text{e}] \]  
(tone association) \[\text{verb} \quad [\text{vesika} + \text{tusom} + \text{e}] \]  
(tonem identity) \[\text{vesikatusome} = \text{vesikatšómé}...\]  

Each of these verbs, as it emerges from the word building process, is, in effect, its own identity rule.

The identity rule provides a very specific SD (and SC) for the output of each cycle. It is the identity rule which acts to exempt the output representation of each cycle from undergoing subsequent internal changes during a later cycle within the same component. This is the Elsewhere account of the Strict Cycle phenomenon.

Let us now build phrases out of the verbs built in (8) and (9). The word vacé 'women' will be used as the subject of the phrases.

(10) **Phrase building**

a. vacé + vénétů́kóma \[\rightarrow\] vacé vénétů́kóma 'the women will kill us'

b. vacé + vesíkatšómé \[\rightarrow\] vacé vesíkatšómé... 'the women won't stab us'

Example (10a) is completely straightforward. In example (10b) we encounter the rule of HTA once again. This time, however, the rule is operating between words (as it did in example (2) earlier).

In the sense that HTA operates phrasally, between words, it must be considered a post-lexical rule. However, it does not operate exceptionlessly across the board. As we observe in (10), ka and ne still resist the effects of HTA. They resist the effects of HTA precisely due to the effects of the Elsewhere Condition. This means that Kipare word building and phrase building must
belong to the same component, which is a requirement of the Elsewhere Condition. In the model of phonology and morphology presented in Kiparsky [1982] it was suggested that lexical rules could not enter into an Elsewhere relationship with post-lexical rules. Furthermore, the post-lexical rules were assumed to be non-cyclic, hence their exceptionless application, because there could be no Elsewhere effects where there was no cyclicity.

My analysis of Kipare, however, suggests that cyclicity may continue beyond word building. It suggests that in Kipare, at least, it would be incorrect to maintain a strict separation between a lexical phonological component and a post-lexical phonological component.

[Mary Marunda graciously supplied the data used in this study.]

REFERENCE

This paper focuses on changes in the use of the locative copula in Swahili spoken as a second language in Nairobi, with additional examples from urban Swahili in Dar es Salaam. In both cases, the Swahili variety is spoken by some children as a first language, as well as by adults as a second language. (No claims are made that the usages described characterize all second language Swahili varieties in these cities. These usages, however, were widespread in over 100 hours of recordings of natural, informal conversations by speakers from diverse ethnic and socio-economic backgrounds in Nairobi; similar usages were also observed in Dar es Salaam or self-reported there.)

A major purpose of this paper is to argue that new uses of the Swahili locative copula to refer to temporary or acquired state may reflect a universal tendency to derive temporal indications from spatial ones and to argue more specifically there is a tendency to use similar derived forms to indicate temporary/acquired state and imperfect aspect. Also the paper argues that the particular socio-psychological conditions surrounding Swahili acquisition in the cases studied are such that language universals are likely to surface. In this way, second language acquisition can resemble creole development [Bickerton 1981].

While these second language Swahili varieties (hereafter SL Swahili), show a number of syntactic differences from native or Standard Swahili, the most striking innovation concerns the locative copula. It is being chosen over simpler and equally accessible strategies as a device to indicate temporary/acquired state. The locative copula consists of a bound stem and a prefix agreeing with its subject in person and number or in noun class. It occurs in the present tense only and, in native or St. Swahili, only before nouns inflected with the locative suffix -ni or before place names.

(1) Watu wangi wa-ko njia-ni. 'Many people are on the road.'

The role of the locative copula is being expanded by SL Swahili speakers so that expressions of state, and especially a temporary or acquired state, increasingly, take a locative copula. See (2-5).

(2) a. U-ko mgonjwa? 'Are you sick?'

b. 1-ko fresh sana. 'It (cl.9) is very fresh.'

(3) Yu-ko worried. 'He is worried'

(4) M-ko na wasiwasi ryingi. 'You (pl.) have much uneasiness.'

(5) Tu-ko confused. 'We are confused.'
This new use of the locative copula (hereafter KO form) occurs whether the form expressing state in native or St. Swahili is invariant ni (before predicate nouns and adjectives), as it would be in (2); associative -na (inflected with a subject prefix), as it would be in (3) and (4); or a stative verb stem, as it would be in (5).

A reaction to this semantic re-analysis and extension of the KO form is to explain it as a simplification. But the simplest possible innovation (using Swahili material) would be to generalize the invariant ni copula. A second reaction might be to explain it as interference from English, since the KO form occurs with many English loans. Yet, for example, while it occurs with many English past participles, other English loan verbs are used with regular Swahili bound prefixes to express other tenses/aspects. Another possibility would be sub-stratum effect from the speakers' first languages. This is not ruled out, but no direct and single source of borrowing is evident. A final explanation is to hypothesize that the preference for a locative form to express a temporary/acquired state may represent a language universal.

To support this universals hypothesis, data are presented making an inferential connection between the KO usage and forms in both some creoles cited by Bickerton [1981:27, 1975:55] and one not cited by him, West African Pidgin English [Eze 1980]. Connections are also seen in data from some non-creoles, Spanish and Amharic, for example. The connection is made through imperfect aspect: I argue that a natural semantic connection exists between non-punctual aspect, a component of imperfectivity, and temporary/acquired state. It is noted that Hawaiian Creole stei and other locative-based forms in other creoles cited by Bickerton are auxiliaries to express non-punctual meanings. Further, in Nigerian Pidgin English, a single form (dey) is used for such meanings and also to express temporary/acquired state. All of these forms have locative sources, along with the KO form.

The argument that universals should surface most easily in linguistic varieties such as creoles seems a good one because the innate tendencies of speakers are not suppressed by an omni-present target. (The fact Bickerton defines creoles very narrowly is irrelevant to my argument since my interest is in the socio-psychological conditions surrounding acquisition, not in whether the variety is a creole.) The link I argue for between Swahili development and creole development is this: it is predicted that when second language acquisition by adult speakers and creole development share social conditions, similar structures result.

The only major social difference between SL Swahili development and that of the creoles which Bickerton cites involves access to models of the target language, it is shown. Although children producing creoles have only partial access to the model, in SL Swahili development there is ample access. Yet, the end result can show similarities because when a model is present, but not highly valued by the language learners, the result is that the prescriptive power of the model is severely mitigated. Such is the case with either native or St. Swahili as a model for language learners in East Africa, especially in Nairobi and other up-country areas. While speakers want to acquire Swahili, they do not necessarily wish to be identified with native speakers of Swahili. In such situations, the conditions of acquisition can resemble those in a creole situation:
innovative structures have more of a chance to survive when there is no specific target. If these structures reflect a universal grammar, they should show similarities to structures found in other varieties emerging without a clear target, but also found in languages in general. This is the argument regarding the locative copula in SL Swahili and its use to express temporary/acquired state.

REFERENCES


This paper explores some aspects of loanword phonology as illustrated by data from OshiKwanyama, a Bantu language spoken in northern Namibia and southern Angola. The principal European source languages for borrowing in OshiKwanyama are English, Afrikaans, and German.

The degree to which a loanword is modified can vary considerably: some loanwords are hardly changed at all and may seem quite foreign-sounding to native speakers of the borrowing language. Other words may be so strongly modified that, to the native speaker, their shape is indistinguishable from that of a native word. Often, the degree to which a loanword is modified seems to depend as much on social, cultural, and historical factors as on purely linguistic ones. My focus in this paper is on loanwords which have been modified, and on the source of these modifications. Those loanwords which have not undergone any of the modification processes which I will discuss are not counterexamples; they are merely loanwords which have not been completely incorporated into the native lexicon.

Loanwords are often modified by sound substitutions, i.e. for any source language sound which the borrowing language lacks, the phonetically closest sound is substituted. The consonant inventory of OshiKwanyama is shown in (1) below (only \[g\]) is not contrastive):

\[
\begin{array}{ccccccc}
p & t & k \\
b & d & g \\
\tilde{\chi} & \tilde{\gamma} & \\
f & \tilde{\xi} & x & h \\
v & m & n & n & n \\
\mid & y & w \\
\end{array}
\]

Since OshiKwanyama has no \(r\), \(l\) is normally substituted, as shown in (2):

\[
\begin{array}{lll}
\text{source language} & \text{source word} & \text{loanword in OshiKwanyama} \\
\text{German} & [\text{radiyo}] & [\text{oladiyo}] & 'radio' \\
\text{English} & [\text{taya}] & [\text{etayela}] & 'tire' \\
\text{Afrikaans} & [\text{frax}] & [\text{ofulaxa}] & 'burden, load' \\
\end{array}
\]
OshiKwanyama also has no s sound; both [f] and [ɔ] are substituted, although [ɔ] seems to be preferred. Some examples are shown in (3):

(3) source language source word loanword in OshiKwanyama  
| Eng  | [sak]   | [oŋako] | 'sack'  |
| Eng  | [vọ́n]  | [ovelisa] | 'verse' |
| Eng  | [sọ́j]  | [ofoloto] | 'sort'  |

One of the major constraints on sounds in OshiKwanyama is the constraint on closed syllables; that is, no syllable may end in a consonant. It follows from this that all OshiKwanyama words must end in a vowel. This appears to be the very strongest constraint in the language. I have found not one single loanword which ends in a consonant. Even loanwords which lack most other signs of incorporation have acquired a final vowel; the examples in (4) are very "foreign-sounding" loanwords, yet all of them have final vowels.

(4) Eng  | [aŋ kai̯m] | [aŋsəkrima] | 'ice cream'  |
| Eng  | [bi:ɔ]  | [obiɾa] | 'beer'  |
| Eng  | [bl̩̊kbo̊d] | [bl̩̊kbo̊də] | 'blackboard'  |
| Eng  | [sətæf] | [ətæf] | 'seraph'  |

The example words meaning 'ice cream' and 'beer' are foreign-sounding in large measure because the r was not replaced by ɾ. The other words do not sound native primarily because they lack noun class prefixes.

OshiKwanyama has nine noun classes. The noun class prefix system (slightly simplified) is: Cl. 1 omu-/ova-, Cl. 2 omu-/omi-, Cl. 3 e-/oma-, Cl. 4 o(N)-/ə:(N)-, Cl. 5 oši-/ol, Cl. 6 olu-/omalu-, Cl. 7 oku-/oma-, Cl. 8 ou-/omau-, Cl. 9 oka-/ou-. Loanword assignment of noun classes corresponds closely to the size of the noun class; that is, the classes which contain the largest number of native nouns also contain the most loanwords. Loanwords belonging to the five classes with largest membership are shown in (5):

(5) Eng  | [læmp] | [ɔlæmpa] | 'lamp'  |
| Eng  | [pən] | [opena] | 'pen'  |
| Eng  | [faɾk] | [efoloka] | 'fork'  |
| Germ | [hau̯tə] | [eha:rə] | 'dune'  |
| Afr  | [taɾfaɾ] | [oʃiɾafula] | 'table'  |
| Eng  | [goɾid] | [oʃiŋolodo] | 'gold'  |
| Eng  | [pelaɾis] | [omupoliɾi] | 'policeman'  |
| Eng  | [biɾɔp] | [omumbiɾɔpa] | 'bishop'  |
An interesting interaction occurs between the phonological constraint against certain kinds of consonant clusters and the sound substitution processes. Native Oshikwanyama words have a constraint against all clusters except consonant plus glide and nasal plus voiced stop. Clusters consisting of obstruents or clusters which contain liquids never occur. When these sorts of clusters are found in borrowings, they are almost always eliminated. A few of the small number of loans which have not eliminated such clusters were shown in (4).

The most common method of cluster simplification is to insert a vowel, as in the first five examples in (6); occasionally one of the two consonants is deleted, as in the last three examples in (6):

(6) Eng [taJm] [ofalama] 'farm'
    Germ [karte] [okalita] 'map'
    Eng [kusmas] [okilišitimiša] 'Christmas'
    Germ [brot] [ombolota] 'bread'
    Afr [spel] [ošipeia] 'pin'
    Eng [powst] [opoša] 'post, mail'
    Afr [kombers] [ošikumbafa] 'blanket'

Note that, while unsimplified clusters of consonant plus r and consonant plus l may occur (as in (4)), simplified clusters, like those in (6) always contain the substituted l sound. Thus, while substitution of a native sound for a foreign sound does not necessarily imply cluster simplification, cluster simplification does imply sound substitution; consequently forms such as *[oøyšikirima] and *[ošipeia] never occur. Thus, there must be a hierarchy of loanword incorporation processes; sound substitution is primary with respect to processes which reflect constraints against consonant clusters.

Finally, I will discuss the processes which modify nasal clusters. In the native system of Oshikwanyama, the voiceless stops are all underlyingly distinct. The voiced alveolar stop, d, is also contrastive. However, it occasionally alternates with t. The b also is contrastive, although it sometimes alternates with v. The g occurs only after [ŋ]. None of the voiceless stops occur after nasals; only voiced stops are found in this environment. When nasals and stops come together as a result of morpheme conjunction, the results are illustrated in (7):

(7) native root non-nasal precedes nasal precedes
    -pat- [olupati] 'rib' [e·mati] 'ribs'
    -pot- [okupota] 'to be rude' [omote] 'good-for-nothing'
    -tan- [okatana] 'little calf' [onana] 'calf'
What, then, occurs in loanwords? If the source language word has an initial voiced stop, normally the loanword is modified in one of two ways (as shown in (8)) either a homorganic nasal is inserted before the voiced stop, as in 'bet' and 'gold', or else the word is assigned the nasal-final form of the Class 4 prefix:

(8) Eng [bet] [mbeta] 'bet, gamble' (verb)
    Eng [gəwl] [oʃingolodo] 'gold'
    Germ [bibel] [ombibela] 'Bible'
    Eng [dæm] [ondama] 'dam'
    Eng [donkiy] [ondongi] 'donkey'

Source language words which contain clusters of nasal plus voiceless stop are normally modified as well, as shown in (9):

(9) Eng [stamp] [ʃitamba] 'stamp' (verb)
    Afr [pomp] [opomba] 'pump'
    Eng [punt] [pelenda] 'print' (verb)
    Germ [tinta] [otinda] 'ink'
    Afr [trunq] [ondoongo] 'prison'
    Eng [ink] [oinga] 'ink'

In these cases the voiceless stops are voiced after nasals. But note, this modification can not be the result of applying a synchronic phonological rule. Oshikwanyama has no rule which voices stops after nasals; instead the Oshikwanyama rule would delete the voiceless stops, as shown in (7). As well, note that nasal plus voiceless stop clusters are, in fact, allowed in Oshikwanyama in certain cases. As a result of certain fast speech rules, clusters of nasal plus any consonant may actually occur, even though there is a deep structure constraint against any but nasal plus voiced stop clusters. (Thus, 'tree' may be pronounced [onmu], [omt], or even [onti] in increasingly more casual and faster speech.) However, the loanwords are modified in spite of
Loanwords and MSC's in OshiKwanyama

it seems clear that in this case the loanword incorporation process is derived not from the surface structure but as a result of the deep structure constraints.

Although OshiKwanyama allows voiceless stops to occur both in stem-initial and word-initial position, nevertheless a small number of loanwords with initial p, t, or k undergo the modifications illustrated in (10):

(10) Afr [papir] [ombapia] 'paper, letter'
Eng [plk] [ombiki] 'pick-axe'
Afr [kerk] [ongeleka] 'church'
Germ [kamel] [ongameilo] 'camel'
Afr [trok] [ondolongo] 'prison'

The word in each case is assigned the nasal-final form of the Class 4 prefix, even though initial op, ot, or ok are perfectly allowable sequences. However, there are no sequences of these sounds in this particular class of nouns. Thus, it is possible to find words like [opo] 'here', [otwa] 'we', and [okakadona] 'girl' (Cl. 9), but no such sequences occur in the native words of class 4 (because, of course, of the rule illustrated in (7)). However, these initial voiceless stops could have been retained without inserting a preceding nasal if the words had been assigned to class 3 or class 5, both of which contain large numbers of loanwords and which allow stem-initial voiceless stops. However, it would appear that these words were simply assigned to class 4 because this is the class that most loanwords are assigned to; however, once in class 4, a class-specific constraint against stem-initial voiceless stops gave rise to the use of the nasal-final form of the prefix. Once this form of the prefix was assigned, the incorporation process illustrated in (9) came into effect and voiced all of these post-nasal voiceless stops.

In conclusion, there appears to be a hierarchy of loanword incorporation processes in OshiKwanyama. As well, some of these incorporation processes derive from the deep structure constraints and not from the surface structure. Finally, some processes may even be specific to a particular subset of words, i.e. nouns of Class 4.

REFERENCES

Swahili relative clause constructions are complex and variable. On the surface, the language has three relative clause constructions, each of which has subvarieties. This paper argues that earlier attempts to analyze Swahili relatives as one construction have masked significant differences between the relatives and presents a Generalized Phrase Structure Grammar (hereafter GPSG) account of Swahili relatives. Transformational accounts [Givon 1972, Vitale 1981] have derived all three types of relative clauses from one deep structure, seeing these surface variations as reflexes of other rules applying to the output of a single relative clause rule. These accounts have obscured differences in constituent structure and conditions on extraction among the types of relative clauses. A phrase structure approach, on the other hand, emphasizes the differences in surface strings. In Swahili relatives these surface variations provide key information about differences in internal constituent structure in the three relative constructions which in turn account for differences in subject post-posing, occurrences of the complementizer KWAMBA, and conditions on extraction. The analysis I present is basically within the framework presented by Gazdar and Pullum [1982].

Consider the following TENSED relative constructions:

(1) a. mtoto a-  li- ye- ki- ona kitabu  'child who saw the book'
     child subj past rel obj see book

b. kitabu a-  li- cho- ki- ona mtoto  'book which the child saw'
     book subj past rel obj see child

c. *kitabu mtoto a-  li- cho- ki- ona
     book     child subj past rel obj see

In (1a) the head noun is followed immediately by the inflected verb, which is expected as the subject noun phrase is relativized. However, in (1b) the head noun is also followed immediately by the verb with the subject appearing after the verb. Sentences in which the full noun subject appears before the verb are ungrammatical as shown in (1c). In short, a canonical SVO sentence construction never appears in TENSED relatives.

AMBA relatives, on the other hand, evidence a rather different pattern. Consider the following AMBA relatives:

(2) a. Mary ambaye a-  li- m- pa mtoto kitabu
     Mary rel     subj past rel obj give child book

'Mary who gave the child the book'
b. mtoto ambaye Mary a- li- m- pa kitabu
child rel Mary subj past obj give book
'the child who Mary gave the book'

c. kitabu ambacho Mary a- li- m- pa mtoto
book rel Mary subj past obj give child
'the book which Mary gave the child'

Sentences (2b) and (2c) show a full noun subject occurring before the verb in these relatives. In fact, AMBA relatives appear to be canonical SVO sentences in which a noun phrase is missing, i.e. in (2a) the subject noun is missing, in (2b) the direct object noun is missing, and in (2c) the indirect object noun is missing.

Traditional analyses have accounted for these differences in surface strings by positing a single deep structure for both relatives. An obligatory subject post-posing mechanism moves the subject noun phrase to a post-verbal position in TENSED relatives. However, another way to look at AMBA and TENSED relatives is to accept the surface strings at face value. Instead of analyzing these strings as representing one deep structure to which different transformations have applied, AMBA and TENSED relatives could be seen as having different constituent structures. AMBA relatives would be analyzed as a head noun followed by a V and TENSED relatives as a head noun followed by a V N. Assuming this analysis, AMBA relatives will be admitted by rule (3):

(3) [ [R AMBA V/N [aclas]] [aclas] [aclas] ]

Rule (3) employs the /category convention developed by Gazdar to represent extraction phenomena. It indicates that there will be a gap at some point in the string. In the case of /NP a gap will appear in the string where one would expect a noun phrase to appear. Rule (3), then, claims that a relative introduced by AMBA is a head noun followed by a sentence which is missing a noun phrase.

TENSED relatives will be admitted by rules (4a) and (4b):


Rule (4a) claims that a TENSED relative in which the subject noun phrase has been relativized consists simply of a verb bearing the proper markings followed by a noun phrase which agrees with the object markings on the verb. TENSED rel-
atives formed on object noun phrases are admitted by rule (4b).

Before accepting such a proliferation of phrase structure rules and the notion that these two relatives are quite different structurally, we would want proof that indeed these strings act differently. The appearance of the complementizer KWAMBA provides such evidence.

KWAMBA is an optional complementizer which appears before complements which seem to be full sentences displaying canonical SVO order. Sentences involving S-complements will be introduced by rule (5):

(5) \[ V \] (kwamba) \[ +\text{comp} \] \[ -\text{comp} \]

This rule specifies that any embedded \( \tilde{V} \) can be optionally introduced by KWAMBA. Thus pairs of sentences such as:

(6) a. Mary a- li- sema John a- li- kwenda
Mary subj past say John subj past go
'Mary said John went'

b. Mary a- li- sema kwamba John a- li- kwenda
Mary subj past say that John subj past go
'Mary said that John went'

are both introduced by the same rule. KWAMBA can occur in AMBA relatives but never in TENSED relatives:

(7) a. mtu ambaye kwamba Mary a- na- m- penda a- li- kwenda sokoni
person rel comp Mary subj present obj like subj past went store to
'the person who Mary likes went to the store'

b. mtu kwamba a- na- ye- penda Mary a- li- kwenda sokoni
person comp subj pres rel like Mary subj past go store to

A plausible explanation for these facts is that the internal constituent structure of the two relatives differs. If TENSED relatives do not have the internal constituent structure of a \( \tilde{V} \), then we would not expect KWAMBA to appear before the verb in these constructions. And indeed, it cannot. If the AMBA relatives, on the other hand, have \( V \) as their constituent structure, we would expect KWAMBA to be able to appear after AMBA, as it does. Simply allowing the KWAMBA rule, which is independently needed in the grammar, to apply in relatives predicts the correct distribution. Thus analyzing the constituent structure of the TENSED relative as \( \tilde{V} \tilde{V} \) and AMBA relatives as \( \tilde{V} \) automatically accounts for the restrictions on KWAMBA in relative clauses.

Post-verbal subject data provide additional support for this analysis. Givon [1972] and others have argued that subjects in AMBA constructions are moved to a post-verbal position by the same process that moves subjects to a post-verbal position in TENSED relatives. Hence the argument says that the
two constructions are essentially the same. Close examination of the data, however, indicates that post-verbal subjects in AMBA relatives do not behave the same as post-verbal subjects in TENSED relatives.

When both the subject and the relativized object of TENSED relatives are of the same noun class, sentences are ambiguous. For instance, sentence (8) can be interpreted as either 'the person who liked Juma' or 'the person who Juma liked.'

(8) mtu a- li- ye- m- penda Juma
   person subj past rel obj like Juma

Keach [1982] points out, however, that AMBA relatives in which the subject appears post-verbally are not ambiguous. This is because a noun phrase following the verb in an AMBA relative must be set off with a pause if it is to be interpreted as the subject.

(9) a. mtu ambaye Juma a- li- m- penda 'the person who Juma liked'
    person rel Juma subj past obj like

b. mtu ambaye a- li- m- penda Juma 'the person who liked Juma'
   person rel subj past obj like Juma *'the person who Juma liked'

c. mtu ambaye a- li- m- penda, Juma
   person rel subj past obj like (pause) Juma
   'the person who Juma liked'

(Keach: 161)

Keach emphasizes that these AMBA constructions are never ambiguous while seemingly parallel TENSED constructions are.

Keach further notes that subjects can appear post-verbally in non-relatives:

(10) a- li- m- penda Juma, mtu 'the person liked Juma'
    subj past obj like Juma (pause) person 's/he liked Juma, the person'
    (s/he = person)

Thus there appears to be an optional rule in Swahili that allows subjects of full sentences to appear post-verbally if the subject noun is set off by a pause. A GPSG type analysis could account for these sentences by a meta-rule as in (11):

(11) if \([\overrightarrow{N} \ [\text{asubj}] \ [\text{+ fin}] \rightarrow [\overrightarrow{V} \ [\text{+fin}] \ [\text{asubj}]])\]

I assume the morphophonemic rules will interpret \(N\) as "pause \(N\)." This rule says that if \(N \ V\) is an acceptable string then \(V\) pause \(N\) will also be an acceptable string.

Without adding any further conditions, we would expect this rule to apply to any \(V\) construction. Thus if the constituent structure of AMBA relatives is
we would expect to find post-verbal subject nouns set off by pauses occurring in AMBA relatives but not in TENSED relatives. The data indicate that this is the case. Again, we find that an independently motivated rule which applies to $\bar{V}$ can automatically account for a difference between TENSED and AMBA relatives if AMBA constructions are analyzed as $\bar{V}$. These representations have the further consequence of predicting differences in the conditions on extraction for each construction.

Gazdar [1981] showed that the /category notation could be used to account for unbounded dependencies such as those formed by wh-movement in English relatives. With no modification of Gazdar’s /category convention, the phrase structure rules introducing AMBA relatives predict that AMBA constructions will allow unbounded dependencies. TENSED relatives, however, are not introduced by the /category notation. Since /category is the only mechanism in GPSG allowing unbounded dependencies, phrase structure rules (4a) and (4b) predict that TENSED relatives will allow only bounded extractions. Consider sentence (12):

(12) $[S_1 mtu [S_2 ambaye John a- na- fikiri [S_3 kwamba a- li- m-] person subj present think comp subj past m- piga Mary]] a- na- ki- penda kitabu] obj hit Mary

'the person who John thinks hit Mary likes the book'

In this sentence a dependency occurs between the relative marker and a gap in $S_3$ which is one cyclic node away. This sentence is perfectly grammatical. However, a parallel sentence using the TENSED strategy is ungrammatical:

(13) $*[S_1 mtu [S_2 a- na- ye- fikiri John [S_3 kwamba a- li- m-] person subj pres rel think comp subj past piga Mary]] a- na- ki- penda kitabu] hit Mary subj pres obj like book

'the person who John thinks hit Mary likes the book'

The only difference between the strings is that the AMBA construction is used in (12) while the TENSED construction is used in (13). These sentences indicate that a noun phrase cannot be extracted from a lower clause in the TENSED construct but can be extracted from a lower clause in the AMBA relative. The conditions on extraction are clearly different for TENSED and AMBA relatives as predicted by the phrase structure rules.

To summarize, we began this analysis by simply examining TENSED and AMBA constructions and noting that TENSED relatives always appear as verb initial strings while AMBA relatives appear as SVO sentences missing a noun phrase. Following a GPSG type analysis which emphasizes surface structure, we posited different constituent structures for AMBA and TENSED relatives. This analysis, in turn, automatically accounted for the distribution of KWAMBA and the differences in ambiguity in subject post-posed relatives. These phrase structure rules had the additional consequence of correctly predicting the conditions on extraction in TENSED and AMBA relatives.
A persistent objection to phrase structure accounts has been that they simply list possible surface strings, thus missing generalizations which could be captured in transformational accounts. In this paper I argue the contrary: a phrase structure account which begins by paying close attention to surface structure reveals important generalizations about Swahili relatives which have been masked in earlier transformational accounts.

[All data, unless otherwise indicated, is from Judy Nyanda who is a native speaker of Swahili from Dar Es Salaam, Tanzania.]

REFERENCES


The methodology of this paper assumes that the existence of grammatical relations (GR's) must be established language-internally on the basis of a proliferation of morpho-syntactic phenomena which make reference to the same set of logical argument positions in a clause. I furthermore reject a methodology which makes a priori initial assignments to GR's based ultimately on semantic roles as in mainstream Relational Grammar (RG), which consequently must often sanction obligatory revaluations to account for surface accessibility to the GR-defining phenomena. Note that this methodological point is a natural extension of recent findings with respect to the Unaccusative Hypothesis [Rosen 1982], which have shown that it is impossible to decide on universal semantic grounds whether the single argument of a monovalent verb will head an unaccusative (2) or unergative (1) arc at the initial level. To say that there is no "universal alignment" between semantic roles and unaccusativity is equivalent to saying that there is no such alignment between semantic roles and 2-hood, which is independent of transitivity. These arguments may be extended to the 3 relation, initial instantiation of which is restricted to semantic recipients in RG.

A consequence of the methodological assumption that GR's are logically independent of semantic roles is that the set of semantic roles which can be mapped onto initial GR's in one language will not necessarily be coextensive with the same set in another language. I will argue that in Gokana, a Benue-Cross language of Nigeria, there are two object relations (2 and 3), accessibility to which is determined inversely along a hierarchy of semantic roles such that:

(1) Recipient/benefactive/possessor - Patient - Cognate object - Instrument

That hierarchies of this kind may characterize GR's at an initial level has been argued for a number of languages, including Huichol [Comrie 1982], Fula [Wagner 1982a], Nahuatl [Wagner 1982b], and Bantu generally [Hyman and Duranti 1982]. Note that whereas the hierarchies corresponding to objects in these languages are not identical in every respect, a consistent generalization underlying them indicates a tendency to grammaticalize a ranking of semantic roles on the basis of the relative discourse prominence (or "topic-worthiness") of the class of NP's most typically instantiating those roles and suggests the need for further research into the explanatory power of such discourse notions on phenomena involving grammatical relations, as in, e.g. Schachter [1977]. Note furthermore that a proliferation of object-types should be regarded as no more problematic (or "unconstrained") for a theory of language than, for example, the extension of a verb form in one language to assume semantic functions encoded in other languages by means of prepositions (as in serial verb languages) or adjectives. The point here is that language-internal evidence must decide category membership.

One property of the non-subject argument of monotransitive verbs in Gokana
is that, under certain discourse conditions, it may appear both in its characteristic immediate post-verbal position and in a serial construction as the object of \( t \) 'take':

\[
\begin{align*}
(2) \text{a. } & \text{aè } fè \text{ nom } \text{ 'he killed the animal'} \\
& \text{3sg-pst kill animal} \\
\text{b. } & \text{aè } tù \text{ nom } fè = (2a) \\
& \text{3sg-pst take animal kill}
\end{align*}
\]

As a serial verb language Gokana prefers a clause type with one object per verb. Nevertheless it has a small number of underived ditransitive predicates which may occur with two overt objects. Note that it is the patient which evinces the 2-hood properties of word order and \( t \) serialization here:

\[
\begin{align*}
(3) \text{a. } & \text{gín } è \text{ nè kpá pábìa } \text{ 'the child gave the book to the woman'} \\
& \text{child pst give book woman} \\
\text{b. } & *\text{gín } è \text{ nè pábìa kpá}
\end{align*}
\]

\[
\begin{align*}
(4) \text{a. } & \text{gín } è \text{ tù kpá nè pábìa } = (3a) \\
& \text{child pst take book give woman} \\
\text{b. } & *\text{gín } è \text{ tù pábìa nè kpá}
\end{align*}
\]

In clauses with an expressed instrumental (a category which includes comitative and manner NP's as well), the suffix -\( \text{mà} \) is added to the verb, increasing its valence by one. Here however it is the instrument which must occur in immediate post-verbal position:

\[
\begin{align*}
(5) \text{a. } & \text{aè } \text{ feèmà } \text{ gìmà nom } \text{ 'he killed the animal with a knife'} \\
& \text{3sg-pst kill-with knife animal} \\
\text{b. } & *\text{aè } \text{ feèmà nom gìmà}
\end{align*}
\]

Furthermore only the instrument may appear as the object of \( t \) :

\[
\begin{align*}
(6) \text{a. } & \text{aè } tù \text{ gìmà feèmà nom } = (5a) \\
& \text{3sg-pst take knife kill-with animal} \\
\text{b. } & *\text{aè } tù \text{ nom feèmà gìmà}
\end{align*}
\]

Note finally that there is no alternative periphrastic means for expressing instrumentality and thus no motivation for any non-obligatory advancement rule (see below).

Gokana also has a number of verbs which take obligatory cognate objects bearing various participant roles with respect to the verb. These occur in the direct object slot:

\[
\begin{align*}
(7) \text{a. } & \text{aè } \text{ bìrì múù gín } \text{ 'he bathed the child'} \\
& \text{3sg-pst wash water child}
\end{align*}
\]
b. *ae bire q' in můů

Nevertheless, either the patient or the cognate object may appear as the object of tů in the serial construction:

\( (8) \)

<table>
<thead>
<tr>
<th>a. ae tů můů bire q' in</th>
<th>( (7a) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3sg-pst take water wash child</td>
<td></td>
</tr>
</tbody>
</table>

\( (8) \) b. ae tů q' in bire můů = \( (7a) \)

Thus tů serialization (as well as passive, not discussed for this construction) does not determine 2-hood unambiguously, suggesting that this object relation is to some degree diffuse; accessibility to the 2-relation diminishes as one moves up the hierarchy from the right. Note that the facts in \( (7) \) and \( (8) \) can be accounted for within a RG framework which assumes that different targets of demotion accompany different promotion rules. Thus tů serialization would be sensitive to working, and word order to final, 2-hood. While this is of course an adequate formalization of the facts discussed, we must look elsewhere, e.g. to discourse factors, for explanatory principles governing such behaviors.

Gokana has a lexically detransitivizing suffix -a which denotes either passive or reciprocal meaning. With monotransitive roots this form may be construed in either way if pragmatically feasible:

\( (9) \)

<table>
<thead>
<tr>
<th>to</th>
<th>tiv - à</th>
<th>'the house has been built'</th>
</tr>
</thead>
<tbody>
<tr>
<td>house build-a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( (10) \)

<table>
<thead>
<tr>
<th>gbopá'mín</th>
<th>ñv - á</th>
<th>a. 'the children are hated'</th>
</tr>
</thead>
<tbody>
<tr>
<td>children hate-a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The facts involving derived instrumental-verbs are more categorical. As in \( (11) \), only the rightmost NP in \( (1) \) above may passivize:
(12) a. gẹ kpọ - ṣ ọm
    knife cut - a meat
    'the knife was used to cut the meat'

b. *nom kpọ - ṣ ọm
    meat cut - a knife
    '*the meat was cut with a knife'

Nevertheless, the subject may be interpreted as coreferential to the patient:

(13) gbopam fn fe- à ọm
    children kill - a knife
    a. 'the children killed each other with a knife'
    b. '*the children were killed with the knife'
    (i.e. not by each other)

Note that the instrumental-2 is not a possible target for reciprocal formation here, though in principle it may be instantiated by a comitative NP with human reference.

Benefactives in Gokana are normally expressed using a serial construction with nè 'give':

(14) bẹ̀b ẹ́ tib to nè ẹ̀m 'they built a house for the child'
    3pl-pst build house give child

(15) gbopam fn fe- èm
    children kill - a animal
    a. 'the children killed animals for each other' benefactive-reciprocal
    b. '*the children were killed an animal' benefactive-passive
    c. 'the children killed each other's animals' possessor-reciprocal
    d. '*the children's animals were killed' possessor-passive

Reciprocal Formation may access the benefactive nominal directly (a), though this NP is not available for advancement to subject via passive (b). Note that (15) also admits a reading in which the possessor of the patient NP is coreferential with the subject (c), though it may not be construed as a passive with a possessor-subject (d). Now a maximally general characterization of the class of possible targets of reciprocal formation under the assumptions of RG, i.e. the initial "oblique" status of non-subcategorized arguments and the relegation of possessor NPs to "pre"-relational status, would require promotion in the case of benefactives and possessives to some GR accessible to reciprocal formation. Since neither NP is a possible passive subject, and since we have seen above that 2's may passivize, it follows that these nominals are promoted to 3 (though note, in the case of the possessor of the initial 2, nom (cf. (c) above) that this violates the Relational Succession Law constraining the class of possible ascensions.). Nevertheless there is no evidence from active clauses for this level of derivation; reciprocal formation w/-a is thus obligatory for non-initial 3's. Yet while we cannot rule out obligatory rules a priori, there is no language internal evidence for such rules independent of their interaction with the grammatical processes we are seeking to explain. In contrast a formulation of GR's as in (1), in terms of ordered sets of participant roles, does not require this unmotivated (and unexplanatory) abstractness.

In conclusion the data discussed here from Gokana provide three morpho-syn-
tactic tests which motivate the category of DO—word order, to-serialization, and passivization. These phenomena consistently isolate the rightmost position in (1), though further up the hierarchy the notion of DO is more diffuse; NP's bearing roles to the left of patient are inaccessible. The 3-relation is defined here by accessibility to reciprocal formation and is instantiated by the NP in (at least strongly preferentially) the leftmost position, stopping to the right of the patient. I have argued against the positing of obligatory rules required in RG by assumptions involving the possible semantic roles of initial GR's, assumptions which we have rejected in our methodology and which, indeed, are untenable if an initial level distinct from a representation of participant relations is to be motivated. Finally, I have suggested that further research may afford an explanation of the structure of hierarchies such as (1) above in terms of the discourse properties of the processes which reference GR's.

REFERENCES


SOME ASPECTS OF LF MOVEMENT IN IRAQI ARABIC

Wafaa Wahba
University of Mansoura, Egypt

0. Introduction

This study investigates certain aspects of the Syntax of S-Structure (SS) and the syntax of Logical Form (LF) as manifested by the behaviour of wh-questions in the grammar of Iraqi Arabic (IA), as spoken in Baghdad and in central and southern Iraq. Wh-operators in IA can appear in the Comp node or in their base position (wh-in-situ). Like English, the wh-operator appears in the Comp and the questioned site is marked by a gap:

(1) a. meno. Mona shaafat e.?  b. weyya meno. Mona xarja.t e.?  
who Mona saw          with whom Mona left  
'who did Mona see?'    'with whom did Mona leave?'

Like Turkish, Japanese [Hankamer 1973], and Chinese [Huang 1982], the wh-operator in IA has the option of occurring in its base position:

(2) a. Mona shaafat meno?  b. Mona xarja.t weyya meno?  
Mona saw whom          Mona left with whom  
'whom did Mona see?'    'with whom did Mona leave?'

Following a long tradition, I assume the existence of a rule of wh-movement in LF. This rule moves the wh-phrase-in-situ to a position c-commanding the S that constitutes its domain. The fact that the grammar of IA allows wh-movement both in the Syntax and in LF provides us with the opportunity to contrast the syntax of SS with the syntax of LF as manifested by the same construction in the same grammar. The question arises as to what extent the properties of the LF rule and the Syntactic rule are the same. The data in this paper provide strong support for the existence of the LF movement rule and hence for the level itself. This rule affords a uniform account of numerous significant generalizations that hold between the two levels of representations of SS and LF, which might otherwise be describable via a conjunction of unrelated properties. This rule will also enable us to account for the fact that wh-in-situ questions observe several locality requirements which are observed by wh-Movement in the Syntax. The data discussed in this paper provide evidence for the following claims with respect to the LF rule: (a) Subjacency holds as a well-formedness condition at LF, i.e. the LF rule in IA observes the various island conditions; (b) there is independent evidence that there is Comp-to-Comp LF movement, hence, LF extraction affects both argument and non-argument positions; (c) successive cyclicity is observed by the SS and LF rules. Due to space limitations, I will condense my arguments, and, for a more detailed and fully analysed version of this paper, the reader is referred to Wahba [forthcoming], where I discuss in detail the logical structure of wh-in-situ questions in IA.

Section (1) deals with some scope as well as distributional properties of
wh-in-situ. Section (2) discusses the weak crossover phenomenon which holds both at SS and at LF in IA. Section (3) provides empirical evidence that Move alpha applies in successive-cyclic steps. Section (4) discusses the behaviour of wh-phrases-in-situ with respect to Subjacency. Section (5) summarizes the major points of this paper.

1. Scope Properties of Wh-phrases-in-situ

This section argues in favor of a movement rule at LF to account for the fact that syntactically unmoved wh-operators exhibit similar scope properties to syntactically moved wh-operators. This rule will also help to account for the scope ambiguities exercised by wh-phrases-in-situ in IA.

I will assume, following Chomsky [1973], that the Comp node is marked [+wh] to mark its complement as interrogative and [-wh] to mark it as indicative. Consider a paradigm where a wh-phrase appears in the complement clause of a verb subcategorized only for indicative complements, as in (3a), for both [+wh] and [-wh] Comp's, as in (3b) and, finally, for [+wh] Comp only, as in (3c):

(3) a. Mona itmannat tishtri sheno? b. Mona nasat tishtri sheno.?/ 
   Mona hoped to-buy what Mona forgot to-buy what
   'what did Mona hope to buy?' 'Mona forgot what to buy'
   or 'what did Mona forget to buy?'

c. Mona se?iat Ali Ro'a ishtarat sheno
   Mona asked Ali Ro'a bought what
   'Mona asked Ali what Ro'a bought'

In (3a), sheno must have a wide scope, since the lower Comp is marked [-wh]. In (3b), it may have a narrow scope, resulting in an indirect question reading. (3b) may also have a direct question reading where eeh has a wide scope over the entire sentence. In (3c), sheno must have a narrow scope since the matrix verb se?iat 'to ask' obligatorily takes an interrogative complement.

The scope properties exhibited by the wh-in-situ sheno in (3a-c) above mirror those of syntactically moved wh-operators. The paradigm in (4) illustrates this point with respect to syntactic wh-movement in IA. The translations demonstrate that similar facts obtain in English:

(4) a. sheno Mona itmannat tishtri e_i? 'what did Mona hope to buy?' sheno
   what i Mona hoped to-buy

   b. Mona nasat [ sheno_i [ tishtri e_i ]] 'Mona forgot what to-buy'
      sheno_i i M. nasat [ tishtri e_i ]

   b'. [sheno_i [ M. nasat [e_i [ tishtri e_i ]]]]? 'what did Mona forget to buy?'
      what M. forgot to-buy

   c. Mona se?iat Ali [ sheno_i [ Ro'a ishtarat e_i ]]
      'Mona asked Ali what Ro'a bought'

Assuming that wh-phrases-in-situ undergo movement in LF in a way similar to wh-movement in SS accounts for the fact that they exercise the same scope proper-
ties as syntactically moved wh-operators. This rule will also account for the scope ambiguities in (3b) by allowing the wh-phrase to move to where it can c-command its scope at LF. Hence (3b) will have two LF structures as in (5) below:

(5) a. \([\text{sheno}_1 [M. \text{nasat}[[\text{tishri}_1 e_i]]]]\) b. \(\text{Mona} \text{nasat}[\text{sheno}_1[\text{tis}\text{tiri}_1 e_i]]\)

2. Wh-movement and the Weak Crossover Phenomenon

The second argument which strongly supports an abstract movement rule at LF has to do with the weak crossover phenomenon which prohibits a coindexed pronoun from preceding a variable [Chomsky 1976]. Thus, (6) is excluded on the grounds that \(e\) is preceded by the coreferential pronoun \(\text{his}\):

(6) \(^*\text{Who}_1 \text{did his}_1 \text{mother see e}_1?\)

One way to exclude (6) is to assume that a moved element cannot cross a coindexed pronoun in its path to Comp. The weak crossover phenomenon holds in the Syntax and LF of IA. Thus, a syntactically moved wh-phrase can not cross a coindexed pronoun as in (7a), whereas a wh-phrase-in-situ can not be preceded by a coindexed pronoun in its domain as in (7b):

(7) a. \(^*\text{meno}_1 \text{?uxt-ha} \text{darbat e}_1?\) b. \(\text{?uxt-a-ha}_1 \text{darbat meno}_1?\)

&(who) sister-her hit e\_i & sister-her hit who\_i

'who did her sister hit?' 'who did her sister hit?'

(7a) is ruled out because the moved wh-operator crosses the coindexed pronoun -ha resulting, thus, in a crossover violation. In (7a), there is no visual movement, but the sentence is ruled out due to the presence of a coindexed pronoun that precedes the wh-phrase-in-situ. The similarity between (7a) and (7b) can be captured if we make the assumption that meno in (7b) undergoes movement at LF. Accordingly, (7b) will have the LF structure in (8):

(8) \([\text{meno}_1 [\text{?uxt-a-ha} \text{darbat} \ e_i ]]\)

In (8), we have a crossover violation since meno crosses a coindexed pronoun in its path to Comp. Thus, (7b) is ruled out on the same grounds as (7a) and (6).

3. Move alpha and Successive-cyclicity in SS and LF

I now provide empirical evidence that Move alpha applies in successive-cyclical steps as opposed to a single step in both SS and LF. Two arguments will be given below towards this end. One argument deals with the behaviour of a wh-phrase-in-situ at SS, where it has the option of appearing in any Comp node that intervenes between its base position and the matrix Comp. A second argument deals with Subjacency as a well-formedness condition at LF. Finally, I discuss the implications of the above claim with respect to the LF movement rule.
In addition to the fact that wh-operators in IA appear in argument (in-situ) as well as non-argument (Comp) positions, wh-phrases-in-situ have the option of appearing in any intermediate Comp that intervenes between their locus and the controlling Comp. In (9a), meno appears in its base position in the most embedded clause. In (9b), it appears in the next higher Comp (Comp3). In (c), it appears in Comp2. In (9d), it appears in the matrix Comp (Comp3):

   Comp1 Comp2 Comp3
Mona tried to-force S. to-help who
'who did Mona try to force Su'ad to help?'

   Comp1 Comp2 Comp3

c. [[[Mona hawlat [ meno_i [ tijbir S. [ e_i [ tisa'ad e_i ] ] ] ] ] ]]
   Comp1 Comp2 Comp3

d. [[[ meno_i [ Mona hawlat [ e_i [ tijbir S. [ e_i [ tisa'ad e_i ] ] ] ] ] ] ]]
   Comp1 Comp2 Comp3

All sentences in (9) have the reading of a direct question, where meno assumes a wide scope over the entire sentence, regardless of where it appears at SS. Thus, (9a-c) will have a single LF structure where meno moves to the matrix Comp, as in (9d).

The above paradigm has a number of significant consequences with respect to the rule Move alpha in general and LF movement in particular. The first has to do with the mode of application of Move alpha, i.e. whether it moves in successive cyclic steps or in a single step from its base position to the controlling Comp. To capture the fact that meno in (9a-c) above exhibits a wide scope over the entire sentence, we must conclude that Move alpha applies in successive steps; otherwise, we will not be able to account for the fact that meno surfaces in the intermediate Comp's in (9b-c), having a wide scope over the entire sentence and not over the domain it c-commands at SS.

The second consequence of the paradigm in (9) is that LF movement mimics SS movement in that extraction at LF takes place from argument as well as non-argument positions. In (9b-c), LF obligatorily moves meno from [Comp1] and [Comp2] respectively, to the matrix Comp where it c-commands its entire domain (contra Aoun et al. [1981]). This also indicates that a wh-operator at S-structure may occupy a [-wh] Comp node. This provides evidence that LF movement applies successive cyclically. Huang [1982] has shown that this is the case with respect to LF movement of adjuncts only since the latter exhibits stricter locality requirements than their argument counterpart at LF. In IA, both arguments and adjuncts can be extracted from Comp position.

The third consequence of the paradigm in (9) is that it provides more compelling evidence for the presence of a nonovert wh-movement rule at LF and hence for the level itself. This rule, which is assumed to be an instance of Move alpha, captures the striking similarities that hold between wh-in-Comp and wh-
in-situ questions in IA that have been shown so far: both constructions behave the same with respect to scope properties, the weak crossover phenomenon, and both allow extraction from argument as well as non-argument positions.

4. **LF Movement and Subjacency**

It has been assumed so far that the LF rule violates Subjacency; hence, a wh-in-situ in Chinese, Japanese, Turkish, English, etc., can have a wide scope if it occurs inside a syntactic island (except for adjuncts, cf. Huang [1982, 1983]). However, in IA, a wh-phrase-in-situ is barred from occurring inside syntactic island. The following examples show that a wh-in-situ is barred from occurring inside a wh-island (10a), a Complex NP (10b), and finally, a co-ordinate structure (10c):

(10) a. *Mona nasat [li-meno, [ tin+i sheno e_i ]]  
Mona forgot to-whom [ to-give what ]
'*what did Mona forget to whom to give?'
(Mona forgot [for which x, for which y [to give x to y ]])

b. *Mona surfut [NP il-bint_i [NP ili [ e_i ishtarat sheno]]]?  
Mona knew the-girl who bought what
'*what did Mona know the girl who bought?'

Mona saw Ali and-whom
'whom did Mona see Ali and?'

The sentences in (10) are ungrammatical because sheno occurs inside syntactic islands.

5. **Conclusion**

To summarize, I have independent evidence that (a) LF Comp-to-Comp exists. As a consequence, the LF rule can extract from Comp positions (cf. 9b-c). It also follows that wh-operators in IA may surface in [-wh] Comp at SS despite the fact that IA has wh-movement in the Syntax. (b) Subjacency holds at LF applying to all operators, arguments and adjuncts. To account for the facts discussed with respect to the behaviour of the LF rule in IA, some SS parameters need to be modified (cf. Aoun et al. [1981], Lasnik and Saito [1984]). Furthermore, we need a new parameter that applies at LF to accommodate the fact that Subjacency is observed by the LF rule in IA, Egyptian Arabic [Wahba 1984], and Tangale [Kidda 1983].
REFERENCES


Wahba, W. forthcoming. ms.
THE INFLUENCE OF ARABIC ON THE SYNTAX OF SWAHILI DISCOURSE
(Part 1: Verbal and Nominal strategies in temporal clause subordination)

Benji Wald
National Center for Bilingual Research

The paper raises the suggestion that Arabic has had an impact on Swahili beyond the lexicon. In particular, it suggests that Arabic has affected or reinforced trends in the syntactic devices used to organize discourse in Swahili at the inter-clause level.

Attention first focuses on syntactic devices for the overt expression of the temporal relations before and after between Swahili clause pairs. It is noted that Swahili has two basic strategies for expressing these relations, characterized as verbal and nominal strategies. The verbal strategies make use of verbal auxiliaries to express the interclausal relationships, e.g.

(1) a. sasa tu-ki-SHA ku-toka...tunakwenda mpaka ile dau
   now we-when-FINISH Inf-leave...we.go to the dhow
   'now, AFTER we leave... we go to the dhow' (AB21mMS)

   b. hata h-a-JA-osha uso a-na-yo yeye
   even Ng-he-'come'-wash face he-with-it him
   'even BEFORE he washes his face he has it' (FZ18mMS)

   This strategy is roughly equivalent to:
   a. (when) we FINISH leaving... (and then) we go to the dhow
   b. (when) he HASN'T YET washed his face (and) he has it

   The nominal strategies treat one clause as a nominal using an Arabic origin "conjunction" to introduce the clause, e.g.

(2) a. BAADA ya ku-wa n-SHA-dung-wa zile sindano ...
   AFTER of Inf-be I-FINISH-stab-Pass those needles ...
   'AFTER I was given the injections...' (CH18mMS)

   b. napajui hapa KABLA nyi ha-m-JA-zal-iwa
   I.knew here BEFORE y'all Ng-y'all-'come'-bear-Pass
   'I knew this place BEFORE you were born' (MS42mMS)

   This strategy approaches the more felicitous BEFORE and AFTER English translations given for all the sentences above. The source of these conjunctions are Arabic verbal nouns semantically used in either temporal (before/after) or locative senses (facing/beyond), and syntactically used to introduce either full clauses (in most varieties of Arabic baŷada/qablə MA: + Clause) or NPs.
In Swahili the borrowed conjunctions are restricted to temporal uses, but preserve the Arabic privileges of occurrence, i.e. introducing a clause or NP. Swahili has a pair of Bantu prepositions (actually NPs) semantically equivalent to baṣada and qabla, in nyuma 'behind' and mbele 'front', respectively, in that the latter pair have both temporal and locative uses. However, these are syntactically restricted to introducing NPs, not full clauses.

One may suspect that baṣada/qabla have simply replaced nyuma/mbele as clause introducers, and therefore have not altered the strategies inherited from Bantu for expressing the temporal relations at issue. However, several considerations suggest that such a summary dismissal of the impact of Arabic in these cases is premature. First, the verbal strategies exemplified in (1) above are an extremely widespread areal phenomenon in subequatorial Africa, especially among the Bantu and Nilotic languages, in many of these languages to the exclusion of the alternative nominal strategy. While not unknown among African languages, the nominal strategies in discourse are more typical of a more northerly area including Semitic and Southern Indo-European. (Space limitations prevent a full discussion of the evidence for this assertion, but particularly noteworthy is the detail of structural parallelism between Classical Arabic and Latin in the use of a preposition followed by a "wh" pronoun, Latin quam , Arabic ma , to form a clausal conjunction). Furthermore, in vernacular Swahili speech (at least among first speakers from traditional Swahili communities in Kenya), the verbal strategies are highly resistant to replacement by the nominal strategies. Thus, a compromise strategy involving elements of both strategies is found (double-marking). With respect to baṣada , the 'after' clause tends to retain the verbal auxiliary sha 'finish', essential to the verbal strategy. The 'before' clause is even more resistant so that, as in (2b) above, kabla is a superfluous "left-bracket" of the clause, which fully retains the verbal strategy. The kabla-clause, then, shows minimal adaption to the Arabic syntactic pattern. If kabla simply replaced mbele introducing a 'before' clause, why does Swahili retain the verbal strategy (involving negation), even when kabla is used?

The paper goes on to mitigate the possibly exaggerated conclusion that Arabic is having a typological effect on Swahili with regard to a shift in preference from verbal to nominal strategies, even in the case of baṣada and kabla. It is important not to overstate the case because much work remains to be done in determining the interplay of structural and social factors which may favor nominal strategies at the expense of verbal ones in Swahili discourse. At the same time, some examples are given of the impact of Arabic in erasing Bantu-derived verbal strategies which have become opaque (i.e. of limited syntactic distribution) in current Swahili speech. Among these examples are:

(3) BADO 'still/yet' (Arabic baṣuda ) as a replacement for i-NGA/K'A-lli ( K'A is the traditional Kenyan Swahili auxiliary).

(4) the use of ḥata (Arabic ḥatta ) 'even/up to' + 'if' clause, paralleling the Arabic construction for 'even if' clauses, coupled with the restriction of historical NGA 'even if/though' to i-NGA-wa , now exclusively used for 'though' clauses.
The simpler case is (3), the restriction of K'A to the syntactic formation i-K'A-li (historically SM-K'A-copula), a verbal strategy. Semantically, BADO and i-K'A-li differ under scope of negation. Thus, BADO in a negative clause signals 'still-not', i.e. 'not-yet' as in the Arabic source (internal negation), while i-K'A-li is only subject to external negation, 'not-still', i.e. 'not-anymore'.

The case of i-NGA-wa in (4) is more complex. Originally, NGA was an auxiliary marker neutral to the belief status of the clause, hence 'even if/though'. This state of affairs is still represented in archaic Swahili poetry (traditional Swahili poetry preserves the archaisms), e.g.

(5) a. u-NGenda mbiu h-ui-ki-tafuta
you-EVEN.IF-go fast Ng-you-TM-it-pursue
'EVEN IF you ran all the way, you would not be able to keep up with it'
(Basheikh: mahaba si tamu )

b. a-NGA-ji-shauza ... ha-i-na kirengo
she-EVEN.THO-R1fx-pride ... Ng-it-with content
'EVEN THOUGH she's very proud of herself ... she has no content'
(op. cit.)

In current Swahili, however, NGA remains only in the expression i-NGA-wa ( SM-NGA-be ) and signals '(even) though', i.e. the truth of the introduced proposition is presupposed. The 'even if' use of NGA has been replaced by haṭa + an 'if' clause. This results in a distinction not often made among Bantu languages, but of ancient vintage in Arabic.

These examples indicate that in the past Arabic has had some effect in altering the strategic base of Swahili by replacing verbal strategies with non-verbal strategies. The least that can be suggested is that Arabic has had some impact on Swahili beyond the "mere" lexical level. Traditional statements about the lack of Arabic impact on Swahili grammar, based on the "inviolable" morphological peculiarity of Swahili's Bantu structure, needs to be reconsidered in view of a broader perspective of grammar encompassing the use of syntax to organize information in discourse.
1. Introduction

The West African city of Niamey, in the Republic of Niger, is a particularly interesting setting for the study of language contact and bilingualism. In this city of nearly 300,000 people, there are two major languages spoken, each of which is extremely important in the day-to-day life of city residents, a large portion of whom are bilingual in the two languages. The languages, Hausa and Zarma-Songhai (hereafter referred to as Zarma in keeping with usage in Niamey), belong to separate language families and are structurally quite distinct from one another. Although there are other languages spoken in the city, it is Hausa and Zarma which serve as languages of intergroup communication. Together, native speakers of these two languages make up approximately 75% of the city's population and, although native Zarma speakers outnumber native Hausa speakers two to one, the two languages nevertheless are on a more or less equal footing in the city (for details, see Yanco [1984]).

This paper examines some of the syntactic consequences of the contact between Hausa and Zarma in the non-native (NN) speech of bilinguals, focussing on word order and the extent to which differing word order patterns in the native and second language are a source of interference or transfer. It is found that the acquisition of a differing order in the adposition phrase is the least problematic, which is similar to findings of a study among Quechua (postpositional) children learning Spanish (prepositional) [Lujan et al. 1983]. Differing order in the genitive construction was found to be the most problematic, native order persisting in this construction more so than in others. Both of these findings bear interestingly on Hawkins' [1983] recent work on the relative stability of noun modifiers in maintaining the noun-modifier serialization of the adposition phrase.

The data which form the basis for this paper were collected through a series of interviews in the 19 major quartiers of the city of Niamey. In each quartier two sets of interviews were conducted, one set with native speakers of Hausa and one set with native speakers of Zarma. The purpose of these interviews was two-fold: to gather information on speakers' linguistic life histories and to collect second language data. The second language data were collected through a translation task in which individual respondents were presented with utterances in their native language and asked how they would say the same thing in their second language. The choice of a translation task was motivated by the desire to elicit as many pertinent (in terms of contrasts between the two languages) constructions as possible. Nor was it an unfamiliar task, as respondents frequently cited interpreting as one of the most important uses to which they put their bilingual skills. Whenever possible, analyses based on this data have been corroborated with data from natural conversation. All interviews were audio recorded, transcribed, and coded for the use of various patterns.
Zarma, part of the Songhai dialect cluster classified by Greenberg [1966] as a branch of Nilo-Saharan, is postpositional; Hausa, of the Afro-Asiatic family, is prepositional. The table below summarizes the serialization of elements in the noun phrase. (The notation is meant to express semantic equivalents across the two languages; thus, for example, N Attr refers to the attributive construction, but entails no claim as to the syntactic category of Attr.)

<table>
<thead>
<tr>
<th>ZARMA (Postpositional)</th>
<th>HAUSA (Prepositional)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N Num</td>
<td>N Num</td>
</tr>
<tr>
<td>N Dem</td>
<td>Dem N/N Dem</td>
</tr>
<tr>
<td>N Attr</td>
<td>Attr N/N Attr</td>
</tr>
<tr>
<td>Poss N</td>
<td>N Poss</td>
</tr>
<tr>
<td>Gen N</td>
<td>N Gen</td>
</tr>
<tr>
<td>N Rel</td>
<td>N Rel</td>
</tr>
</tbody>
</table>

As can be seen above, relative clauses and numeral quantifiers pattern the same way in both languages. Hausa has two doubling constructions, demonstratives and attributives; in each case the more basic is indicated first. (For a discussion of the term "basic", see Hawkins [1983:12-16].) For a number of constructions, the basic order in one language is the opposite of that in the other. This is true for adpositions, attributives, possessives, and genitives. (I have distinguished possessive constructions in which the possessive element is a pronominal form from those in which it is a nominal, as they pattern quite differently in the non-native speech data.) It is also of interest to note that, while Hausa is strictly SVO, Zarma has dominant SOV order, with a limited number of verbs requiring that their objects follow. This is not a stylistic variation, but is determined by the verb itself. For a few verbs, either order (OV or VO) appears to be acceptable; however, informants varied widely in their interpretations of what the difference was.

2. The Data

In this section, word order in non-native speech is summarized in terms of percentages for each of the following constructions: numeral quantification, adposition phrase, attributive construction, possessive construction (with a pronominal form), and genitive construction (with a nominal form). Left hand columns refer to non-native Zarma (NN ZARMA), i.e. Zarma spoken by native Hausa speakers. Right hand columns refer to non-native Hausa (NN HAUSA). Examples labeled (a) reflect the standard or basic word order used by native speakers of the language in question; examples labeled (b) reflect reversed order; and (c), "other", includes all tokens which were not analyzable in terms of the construction in question, including omission of the relevant item. I was not concerned with whether or not people used the "correct" forms; using English as an example, I would accept 'men' and 'mans' or 'under' and 'beneath' as being the same item for the purposes of this study. The number of tokens serving as the basis for percentage calculations varies with the different constructions. The smallest number is 90, the largest, 384.

As an example of a construction in which the order of meaningful elements is the same in both languages, the word order percentages for numeral quantifica-
Studies in African Linguistics, Supplement 9

tion are presented below:

<table>
<thead>
<tr>
<th>% NN ZARMA tokens</th>
<th>% NN HAUSA tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) boro hinka man two</td>
<td>(a) mutum biyu man two</td>
</tr>
<tr>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>(b) hinka boro</td>
<td>(b) biyu mutum</td>
</tr>
<tr>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>(c) other</td>
<td>(c) other</td>
</tr>
<tr>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

'two men'

Not surprisingly, there are no reversals of basic word order in this construction. In relative constructions, the situation is somewhat more complex since what is most commonly expressed in one language by a relative may be most commonly expressed in the other by a different construction. However, in no cases was the order N Rel reversed by non-native speakers.

In those construction which have different word order in the two languages, reversals are made by non-native speakers, but not nearly as frequently as one might expect. The first construction to be examined is the adposition phrase:

<table>
<thead>
<tr>
<th>% NN ZARMA tokens</th>
<th>% NN HAUSA tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) tabalo boŋ table on</td>
<td>(a) bisa tebur on table</td>
</tr>
<tr>
<td>91.0%</td>
<td>89.5% 'on the table'</td>
</tr>
<tr>
<td>(b) boŋ tabalo less than</td>
<td>(b) tebur bisa</td>
</tr>
<tr>
<td>0.5%</td>
<td>0.0%</td>
</tr>
<tr>
<td>(c) other</td>
<td>(c) other</td>
</tr>
<tr>
<td>8.5%</td>
<td>10.5%</td>
</tr>
</tbody>
</table>

The low percentage of reversals in adposition order is interesting in light of Lujan et al.'s findings. They found that the acquisition of prepositions was not problematic, suggesting that it represents "...an exception to the massive word order transfer simply because they [prepositions] correspond to Quechua postpositions, which are not words, but inflexional suffixes" [1983:3]. However, this argument could not be used to explain the case of Hausaphones non-problematic acquisition of postpositions in Zarma.

In attributive construction once again, Zarma and Hausa have contrasting orders. Hausa, however, has an alternative order as well.

<table>
<thead>
<tr>
<th>% NN ZARMA tokens</th>
<th>% NN HAUSA tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) mooto beeri car big</td>
<td>(a) baban mota big car</td>
</tr>
<tr>
<td>92.0%</td>
<td>79.0% 'a big car'</td>
</tr>
<tr>
<td>(b) beeri mooto</td>
<td>(b) mota babba</td>
</tr>
<tr>
<td>1.0%</td>
<td>18.0%</td>
</tr>
<tr>
<td>(c) other</td>
<td>(c) other</td>
</tr>
<tr>
<td>7.0%</td>
<td>3.0%</td>
</tr>
</tbody>
</table>

The relatively high percentage of NN HAUSA tokens using the N Attr order does not necessarily reflect interference or transfer of native Zarma patterns, since both orders are attested in Hausa. The N Attr order, however, is infrequently used by native speakers and has a special meaning.
Possessive constructions have been divided into two groups: those in which the element expressing the possessor is a pronominal form (Poss N), and those in which it is a nominal (Gen N). First, the constructions with pronominal forms:

<table>
<thead>
<tr>
<th>% NN ZARMA tokens</th>
<th>% NN HAUSA tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) ay cora</td>
<td>(a) aboki-na</td>
</tr>
<tr>
<td>1 sg. friend</td>
<td>friend-1 sg.</td>
</tr>
<tr>
<td>66.5%</td>
<td>76.0% 'my friend'</td>
</tr>
<tr>
<td>(b) cora ay</td>
<td>(b) na aboki</td>
</tr>
<tr>
<td>0.0%</td>
<td>0.5%</td>
</tr>
<tr>
<td>(c) other</td>
<td>(c) other</td>
</tr>
<tr>
<td>4.5%</td>
<td>6.0%</td>
</tr>
<tr>
<td>(d) poss element</td>
<td>(d) poss element</td>
</tr>
<tr>
<td>omitted</td>
<td>omitted</td>
</tr>
<tr>
<td>29.0%</td>
<td>17.5%</td>
</tr>
</tbody>
</table>

In the above, although there are few cases of reversal, there is a high percentage of tokens in which the possessive element is omitted altogether. This may have to do with pragmatic considerations; since the antecedent is known, omitting the pronoun does not greatly impede comprehension. What is most interesting here, however, is the difference in the figures for NN ZARMA and NN HAUSA. Not only are there substantially fewer cases of omission in NN HAUSA, but there are also substantially more cases of reproduction of the normal or basic word order. Whereas in Zarma the possessive element is an independent form, in Hausa, it is a syntactically bound form. This may explain why it is less frequently omitted.

When the possessive element is a nominal, the percentage of reversals is higher in both groups:

<table>
<thead>
<tr>
<th>% NN ZARMA tokens</th>
<th>% NN HAUSA tokens</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Mintu zaara</td>
<td>(a) zanen Mintu</td>
</tr>
<tr>
<td>Mintu &quot;pagne&quot;</td>
<td>&quot;pagne&quot; Mintu</td>
</tr>
<tr>
<td>79.0%</td>
<td>74.5% 'Mintu's</td>
</tr>
<tr>
<td></td>
<td>&quot;pagne&quot;</td>
</tr>
<tr>
<td>(b) zaara Mintu</td>
<td>(b) Mintu zanen</td>
</tr>
<tr>
<td>15.0%</td>
<td>13.5%</td>
</tr>
<tr>
<td>(c) other</td>
<td>(c) other</td>
</tr>
<tr>
<td>6.0%</td>
<td>12.0%</td>
</tr>
</tbody>
</table>

Of the 6% "other" in NN ZARMA, 2% used a doubling construction of the form Gen N Gen. In NN HAUSA, 8.8% of the tokens under "other" were of the type Gen N Poss, a sort of doubling in which the respondent used both nominal and pronominal forms. This construction seems to be the most difficult to master in terms of word order, native language order persisting in this construction more so than in others.

3. Summary

One of the most interesting things about these data is the fact that the percentages of 'faithful' word order reproduction as well as the percentages of reversed word order remain about the same in each group, as shown in the table on the next page:
Overall, basic word order patterns in these constructions appear to be quite faithfully reproduced by non-native speakers. One wonders why this should be the case. I would suggest that extra-linguistic considerations are important here. Respondents claim that the most important uses for their bilingual skills are in speaking with monolinguals and interpreting for them. When one is speaking with a monolingual, one cannot stray too far from standard patterns without risking a breakdown in communication. This is in contrast to other situations where the languages in contact have had a greater influence on one another.

REFERENCES


1. Introduction

The aim of this paper is to investigate the socio-linguistic features which distinguish Juba Arabic from Khartoum variety. A descriptive analysis involving a comparison of the linguistic features of Juba Arabic and Khartoum Arabic is presented on the basis of phonological, syntactic, and semantic data. The paper concludes with a number of observations on the future development of Juba Arabic in relation to Khartoum variety in the field of education in the region.

Juba Arabic (JA) is a variety of the Sudanese Colloquial Arabic spoken in Khartoum, the capital of the Sudan. Juba Arabic has been classified as a Pidgin or pidgin-creole [Nhial 1975, Mohmud 1983] respectively. It is used as lingua franca by the different ethnic groups of the Southern Sudan who reside in Juba. This paper investigates the socio-linguistic features which distinguish JA from Khartoum Arabic (KhA) phonologically, syntactically, and semantically. It concludes with a speculative statement that JA will in the far future give way to Khartoum variety given the factors of education and mobility of the people.

2. Phonology

2.1. Consonants. Arabic language in general and KhA in particular contains phonological characters such as the pharyngeals (emphatic sounds) (t, d, z, b, s), velarized (kh, gh), and laryngeal (ʔ) consonants used as in (1a-b).

\[(1)\]
\[
\begin{align*}
a. & \quad \text{t̬er} & \text{'bird'} \\
& \quad \text{ģeef} & \text{'visitor'} \\
& \quad \text{ 자리f} & \text{'envelope'} \\
& \quad \text{ģuwa} & \text{'he'} \\
& \quad \text{šabuun} & \text{'soap'} \\

b. & \quad \text{khamṣa} & \text{'five'} \\
& \quad \text{ģarib} & \text{'west'} \\
& \quad \text{šāma} & \text{'work'}
\end{align*}
\]

These characters are non-existent in the native languages, and they present great difficulty to the non-Arab speakers. As a result the speakers of JA substitute the obstruents t, d, k, g, s, and the glottal ʔ, for those Arabic characters as in (1c-d).

\[(1)\]
\[
\begin{align*}
c. & \quad \text{t̬er} & \text{'bird'} \\
& \quad \text{dif̪an} & \text{'visitor'} \\
& \quad \text{jar̪f} & \text{'envelope'} \\
& \quad \text{ūwa} & \text{'he'} \\
d. & \quad \text{kám̪̂sa} & \text{'five'} \\
& \quad \text{ģarib} & \text{'west'} \\
& \quad \text{šāma} & \text{'work'}
\end{align*}
\]
2.2. Stress. In KhA the rule of stress places stress on the heavy syllable (a syllable is heavy when it contains a long vowel or when the syllable is followed by two consonants) of a word beginning with the last, the penultimate, and the antepenultimate syllables respectively; otherwise, stress is placed on the first syllable of the word. For example, in the bisyllabic words ṣābuun and khāmsa, stress is placed according to where the heavy syllable of the word is. In ṣābuun stress is placed on the last heavy syllable while in khāmsa, stress is placed in the penultimate syllable which is the heavy syllable of the word. Otherwise stress is placed on the first syllable of the rest of the words such as ḡārif, ḥūwa, ḡārib.

However, JA substitutes tone for stress and so the stress-tone (') is placed on the words according to the tone pattern of the native languages. Thus in KhA the words thalātha 'three' and khāmsa 'five' each bear a single stress on the heavy syllable. The same words in JA, ṭālāṭā and kāmsa, bear two high tones each, with low tones on the unmarked syllables.

These phonological changes in JA indicate what Bokamba [1977] referred to as "reanalysis" of data resulting from languages in contact. In this case KhA is being reanalyzed by the JA speakers. These changes may be explained in terms of tendency toward simplification of complex structures, mother tongue interference, and socio-economic factors involving jobs, education and ethnicity.

3. Morphology

3.1. Gender. KhA strictly speaking distinguishes between masc/fem gender which is reflected in personal pronouns and case inflections. For example, the 2 and 3 persons singular inta/inti 'you (masc/fem)', ḥūwa/hiya 'he/she' are distinct forms. In each pair the feminine gender is represented by the morphemes -i, as in (int-i) and -iy-, as in (ḥ-iy-ā). In the pair ḥum/ḥunna 'they (masc/fem)', the feminine gender is marked by -nna as in (hu-nna).

In JA, however, all the fem. gender markers are lost and replaced by the masc. gender as shown in (2).

(2) ita 'you' (for both sexes) itakum 'you (pl)' (for both sexes)
    uwa 'he/she' umen 'they' (for both sexes)

3.2. Case. In KhA, the nominative, accusative, and possessive case forms are strictly speaking inflected word finally. For example, the nominative suffix forms -ta,-ti,-na,-tum,-tuna,-u, represent the 1, 2, 3 persons singular and plural respectively. These forms are inflected verb finally as in (3).

(3) a. (inta) akal-ta 'you (masc) ate'
    b. (inti) akal-ti 'you (fem) ate'
    c. (huna) akal-tuna 'they (fem) ate'

Likewise, in KhA, the accusative case forms -nā, -ak, -īk, -u, -ḥa, -nā, -kum, -ḥum, -ḥunna of 1, 2, 3 persons singular and plural respectively and are
inflected verb (transitive) finally as in (4).

(4) a. (hûwa) darâb-ni 'he hit me'
    b. (hûwa) dârab-ak 'he hit you'
    c. (hûwa) gârab-îk 'he hit your (fem)'

The words in parentheses are optional. In KhA they are commonly deleted.

Possessive cases are also inflected word finally in KhA. Thus the posses­
sive forms -î, -ak, -îk, -hu, -ha, -nâ, -kum, -un, -huna are inflected to the
noun 'pencil' as in (5).

(5) a. gâlam-i 'my pen'
    b. gâlam-ak 'your pen'
    c. gâlam-îk 'your (fem) pen'
    d. gâlam-un 'your (masc) pen' etc.

The case inflections in KhA do not exist in Juba Arabic. The inflection
forms are reanalyzed into independent nominative, accusative and possessive
pronouns ana, ita, uwa, anina, itakum, umen 'I, you, he/she, we, you (pl),
they' respectively. The sentences in (6) illustrate the nominative case struc-
tures.

(6) a. ãna akûlu 'I ate'
    b. ëta akûlu 'you ate'
    c. ûwa akûlu 'he ate'
    d. ûmen akûlu 'they ate' etc.

Thus the pronouns ana, ita, uwa, umen...appearing sentence initially in (6)
represent the nominative case in JA. Note that the verb has no agreement suf­
fix.

When the same pronouns follow a verb as in (7) below, then they stand for
the accusative case.

(7) a. ûwa dûgu ãna 'he hit me'
    b. ûwa dûgu ëta 'he hit you'
    c. ûwa dûgu ûman 'he hit them'
    d. ûwa dûgu ëtakum 'he hit you (pl)' etc.

JA has adopted the possessive word morpheme pîta 'of' from KhA and used
it to represent possessive case throughout as shown in (8).

(8) a. gûlam pîtái 'my pen'
    b. gûlam pîtak 'your pen'
c. gálam pítáu 'his pen'

d. gálam pítá umén 'their pen' etc.

This phenomenon is dealt with in Section (5).

In KhA, an adjective must agree with the noun it modifies in number and gender. Gender agreement is lost in JA.

(9)     |      KhA       |      JA        |
--------|----------------|---------------|
1.sg.   | ₁ana tawili    | ₁ana tówil    | 'I am tall'
2.sg.   |  žtawili -a     | ₁tawili       | 'you are tall'
3.sg.   | ḥuwa tawili    | ḥuwa tówil    | 'he is tall'
        | ḥuwa tawili -a |               | 'she is tall'
1.pl.   | nínna tuwáal   | nínna tówilín | 'we are tall'
2.pl.   | nínna tuwáal   | nínna tówilín | 'you are tall'
3.pl.   | ḥum tuwáal     | ḥum tówilín   | 'they are tall'

4. Syntax

4.1. Declaratives. Khartoum Variety maintains verb agreement in declarative sentences. Secondly, there is an optional rule in KhA that deletes the subject noun phrase at the beginning of the sentence so that (10a) is, for example, derived from (10b).

(10)    a. fátar-ta as saa atamánvia 'I breakfasted at eight o'clock'
        b. ₁ana fátar-ta as saa atamánvia = (10a)

Here the suffix -ta is the nominative form of ₁ana 'I' inflected to the verb fátar 'breakfasted' to mark the verb agreement. For the feminine it would be fátar-ti, with suffix morpheme -ti 'you (fem)' marking the gender, person, and number agreement. The nominal pronoun ₁ana in (10b) has been optionally deleted to yield (10a), the form that is widely used by the Khartoum speakers. This pattern applies to almost all the declaratives in KhA.

In JA the verb-agreement is lost and the nominative pronoun is maintained as in (10c).

(10)    c. ₁ana aqulu fátúr fi sá támánvia = (10a)

I ate breakfast in time eight

4.2. Interrogatives. Like the declaratives, interrogatives in KhA delete or omit the nominal pronouns, as in (11a-c).

(11)    a. saama si̇nu? 'what are you listening to?'

listening what
b. akal-ta sīnu? 'what did you eat?'
   ate what (masc)
c. akal-ī sīnu 'what did you eat?'
   ate what (fem)

The rule of Verb Agreement is also maintained in (1lb-c) just in the same way as it has been maintained in the declaratives, etc.

JA tends to be analytic throughout and maintains the subject pronouns without verb agreement as in (1ld-e).

(11) d. ṫa ḍe aṣūma sīnu? 'what are you hearing?'
   you marker hearing what
e. ṫa akūlu sīnu? 'what did you eat?'
   you ate what

The element ge functions as a present/future tense marker which is a creation of JA for it does not exist in KhA.

4.3. Negation. Even in negation, the aspects of Verb Agreement and inflection are observed in KhA as in (12a-c).

(12) a. ma ṭa-sma kāla-mu 'don't listen to his talk'
   not, you hear talk his
b. ma ṭa-sma -ī akāl-mu = (12a)
   not you hear (fem) talk his
c. ma akal-ta al fatūr 'I did not eat breakfast'
   not ate I the breakfast

Again Verb Agreement and nominal inflections are lost in JA as shown in (12d-e).

(12) d. matā aṣūma kalām pītāu = (12a)
   not listen talk his
e. āna ma akūlu fātur = (12c)
   I not ate breakfast

5. Vocabulary

Interestingly JA has borrowed substantially from the native languages in contact with Arabic language, particularly Bari language (spoken by Bari people among whom Juba City is located). The following are examples of borrowed words:

(13) kūrju (Bari) 'to dig or cultivate'
    korūpo (Bari) 'leaves (used in contexts like korūpo lubia 'bean leaves', korūpo shai 'tea leaves')
    tōŋga (Bari) 'to pinch'
    'dōŋŋa (Bari) 'back part of the head'
    gwangīrī (Bari) 'cheek'
Another semantic aspect of JA that is worth mentioning is the extraction of single words from the Arabic synonyms and their invariable usage. The words ruh/maashi, salaam/izaeyyk, kwes/zeyyid, etc. are pairs of synonyms which exist in KhA and are used according to shades of meanings and situations. JA picks up single members of those sets and uses them invariably as shown in (14a-b).

(14) a. íta gé ruwa weén? 'where are you going?'
    you marker go where

           b. ñna gé ruwa fi sük 'I am going to the market'
    I marker go to market

In KhA the word salaam 'peace be with you' is used for greetings in formal occasions such as funeral gatherings or political rallies or when entering the house of someone who is a non-acquaintance. The word izeyyek is the informal one for greeting friends, relatives, age mates, etc. JA picks up the word salaam and employs it in every situation.

6. Conclusion

The above descriptive analysis has shown and illustrated how JA differs from KhA on phonological, syntactic, and semantic levels. Children born in Juba today tend to pick up JA faster than they do their native languages [Ushari 1983]. It is speculated that due to influence of education where children read Arabic in schools, JA may in the distant future be replaced by the Khartoum variety.

REFERENCES


A FUNCTIONAL EXPLANATION FOR
THE NF-NP CONSTRUCTION IN YORUBÁ

Qrẹ Yusuf
University of California, Los Angeles

1. Introduction

There is sparse grammatical marking in Yorùbá. Word order usually carries the sole burden of differentiating grammatical relations. Of the few markers is \( nf \), which is homophonous with both the verb meaning 'to have' and the preposition that marks LOCATION, TIME, DIRECTION, or MANNER, but which is clearly different from the verbal \( nf \) and seemingly also from the preposition in that it often has no meaning, however vague. For instance, the meaning of \( nf \) is discernible in (1) and (2), but not in (3):

(1) \( mo ~ nf ~ owó \)
    I have money

(2) \( Adé ~ wà ~ nf ~ ilé \)
    be at house

(3) \( ọ ~ ya ~ mf ~ nf ~ asọ \)
    3sg. tear me (?!) cloth

Primarily because no meaning is available for \( nf \) in sentences like (3) Awobuluyi [1969, 1978] has hypothesised that the \( nf \) has been transformationally inserted, i.e. where the \( nf \) appears, the basic word order has been tampered with. Awobuluyi's work has been the only analysis that draws any serious attention to this morpheme. Other Yorùbá scholars have dismissed it rather casually, treating it variously as "verb", "preposition", or "particle" (see Awobuluyi [1969].

Despite the rigors with which Awobuluyi pursues the transformational analysis, the result has not been entirely satisfactory, and no new formulation has been proposed. This paper re-examines the phenomenon and comes to the conclusion that the \( nf \) is an Oblique case marker, not a transformational marker. Our analysis also takes care of the supposedly different \( nf \) as in (2). We should note that Madugu [1982] has analyzed \( nf \) with a similar result, but unlike our analysis, he assumes the correctness of the transformational analysis, and he treats adverbial \( nf \) as homophonous only with \( nf-NP \).

2. The Transformation Analysis

2.1. Awobuluyi's analysis. Awobuluyi [1969] has posited a tripartite source for the \( nf-NP \) construction, namely, the genitival, the adverbial, and the causative constructions, illustrated with (4), (5), and (6) respectively, with the (b) set as the hypothesised bases.

(4) a. \( Dàdà ~ jf ~ Ọjọ ~ nf ~ owó \)
    steal \( Ọjọ \)'s money
Studies in African Linguistics, Supplement 9

(4) b. Dada jí owó Ọjọ
   steal money
   'Dada stole Ojo's money'

(5) a. Dàda ja Ọjọ nif iyàn
   wage (? ) argument
   'Dada argued with Ojo'

b. Dàda bá Ọjọ ja iyàn
   with wage argument
   'Dàda argued with Ojo'

(6) a. oúnje náà kún Ọjọ nif inú
   food the fill (? ) stomach
   'the food gave Ojo indigestion'

b. oúnje náà mú inú kún Ọjọ
   food the cause stomach fill
   'the food gave Ọjọ indigestion'

2.2. Problems of the transformational analysis. There are a number of problems arising from the transformational analysis above.

a. The first argument against the analysis is that the nif construction is derived from heterogenous sources, amounting to a taxonomic listing of the suppos ed sources based on our ability to find semantic (or quasi-semantic) paraphrases. Here, one cannot objectively predict which structure any given nif-NP would be derived from.

b. The adverbial and causative sources are akin to serial verb constructions, and the syntax of the serial verb is far from settled. In other words, an analysis that uses a poorly understood phenomenon to solve another cannot be neat.

c. Apparently, the only plausible motivation for introducing nif by transformation is that it has no meaning. But there are other nif's that equally have no meaning (as we shall soon show) and yet the transformational insertion hypothesis is not extended to them.

b. Not only will some of the supposed "bases" not have corresponding nif-NP counterparts, examples could be multiplied where nif-NP sentences could not be traced to any other underlying sources. Thus, (7b) cannot possibly be related to (7a) and all of (7b-d) simply do not have non-nif counterparts:

(7) a. mo mọ ajá Adé
    I know dog
    'I know Adé's dog'

b. Mo mọ Adé nif ajá
    I know (?) dog
    'I know Adé as a dog'

c. mo nif owó nif owó
    I have money (?) hand
    'I have money'

d. olè fì ọdá sà Adé nif ogbè
    theft use machete cut (?) wound
    'the thief wounded Adé with a machete'

These shortcomings call for a reanalysis. We find that a more economical one would be the one that takes the nif-construction as basic too. As we shall soon show, our non-transformational account offers a unitary explanation for the
The nf-NP Construction in Yorùbá sentences supposedly derived from heterogeneous sources. We also observe that while the non-nf forms are grammatical, the nf types are definitely preferred, more commonly used, and more idiomatic. In fact, the adverbial and causative types are hardly heard, either in formal, or colloquial usage. In addition, our analysis would account for all non-verbal nf.

3. The Functional Explanation

The key to the explanation of nf is the fact that the "enigmatic" form appears only in bitransitive sentences, apparently functioning to keep the arguments apart. As the verb separates the subject and the object (Yorùbá is SVO), the distribution of nf is invariably between the object and the additional argument(s) as in (8). NPs are underlined for easy recognition:

(8) a. mo gba Adé nf iyanjú
   I take (?) trial
   'I encouraged Adé'

b. owó n je tàlákà nf iyà
   money ASP eat poor-man (?) punishment poor
   'money is a big bother for the poor'

Removing nf in these sentences automatically renders them ungrammatical.

3.1. Language universals. Returning to the issue of the basicness of the nf construction, we note that this view correlates in a most simple and economical manner with a number of independent language universals dealing with the syntax of body parts [Hyman 1977, Fox 1981], transitivity [Hopper and Thompson 1980] and Indirect Objects [Faltz 1977]. For example, Hyman [1977] has shown that grammars tend to code as nuclear terms the NPs referring to the possessor of a body part rather than the NPs referring to the body part. It is interesting that Hyman treats "possessor promotion" as a transformational process (like Awobuluyi does). For instance, Hyman says:

"Possessor promotion in Haya, as well as every Bantu language that I have studied transforms the possessor into a direct object if the verb is transitive" [emphasis mine] (p. 101).

The transformational process in possessor promotion is given as:

\[
\text{SUBJECT-VERB COMPLEX- X of Y} \quad \Longrightarrow \quad \text{SUBJECT-VERB COMPLEX-YX}
\]

However, what is more important for our study is the pragmatic explanation for the transformation. Hyman says that the transformation "is brought about in a situation in which the possessor is at the same time the experiencer of the action of the verb. The object created by possessor promotion experiences in other words is affected by the action of the verb" [emphasis Hyman's] (p. 104). The same conclusion is reached (though not by transformational means) by Fox [1981].

Hyman's grammar is interesting for another reason. He claims that the construction with Prepositional phrase is the basis of Possessor promotion; an exact opposite claim to Awobuluyi's:
It would appear that the conflicting claims derive from the fact that there are no objective criteria for positing one construction as the basis of the other.

In their discussion of the transitivity hypothesis, Hopper and Thompson [1980] give us an insight into which NPs would be marked object which coded as oblique. For instance, they suggest that the owner of the body is more "highly individuated" than the parts of the body and is more likely to be coded as direct object. Our examples of the sentences which contain body parts amply illustrate this. Witness (6a) and (10):

(10) ọ̀ yà mì nì ẹ̀nu ‘it surprised me’
3sg.open me (?) mouth

The same principle applies in POSSESSOR-POSSESSED relationship with the possessor being more highly individuated, i.e. more salient. Where the relationship of the NPs is not strictly Possessor-possessed or part-whole, we may note that the individuation hypothesis is not violated either. For instance in

(5a) Dàda ja Ọjọ́ nì iyàn ‘Dada argued with Ọjọ́’
wage (?) argument

(11) wọn fún Àsùnlé nì oyún ‘they made Àsùnlé pregnant’
they give (?) pregnancy

It is the more individuated personal NPs, Ọjọ́, Àsùnlé that get coded as direct objects (the first NPs in the sentence are uncontroversially subjects) and the less individuated NPs coded with nì .

It is equally interesting that the universals hold where prepositional phrases with adverbial "meaning" as in (12) are involved. Where three arguments are called for, the preposition codes the third arguments, and these could be seen to be less individuated.

(12) a. ọ̀ pè mì nì aíẹ̀ ‘he called me at night’
3sg.call me TIME night

b. Ọgá na Ọmọ̀ nì inàkùnà ‘the teacher flogged the child badly’
teacher flog child MANNER beating-bad-beating

It would be interesting to find instances of such a violation as where nì codes NPs other than the third.

3.2. Intransitive sentences. Finally, we look at intransitive sentences. Briefly, where the verb does not select an object, any NP following it will be coded as oblique. Sentences like (2) (p. 329) and (13) illustrate this point.

(13) obí pò nì ápò mì ‘there is plenty of colanut in my pocket’
colanut plenty in pocket my
Note that \( n^{f} \) marks the oblique NPs. In fact I would want to posit that the \( n^{f} \) variously glossed as 'in', 'at', 'on', 'as', etc. is equally empty of any semantic content. Any meaning attributed to it is nothing inherent but derives from our knowledge of the NPs which follow it. In short, there is only one non-verbal \( n^{f} \) which is an empty form that marks all oblique NPs. It is not a marker of a derived order.

4. Conclusion

In sum, we have shown that \( n^{f} \) functions to case mark any NP that cannot otherwise receive Nominative, Accusative, or any other case. Also, there is no need to distinguish more than one non-verbal \( n^{f} \). And although this idea was not explicitly developed in the paper, we want to note that the most recent formulations of the transformational syntax would support our case-assigning function of \( n^{f} \). And by way of a summary vindication we offer the following sentence:

\[
\text{(14) ógá fi ógbà na ọmọ iṣẹ rẹ \( n^{f} \) ìnàkùnà \( n^{f} \) ojú gbogbo master INSTR. belt flog child work his (?) badly (?) eye all}
\]

\[
\begin{align*}
\text{bèyàn \( n^{f} \) etí odù \( n^{f} \) ìnà people (?) ear river (?) yesterday}
\end{align*}
\]

'\text{the master flogged his apprentice badly with a belt in front of all the people at the river bank yesterday}'}

where \( ógá 'master' \) receives Nominative Case, \( ógbà 'belt' \), Instrumental, and the post verbal \( ọmọ'iṣẹ 'apprentice' \) Objective, and the other NPs were unexceptionally coded as Oblique, with \( n^{f} \). In short, \( n^{f} \) is a case marker, not a transformational marker.

[Our most sincere appreciation for Professor Sandy Thompson's guidance, encouragement, and comments on the longer work from which this paper is extracted. We are also grateful to Carol Lord, Lynell Marchese, and John Singler for their useful comments. Also, our thanks to Prof. Tim Stowell for the lengthy discussions on a possible transformational account of the topic. All errors remain mine.]
REFERENCES


