STUDIES IN AFRICAN LINGUISTICS

Volume 18, Number 1
April 1987

TABLE OF CONTENTS

GENDER AGREEMENT IN CHICHEWA*

Greville G. Corbett
University of Surrey

Alfred D. Mtenje
University College London

Gender in Chichewa is described as a complete system.
First the basic data on gender agreement are presented
and it is shown how the available agreement markers cor-
relate with the noun genders (and how the system has
changed in the recent past). There follows a discussion
of interesting phenomena which do not fit easily into the
main gender system. Next structures involving conjoined
noun phrases headed by nouns from various genders are
analysed in detail. The rules required to account for
the Chichewa system prove particularly complex; rules
proposed for other Bantu languages do not cover all the
Chichewa facts. The data are important for comparative
work within Bantu and for typological claims which go be-
yond.

0. Introduction

     In this paper we give an overall view of the gender system of Chichewa.
We distinguish between the genders into which nouns are divided and the agree-
ment markers used to agree with them, and we show that the relation between
them is not straightforward. We also cover phenomena on the fringe of the gen-
der system, such as nouns which do not fall completely into a single gender,

---

and we relate these data to more general typological claims. The most inter-
esting problem is that of gender agreement with conjoined noun phrases (gender
resolution). This question has been the subject of several studies on differ-
ent Bantu languages. Chichewa is particularly instructive and we relate our
findings both to work in Bantu and to relevant data outside.

The paper is structured as follows: after brief background notes (section
1), we outline the gender agreement system (section 2), then discuss problems
at the margin of the system (section 3), before turning to the major problem,
that of gender resolution (section 4).

1. Background

This study is essentially a description of the gender system of one native
speaker of Chichewa (the second author) as spoken in Malawi in Central Africa.
The dialect examined here corresponds fairly closely to the "Standard" dialect
of Malawian Chichewa; it is one of the dialects of what is called "Chinyanja"
in neighbouring countries like Zambia, Zimbabwe, and Mozambique. Since, for
socio-political as well as linguistic reasons, the term "Chichewa" is used on-
ly in Malawi, we will continue in this paper to employ this name to refer to
the Malawian dialect, as opposed to other terms (such as Chinyanja/Chichewa as
used by some other authors, including Mapanje [forthcoming]).

Chichewa belongs to the Bantu group of languages of Africa; it is classi-
fied by Guthrie as N31B. Like many other Bantu languages, Chichewa is a tone
language; lexical (and, in some cases, grammatical) contrasts may be signalled
by variations in pitch levels. There are two level tones in Chichewa, high
(H) and low (L). One may also find contour (gliding) tones which are obtained
only as a combination of the two level tones. Thus, a low tone followed by a
high tone on one syllable ($\widehat{LH}$) yields a rising tone while the reverse sequence
($\widehat{HL}$) gives a falling tone.

Since our interest in this study is in the system of gender in Chichewa,
we will ignore details of tone; examples will therefore be given without tone
markings. Besides, none of the arguments given crucially hinges on tone. A
detailed discussion and analysis of tone in Chichewa within an autosegmental
model is available in Mtenje [1986].

Another salient structural property which Chichewa shares with other Bantu languages is the division of nouns into genders, traditionally referred to as noun classes. Nouns can be assigned to genders in part on semantic grounds but phonological and morphological factors also play a part. It is this gender system which is the focus of our paper.

2. Basic Agreement Facts

We take as our starting point the clear account given by Watkins [1937]; data from Price [1943] have also proved useful, particularly as the basis for examples. First we present the forms acceptable to Mtenje within the descriptive scheme of Watkins, then we give the differences in detail between this idiolect and that described by Watkins; we also relate the description to the traditional Bantu schemes of noun classes. Watkins recognises ten classes (we shall call them genders) and a similar scheme is adopted by Mapanje [forthcoming]. We take agreement evidence to be crucial in setting up the genders. The following sentences illustrate the ten genders in turn, using subject agreement forms as a diagnostic; we label agreements simply as "ag":

(1) a. munthu a-ku-thamanga   'the person is running'
        person ag-pres-run

    b. anthu   a-ku-thamanga   'the people are running'
        people ag-pres-run

In this first gender, the subject agreement marker a is the same for singular and plural. However, object markers are distinct; (1c) and (1d) are continuations of (1a) and (1b) respectively:

    c. ndi-ku-mu-ona   'I see him/her'
        1st sg-pres-ag-see

    d. ndi-ku-wa-ona   'I see them'
        1st sg-pres-ag-see

The first gender is exceptional in having this coincidence of singular and plural for subject agreement, as well as in having considerable allomorphic variation for different agreeing elements. In the other genders, the singular and plural subject agreement forms are distinct, as in the following exam-

ples (from genders 2, 3 and 4):

(2) a. mudzi   u-ku-kula   'the village is growing'
        village ag-pres-grow

    b. midzi   i-ku-kula   'the villages are growing'
        villages ag-pres-grow

(3) a. tsamba li-ku-bvunda   'the leaf is rotting'
        leaf   ag-pres-rot

    b. masamba a-ku-bvunda   'the leaves are rotting'
        leaves ag-pres-rot

(4) a. ulendo   u-dza-tha   'the journey will end'
        journey ag-fut-end

    b. maulendo a-dza-tha   'the journeys will end'
        journeys ag-fut-end

In the fifth gender, number is again differentiated in the agreement markers, but not in the noun itself:

(5) a. njoka i-ku-gona   'the snake is lying flat'
        snake ag-pres-lie flat

    b. njoka zi-ku-gona   'the snakes are lying flat'
        snakes ag-pres-lie flat

In the sixth gender, number is typically marked on noun and verb in the same way:

(6) a. chipatso chi-ku-bvunda   'the fruit is rotting'
        fruit   ag-pres-rot

    b. zipatso zi-ku-bvunda   'the fruits are rotting'
        fruits ag-pres-rot

   This phenomenon, sometimes termed "alliterative concord", is often treated as the norm in Bantu languages, even a defining feature. When we view the Chichewa gender system as a whole, however, we see that while there is often some relation between the morphology of nouns and the agreements they take, the two by no means coincide.[1] (The situation is not so different from Indo-

---

[1]Semantic factors clearly have a role in that most nouns denoting humans

European languages like Latin or Russian.)   The seventh gender also shows al-
literative concord. It is a special gender in that it comprises mainly nouns
formed from nouns of other genders. Original prefixes are retained, and a new
prefix is added to give diminutive meaning. Thus besides the noun   mwana/ana
'child/children' (gender 1), we find kamwana/tiana 'infant/infants':

(7) a. kamwana ka-li bwino   'the infant is well'
        infant ag-be in good order

    b. tiana   ti-li bwino   'the infants are well'
        infants ag-be in good order

The remaining three genders are all to be distinguished from the genders
we have discussed so far. They are the so-called locative genders. They each
have one agreement form, not a singular-plural pair like the other genders.
And there are very few nouns which are restricted to these genders. Gender 8
is used to show position (roughly 'at/on'):

(8) panyumba pa-ku-tentha   'it is hot at the house'
     at house ag-pres-hot

Nyumba 'house' is an ordinary gender 5 noun. There are a few nouns, often
body parts, which typically occur in this gender, such as   pakhosi 'neck' and
pakamwa 'mouth'; pa is felt to be the normal prefix and similarly pa the
normal agreement marker. It is not, however, possible to add a second pa to
ensure locative meaning (*papakhosi) .

     Gender 9 is for location within:

(9) munyumba mu-ku-tentha   'it is hot in the house'
     in house ag-pres-hot

The final gender, number 10, can designate more general location:

---

belong to gender 1. But phonological and morphological factors also play a
part in gender assignment. For instance, it is generally the case that nouns
whose initial prefixes have the palatal affricate ([č]) belong to gender 6.
Similarly, most nominals with initial nasals belong to gender 5, just as most
nouns beginning with aspirated plosives are members of gender 3.

(10)   kudambo   ku-ku-nunkha   'it smells about the marsh'
       about marsh ag-pres-smell

The tenth gender also includes verbal nouns:

(11)   kuthamanga ku-ma-pweteka   'running hurts'
       running   ag-habitual-hurt

These last three genders are clearly different from the others and there might be some question as to whether they should be called genders at all. For the present, our purpose is to record the forms. For a convincing account of the reasons for special treatment of these genders in ChiBemba see Givón [1972:12-13, 28-34]; for a discussion of locatives in OluLuyia, see Dalgish [1976].

We now bring together the data on Watkins' ten classes (as illustrated in the sentences above) and relate them to the traditional Meinhof numbering scheme. We thus follow Doke's recommendation [1935:64] of grouping singular and plurals together for studying an individual language, while also referring to the Meinhof numbering.

Table 1:   Subject concord markers in Chichewa

Watkins class	Common Bantu class	agreement markers singular	plural
1	1/2	a	a
2	3/4	u	í
3	5+11/6	lí	a
4	14/6	u	a
5	9/10	í	zi
6	7/8	chi	zi
7	12/13	ka	tí
8	16	pa	
9	18	mu	
10	15+17	ku	

The Watkins numbers are straightforward; the concords are illustrated in sentences (1)-(10), gender 10 also being illustrated in sentence (11). The tra-

ditional Bantuist numbering separates singular and plural, hence Watkins'
class 1 would be class 1 singular and class 2 plural. This scheme has advan-
tages for historical and comparative work. Nouns of class 11 (singular) ap-
pear to have joined class 5, and the two functions of Watkins' class 10 were
originally fulfilled by two different classes.

In Table 1 we have given only subject agreement markers, rather than a
large table with all agreement markers listed separately (for a particularly
impressive example of this genre see Laws [1885]). The point is that the dif-
ferent agreement markers (for different agreeing elements) almost all involve
automatic alternations,[2] and listing these variants would obscure rather than
clarify the issues we wish to address. The gender where alternations tend to
be irregular is the first, as we saw in sentences (1c) and (1d). Besides the
main allomorph a and the allomorph mu for object agreement, there is anoth-
er allomorph u found, for example, with certain pronouns and with the per-
fect tense.

We now turn to the differences between Mtenje's idiolect and that de-
scribed by Watkins, which are as follows:

1. Watkins gives βa (orthographic ŵa ; ŵ is an unrounded bilabial glide) as
   the plural for genders 1, 3, and 4. Mtenje has ŵa as a more marked
   alternative to a in all three cases.

2. Watkins has bu as an alternative to u for singular agreement with
   gender 4. He states that bu is used more frequently by older people
   [1937:34]. This alternative is not used by Mtenje, but only u .

3. For singular object agreement with gender 4, besides the alternative
   bu as for subject concord, Watkins gives only u , while for singular
   object concord with gender 2 he gives (w)u . In fact the appearance
   or not of w is a phonetic problem; there is variability as to when
   w is possible, but u and wu cannot constitute a minimal pair.
   Hence there are no grounds for differentiating between the singular
   concords of the second and fourth genders, since when w is possible
   in one it will also be possible in the other.

---

[2]The morphological processes involved are described with relation to spe-
cific examples in footnotes 4, 9, and 10 below.

4. For the plural concord of gender 6, Watkins has vì while Mtenje has
   zì . Vì is maintained in some dialects, particularly in the northern
   part of the Chichewa speaking area, but for Mtenje the plural concords
   of genders 5 and 6 are identical.

5. For the plural concord of gender 7 Watkins has tu while Mtenje uses
   tì .

It is also worth recording a difference as compared to the table at the end of
Price [1943], which links to point 1 above. For the perfect tense prefix for
the plural of gender 1, Price gives wa , while for the plurals of 3 and 4, he
gives a . This is probably just an error. For Mtenje the unmarked form for
all three is a , while ŵa is an alternative; wa is not possible for any of
the three.

    If we now look back to Table 1, we can claim that the three plural markers
a are indeed identical. Their allomorphs for the different agreeing elements
have identical distribution (recall, however, that this is not the case for
the singular a). Similarly the two singular markers u are identical, as
are the plural markers zì . We can therefore redraw Table 1 to give a more
accurate picture of the gender system of Chichewa, by including agreement
markers (or target gender forms) once only, and by representing sets of nouns
taking identical agreements (controller genders) by lines drawn linking the
relevant agreements:

Table 2: The gender system of Chichewa
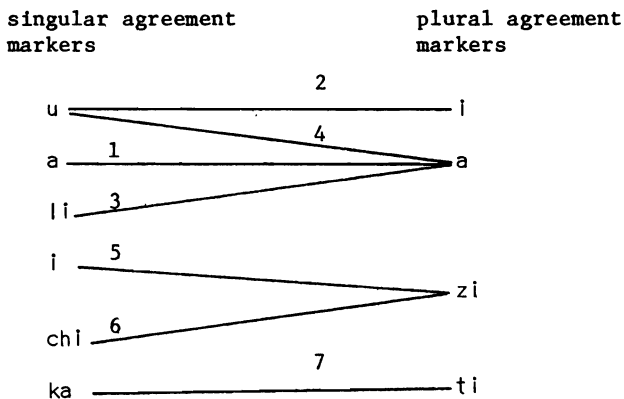
The numbers of course represent the genders according to Watkins' scheme. It is interesting to note that there are seven genders (we omit genders 8-10 here since they are outside the singular-plural opposition), though there are not seven distinct agreement forms in either number. There are in fact six singular forms and four plural forms. This pattern is in accord with Greenberg's [1963:112] Universal number 37: "A language never has more gender categories in nonsingular numbers than in the singular." The relation between the singular and plural forms is also of interest; given the concord taken by a noun in the singular, one cannot determine unambiguously the form it takes in the plural; similarly, given the plural, the singular cannot be uniquely determined. Such systems have been termed "crossed systems" [Heine 1982:197]. The appearance of this complex system is quite recent; as long as the singular agreements of gender 4 were distinct from those of gender 2, the system was not crossed. Given the singular agreement form, the plural could be predicted (though not vice versa). Heine calls systems like this earlier one "paired", though we would prefer the term "convergent". If we postulate an even earlier system in which the Common Bantu classes given in Table 1 had distinct agreement forms, then the system was even simpler in the sense there was a clearer matching of singular and plural agreement forms (though conversely it was more complex in the sense that there were more actual agreement markers).

3. Marginal Gender Phenomena

Having described the core of the gender system, we now turn to two problems: the first concerns nouns whose semantics and morphology are in conflict and the second is the problem of agreement in gender when there is no noun as head of the noun phrase controlling the agreement (neutral agreement).

3.1. Semantic agreement. The vast majority of nouns belong to a particular gender and consistently take all the expected agreements. However, as is normally the case in gender systems, there is a small number of nouns which are not completely consistent. These are nouns for which the different principles of gender assignment come into conflict.[3] There is a semantic principle ac-

---

[3]There are also instances in which nouns do not take the expected number

cording to which nouns denoting humans are in gender 1. And, as mentioned
earlier, there is a morphological principle which determines the gender of a
noun according to the prefixes it takes. Nouns which take no prefix in either
singular or plural (and which therefore do not differentiate the numbers) nor-
mally belong in gender 5. There are a few nouns which denote humans (and so
would be expected to be in gender 1), but which have the morphological form of
gender 5. In Chichewa these normally take gender 5 agreements. Similarly,
there are nouns denoting humans with the prefixes chi-/zi- (the morphologi-
cal form of gender 6) which normally take gender 6 agreements; in addition
there are diminutives formed from nouns denoting humans which normally take
the agreements of gender 7.

---

agreement. Singular nouns denoting humans may take plural agreement, which
indicates respect:

(i) bambo a-ku-yenda   'father is walking'
     father ag-pres-walk

Bambo belongs to the first gender and so the a is ambiguous; object agree-
ment (wa) is unambiguously plural:

(ii) ndi-ku-wa-ona   'I see him' (literally 'them')
      1st sg-pres-ag-see

The singular object marker mu would be grossly impolite. In fact all agree-
ments will be plural:

(iii) bambo anga (< a + a + nga)   'my father'
       father my

Again the singular wanga (< u + a + nga) would be inappropriate. Of course,
the use of the plural for politeness is a widespread phenomenon in Bantu, in
Dravidian languages, and in Indo-European. What is particularly interesting
here is that plural agreement is usual for all agreement targets, including
the nominal predicate:

(iv) bambo ndi aphunzitsi   'father is a teacher'
      father is teachers

The use of the singular (mphunzitsi) would be a bald statement of fact,
while (iv) identifies with father and is polite. The singular in (iv) is,
however, less bad than in (iii). Note that the copula ndi does not inflect
for number. The interest of (iv) is the plural nominal predicate as a polite-
ness marker; the nominal predicate thus shows syntactic agreement with the
subject, a phenomenon which Comrie [1975:410-412] in his survey of different
predicate types, found extremely rare.

Let us begin with the diminutives. Kamwana is a small child or infant. It takes agreements of gender 7:

(12)   kamwana   ko-kongola ka-ku-gona
        small child ag-pretty ag-pres-sleep

        'the pretty small child is sleeping'[4]

Gender 1 agreements are ungrammatical:

(13)   *kamwana   a-kongola a-ku-gona
         small child ag-pretty ag-pres-sleep

         'the pretty small child is sleeping'

Similarly with the relative pronoun:

(14)   kamwana   ka-mene ka-ku-gona
        *kamwana   a-mene a-ku-gona
        small child ag-who ag-pres-sleep

        'the small child who is sleeping'

Subject pronouns are normally dropped; however, the form of emphatic pronouns, and of the subject agreement marker when no pronoun is included, are both normally of gender 7. Yet gender 1 agreements are also possible when sufficiently separated from the noun:

---

[4]The form ko on kokongola is derived as follows. The initial stage is the attachment of the infinitive prefix ku (which is justified by the fact that it appears on the surface in the case of monosyllabic roots). Then the adjectival prefix a and the agreement marker ka are attached:

    ka   +   a   +   ku   +   kongola
    ag   adj   infin   verb root
                prefix

Vowel coalescence reduces the identical vowel sequence a + a to a and the resulting structure ka + ku undergoes a regular morpheme fusion rule which changes a + ku into o if the verb is not monosyllabic, yielding the required surface form kokongola . For details of the (morpho-)phonological rules referred to here, see Mtenje [1986; forthcoming]. Note that, for simplicity, we do not separate the final vowel from the stem throughout.

(15)   kamwana   ka-mene ka-ma-gona   mu-nyumba umu ka-mene ka-ma-pita
       small child ag-who ag-habit-sleep in house this ag-who ag-habit-go

                                     ⎧ka ⎫   ⎧iko(ko)⎫
       ku sukulu ku London, mai   ake a-ma- ⎨mu ⎬ -konda ⎨iye(ko)⎬
       to school in London, mother his ag-habit-ag-love   him ⎩   ⎭

       'the small boy who sleeps in this house who goes to school in London,
       his mother loves him'

In (15), ka and iko (which has the optional extension ko) show agreement
as for gender 7, while mu and iye(yo) are gender 1 forms. A similar effect
can be observed in the plural (we change one of the nouns to avoid ambiguity
in the agreement markers):

(16)   tiana   ti-mene ti-ma-gona   mu-nyumba umu ti-mene ti-ma-pita
       small children ag-who ag-habit-sleep in-house this ag-who ag-habit-go

                                          ⎧ti ⎫   ⎧ito(to)⎫
       ku sukulu ku London, galu wa-wo   a-ma- ⎨wa ⎬ -konda ⎨iwo(wo)⎬
       to school in London, dog ag-their ag-habit-ag-love   them ⎩   ⎭

       'the small children who sleep in this house who go to school in London,
       their dog loves them'

Again the emphatic pronoun and the object agreement marker can be of gender 7
(syntactic) or of gender 1 (semantic), provided they are sufficiently far from
the noun.

     When we substitute the word chitsilu 'fool' (morphologically gender 6),
then the switch to semantic agreement is not possible in an example similar to
(15):

(17)   chitsilu chi-mene chi-ma-gona   mu-nyumba umu chi-mene chi-ma-pita
       fool   ag-who   ag-habit-sleep in-house this ag-who ag-habit-go

                                          ⎧*mu ⎫   ⎧*iye(yo) ⎫
       ku sukulu ku London, galu wa-ke   a-ma- ⎨chi ⎬ -konda ⎨icho(cho)⎬
       to school in London, dog ag-his ag-habit-ag-love   him ⎩   ⎭

       'the fool who sleeps in this house who goes to school in London,
       his dog loves him'

However, when the pronoun is yet further removed from the noun, semantic agree-
ment becomes a possible alternative:

(18) chitsilu chi-mene chi-ma-gona   mu-nyumba umu chi-mene chi-ma-pita
      fool   ag-who   ag-habit-sleep in-house this ag-who   ag-habit-go

      ku sukulu ku London, galu wa-ke a-ma-chi-konda   ndipo mai   a-ke
      to school in London, dog ag-his ag-habit-ag-love and   mother ag-his

                                     ⎧mu ⎫   ⎧iye(yo) ⎫ ⎧a ⎫
      a-ma-bvomera-di   kuti a-na-⎨chi⎬ -ona ⎨icho(cho)⎬ ⎨chi⎬ -ku-yenda
      ag-habit-agree-indeed that ag-past-ag-see   him   ag-pres-walk

      ndi galu yo
      with dog that

      'the fool who sleeps in this house who goes to school in London,
      his dog loves him and indeed his mother agrees that she saw him
      walking his dog'

Not surprisingly, the switch to semantic agreement in longer sentences is also
possible with kamwana 'infant'. With a noun like nkhalamba 'old person',
the switch from gender 5 to gender 1 is acceptable in sentences like (18), and
is marginally possible in sentences like (15), but the result in the latter is
less good than with kamwana . Thus the switch is easiest with kamwana (gen-
der 7), then with nkhalamba (gender 5), and most difficult with chitsilu
(gender 6).

     These examples provide interesting support for the claim of the Agreement
Hierarchy. This hierarchy, for which see Corbett [1979, 1983a:8-41, forth-
coming] and Cornish [1986:203-213], consists of four basic positions:

     attributive modifier < predicate < relative < personal
                                            pronoun   pronoun

The claim made is as follows:

     For any controller that permits alternative agreement forms, as we move
     rightwards along the Agreement Hierarchy, the likelihood of agreement
     forms with greater semantic justification will increase monotonically.

The Chichewa data support this claim. Though semantic agreement (gender 1
agreement with nouns in other genders which denote humans) is only a marginal
phenomenon in Chichewa, the one position where it is possible is in the (em-
phatic) personal pronoun (and of course in agreements dependent on it or on a
dropped pronoun). Though the hierarchy was postulated on the basis of evi-
dence from a range of languages, most of the data were from Indo-European;

thus data from Bantu are of special interest. Furthermore, it was claimed
that within any one position on the hierarchy, the choice of agreement form
may be influenced by "real" distance, the degree of separation of the target
from its controller measured in words. And indeed, we have seen convincing
evidence that semantic agreement becomes more likely the further the pronoun
is separated from the controlling noun.

3.2. <u>Neutral agreement</u>. All gender systems require a strategy for agreement
with elements which do not carry gender and number features in the normal way.
Such elements vary from infinitives to interjections. There are two strate-
gies: either one of the normal agreement forms is used (often the third singu-
lar neuter), or there may be a special form solely for this purpose [Corbett
1980]; even when a normal form is used it can usually be distinguished by some
syntactic means.

In Chichewa, the situation is particularly interesting in that the equiva-
lent of the infinitive is marked by a special prefix ku and also takes ku
as an agreement marker, as illustrated in example (11) above; there is a spe-
cial gender (gender 10) which infinitives share with locatives. There is
still the question of interjections. Consider the following example:

(19) 'aaa' a-na-mv-eka   'an "aaah" was heard'
      'aaa' ag-past-hear-passive

We must ask what form of agreement the a is. It is not a simple allitera-
tive agreement as the next examples prove:

(20) 'ooo' a-na-mv-eka   'an "oooh" was heard'
      'ooo' ag-past-hear-passive

(21) 'mayo'   a-na-mv-eka   'a crying sound was heard'
      crying-sound ag-past-hear-passive

To determine which a marker is involved we should consider object agreement
(cf. (1c) and (1d)):

(22) a-na-mu-mva   'mayo'   'he heard a crying sound'
      ag-past-ag-hear crying sound

In this example the initial a represents subject agreement; mu is the form

of interest; it is the object agreement marker which agrees with, or, more ac-
curately, shows the failure of agreement with, the element in object position,
mayo . Note that in all these examples the form ku , which is found with in-
finitives, is unacceptable. The last example demonstrates that the acceptable
form is in fact the singular of gender 1; this is surprising, since the first
gender is associated with humans (though not exclusively). The syntactic ef-
fect with such neutral agreement mentioned earlier, is that conjoining forms
which take neutral agreement usually does not produce a plural. Indeed, the
same effect is observed in Chichewa:

(23) 'mayo'   ndi 'aaa' wa-ke ti-na-mu-mva   kutali
      crying-sound and 'aaah' ag-his ag-past-ag-hear far away

      'we heard his crying sound and "aaah" far away'

The wa (u+a) of wake is a gender 1 singular marker; ti is first plural
subjective concord; the mu is again gender 1 singular objective concord.
Thus elements which are outside the gender system take gender 1 singular
agreement as the neutral agreement form; this remains the only acceptable
form even when such elements are conjoined. Chichewa (and, we assume, various
other Bantu languages) is interesting typologically in this respect, since
there is a special agreement form for infinitives, but this form does not
function as the neutral agreement form.[5] Whereas, from comparisons beyond
Bantu, we might have expected elements outside the gender system to take the
same agreements as infinitives, the role of neutral agreement marker is, rath-
er surprisingly, filled by the marker for the singular of gender 1.

---

[5]This construction should be distinguished from that in which there is an
implied locative such that the noun phrases present in surface structure do
not control verb agreement, as in the following:

(i) ku-li mphepo ndi mitambo   'there is wind and clouds'
     ag-be wind   and clouds

Ku is the gender 10 locative marker, indicating general location. In suita-
ble contexts, similar examples with pa and mu are also possible. The
word order is fixed as in (i); noun phrases which precede the verb must con-
trol subject-verb agreement.

4. <u>Gender Resolution</u>

Gender resolution is one of the most interesting areas of Bantu syntax; given the relatively high number of genders it is natural to ask what happens in conjoined structures. Provided agreement is with all the conjuncts, then rules are required to specify both number and gender. The number rule is straightforward (there is no complication of the dual number, for example): conjoined noun phrases take plural agreements. (However, a restriction will be imposed on this rule at the end of section 4.1.)[6]

Gender is, of course, much more complex. Interest in this area was stimulated by the important work of Givón [1970; 1972:80-93]. Givón coined the term "gender resolution", though it is important to realise that gender reso-

---

[6]Note too that resolution does not always apply; the usual alternative is for agreement to be controlled by the nearest conjunct (see Brauner [1979: 425], Corbett [1983b:179-182], and Bokamba [1985:44-45]). Resolution or lack of it is an instance of semantic versus syntactic agreement, resolution representing semantic agreement. Again Chichewa provides data to support the Agreement Hierarchy (cf. section 3.1). In attributive position, resolution is not possible. Both conjuncts must take identical modifiers:

(i)   lalanje la-bwino ndi tsamba la-bwino   'a good orange and a good leaf'
      orange ag-good and leaf   ag-good

The adjectives (labwino < li + a + bwino) are singular. If either is omitted then the referent of the noun phrase without the adjective is not understood to be "good". The more important point for our purposes is that the plural (resolved form) is ungrammatical:

(ii) *lalanje ndi tsamba za-bwino   'a good orange and leaf'
       orange and leaf   ag-good

Placing the plural adjective after lalanje is equally bad, as is the use of the gender 3 plural form a-bwino in place of za-bwino . However, in all other positions on the Agreement Hierarchy, plural forms are normal:

(iii) lalanje ndi tsamba zi-li apo   'the orange and leaf are there'
      orange and leaf   ag-be there

(iv) Lalanje ndi tsamba zi-mene zi-li apo   za-gul-idwa   ku msika.
      orange and leaf   ag-which ag-be there ag-buy-passive at market

      Izi   ndi zakudya   za-ana.
      pronoun-ag be   plural-food   ag-children

      'The orange and leaf which are there were bought at the market.
       They are foodstuffs for children.'

lution rules may operate even when there is no "clash"; as we shall see below, if all conjuncts are headed by nouns of the same gender it does not follow that agreement will be in that gender. Givón's work has led to several publications on the problem, and we too will use his account as the framework for our analysis. Givón analyses Luganda and ChiBemba. Voeltz [1971] considers problematic data in Xhosa, where gender resolution is more restricted than in Luganda and ChiBemba. Roberts and Wolontis [1974] modify and extend the rules proposed by Voeltz, and also widen the range of languages to include Tswana and Zulu. Brauner [1979] gives examples of gender resolution (mainly for non-human conjuncts but also cases of human and non-human conjoined) from written Swahili sources. The brief description of gender resolution in Mateene [1980: 332-333] suggests that Nyanga is broadly similar to ChiBemba. Givón's data were considered in a general typology of resolution rules in Corbett [1983b]. Finally Bokamba [1985] extends the languages considered to include Dzamba, Likila, and Lingala.

Given the different patterns of gender resolution found in the languages investigated to date there is clearly some challenging comparative and typological work to be done. As preparation for this we require data on as many different languages as possible. Chichewa appears to be an interesting case, being more restrictive than ChiBemba but less so than Xhosa. In some cases the judgements are not clear-cut; various combinations cause problems. And there is variation between speakers in the acceptability of some sentences. In view of the need for comparative data, we have concentrated on trying to describe the system of one speaker (Mtenje) as accurately as possible. It is perhaps worth pointing out that for those examples for which we have the judgements of three speakers, Mtenje's judgement is "average", in that he usually accepts more examples involving gender resolution than one of our three main informants and fewer than the third. In this description we include

---

izi is an emphatic personal pronoun, consisting of i (root) + zi (plural agreement marker) + i (proximate marker); $i + i$ becomes i by vowel coalescence. Thus, semantic (resolved, plural) agreement is found in the predicate and in the positions to the right on the Agreement Hierarchy (the relative and personal pronouns), but not in attributive position. This is fully in accord with its predictions.

problems which are often disregarded in accounts of gender, such as the agree-
ments used with conjoined infinitive phrases.

We begin with Givón's scheme of rules, which may be reformulated as fol-
lows:

1. If all the conjuncts are semantically [+human], the plural of gen-
    der 1 (a in Chichewa) will be used.

2. If none of the conjuncts is semantically [+human], the plural of
    gender 6 (zi in Chichewa) will be used.

3. If the conjuncts are semantically mixed, the comitative construc-
    tion is preferable; if gender resolution is forced, the form will
    be as in rule 2.

It is worth pointing out that similar rules (but without any mention of diffi-
culty in conjoining animates with inanimates) were proposed by Horton [1949:
189] for Luvale. The important thing about these rules is that they have a
semantic basis: it is the noun's meaning rather than its grammatical gender
which counts. For ease of exposition we shall consider in turn examples in
which none of the conjuncts denotes a human (section 4.1), those where only
humans are involved (section 4.2), and those in which the conjuncts are seman-
tically mixed (section 4.3); finally we summarise the rules for gender resolu-
tion (section 4.4).

4.1. <u>No conjuncts denoting humans</u>. When none of the conjuncts denotes a hu-
man, we would expect the agreement marker to be zi (traditional class 8).
Since this form serves as the plural of gender 5 as well as of gender 6 in
Chichewa, it would be reasonable to suppose that it would be, if anything,
more generally used for gender resolution in Chichewa than in ChiBemba. How-
ever, though zi is indeed the main form for agreement with conjoined noun
phrases denoting non-humans, there are several complications, as we shall see
below. Let·us consider first the straightforward cases:

(24)   ukonde ndi chipatso zi-ku-bvunda
        4 sg   6 sg   ag-pres-rot
        net   and fruit   are rotting

(25)   mtengo ndi masamba zi-ku-kula
        2 sg   3 pl   ag-pres-grow
        tree   and leaves are growing

(26)   malalanje ndi zikuni   zi-ku-bvunda
       3 pl   6 pl   ag-pres-rot
       oranges   and pieces of wood are rotting

In each of these there is a clash of genders. But this is certainly not a re-
quirement for gender resolution, resulting in the use of zi . In the next
example we have two nouns from the second gender, but zi is still the re-
quired form:

(27)   mpeni ndi mphika zi-ku-sowa
       2 sg   2 sg   ag-pres-missing
       knife and pot   are missing

The plural of gender 2 (*ikusowa) is quite unacceptable. Similarly with
two nouns from gender 3:

(28)   lalanje ndi tsamba zi-ku-bvunda
       3 sg   3 sg   ag-pres-rot
       orange and leaf   are rotting

The plural of gender 3 (*akubvunda) is not possible. There are thus good
grounds for the claim that non-human conjuncts require the zi form. Not all
combinations are so readily accepted. There was some hesitation over the fol-
lowing, but zi was the only possible form:

(29)   munda, ng'ombe   ndi khasu zi-li uko
       2 sg   5 sg (or pl)   3 sg ag-be
       garden, cow   and hoe   are   there

There was also considerable uncertainty when a diminutive was included, wheth-
er singular or plural:

(30) (?)   ukonde ndi kakhasu   zi-li uko
           4 sg   7 sg   ag-be
           net   and little hoe are   there

(31) (?)   ukonde ndi timakasu   zi-li uko
           4 sg   7 pl   ag-be
           net   and little hoes are   there

Though these combinations were less readily accepted than those given earlier,
zi was still the only possible form. We suggest the following reason for
the difficulty with (30) and (31). The overall scheme of the resolution rules

is semantic (based on the feature [+/-human]). However, gender 7 has the
closest correlation to semantics, its members being diminutives. It is the
ignoring of this semantic feature in the agreement in (30) and (31) which
gives the problem.

      So far we have considered combinations including at least one singular
conjunct. When all conjuncts are plural, the picture is more complex:

(32)   mipeni ndi miphika $\left\{ \begin{array}{l} i \\ zi \end{array} \right\}$ -ku-sowa

       2 pl   2 pl   ag-pres-missing
       knives and pots   are missing

Both i , the agreement form for the plural of gender 2, and the zi form
are acceptable. With gender 3, however, a , the agreement form for the plural
of 3, seems normal, while zi is less good:

(33)   malalanje ndi masamba $\left\{ \begin{array}{l} a \\ ?zi \end{array} \right\}$ -ku-bvunda

       3 pl   3 pl   ag-pres-rot
       oranges   and leaves   are rotting

With gender 4, zi is less good again:

(34)   maukonde ndi mauṭa $\left\{ \begin{array}{l} a \\ ??zi \end{array} \right\}$ -li uko

       4 pl   4 pl   ag-be
       nets   and bows   are   there

There are two hypotheses which would fit these data:

      1. If plural nouns of the same gender are conjoined they take the
          plural agreement form for that gender.

This hypothesis is based on the gender of the nouns (the controller gender);
it covers, for example, a sentence in which two gender 3 nouns are conjoined.
The second hypothesis refers not to the gender of the nouns but to the agree-
ment form (the target gender):

      2. If plural nouns which would take the same plural agreement form are
          conjoined, then that form will be used.

(In either case zi- will be an alternative with varying degrees of accepta-
bility.) To choose between the two hypotheses, we conjoin plural nouns of

genders 3 and 4, which both have a as the plural marker. The second hypo-
thesis predicts that a will be used, while the first makes no prediction
(and so we would expect the regular form zi):

(35)   malalanje ndi maukonde $\left\{\begin{matrix} a \\ *zi \end{matrix}\right\}$-li uko

     3 pl   4 pl   ag-be
     oranges   and nets   are   there

This demonstrates that the second hypothesis is correct: when plural nouns
which take the same agreement form (target gender form) are conjoined, that
form will be used (zi may be an alternative). We can confirm this view by
considering non-humans in gender 1. First an example with two nouns of that
type:

(36)   amphaka ndi agalu $\left\{\begin{matrix} a \\ *zi \end{matrix}\right\}$-ku-thamanga

     1 pl   1 pl   ag-pres-run
     cats   and dogs   are running

Then a non-human from gender 1 together with a plural noun from gender 3
(which would also take a):

(37)   amphaka ndi malalanje $\left\{\begin{matrix} a \\ *zi \end{matrix}\right\}$-li uko

     1 pl   3 pl   ag-be
     cats   and oranges   are   there

Note that when animate class 1 nouns are conjoined, as in (36), *zi was not
accepted. With the few inanimates, zi was not excluded:

(38)   akatundu   ndi akabudula $\left\{\begin{matrix} a \\ ?zi \end{matrix}\right\}$-ku-sowa

     1 pl   1 pl   ag-pres-missing
     the pieces and the pairs   are missing
     of baggage   of shorts

Gender 5 has the added complication that the noun does not mark number. For
this reason we include modifiers. Not surprisingly, all possibilities, singu-
lar and plural, take zi , since the general rule and the special rule for
plurals both predict the zi form:

(39) mvuu ĺ-modzĺ ndĺ njobvu ĺ-modzĺ zĺ-ku-meny-ana
      5 sg ag-one   5 sg   ag-one ag-pres-hit-recip

      'a hippo and an elephant are fighting'

(40) mvuu ĺ-modzĺ ndĺ njobvu zĺ-tatu   zĺ-ku-meny-ana
      5 sg ag-one   5 pl   ag-three ag-pres-hit-recip

      'a hippo and three elephants are fighting'

(41) mvuu zĺ-tatu ndĺ njobvu ĺ-modzĺ zĺ-ku-meny-ana
      5 pl ag-three   5 sg   ag-one ag-pres-hit-recip

      'three hippos and an elephant are fighting'

(42) mvuu zĺ-tatu ndĺ njobvu zĺ-tatu   zĺ-ku-meny-ana
      5 pl ag-three   5 pl   ag-three ag-pres-hit-recip

      'three hippos and three elephants are fighting'

We must also consider conjoined plurals from the diminutive gender:

(43) tĺmĺpenĺ   ndĺ tĺmĺphĺka $\begin{Bmatrix} ti \\ ?zi \end{Bmatrix}$ -ku-sowa

      7 pl   7 pl   ag-pres-missing
      small knives and small pots   are missing

This confirms our hypothesis: each noun individually would take tĺ , and
this is the preferred form when the nouns are conjoined.

    When the conjoined noun phrases are headed by plural nouns which would re-
quire different agreement forms, then zĺ is assigned by the usual rule:

(44) maukonde ndĺ mĺpenĺ zĺ-lĺ uko
      4 pl   2 pl   ag-be
      nets   and knives are   there

    Let us now consider the implications of the data in this section. We took
as a working hypothesis the suggestion (from analyses of other Bantu lan-
guages) that any example all of whose conjuncts denote non-humans would take
the agreement zĺ . The situation in Chichewa turns out to be considerably
more complex than that. There are cases where zĺ is not possible or is not
the preferred form. The most consistent and also the most interesting cases
are those involving plural conjuncts: if each individually would take the
same target gender form, then this will be preferred (zĺ may also be pos-

sible).

These data recall the well-known analysis of Xhosa by Voeltz [1971]. Agreement with conjoined noun phrases is highly restricted in Xhosa, but is possible if plural noun phrases require phonologically identical markers. Chichewa has the same possibility, though without the theoretically important complication of Xhosa introduced by degemination. However, the rule has wider application in Chichewa in another respect: There are more types of noun phrase to which it can apply since Chichewa has greater syncretism of plural markers than Xhosa. Our analysis of Chichewa shows some similarities to the reanalysis of the Xhosa data by Roberts and Wolontis [1974].

The data in question are also comparable to data from Serbo-Croat, and the account given for Serbo-Croat applies in large measure to Chichewa. It will be necessary to give a brief account of resolution in Serbo-Croat in order to demonstrate the relevance of the data. Serbo-Croat is a South Slavonic language spoken in Yugoslavia. It preserves the three Indo-European genders, with distinct agreement markers in singular and plural. The basic gender resolution rules are as follows:

1. If all conjuncts are feminine, the feminine form is used.
2. Otherwise the masculine form is used.

For examples, and for complications which need not concern us here, see Corbett [1983a:187-191]. As can be deduced from these rules (together with the number resolution rule, which specifies the plural), conjoining masculine and feminine, masculine and neuter, feminine and neuter, and even neuter and neuter gives rise to masculine plural agreements. However, and this is the relevant point, if all the conjuncts are neuter plural, then neuter plural agreement is required:

(45) ... ta   sećanja (neut pl) i   razmatranja (neut pl) sve su   više
           those memories   and reflections   ever have more

      ustupala (neut pl) mesto novim utiscima ...
      yielded   place to new impressions (i.e. made way more and
                                      more to new impressions)

(Andrić, *Travnička Hronika*)

Since Serbo-Croat has only three genders, the neuter plurals appear as exceptional. Sets of masculine plural or feminine plural conjuncts behave like the singulars and can, at least at first sight, be covered by the ordinary rules.

The analysis offered for the neuter plurals was as follows [Corbett 1983a: 208-209]. We cannot claim that gender resolution rules operate only when the conjuncts are of different genders, since two neuter singulars take a masculine plural verb. Nor can we claim that gender resolution operates only as a consequence of number resolution, since the combination feminine plural plus neuter plural requires a masculine plural verb, which must result from gender resolution. The correct generalization appears to be that gender resolution can be triggered in two ways. First it must operate if number resolution operates. There is a very general principle that if one resolution rule operates, all must operate if possible.[7] This solution requires a stipulation in the number resolution rule that number resolution can operate only provided there is at least one singular conjunct. (This is the restriction mentioned at the beginning of section 4.) The second possible trigger for gender resolution is the presence of different genders in the subject. It can be seen that (45) above meets neither condition: number resolution could not operate, since there is no singular conjunct, and gender resolution could not operate without this trigger since the conjuncts are of the same gender. It followed as a logical consequence of this analysis that examples with conjoined masculine or conjoined feminine plurals also do not undergo resolution; however, in Serbo-Croat there was no way to prove this claim. In Chichewa, on the other hand, there are two agreement forms (i̇ and ti̇) which are not specified in

---

[7]There is often the option not to apply the resolution rules and have agreement with one conjunct only, typically the nearest, but it is not possible to select from among the resolution rules. Thus, 1st singular and 3rd singular will not produce 3rd plural, which would result from the operation of number resolution but not of person resolution, agreement in person being with the nearer conjunct. Similarly, masculine plural plus neuter singular will not give rise to masculine singular, by the operation of gender resolution, with agreement in number being with the nearer conjunct.

the output from any of the gender resolution rules and yet, as we have already seen, any of the plural agreement forms can occur, provided each conjunct individually would require that form. For Chichewa, more obviously than for Serbo-Croat, the general conditions on the operation of resolution rules are a more economical solution than a long list of individual rules specifying a large number of possible combinations of conjuncts and the required agreement forms.

There are two differences between Chichewa and Serbo-Croat which are of importance here. Serbo-Croat has a "parallel" gender system, i.e. there is a one-to-one mapping of the (three) singular target genders onto the target genders in the plural. Nouns are divided into three controller genders, matching the three target genders. As we saw in section 2, the situation in Chichewa is considerably more complex. Hence, it makes a difference whether we refer to controller (noun) genders or target genders. Thus, for Chichewa the possible triggers for gender resolution are as follows:

1. Number resolution (number resolution specifies plural agreement; it can operate provided there is at least one singular conjunct).

2. The presence of conjuncts which would require different target gender forms.

Note that the examples considered at the beginning of this section involving singular conjuncts of the same gender are covered by the first trigger: the presence of a singular conjunct triggers number resolution, which dictates the plural; number resolution in turn triggers gender resolution, which specifies zi- when all conjuncts denote nun-humans. The second difference between the two languages is that in the cases of plural conjuncts requiring the same agreement form, resolution is excluded in Serbo-Croat while, as we have seen, it is sometimes an alternative, less favoured, possibility in Chichewa. There is no unambiguous evidence as to where this option should be stated. One possibility is merely to state that gender resolution may apply optionally even when not obligatorily triggered. An alternative would be the optional dropping of the condition on number resolution (hence simplifying that rule); number resolution could then apply even if all conjuncts are plural (still speci-

fying the plural form of course); it would then trigger gender resolution, which would require the zi form in the examples in question.

We have seen that plural conjuncts which require the same agreement form are not normally involved in resolution. Apart from this, Givón's rule requiring the zi form covers many of the examples of conjoined noun phrases denoting non-humans. There are some cases where this form is not felicitous,[8] the most systematic group being examples involving gender 7. We suggest that this is because gender 7 has a firm semantic basis (the nouns are diminutives), and the resolution rule, itself semantically based, takes no account of this semantic feature. There are two further types of exception to the rules just

---

[8]There is a curious exceptional case, which we leave for future research. It concerns the first gender, which includes humans, several animals, and a very few inanimates. When humans from this gender are conjoined, the agreement marker is a , which is the expected form (as discussed in section 4.2). When, however, noun phrases denoting animals are conjoined, Mtenje still accepts only a :

(i)   mphaka ndi galu a-ku-thamanga
      1 sg   1 sg ag-pres-run
      cat   and dog are running

The expected form zi is excluded. However, in constructions in which an object marker for the conjuncts would be required, no acceptable form could be found, a , mu , zi , and wa all being rejected. When two of the rare inanimates are conjoined, both a and zi are possible.

(ii)   katundu   ndi kabundula $\begin{Bmatrix} a \\ zi \end{Bmatrix}$ -ku-sowa

       1 sg   1 sg   ag-pres-missing
       piece of baggage and pair of shorts   are missing

Though this requires further investigation, it appears that the human/other animate distinction plays a role. Finally, we conjoin noun phrases denoting a human and an inanimate:

(iii)   mnyamata ndi katundu $\begin{Bmatrix} a \\ ??zi \end{Bmatrix}$ -ku-sowa

        1 sg   1 sg   ag-pres-missing
        boy   and piece of baggage   are missing

The fact that a is a marker both of the singular and of the plural for gender 1 nouns is probably at least a part of the explanation for this irregularity. And (iii) is likely to involve agreement with mnyamata only (cf. section 4.3).

given, usually ignored in work on gender resolution. We consider them in sub-sections 4.1.1. and 4.1.2.

4.1.1. <u>Gender 10 (infinitives)</u>. Studies of gender resolution in Bantu often ignore the question of conjoined infinitives. In other language groups it is frequently the case that conjoined infinitives require neutral agreement (cf. section 3.2.) just as a single infinitive does. However, the fact that Bantu infinitives have a special agreement form and a special prefix means that the situation is potentially different, and so worth investigating. The follow-ing sentence has conjoined infinitives:

(46) kubvina ndi kuimba $\begin{Bmatrix} ku \\ zi \end{Bmatrix}$ -ku-chitikira uko

    to dance and to sing ag-pres-take place there

    'dancing and singing are going on there'

Both alternatives, gender 10 agreement and the zi form, are fully accept-able. (Note that this was the case even for the informant who was least ready to accept conjoined structures, querying examples which others immediately found unexceptional.)

    There is a semantic difference between the options; given the right con-text, one of the options can be excluded:

(47) kudya ndi kuyankhula nthawi yomweyo ndi $\begin{Bmatrix} koipa \\ *zoipa \end{Bmatrix}$[9]

    to eat and to talk time same is bad

Note the two occurrences of ndi in this example. Ndi with high tone means 'and' (as we have seen in numerous examples) and also 'with'. With low tone,

---

[9]The adjectival forms koipa and zoipa are derived as discussed in footnote 4. Koipa is derived from the verb ipa :

    ku   +   a   +   ku   +   ipa
    ag   adj   infin   verb root
               prefix

The sequence a + ku undergoes morpheme fusion giving o , and the resulting structure ku + o is changed into ko by morpheme final vowel deletion. Zoipa is similarly derived from zi + a + ku + ipa . Details of these rules and conditions on their application are given in Mtenje [1986:1-50].

ndi is the copula 'is'/'are' (which does not take prefixed agreement markers)
There is a further ndi which is the first person singular agreement marker;
this has low tone, unless high is assigned by a tense marker.

    In the next example the zi form is also unacceptable:

(48)   kukwatira   ndi kumanyenga akazi ena   nthawi yomweyo ndi $\left\{ {koipa \atop {}^*zoipa} \right\}$

       to be married and to flirt   women other time   same   is   bad

       'being married and flirting with other women at the same time is bad'

In both these examples, in place of koipa , we could have chinthu choipa
'a bad thing', but not *zinthu zoipa 'bad things'. The point, of course, is
that in both sentences it is doing the two things together which is considered
bad.

    When it is clear that the actions are considered separately, then the
judgements are reversed:

(49)   mai   a-ku-dwala   ndipo kuonera TV ndi kudya ndi
       mother ag-pres-sick and so to see TV and to eat are

       $\left\{ {zomwe \atop {}^*chomwe} \right\}$ a-ku-chita basi[10]
         ag-what ag-pres-do only

       'mother is sick, and watching TV and eating are the only things
        she can do'

(50)   kuphunzira sukulu ndi kulemera   ndi $\left\{ {zobvuta \atop {}^*chobvuta} \right\}$
       to learn   school and to be rich are   difficult

       'being educated and being rich are difficult (to achieve)'

(Zobvuta <   zi + a + ku + bvuta ; see footnote 9.) Thus conjoined infini-
tives are an exception to the general rule, in that when they are construed
as a joint activity the gender 10 form is used; otherwise the zi form is
used as normal. Some examples (like (46) above) allow either interpretation.

_____

[10]The forms chomwe and zomwe are derived as follows: to the basic
form omwe are added the agreement forms chi and zi . The rule of final
vowel deletion referred to in the previous footnote then deletes the i of
the agreement markers resulting in chomwe and zomwe respectively.

It is relatively unusual to conjoin an infinitive with another noun, but when this occurs the regular zi form is used:

(51) kulpa   ndi imfa ndi { zoopsya }
                                    { *koopsya }

      10   5 sg
      (to be) evil and death are   frightening

This example would come under the general rules for inanimates.

4.1.2. Locatives. A related problem, also rarely addressed, is that of con-
joined locatives (but cf. Givón [1972:101]). Gender 8, pa 'at', is the most
specific in terms of location; it cannot be conjoined even with another gen-
der 8 form:

(52) *panyumba ndi padambo pa-ku-tentha
       at house and at marsh ag-pres-be hot

         'it is hot at the house and at the marsh'

The alternatives, *mukutentha , *kukutentha , *zikutentha are all excluded.
Not surprisingly, when pa is conjoined with the other locatives, there is
no acceptable form:

(53) munyumba ndi panja   (*pa) -ku-tentha
                              (*mu)
                              (*ku)
                              (*zi)
      in house and outside   ag-pres-be hot
      'it is hot in the house and outside (in the yard)'

(54) kudambo ndi panja   (*pa) -ku-tentha
                              (*mu)
                              (*ku)
                              (*zi)
      at marsh and at outside   ag-pres-be hot
      'it is hot at the marsh and outside'

Gender 9, mu 'inside' allows conjunction with other gender 9 forms:

(55) munyumba ndi mugalaja mu-ku-tentha
      in house and in garage ag-pres-hot
      'it is hot in the house and in the garage'

Other agreements, including zi, are excluded. Gender 9 will not permit con-
junction with other genders as shown by (53) above and by (56):

(56)   munyumba ndi kunja $\left\{\begin{array}{l} {}^*\text{pa} \\ {}^*\text{mu} \\ {}^*\text{ku} \\ {}^*\text{zi} \end{array}\right\}$ -ku-tentha

        in house and at outside   ag-pres-be hot

        'it is hot in the house and outside'

Gender 10, the least specific of the locatives, also allows conjunction with
forms of the same gender:

(57)   kudambo ndi kuthengo ku-ku-tentha
        at marsh and at bush ag-pres-be hot

        'it is hot at the marsh and in the bush'

No other agreement, including zi , is acceptable; thus locative gender 10
differs from infinitive gender 10 in this respect. Locative gender 10 cannot
be conjoined with other locatives as (54) and (56) show.

    We have seen that the locative genders do not follow the general rule for
inanimates, since zi is excluded. Gender 8 does not allow conjoining, 9
and 10 allow conjoining only within the gender. Thus the locative genders,
once again, behave differently from the normal noun genders; nevertheless, a
full account of the agreement system must include the conditions for conjoin-
ing them and the form to be used where conjoining is possible.

4.2. Conjuncts denoting humans. When all conjuncts denote humans, then on
the basis of patterns established elsewhere in Bantu, we expect the target
gender form used for the plural of gender 1, that is the form with a in Chi-
chewa. When both nouns are from the first gender then indeed not surprising-
ly, the a form is accepted without query:

(58)   mkazi ndi mwana a-ku-yenda
        1 sg   1 sg ag-pres-walk
        woman and child are walking

Of course, the more interesting cases are those involving at least one noun
from a different gender. The following includes a noun from gender 7; again

three informants had no difficulty in selecting the a form:

(59) mkazi ndi kamwana a-ku-yenda
     1 sg   7 sg   ag-pres-walk
     woman and infant are walking

Even with two nouns of gender 7, a is found:

(60) kamkazi   ndi kamwana a-ku-gona
     7 sg   7 sg   ag-pres-lie
     little woman and infant   are lying down

(61) kamkazi   ndi tiana   a-ku-gona
     7 sg   7 pl   ag-pres-lie
     little woman and infants are lying down

In both of these sentences Mtenje rejects the ti form (plural of gender 7).
When both are plural the picture changes:

(62) tiakazi   ndi tiana   ti-ku-gona
     7 pl   7 pl   ag-pres-lie
     little women and infants are lying down

Here a was accepted with different degrees of reluctance; ti is considered
the normal form. This links immediately to the similar examples involving in-
animates; both conjuncts require the same plural gender form and so this is
used. The following includes a noun from gender 5:

(63) mkazi ndi mfumu a-ku-yenda
     1 sg   5 sg ag-pres-walk
     woman and chief are walking

(Note that for Mtenje mfumu could also be plural; some speakers have a plu-
ral form mafumu .) While this example is unproblematic, others involving
nouns from genders 5 and 6 often prove more difficult:

(64) ??mbala   ndi kamwana a-li uko
       5 sg (or pl)   7 sg   ag-be
       thief   and infant are there

While a is questionable, there is no better alternative. Since nouns in
gender 5 have no distinct plural form, we include numerals in the following
examples of genders 5 and 6 conjoined:

(65) mbala í-modzí ndí chítsílu chi-modzí $\begin{Bmatrix} a \\ ??zí \end{Bmatrix}$ -ku-meny-ana

      5 sg ag—one   6 sg   ag—one   ag-pres-hit-recip

      'one thief and one fool are fighting'

(66) mbala í-modzí ndí zítsílu zi-tatu $\begin{Bmatrix} a \\ ?zí \end{Bmatrix}$ -ku-meny-ana

      5 sg ag—one   6 pl   ag-three   ag-pres-hit-recip

      'one thief and three fools are fighting'

(67) mbala zí-tatu ndí chítsílu chí-modzí $\begin{Bmatrix} a \\ ?zí \end{Bmatrix}$ -ku-meny-ana

      5 pl ag-three   6 pl   ag-one   ag-pres-hit-recip

      'three thieves and one fool are fighting'

(68) mbala zí-tatu ndí zítsílu zí-tatu $\begin{Bmatrix} zí \\ ??a \end{Bmatrix}$ -ku-meny-ana

      5 pl ag-three   6 pl   ag-three   ag-pres-hit-recip

      'three thieves and three fools are fighting'

The most obvious point is that, as with non-humans, plural conjuncts which
would each require the same target agreement marker take that marker (as in
(68)); note once again that it is not a question of the nouns themselves be-
longing to the same controller gender. In all the other examples we find a ,
as expected. But it is surprising that zí is considered to be an alterna-
tive, albeit not a favoured one. This option should probably be seen in the
light of the data presented in section 3.1, where we saw that semantic agree-
ment with single nouns is a very marginal phenomenon in Chichewa; it is syn-
tactic gender which counts. In conjoined expressions, where the overriding
rule is a semantic one (based on the human/non-human distinction), syntactic
gender is not totally excluded. The fact that zí is better in (66) than in
(65) or (67) is to be explained by the fact that in (66) the conjunct nearer
to the verb would require zí .

     Summarizing our discussion of conjoined noun phrases denoting humans, we
may say that the situation is not so clear-cut as in some other Bantu lan-
guages. Nevertheless, once plural conjuncts requiring the same agreement
form are left out of account, the basic rule requiring the a form for hu-
mans is confirmed.

4.3. <u>Conjuncts denoting humans and non-humans</u>. Givón found that for seman-
tically mixed noun phrases, informants prefer to avoid conjoining the noun
phrases and to use the comitative construction instead. If conjoining is
forced, however, then the same form as for non-humans is used. In view of
the results obtained with plural nouns which require identical agreement
forms, it is worth investigating what happens with such nouns when they are
semantically mixed. Our first example includes two diminutives, one human
and one not:

(69) tiana   ndi timiphika   ti-li apo
      7 pl   7 pl   ag-be
      infants and little pots are   there

This sentence was found fully acceptable, while zi was excluded. The same
was true of the next example:

(70) anthu ndi abakha $\left\{ \begin{matrix} a \\ *zi \end{matrix} \right\}$ -ku-thamanga

      1 pl   1 pl   ag-pres-run
      people and ducks   are running

Though these two examples include nouns of the same gender, this is not re-
quired; the requirement once again is that the nouns should be plural and re-
quire the same plural target gender agreement marker, as the following exam-
ple proves:

(71) ana   ndi malalanje a-ku-sowa
      1 pl   3 pl   ag-pres-missing
      children and oranges   are missing

Though of different genders, both take the plural marker a and the sentence
is acceptable; zi is totally rejected. The phenomenon is illustrated nicely
by this example involving a human noun of gender 5:

(72) mbala   ndi zinthu zi-li apo
      5 pl   6 pl   ag-be
      thieves and things are   there

Recall that nouns of gender 5 do not distinguish number. Mtenje accepted
this sentence, but when asked how many thieves were involved, he said that
there must be more than one. If imodzi 'one' is added after mbala to

make it singular, then the sentence is unacceptable, whether zi or a is
chosen as the agreement marker.

Other examples involving singular nouns were also rejected:

(73)   *munthu ndi ng'ombe   zi-ku-yenda
       1 sg   5 sg or pl ag-pres-walk
       person and cow(s)   are walking

A different informant, however, accepted this sentence as unproblematic, irre-
spective of the number of cows involved (he was the informant who, in general,
was most willing to accept conjoined structures). The interesting point about
(73) is Mtenje's reaction to the alternative akuyenda . There was some un-
certainty at first, but it was considered bad. The use of a gets better if
a numeral is included after ng'ombe (whether imodzi 'one' or a higher nu-
meral). Consider then the following variant:

(74)   munthu ndi ng'ombe zi-tatu a-ku-yenda
       1 sg   and 5 pl   ag-three ag-pres-walk

       'a person and three cows are walking'

Zi is again rejected by Mtenje. The sentence as it stands is accepted or re-
jected, depending on its interpretation. If the person and the cows are walk-
ing separately, then a is also rejected. But if they are together, say the
cows are on ropes being pulled along by the person, then a is fully accept-
able. The inclusion of a possessive, which favours this type of interpreta-
tion, also makes a acceptable:

(75)   anthu ndi ng'ombe zawo a-ku-yenda
       1 pl   5 pl   poss ag-pres-walk

       'the people and their cows are walking'

Now a could be a singular or a plural marker; the crucial example is the
following:

(76)   chitsilu ndi ng'ombe yake chi-li uko
       6 sg   5 sg   poss ag-be

       'the fool and his cow are there'

The agreement marker chi is unambiguously singular; both zi and a are

rejected by Mtenje. This example demonstrates that agreement is with the first noun phrase only (and this must also be the analysis for the examples involving a). We might be tempted to think that this is a case of agreement with the first conjunct, or "distant" agreement, which is rare but which nevertheless does occur [Corbett 1983b:180]. However, given the data on the interpretations which make this agreement possible in Chichewa, we must conclude that these examples do not in fact include conjoined noun phrases. Recall that ndi means both 'and' and 'with'. It appears that in the examples where agreement is with the first noun phrase, ndi should in fact be glossed as 'with'. We are dealing with a comitative construction (and so agreement with the first noun phrase is accepted). Note that the construction is possible without the need to postpose the second noun phrase (for comparable data from Xhosa see Roberts and Wolontis [1974:240-241]).

To sum up our analysis of noun phrases denoting humans and non-humans, we may say that if both are plural and each requires the same plural gender agreement form, that form will be used giving a fully acceptable sentence. In other examples the form zi , which was expected to be marginally possible, is in fact excluded (though one informant accepts it). Instead the comitative construction is used, in appropriate contexts, though this is not immediately obvious since both noun phrases may still stand in front of the verb.

4.4. <u>Summary of gender resolution</u>. The essentials of the strategies for dealing with conjoined noun phrases in Chichewa are as follows:

1. If all conjuncts are plural and require the same plural agreement marker, then this form will be used.

As we saw, this principle operates irrespective of the type of noun involved. In section 4.1 we treated instances of this type as not coming under the jurisdiction of gender resolution (though for some examples gender resolution was a less good alternative). Nevertheless it is surprising that this syntactic principle takes precedence over the semantically based gender resolution rules which follow.

2. If all conjuncts denote humans then the a form will be used.

While there are numerous examples to support this rule, it is not quite
straightforward. In section 4.2 we noted several examples (typically involv-
ing gender 5 nouns) in which the a form was not fully acceptable (though
there was no better alternative).

   3. If none of the conjuncts denotes a human then the zí form
       will be used.

Again this holds for many examples, but there are also exceptions. Nouns
from the diminutive gender (gender 7) are problematic. Genders 8, 9 and 10
also require special provision as follows:

   3a. If infinitive phrases are conjoined and the interpretation is of
        simultaneous action, then ku will be used; otherwise zí as in
        the main rule.

   3b. Conjoined structures involving the locative genders are unaccept-
        able unless all are from gender 9 and the form mu is used, or
        all are from gender 10 and the form ku is used.

An interesting point here is that though there is a single agreement form for
gender 10, locatives and infinitives do not always behave in the same way
with regard to agreement. The remaining possibility is the conjoining of hu-
mans and non-humans:

   4. Conjoined structures involving noun phrases denoting humans and non-
       humans are unacceptable (unless covered under 1 above). Provided
       that the comitative interpretation is possible, a comitative con-
       struction may be used, in which case the head noun phrase controls
       the agreement. (Other noun phrases are not necessarily postposed.)

5. Conclusion

We have attempted to describe the gender system of one native speaker of
Chichewa as a whole, including phenomena on the fringe of the system, which
are often omitted. This approach has considerable advantages. For example,
the study of the interrelation between agreement markers which express gender
and the genders into which nouns are divided facilitates the analysis of gen-
der resolution. By analysing plural gender markers like a and zí as sin-
gle forms we can make sense of the rule which allows the conjoining of some
plural noun phrases but not others.

We have seen that Chichewa has a particularly interesting gender system, especially in its rules for gender resolution. These do not coincide with any of the sets of rules which have been proposed for other Bantu languages, though there are areas of overlap. Throughout the paper the relevance of the data to typological claims (based on languages within and beyond Bantu) has been amply demonstrated.

REFERENCES

Bokamba, E.G. 1985. "Verbal agreement as a noncyclic rule in Bantu." In D.L. Goyvaerts (ed.), *African Linguistics: Essays in Memory of M.W.K. Semikenke*, pp. 9-54. Studies in the Sciences of Language, 6. Amsterdam: Benjamins.

Brauner, S. 1979. "Actuelle Tendenzen der Entwicklung der Konkordanzbeziehungen im Swahili." *Zeitschrift für Phonetik Sprachwissenschaft und Kommunikationsforschung* 32:422-28.

Comrie, B. 1975. "Polite plurals and predicate agreement." *Language* 51:406-418.

Corbett, G.G. 1979. "The agreement hierarchy." *Journal of Linguistics* 15:203-224.

Corbett, G.G. 1980. "Neutral agreement." *Quinquereme - New Studies in Modern Languages* 3:164-170.

Corbett, G.G. 1983a. *Hierarchies, Targets and Controllers: Agreement Patterns in Slavic*. London: Croom Helm.

Corbett, G.G. 1983b. "Resolution rules: agreement in person, number and gender." In G. Gazdar, E. Klein and G.K. Pullum (eds.), *Order, Concord and Constituency*, pp. 175-206. Linguistic Models, 4. Dordrecht: Foris.

Corbett, G.G. forthcoming. "The morphology-syntax interface: evidence from possessive adjectives in Slavonic." To appear in *Language* 63.

Cornish, F. 1986. *Anaphoric Relations in English and French: a Discourse Perspective*. London: Croom Helm.

Dalgish, G.M. 1976. "Locative NP's, locative suffixes, and grammatical relations." *Proceedings of the Second Annual Meeting of the Berkeley Linguistic Society*, pp. 139-148. Berkeley: Berkeley Linguistics Society, University of California.

Doke, C.M.   1935.   *Bantu Linguistic Terminology*.   London: Longmans, Green and
    Co.

Givón, T.   1970.   "The resolution of gender conflicts in Bantu conjunction:
    When syntax and semantics clash."   *Papers from the Sixth Regional Meet-
    ing, Chicago Linguistic Society*, pp. 250-261.   Chicago: Chicago Linguis-
    tic Society, University of Chicago.

Givón, T.   1972.   *Studies in ChiBemba and Bantu grammar*.   *Studies in African
    Linguistics*, Supplement 3.   Los Angeles: Department of Linguistics and
    African Studies Center, University of California.

Heine, B.   1982.   "African noun class systems."   In H. Seiler & C. Lehmann
    (eds.), *Apprehension: Das sprachliche Erfassen von Gegenständen: I:
    Bereich und Ordnung der Phänomene*, pp. 189-216.   Tübingen, Gunter Narr.

Horton, A.E.   1949.   *A Grammar of Luvale*.   Johannesburg: Witwatersrand Univer-
    sity Press.

Laws, R.   1885.   *Table of Concords and Paradigm of Verb of the Chinyanja
    Language as spoken at Lake Nyasa*.   Edinburgh: James Thin.

Mapanje, J.   forthcoming.   "A case study of agreement phenomena in Bantu."
    Paper read at the International Seminar concerning current problems of
    linguistic research in African and Caribbean countries (UNESCO), Paris,
    September 1985.

Mateene, K.   1980.   *Essai de grammaire générative et transformationnelle de
    la langue nyanga*.   Kinshasa: Presses Universitaires du Zaïre.

Mtenje, A.D.   1985.   "Arguments for an autosegmental analysis of Chichewa
    vowel harmony."   *Lingua* 65:21-52.

Mtenje, A.D.   1986.   "Issues in the nonlinear phonology of Chichewa."   Unpub-
    lished Ph.D. thesis, University College London.

Mtenje, A.D.   forthcoming.   "Compensatory lengthening in Chichewa: a nonlin-
    ear analysis."   To appear in *Glossa*.

Price, T.   1943.   *The Elements of Nyanja for English-speaking Students*, 2nd
    edition (first edition 1941).   Blantyre, Nyasaland: Church of Central
    Africa Presbyterian.

Roberts, L. and Wolontis, M.   1974.   "Conjunction and concord in Bantu."   In
    E. Voeltz (ed.) *Third Annual Conference on African Linguistics*, 7-8
    April 1972, pp. 231-242.   Indiana University Publications, African Ser-
    ies, 7.   Bloomington: Research Center for the Language Sciences, Indi-
    ana University.

Voeltz, E.   1971.   "Surface constraints and agreement resolution: some evi-
    dence from Xhosa."   *Studies in African Linguistics* 2:37-60.

Watkins, M.H.   1937.   *A Grammar of Chichewa: a Bantu Language of British
    Central Africa*.   Philadelphia: Linguistic Society of America.

AN OUTLINE OF LULUBO PHONOLOGY*

Torben Andersen
University of Aalborg

This article outlines the phonology of Lulubo, a little known
Central Sudanic language spoken in the southern Sudan. An
account is given of the phonemic inventory (vowels, conso-
nants, and tones), vowel harmony, syllable structure, special
features of ideophones, and vowel elision.

1. Introduction

    The Lulubo language is spoken by the Lulubo people inhabiting the area
around the Lulubo Hills in the Eastern Equatoria Province of the Sudan. The
Lulubo call themselves /ólúbó/ and their language /ólúbó-ʈ̇ì/ , lit. 'Lulubo-
mouth'. According to the latest census (from 1983) they number approximately
eight thousand. The only data from the Lulubo language previously published
are included in Tucker [1940] in the form of a word list (comprising about 500
items) and a few scattered grammatical observations.

    Tucker [1940] and Tucker and Bryan [1956] classified Lulubo as a member of
the Madi dialect cluster, Lulubo being its northernmost dialect. Certainly Lu-
lubo belongs to the Moru-Madi group, which is a genetic unit within the Cen-
tral Sudanic language family suggested by Greenberg [1966]. However, the Lu-
lubo do not conceive of themselves as being Madi, and their main social con-
tacts are with Eastern Nilotic speaking peoples (the Bari and the Lokoya) rath-

er than with the Madi.

The present article deals with the phonology of Lulubo. An account is given of the phonemic inventory (vowels, consonants, and tones), the syllable structure, special features of ideophones, constraints applying across syllable boundaries (vowel harmony), and processes operating across morpheme boundaries (vowel elision and tone deletion or retention).[1] The account is based on the variety spoken in Lokiliri, a village in the central part of the Lulubo area.

2. Vowels and Vowel Harmony

Lulubo has nine vowel phonemes belonging to two vowel harmony sets which are distinguished by the feature Advanced Tongue Root (ATR):[2]

(1) -ATR   ɩ   ɛ   a   ɔ   ɒ

    +ATR   i   e   o   u

The phonemic status of the nine vowels is shown by the subminimal sets of monosyllabic words in (2-4). The sets are minimal if tone is disregarded.

(2) /i/   t̪ì   'to give birth'[3]   /u/   t̪ū   'to climb up'

    /ɩ/   t̪ì   'to pick'   /ɒ/   t̪ō   'to bend or straighten'

    /e/   t̪è   'to wait for'   /o/   t̪ō   'carefully'

    /ɛ/   t̪ɛ̄   'to swallow'   /ɔ/   t̪ɔ̄   'to dance'

               /a/   t̪ā   'tiredness'

(3) /i/   wī   'to skin'   /u/   wù   'to remove'

    /ɩ/   wī   'to spit'   /ɒ/   (ɔ̄wō   'to cry')

    /e/   wē   'to do'   /o/   wò   'to bend'

    /ɛ/   wɛ̄   'to rub'   /ɔ/   wɔ̄   (adverb indicating resulting

             /a/   wā   'to throw'   situation)

---

[1]For a similar account of Madi, see Andersen [1986c].

[2]For a comparison of the vowel systems in the Moru-Madi languages, see Andersen [1986b].

[3]Verb forms translated 'to (do something)' are stem forms. A monosyllabic stem is identical with the 2nd person singular subjunctive (= imperative) form. A polysyllabic stem is identical with the 3rd person perfective form.

(4)   /i/   kī   'to go'   /u/   kū   'to weed'

　　　/ɪ/   kī̄   'to affect'   /o/   kō   'not'

　　　/e/   ké   'non-existent' (PTC) /o/   (ōkō 'to finish')

　　　/ɛ/   kɛ̄   'to tear'   /ɔ/   kɔ̄   'to catch'

　　　　　　　　　　　/a/   kā   'to ripen'

　　In monomorphemic words with more than one syllable, all vowels which are
not /a/ belong to the same harmonic set:

(5) a.   -ATR words   b.   +ATR words

　　　　 èbɔ́   'basket'   èbì   'lion'

　　　　ēmbɛ̄   'neck'   élé   'egg'

　　　　fɔ̀lɪ̀   'buttock'   ó'jú   'horn'

　　　　íɡɛ́   'rat'   kólólí 'adze'

　　　　ómbɪ́   'locust'   úɡó   'liver'

The vowel /a/ , which phonetically is -ATR, can cooccur with any vowel. It
occurs both before and after each of the other -ATR vowels, as in (6) and (7),
and also in initial syllables before each of the +ATR vowels, as in (8).

(6)   àsɪ̄   'fire'

　　　àŋgwɛ̄   'white'

　　　àgɔ́   'pumpkin'

　　　àlō   'one'

(7)   ìbā   'rope'

　　　ɛ́dʒá   'pot'

　　　ɔ́ʔā   'fruit'

　　　ōkwā   'wealth'

(8)   àlí   'short'

　　　àgwē   'to be beautiful'

　　　rābōlō   'banana'

　　　ámbúrú   'mosquito'

While /a/ has a high lexical frequency in these positions, it never occurs
between +ATR vowels and rarely after +ATR vowels, all the attested instances

probably being loanwords. See the examples in (9).

(9) ìmòkwà   'to iron' (from Arabic)

    gómá   (kind of hoe)

    tʃópà   'house pole'

    àgúrá   'horn' (musical instrument)

    Derivational and inflectional prefixes agree with the root in terms of the ATR feature, i.e. the root controls the ATR value of the prefix. Cf. the pairs of verb forms in (11-14), whose prefixes have the respective meanings indicated in (10).

(10) a. i- ∿ ɪ-   causative

    b. e- ∿ ɛ-   direction towards the deictic center

    c. o- ∿ ɔ-   3rd person subject, perfective aspect

    d. u- ∿ ʊ-   multiple performance

(11) ì-mì   'to teach'   (cf. 10a)

    ɪ̄-ndɔ̄   'to suckle'

(12) ē-kɪ̄   'to come'   (cf. 10b)

    ɛ̄-ŋgwɪ̄   'to come back'

(13) ō-mbū   'he drank'   (cf. 10c)

    ɔ̄-bɔ̄   'he threw'

(14) ū-mbū   'to drink'   (cf. 10d)

    ʊ̄-bɔ̄   'to throw'

When the root vowel is /a/ , the prefix vowels are -ATR:

(11') ɪ̀-ḍà   'to kill'

(12') ɛ̄-sā   'to arrive'

(13') ɔ̄-ṇā   'he ate'

(14') ʊ̄-ṇā   'to eat'

These facts show that /a/ belongs to the set of -ATR vowels.

    For suffixes there is no general rule of harmonization. Some suffixes are

inherently either +ATR or -ATR, and they are not harmonized by the root, which is not harmonized by them either. This class includes for instance the +ATR suffix /-gó/ , which is a masculine singulative morpheme, and the -ATR suffix /-kì/ , which is a 2nd person singular possessive morpheme in kinship terms. Examples (15-16) show that they can combine with both +ATR and -ATR roots.

(15) ólúɓó-gó   'Lulubo man'   cf. ólúɓó   'Lulubo people'

    ɔ̀nɔ́dò̀-gó   'boy'   cf. ɔ̀nɔ́dò̀   'children'

(16) àɓì-kì   'your grandfather' cf. áɓí   'grandfather'

    àmbò̀-kì   'your sister'   cf. āmbō̄-ŋá   'sister'

In some exceptional cases, however, /-gó/ harmonizes the root:

(17) éró-gó   'enemy'   cf. ɛ̀rɔ̄   'enemies'

Other suffixes are harmonized by the root, for instance the locative suffix /-lɛ̀/ ∿ /-lè/ , which is added to verbal nouns in non-finite relative clauses (RC) with a relativized locative adverbial:[4]

---

[4]In literal translations the following abbreviations are used:

1/2/3	= first/second/third person	MASC	= masculine
1PEX	= first person plural exclusive	MULT	= multiplicative
1PIN	= first person plural inclusive	N	= noun
2P	= second person plural	NF	= non-finite
1S/2S/3S	= first/second/third person singular	OBJ	= object marker
		P	= plural
ADV	= adverb	PF	= perfective
C	= completion	POSTP	= postposition
CAUS	= causative	PTC	= particle
CONJ	= conjunction	REFL	= reflexive
COP	= copula	S	= singular
CP	= centripetal, i.e. movement towards the deictic center	SG	= singulative
		SUBJ	= subjunctive (including imperative)
FOC	= focus	TR	= transitive
GEN	= genitive	UNSP	= unspecified subject
IMPF	= imperfective	VA	= verbal adverb
LINK	= linker	VN	= verbal noun
LOC	= locative		

An oblique stroke (/) indicates that the following morpheme(s) is/are manifested solely or partly by the tone pattern of the word.

(18) a. wí   kɔ́-ŋgwī   òfó [RC àwī̄ rí ē-ŋgá-lὲ]   rá⁻ wɔ̄
         monster 3-return/IMPF place   3S GEN CP-get-up/VN-LOC to PTC
         'the monster returns to the place it has come from'

     b. kɔ̄-zwὲ   ṭàká   [RC kὲní rí ò-ṭū-lὲ   rī̄] ínì nī̄
         3-push/IMPF platform   Kenyi GEN NF-climb/VN-LOC PTC this OBJ
         'it pushes the platform upon which Kenyi has climbed'

Phonetically, there is a tenth vowel quality, viz. an unrounded, mid, back, +ATR vowel [ɤ]. This vowel only occurs after the +ATR vowel /u/, as in

(19) [úgɤ́]   'liver'
      [ṭúṭwɤ̀]   'ostrich'

When /u/ and [ɤ] are heteromorphemic, [ɤ] varies freely with the rounded, mid, back, +ATR vowel [o]:

(20) [òɡú-ɡɤ́] ∿ [òɡú-ɡó]   'male thief' (thief/S-MASC+SG)
      [ū-ŋgɤ́]   ∿ [ū-ŋgó]   'singing' (MULT-sing/VN)

The morphemes containing [ɤ] ∿ [o] invariably have [o] when they are preceded by a vowel other than /u/:

(21) [màdí-ɡó]   'Madi man' (Madi-MASC+SG)
      [ō-ŋgó]   'singing' (NF-sing/VN)

Thus we can infer that [ɤ] is an allophone of the phoneme /o/.

3. Consonants

Lulubo has 34 consonant phonemes, whose phonetic properties are indicated in Table 1. The stops, the affricates and the nasals form a major subsystem with six contrasting points of articulation. The interdentals and the alveolars together form one series, and so do the palato-alveolars and the palatals. There are no retroflexes and labio-velars among the preglottalized stops and the nasals, nor is there a preglottalized velar stop.

Examples of words with each of the consonants are given sectionwise in (22-29) below. In the first column of each section an example is given in which each consonant is followed by the vowel /a/. Gaps indicate the lack of such words in the present author's corpus. The second column shows other,

Table 1. Consonants

		bilabial	labiodental	interdental	alveolar	retroflex	palato-alveolar	palatal	velar	labiovelar	glottal
stop	unvoiced	p		t̪	t̩				k	kp	ʔ
	voiced	b		d̪	d̩				g	gb	
	prenasalized	mb		nd̪	nd̩				ŋg	ŋmgb	
	preglottalized	ɓ			ɗ			'j			
affricate	unvoiced						tʃ				
	voiced						dʒ				
	prenasalized						ndʒ				
fricative	unvoiced		f		s						
	voiced		v		z						
nasal		m			n			ɲ	ŋ		
lateral					l						
trill					r						
glide								y		w	

preferably (sub)minimal, sets, in which the consonants occur before another vowel.

(22)   Unvoiced stops and affricates

/p/ pā 'to shape (wood with an adze)' pὲ 'to defeat'

/t̪/ t̪ā 'tiredness' t̪ē̄ 'to swallow'

/t̩/ t̩ā 'to put down abruptly' dὲmὲt̪ὲt̩ὲ 'swallow' (bird)

/tʃ/ tʃà 'to plait (a door)' tʃérὲpέ 'calabash'

/k/	kā	'to ripen'	kɛ̄	'to tear'
/kp/	-		kpɛ́	'in vain'

(23) Voiced stops and affricates

/b/	ìbā	'rope'	bómú	'maize'
/ɖ/	ɖà	'to cut (finger-millet)'	ɓóɖó	'half (of something like a fruit)'
/ḍ/	ḍà	'to die'	ōḍōḍú	'frog'
/dʒ/	ɛ́dʒá	'pot'	dʒó	'house'
/g/	gā	'to cut'	ágó	'male'
/gb/	ìgbà	'to barter'	gbóróró	'throat'

(24) Prenasalized stops and affricates

/mb/	mbā	'to grow up'	mbū	'to drink'
/nḍ/	ìnḍá	'good'	-	
/nḍ/	ānḍā	'face'	nḍù	'to look for'
/ndʒ/	ndʒà	'protruding'	ndʒū	'to suck'
/ŋg/	ŋgā	'to get up'	ŋgù	'to smell'
/ŋmgb/	ŋmgbā	'to beat'	ŋmgbú	'nothing being left' (ADV)

(25) Glottal and preglottalized stops

/ɓ/	ɓā	'home' (N)	ɓì	'to taste'
/ɗ/	àdā	'to lower'	ɗī	'to spear'
/'j/	'jà	'to quarrel'	-	
/ʔ/	ʔā	'stomach'	ʔī	'to grind'

(26) Nasals

/m/	mā	'I'	ɛ̀mɛ̄	'to be warm'
/n/	nā	'three'	nɛ̄	'to cook'
/ɲ/	ɲā	'to eat'	ɛ̀ɲɛ̄	'to turn upside down'
/ŋ/	ŋā	'to speak'	tʃíŋɛ́	'sand'

(27) Fricatives

/f/	-		fɔ̀	'to appear'
/s/	sā	'to arrive'	ɔ̀sɔ̄	'to destroy'
/v/	-		vɔ̄	'to blow'

/z/   ìzā   'meat'   ȯzō   'to straighten'

(28)   Lateral and trill
       /l/   lā   'to lay down'   lᵀ   'to cut'
       /r/   rā   'to' (POSTP)   rì   'two'

(29)   Glides
       /y/   yà   'but' (CONJ)   āyᵀ   'rainy season'
       /w/   wā   'to throw'   wᵀ   'to skin'

The only consonants that have a systematically defective distribution are the labiodental fricatives [f] and [v]. They only occur before rounded vowels. However, since [f] and [v] are not in complementary distribution with any other consonants, they have to be ascribed phonemic status. The phonetically most likely candidates for complementary distribution with the labiodental fricatives would be the bilabial stops [p] and [b]. But they also occur before rounded vowels, although lexically infrequently:

(30)   pótò   'tendon'
       pùtá   'whip'

(31)   bòŋgó   'clothes'
       bɔ̀ndɔ̀   'tamarind'

For [b] and [v] there even is a minimal pair:[5]

(32)   bō   'to throw'
       vǭ   'to blow'

---

[5]Historically, the labiodental fricatives must have developed from labial stops before rounded vowels. This sound change seems to have affected all the Moru-Madi languages, but in Lulubo it has not been carried through for all instances of original *b and not at all for original *mb. Cf. the following examples from some of the languages, the dialect in question being indicated in parentheses:

Moru (Miza)	Lugbara (Terego)	Madi (Burulo)	Lulubo (Lokiliri)	Proto-Moru-Madi	
-	ȯvó	ɛvɔ́	ɛbɔ́	*ɛ̀bɔ́	'basket'
vǫ̀	vǫ̀	vǫ̀	bō	*bɔ̀	'place, land'

The only consonantal phonemes with allophonic variation are the velars
/g ŋg ŋ/ . These consonants have palatal coarticulation after the mid front
vowels /e ɛ/ :

(33) /gē/   [gʸē]   'to buy'
      /ŋgē/   [ŋgʸɛ]   'to pull out'
      /tʃíŋɛ́/   [tʃíŋʸɛ́]   'sand'

Palatalization does not occur before the high front vowels /i ι/ , and it
does not affect the unvoiced velar /k/ .

4. Tones

Lulubo has three tones: low (L), mid (M), and high (H). The phonemic
status of the tones is shown by the following minimal triplets:

(34) L   tì   'to give birth'   ʔà   'existent' (PTC)
      M   tⁱ   'mouth'   ʔā   'stomach'
      H   tí   'cow'   ʔá   'we' (1PIN)

Any two tones may be combined within a syllable. Thus there are six com-
pound tones:

(35) L̂M   tɔmbɛ̄   'it is Tombe' (Tombe/FOC)

      L̂H   ɓàʸ   '(at) home' (home/LOC)

      M̂L   ndɛ̀ kwɛ̄ˋ 'look at the tree!' (see/SUBJ+2S tree/SUBJ)

      M̂H   ʔā'   'in the stomach' (stomach/LOC)

      ĤL   î-ɗà   'we killed' (CAUS/1PIN-die/PF)

      ĤM   ízî¯   'his wife' (female/3S)

Moru (Miza)	Lugbara (Terego)	Madi (Burulo)	Lulubo (Lokiliri)	Proto-Moru-Madi	
vō	vō	—	bō	*bō	'to throw'
ʌ́mvú	ámvú	ámvú	ámbú	*ʌ́mbú	'field'
kòmvō	òmvō	ɔmvō	ɔmbō	*(k)ɔmbō	'nose'
cèmvó	òmvó	ɛmvó	ɛmbɔ́	*(k)ɛmbɔ́	'orphan'

Except for H͡L, compound tones are restricted to occurring in morphemically
complex words, in which their two components belong to or manifest two differ-
ent morphemes, as in (35). H͡L is fairly common in monomorphemic words, but
only on final syllables. Such words appear all to be recent borrowings, for
instance the following:[6]

(36) ɗɔ̂kɛ̂   'gift'   (cf. Bari ɗɔk-ɛt 'gift')
      ŋwàkɛ̂   'tongs'   (cf. Bari ŋwak-ɛt 'pincers')
      ɖìn�î   'time'   (cf. Bari dɪnɪt 'time')
      ɖɪ̂   'chisel'   (cf. Bari dɪ 'chisel')
      gùrû   'money'   (cf. Sudanese Arabic giriʃ, pl. guruuʃ 'money')
      sâ   'watch'   (cf. Sudanese Arabic saaʕa 'hour, watch')
      kòmìʈî   'committee' (from English)

Exceptions are

(37) ɗɛ̂ɗɛ̌   'all'
      ʈîʈɤ̌   'quickly'

which are monomorphemic but have H͡L on non-final syllables. However, these
words are also aberrant in having L͡H on their final syllables, and since they
are furthermore characterized by segmental reduplication, they should prob-
ably be considered ideophones (cf. section 6).

     Tone distinguishes many lexical items, especially mono- and disyllabic
ones. In the following examples, the words of each pair belong to the same
word class (noun, verb, or adjective):

(38) rū   'body'   (39) gà   'to refuse'
      rú   'name'   gā   'to cut'

(40) ɓì   'to taste'   (41) àlf   'short'
      ɓɪ   'to twirl'   álf   'deep'

---

[6]The Bari and Sudanese Arabic data have been taken from Spagnolo [1960]
and Persson and Persson [1980], respectively. Although Bari is a tone lan-
guage, Spagnolo does not indicate tones.

(42) èrí   'blood'   (43) èrì   'to come down'
      ērí   'mahogany'   èrī   'to hear'

Tone also has a high functional load in inflection and derivation. Examples
given below will illustrate the use of tone for expressing contrasts within
various grammatical categories.

   Most nouns are not inflected for number, but nouns denoting human beings
are, and one class of them distinguishes between singular and plural by means
of tone. The examples in (44) illustrate two morphophonemic tone rules.
Nouns with the tone pattern HH in the singular have LM in the plural, and
those with LH in the· singular have HH in the plural.

(44)   Singular   Plural
     a.   HH   LM
          ágó   àgō   'husband'
          ɔ́zɔ́   ɔ̀zɔ̄   'witch-doctor'
          ɔ́ndɔ́   ɔ̀ndō   'sterile person'

     b.   LH   HH
          èmbɔ́   émbɔ́   'orphan'
          ògú   ógú   'thief'
          òrí   órí   'coward'

   Verbs are inflected for mood, aspect, and person and number of the sub-
ject. Together, these categories are expressed by prefixes and by particular
tone patterns. The actual tone patterns depend on phonological and grammati-
cal properties of the stem, including its number of syllables, its inherent
tone pattern, its transitivity, and its morphological composition. Two exam-
ples are given in Table 2, which shows the inflection of the stems /ɲā/ 'to
eat' and /ōmbɛ̄/ 'to tie'. The first paradigm is representative of transi-
tive stems that are monosyllabic and have an inherently mid tone. The second
paradigm is representative of disyllabic stems with the inherent tone pattern
MM. Notice for instance that, in the perfective aspect, forms which express
first or second person singular have the tone pattern HM, while forms which
express first person plural exclusive or second person plural have LM. Anoth-

Table 2. <u>Two paradigms of verbal inflection</u>

	/ɲā/ 'to eat'			/ɔ̄mbɛ̄/ 'to tie'		
	sub-junc-tive	indicative		sub-junc-tive	indicative	
		per-fec-tive	imper-fec-tive		per-fec-tive	imper-fec-tive
1S	á-ɲā		mɔ̄-ɲá	m-ɔ́mbɛ̄		m-ɔ̄mbɛ́
1PEX	à-ɲā		mò-ɲá	m-ɔ̀mbɛ̄		m-ɔ̀mbɛ́
2S	ɲā í-ɲā		nɔ̄-ɲá	ɔ̀mbɛ̀	n-ɔ́mbɛ̄	n-ɔ̄mbɛ́
2P	ì-ɲā		nò-ɲá	ɔ̀mbɛ̄	n-ɔ̀mbɛ̄	n-ɔ̀mbɛ́
3	kò-ɲā ɔ̄-ɲā		kɔ̄-ɲá	k-ɔ̀mbɛ̄	ɔ̄mbɛ̄	k-ɔ̄mbɛ́
1PIN	ɔ́-ɲā		ɔ́-ɲá	ɔ́mbɛ̄		ɔ́mbɛ́
UNSP	ɔ́-ɲā		ɔ̄-ɲá	ɔ́mbɛ̄		ɔ̄mbɛ́

er generalization is that the final stem vowel keeps its inherent tone in the perfective aspect but replaces it with a high tone in the imperfective aspect.

As an example from derivational morphology, consider the formation of in-gressive verb stems from adjectives with an initial /V/ syllable. No affixes are used, but the tones of the adjectives are replaced by low tones in the verbs:

(45)   Adjective   Verb
       àzɔ̄   'long'   àzɔ̀   'to become long'
       ìnɡá   'good'   ìnɡà   'to become good'
       ōsú   'good'   òsù   'to become good'
       áIí   'deep'   àIì   'to become deep'
       àkēIí 'red'   àkèIì 'to become red'
       áfɔ́rɔ̀ 'yellow'   àfɔ̀rɔ̀ 'to become yellow'

Combining the lexical and the morphological functions of tone, the number of disyllabic words distinguished tonally is often quite large, as in the following example:

(46) ò-gù   'he laughed at' (3-laugh=at/PF)

      ògū̄   'back' (N)

      ògú   'thief' (N/S)

      ō̄-gū̄   'he carried' (3-carry/PF)

      ō̄-gú   'is being laughed at' (UNSP-laugh=at/IMPF)

      ó-gù   'was laughed at' (UNSP-laugh=at/PF)

      ó-gū̄   'was carried' (UNSP-carry/PF)

      ógú   'thieves' (N/P)

      ô-gù   'we laughed at' (1PIN-laugh=at/PF)

      ô-gū̄   'we carried' (1PIN-carry/PF)

      ô-gú   'we are laughing at' (1PIN-laugh=at/IMPF)

In addition to its lexical and morphological functions, tone also has syn-tactic functions in Lulubo. Thus, certain sentence types end in a floating tone (T̞), which is manifested on the sentence-final syllable as an addition to the inherent tone of that syllable. The result is a compound tone, unless the floating tone is identical with the inherent tone. All three tones, L̞, M̧, and Ḩ, have syntactic functions in this sense.

L̞ occurs after an unfocused object if the sentence has the subjunctive mood. Cf. the examples in (47), where the subjunctive verb form is identical to the perfective indicative verb form for first person subjects.

(47) L+L̞ → L   á-ndè èbì   'let me see the lion!'

      M+L̞ → M̂L   á-ndè àrī̀   'let me see the bird!'

      H+L̞ → ĤL   á-ndè t̞í   'let me see the cow!'

      (á-ndè '1S-see', èbì 'lion', àrī 'bird', t̞í 'cow')

Ḩ occurs after a sentence-final object if the sentence has the indicative mood and indicates that the scope of the focus is either the object or the verb and the object. Cf. the examples in (48), which are minimally different from those in (47).

(48) L+Ḩ → L̂H   á-ndè èbí̇   'I saw a lion'

      M+Ḩ → M̂H   á-ndè àrī́   'I saw a bird'

      H+Ḩ → H   á-ndè t̞í   'I saw a cow'

That this final floating Ḫ is actually a focus morpheme is shown by the fact
that it contrasts with other focus markers, which also occur after the object,
and with adverbials, which are inherently focused in that position:

(49) á-ndɛ̀ èbì ʔō   'I *did* see the lion'   (truth value focus)

     á-ndɛ̀ èbì mā   '*I* saw the lion'   (subject focus)

     á-ndɛ̀ ·èbì ɔ̀-ndɛ̀   'I *saw* the lion'   (verbal focus)

     á-ndɛ̀ èbì ágínī̄   'I saw the lion *yesterday*' (adverbial focus)

     (/ʔō/ indicates completion; /mā/ is the first person singular per-
      sonal pronoun; /ɔ̀-ndɛ̀/ is a verbal adverb; /ágínī̄/ means 'yesterday')

The same morpheme Ḫ is used after the predicate in a non-verbal clause of
identification, which is what native speakers will normally provide when
asked to translate a noun into Lulubo:

(50) èbÝ   'it is a lion'

     àrī̄'   'it is a bird'

     tÍ   'it is a cow'

    Ṃ is used instead of Ḫ in certain cases, for instance when the object is
a proper noun. Cf. the examples in (51), which are minimally different from
those of both (47) and (48).

(51) L+Ṃ → L̂M   á-ndɛ̀ èbⁱ⁻   'I saw Mr. Lion'

     M+Ṃ → M   á-ndɛ̀ àrī̄   'I saw Mr. Bird'

     H+Ṃ → ĤM   á-ndɛ̀ tⁱ⁻   'I saw Mr. Cow'

    Lulubo is a discrete-level tone language, i.e. there is (almost) no down-
drift, nor is there anything like downstep. The pitch value of a tone remains
by and large the same throughout an utterance, except on the utterance-final
syllable. In utterances that are not yes/no-questions, prepausal L is real-
ized as an extra low pitch, and L̂M, L̂H and M̂H rise to a lower pitch level pre-
pausally than non-prepausally. Yes/no-questions are uttered in a higher reg-
ister than the corresponding statements or orders, and a rising pitch is add-
ed after their final tone. Apart from the intonation, there is no formal dif-
ference between yes/no-questions and statements or orders.

5. Syllable Structure

    The segmental structure of a syllable in Lulubo is /V/, /CV/, or /CCV/,
but only initial syllables can be /V/:

(52)   (V)CV   bí   'ear'   ómbí   'locust'

       (V)CVCV   fòlì   'buttock'   ɔ̀nɔ́dɔ̂   'children'

       (V)CVCVCV t̪ɔ́rɔ̀mɛ̀ 'nine'   ìrùgùnɔ̀ 'to kneel'

Apparent exceptions are found in words like the following, all of which are
monomorphemic:

(53) /làwō/   → [làō]   'time'

     /t̪òwú/   → [t̪òú]   'five'

     /gòwɛ́/   → [gòɛ́]   'jackal'

     /āyí/   → [āí]   'grass'

     /àràbíyà/ → [àràbíà] 'car'

Phonetically, these words contain a non-initial /V/ syllable. But in such
cases either that vowel itself or the vowel preceding it is high. Therefore,
these words can easily be analyzed as having a homorganic glide, i.e. a con-
sonant, between the vowels: /w/ if the high vowel is back, /y/ if the high
vowel is front.

    In /CCV/ syllables, the second consonant can only be a glide. Further-
more, only non-rounded vowels have been attested in such syllables. The con-
trast between /w/ , /y/ , and zero after the first consonant is shown by the
following subminimal triplet:

(54) Iwà   'to count'

    Iyà   'to shiver'

    Iā   'to lay down'

While the syllable structure /CyV/ is rare, /CwV/ is quite common (espec-
ially when /C/ is a velar stop), so that there are many minimal pairs like
the following:

(55) ìɓwà   'to let (liquid) run out'   (56) gwē   'to burn'

    ì-ɓà   'you put' (2P-put/PF)   gē   'to buy'

(57) kwā   'bone'   (58) àmbwɛ̄   'to hide'

      kā   'to ripen'   à-mbɛ̄   'we licked' (1PEX-lick/PF)

(59) ŋgwá 'child'

      ŋgá   'thing'

The glide /w/ has been attested after about two thirds of the consonants, as in the examples in (60), which are grouped according to the point of articulation of the consonants. There are no examples of /w/ after plain labial stops and after labiovelars and no reliable examples of /w/ after interdentals and palatals. The absence of /w/ after the bilabial stops /p/ and /b/ and its corresponding presence after the labiodental fricatives /f/ and /v/ is probably systematic, considering the facts that /w/ has lip rounding, that /p/ and /b/ tend not to occur before rounded vowels, and that /f/ and /v/ only occur before such vowels (cf. section 3 and footnote 5).[7]

(60) <u>bilabial</u>

    /p/   -

    /b/   -

    /mb/ mbwɛ̀ 'to be ill'

    /ɓ/   ɓwῑ   'to shoot'

    /m/   mwà   'to rot'

   <u>labiodental</u>

    /f/   fwɛ̀   'to burst'

    /v/   vwí   'hunger'

---

[7]In the southern part of the Lulubo area (Aru), /w/ has disappeared after some consonants, e.g. after /v/ , as in the following examples:

Lokiliri   Aru

vwῑ   vῑ   'to follow'

àvwá   àvá   'infertile soil'

Thus, in Aru, /v/ does not have the defective distribution that it has in Lokiliri (see section 3). In Aru, /v/ now occurs before unrounded vowels as well as before rounded vowels.

<u>interdental</u>

/t̪/   —

/d̪/   —

/nd̪/   —

<u>alveolar</u>

/ɗ/   ɗwà   'to pick'

/s/   swɛ̀   'to become fat'

/z/   zwā̄   'to cross'

/n/   ónwɛ́   'new'

/l/   lwē̄   'to whistle'

/r/   rwɛ̀   'to become thin'

<u>retroflex</u>

/ʈ/   ʈwà   'to become sour'

/ɖ/   íɖwé   'urine'

/nɖ/   nɖwɛ̀   'to squeeze into'

<u>palato-alveolar</u>

/tʃ/   tʃwɛ̄   'to prune'

/dʒ/   ùdʒwè   'to wash'

/ndʒ/   ìndʒwí   'bad'

<u>palatal</u>

/'j/   —

/ɲ/   —

/ɣ/   —

<u>velar</u>

/k/   kwɛ̄   'tree'

/g/   ágwí   'big'

/ŋg/   ŋgwī̄   'to go back'

/ŋ/   ɔ̀ŋwíŋwīní   (kind of ant)

<u>labiovelar</u>

/kp/   —

/gb/   —

/ŋmgb/ -

/w/ -

glottal

/ʔ/   ʔwè   'to scoop (water)'

Whether there are any constraints on /C/ in /Cy/ combinations is difficult to determine, due to the lexical infrequency of such sequences. A few additional examples are given in (61).

(61)   ɔ̀myā   'to mix'

     ìzyē   'excrement'

     mílyé   'oath'

The combinations glide + vowel in /CwV/ and /CyV/ should not be analyzed as heterosyllabic sequences of equi-tonal high vowel + vowel, since such an analysis would result in violations of three well-established constraints on the phonological structure of verb stems. Firstly, disyllabic verb stems cannot begin with a consonant, hence not */lòà/ 'to count' and */lìà/ 'to shiver', but /lwà/ and /lyà/ . Secondly, trisyllabic verb stems can only begin with a high, unrounded vowel (unless they are multiplicative stems (see section 7) or have been derived from an adjective (see section 4)), hence not */ùdʒùè/ 'to wash' and */èguè/ 'to get lost', but /ùdʒwè/ and /ègwè/ . Thirdly, the tone pattern of trisyllabic verb stems can only be LLL, hence not */àmbōɛ̄/ 'to hide' and */ɔ̀mīā/ 'to mix', but /àmbwɛ̄/ and /ɔ̀myā/ .

6. Ideophones

Some ideophones deviate from the phonological system outlined in the previous sections. The deviations concern vowels, consonants, and syllable structure. The features mentioned below are probably not exhaustive.

In ideophones, syllables can have the structure /CVC/, i.e. they can end in a consonant:

(62)   ɖít̪   (having large volume)

     rùyúk   (stone striking something)

pɛ́r-pɛ́r (very clean)

t̪ú?-t̪ú? (sound of hammering)

In light of this, the following adverbs, which are aberrant by ending in a glottal stop, may also be considered ideophones:

(63) líʔ ~ líbídíʔ   'all' (countable)

píʔ   'all' (uncountable)

t̪íʔ   'only'

   In addition to the nine vowels of ordinary words, ideophones can have a syllabic [r̩] after a vowel. In some cases, [r̩] varies freely with non-syllabic [r] plus a vowel. Cf. the following examples:

(64) t̪ír̩   (iron or glass hitting something)

dèr̩-dèr̩   (running with raised neck)

ndɔ́r̩ ~ ndɔ́rɔ́ (high degree of redness)

   While in ordinary words there is no length contrast in the vowels, vowels of ideophones are either short, like those of ordinary words, or long. Examples with long vowels are given in (65).

(65) wúur (movement away from the speaker)

d̪ùut̪ (movement towards the speaker)

   Finally, at least one additional consonant can occur in ideophones, viz. an unvoiced velar fricative [x] , as in

(66) xɔ́xɔ́xɔ́ (sound made by a ratel when attacking people)

   The examples of ideophones listed above should probably all be classified as adverbs. But there are also a few nouns that deviate phonologically in the same ways. Cf. the nouns in (67), one being related to [dèr̩dèr̩] in (64), the other to [xɔ́xɔ́xɔ́] in (66).

(67) àdɛ́r̩   'mongalla gazelle'

xɔ́xɔ́lɛ́ 'ratel'

   Many ideophones do not have any of the aberrant features mentioned above and thus do not differ phonologically from non-ideophones, as in the follow-

ing examples:

(68) ɓòlú   (sound of a stone falling into water)

róɓó   (sound of a calabash breaking into pieces)

médū   (breathing heavily)

However, many of them are characterized by total reduplication, as in the examples in (69) and (70).

(69) ʈó-ʈó   (sound of cutting a tree)

ʈù-ʈù   (sound of footsteps)

bù-bù   (sound of heavy wind)

kpá-kpá   (sound of a person walking)

(70) wàlà-wàlà   (moving of snake)

púʈú-púʈú   (nothing being left)

kwílí-kwílí   (sound of small bell)

kòsè-kòsè   (sound of small animal moving in grass)

fòlà-fòlà   (bubbles coming up)

7. Vowel Elision

In certain cases, a word-final vowel is optionally (but normally) elided before a vowel-initial word. Depending on the (type of) word whose vowel is elided, the tone of the elided vowel is either retained or deleted. If retained, the tone is added to the tone of the adjacent vowel, but only realized phonetically if it is different from the latter. In the following, some of the most common types of elision are illustrated.

One type of elision occurs when a finite verb is followed by a cognate object or by a cognate adverbial. In that case the final vowel of the verb is normally elided, but its tone is always retained:

(71) kō-mbú   ō-mbú   nī   [kōmbōmbú nī]   'he is drinking'
3-drink/IMPF NF-drink/VN OBJ

(72) ɲā   ɔ̄-ɲâ   [ɲɔ̄ɲâ]   'eat!'
eat/2S+SUBJ NF-eat/VN+SUBJ

(73) k-āwέ   àwὲ   [kāwâwὲ]   'he is walking'
3-walk/IMPF walk/VA

If the final syllable of the verb is /CwV/ or /CyV/, the glide is elided along
with the vowel. See (74)-(76) for /CwV/ and (77)-(78) for /CyV/.

(74) kɔ́-mwā   ɔ̀-mwà   [kɔ́mɔ̀mwà]   'it is getting rotten'
      3-rot/IMPF NF-rot/VA

(75) k-ɔ̄ʔwɛ́   ɔ̀ʔwɛ̀   [kɔ̄ʔɔ́ʔwɛ̀]   'it is drying'
      3-dry/IMPF dry/VA

(76) k-úɗ-ūɗwē   ùɗ-ùɗwè   [kúɗùɗùɗùɗwè]   'he vomits'
      3-MULT-vomit/IMPF MULT-vomit/VA

(77) ɔ̀myā   ɔ̀myā   [ɔ̀mɔ̀myā]   'he *mixed* them'
      mix/3+PF mix/VA

(78) ɔ̀-lyà   ɔ̀-lyà   [ɔ̀lɔ̀lyà]   'he *shivered*'
      3-shiver/PF NF-shiver/VA

This behaviour of a glide after another consonant clearly shows that /Cw/ and
/Cy/ are clusters and not unitary consonants like, for instance, the affri-
cates and the prenasalized stops. Note that this also holds for /ʔw/ , as
in (75), i.e. /ʔw/ is a cluster rather than a unitary consonant, and there-
fore it does not belong to the class of preglottalized consonants.

     The vowel /ɩ/ is often elided in certain monosyllabic function words
when they occur in particular syntactic environments. The tone is deleted in
the case of /nɩ̄/ , which either links an adjective to a preceding noun (79)
or a non-finite verb to its preceding object (80).

(79) t̪ɔ́rɛ́ nɩ̄ àzɔ̄   [t̪ɔ́rɛ́ nàzɔ̄]   'a long stick'
      stick LINK long

(80) Tkōzó   nɩ̄ ɩ̀-ŋgwɩ́   ōsú nā'   'it is good to re-
      property LINK CAUS-return/VN good COP+S/FOC   turn property'
      [Tkōzó nɩ̀ŋgwɩ́ ōsú nā']

The tone is retained in case of the genitive postposition /rɩ́/ , the reflex-
ive particle /rɩ̄/ , and the particle /rɩ̄/ that follows a transitive verb
in the imperfective aspect. Cf. (81), (82), and (83), respectively.

(81) TnT rɩ́ ɛ̀rɩ́   [TnT rɛ̀rɩ́]   'snake blood'
      snake GEN blood

(82)　mó-ŋgⓉ　　　rⓉ　ò-ŋgì　　　　　　　'I am falling'
　　　1S-fall/IMPF REFL NF-fall/VA
　　　[móŋgⓉ rō̄ŋgì]

(83)　ó-ɗwā̄　　　　　rⓉ ò-ɗwà　　　　'they are being picked'
　　　MULT-pick/UNSP+IMPF TR MULT-pick/VA
　　　[óɗwā̄ rō̄ɗwà]

　　　　Vowel elision also affects the locative postposition /a/ , which governs
a noun phrase (NP), as in the following sentences with a postpositional
phrase (PostpP):

(84)　íkéré̄ kó-kⓉ　　[PostpP [NP góló̀] á]
　　　Ikere 3-go/IMPF　　　　　river LOC

　　　'Ikere goes to the river'

(85)　àzⓉ kìdề kó-kwⓉ　　[PostpP [NP àwⓉ rí dʒó] ā]
　　　then Kide 3-enter/IMPF　　　3S　GEN house LOC

　　　'then Kide enters her house'

(86)　ɔ̀ɲóɗò̀　　ìlékì k-ɔ̄lá　[PostpP [NP òfó　nⓉ　àlūbé] ā]
　　　children those 3-play/IMPF　　place LINK one　　LOC

　　　'those children played together (lit. in one place)'

(87)　óní　[PostpP [NP dí　nⓉ　ìndá] ā]
　　　stone　　　　hand LINK good　LOC

　　　'the stone is to the right'

The tone of this postposition, which must have developed from the noun /ʔā/
'stomach', is determined by the last tone of the noun phrase: it is mid af-
ter a high tone but high after a low or mid tone. The postposition is enclit-
icized to the last word of the noun phrase whether that word is the head of
the noun phrase, as in (84) and (85), or a modifier, as in (86) and (87).
The vowel of the postposition is retained after a high vowel, but it under-
goes harmonization in that context: it remains /a/ after the -ATR vowels
/ɪ/ and /ʊ/ , but it is changed to /o/ after the +ATR vowels /i/ and
/u/ .[8] The high vowel preceding the encliticized vowel is, in turn, realized

_____

　　　[8]Alternation between /a/ and /o/ is also found in some other Central

phonetically as a semivowel, i.e. it keeps its vocalic quality and its tone, but together with the encliticized vowel it forms a diphthong with the length of a monophthong. (The semivowels cannot be analyzed phonemically as the glides /y/ and /w/ since the latter phonemes do not carry separate tones. Another reason for not analyzing the semivowel [i̯] as /y/ is that it does not become phonetically rounded before the rounded vowel /o/). The opera- tion of these rules is exemplified in (88) for each of the four high stem-fi- nal vowels. Encliticization and harmonization turn the underlying forms into phonemic forms, which are themselves turned into phonetic forms by semivoca- lization and by optional de-rounding of /o/ .

(88)   underlying   phonemic   phonetic
    /ɪ/   àsɪ̄ á   /àsɪ̄-á/   [àsɪ̯á]   'on the fire'
    /o/   íʈó ā   /íʈó-ā/   [íʈo̯ā]   'on the ant-hill'
    /i/   àdɪ̄ á   /àdɪ̄-ó/   [àdɪ̯ó]   'in the well'
    /u/   ògū á   /ògū-ó/   [ògu̯ó] ~ [ògu̯ɣ]   'on the back'

After non-high vowels, the vowel of the postposition is elided, but its tone is retained. This is shown by the following examples for each of the five non- high vowels:

(89)   underlying   phonemic
    /e/   ìzyē á   /ìzyē'/   'in the dung'
    /ɛ/   ʈórɛ ā   /ʈórɛ̄/   'on the stick'
    /a/   ɔ́swà á   /ɔ́swǎ/   'in the fats'
    /ɔ/   ɛ̀bɔ́ ā   /ɛ̀bɔ̄/   'in the basket'
    /o/   dʒó ā   /dʒō/   'in the house'

Note that in this case it is the second of the two contiguous vowels that is elided, not the first one as in the other constructions mentioned above.

Sudanic languages with vowel harmony based on the ATR feature. In Moru, another language of the Moru-Madi group, the alternation occurs in verbal pre- fixes [Andersen 1986a], and in Modo, which belongs to the Bongo group, it oc- curs in the first person singular possessive suffix /-ma/ ~ /-mo/ .

There is evidence that vowel elision also occurs within the boundaries of a word. Consider for instance the formation of multiplicative verb stems from verb stems with more than one syllable, such stems being formed by means of reduplication. The formation of multiplicative stems from disyllabic stems is illustrated by the following examples:

(90) basic stem   multiplicative stem

VC(C)V   V̀C-V̀C(C)V̀
12 3 4   12 12 3 4

ìdɨ̄   ìd-ìdì   'to burn'

ōmbɛ̄   ɔ̀mb-ɔ̀mbɛ̀   'to tie'

ìgbɔ̀   ìgb-ìgbɔ̀   'to bark'

ìḍà   ìḍ-ìḍà   'to kill'

ìgwē   ìg-ìgwè   'to burn'

ūd̠wē   ùd̠-ùd̠wè   'to vomit'

ɔ̄ʔwɛ̄   ɔ̀ʔ-ɔ̀ʔwɛ̀   'to dry'

ɔ̀myā   ɔ̀m-ɔ̀myà   'to mix'

ɔ̀lyà   ɔ̀l-ɔ̀lyà   'to shiver'

The multiplicative stems in (90) consist of the first two segments of the corresponding basic stem followed by all the segments of that stem. (All the tones of the multiplicative stems are low, whatever the tones of the corresponding basic stems.) Notice, however, that this sequence of segments is exactly the same as the one that results from vowel elision applied to a verb before a cognate object or cognate adverbial. Cf. (71-78) above and the following examples:

(91) a. /ɔ̀mb-ɔ̀mbɛ̀   ʔɔ̄/   'he *did* tie them (one by one)'
         MULT-tie/3+PF C

     b. /ōmbɛ̄   ɔ̀mbɛ̀/   → [ōmbɔ̄mbɛ̀]   'he *tied* it'
         tie/3+PF tie/VA

(92) a. /ìg-ìgwè   ʔɔ̄/   'he *did* burn them (one by one)'
         MULT-burn/3+PF C

     b. /ìgwē   ìgwē/   → [ìgīgwē]   'he *burnt* it'
         burn/3+PF burn/VA

Note in particular that a glide following another consonant in the basic stem is absent from the first part of the corresponding multiplicative stem (92a), just as it is when the basic stem has been exposed to vowel elision (92b). This similarity can be accounted for by positing underlying forms with total reduplication, to which vowel elision is applied:

(93)   ɔ̀mbɛ̀-ɔ̀mbɛ̀   →   ɔ̀mb-ɔ̀mbɛ̀   'to tie'

    ìgwè-ìgwè   →   ìg-ìgwè   'to burn'

Independent evidence for such underlying forms comes from multiplicative stems formed from stems with more than two syllables.[9] In such stems the second vowel is retained in the first part of the stem, while the initial vowel of the second part of the stem is elided:

(94)   basic stem   multiplicative stem

    V̀C(C)V̀C(C)V̀X   V̀C(C)V̀-C(C)V̀C(C)V̀X

    12 3 45 6 78   12 3 4 2 3 45 6 78

    ìlèmì   ìlè-lèmì   'to curse'

    ìgbɛ̀là   ìgbɛ̀-gbɛ̀là   'to distribute'

    ìɲàrà   ìɲà-ɲàrà   'to help'

    ìrùgùŋɔ̀   ìrù-rùgùŋɔ̀   'to kneel'

    ìt̪ɔ̀mɔ̀rɔ̀dʒà   ìt̪ɔ̀-t̪ɔ̀mɔ̀rɔ̀dʒà   'to join'

That is, the stem formation rule reduplicates the first two syllables of the basic stem, whether the latter has two or more than two syllables:

(95) $V_1C_2(C_3)V_4X$ → $\dot{V}_1C_2(C_3)\dot{V}_4-\dot{V}_1C_2(C_3)\dot{V}_4X$

    (where X may be zero)

This analysis implies that vowels can be contiguous within a word in its underlying form and hence that the absence of such sequences from the phonetic

---

[9]All non-reduplicated non-derived verb stems with more than two syllables appear to be loanwords which have been adapted to the verbal system of Lulubo by addition of an initial /ì/ ∿ /ì/ . Compare the basic stems in (94) with the corresponding forms in Bari, which seems to be the most common source language: /lɪm/ 'to insult', /gwɛl-/ 'to distribute', /ŋar/ 'to do something together with somebody', /ruguŋʌ/ 'to kneel down', /tɔmɔr-ja/ 'to join'.

form of a word (cf. section 5) is due to a surface constraint. However, even in underlying forms, contiguous vowels always belong to different morphemes.

REFERENCES

Andersen, Torben. 1986a. "Verbal inflexion in Moru." *Afrika und Übersee* 69:19-43.

Andersen, Torben. 1986b. "Tone splitting and vowel quality: evidence from Lugbara." *Studies in African Linguistics* 17:55-68.

Andersen, Torben. 1986c. "The phonemic system of Madi." *Afrika und Übersee* 69:193-207.

Greenberg, Joseph H. 1966. *The Languages of Africa*. The Hague: Mouton & Co.

Persson, Andrew and Janet Persson. 1980. *Sudanese Colloquial Arabic for Beginners*. Horsleys Green: Summer Institute of Linguistics.

Spagnolo, Fr. Lorenzo M. 1960. *Bari English Italian Dictionary*. Verona: Editrice Nigrizia.

Tucker, A.N. 1940. *The Eastern Sudanic Languages*, Volume I. London: Oxford University Press.

Tucker, A.N. and M.A. Brian. 1956. *The Non-Bantu Languages of North-Eastern Africa*. London: Oxford University Press.

YORUBA RELATIVISATION AND THE CONTINUOUS SEGMENT PRINCIPLE*

Nike S. Lawal
University of Ilorin

This paper examines the strategies for relativisation and the
Noun Phrase accessibility hierarchy and constraints in Yoruba
[Keenan and Comrie 1977]. The various positions relativisable
are examined. It was found that contrary to what Keenan and
Comrie thought, Yoruba relativises all positions except the
Object of Comparative. Attention is also focussed on the
status of the coreferential pronoun found in subject relativ-
isation. From the presence of this pronoun it appears as if
Yoruba violates the continuous segment principle. However,
our analysis shows that the pronoun is a surface structure
phenomenon which obscures the underlying strategy for rela-
tivisation. The Yoruba data therefore supports the Hierarchy
Constraints.

1. Introduction

     One important issue which often comes up in the description of relative

clauses (RC) in various languages is the relativisation strategy proposed by

Keenan and Comrie [1977]. Keenan and Comrie observed that languages vary with

respect to the way relative clauses (RC) are formed. They also noted that

even within a single language there is often more than one distinct way of

forming a RC. The different ways of forming RC is what they refer to as dif-

ferent relative clause forming Strategies. Different strategies differ with

regard to which NP positions they can relativize. Their data covers fifty

languages. Essentially, their observation distinguishes two main parameters

of variation. The first concerns the position of RCs with respect to the main

---

clause. For this, three distinct types are recognised. They are postnominal, prenominal, and internal. A postnominal RC is one which is positioned immediately following the matrix constituent within the main clause, a prenominal RC is positioned immediately to the left of its head, while an internal RC is embedded internally with the main clause, and it has no matrix constituent.

Their second parameter classifies RC according to whether they are case-coding or non-case coding. A RC is case-coding if it overtly marks the grammatical role of the relative within the RC. This may be achieved either by morphological marking, the use of pronouns, prepositions or postpositions, or by position. They also observed that some syntactic positions are more accessible to relativisation than others. This they called the accessibility hierarchy. First in the hierarchy is subject position, followed by direct objects, indirect objects, obliques, and lastly, objects of comparison.

Keenan and Comrie proposed a number of universal principles which make crucial reference to this hierarchy. The most important of these as far as Yoruba is concerned is what is described as "the continuous segment principle", and I quote:

> "For unmarked simple sentences, any relative clause-forming strategy must operate on a continuous segment of the accessibility hierarchy."

This in effect means that a given RC-forming strategy can be used to relativize only continuous segments of the hierarchy. The continuous segment principle has been tested on many languages, including Yoruba. This paper discusses the proposal as it relates to Yoruba data. The focus is on the coreferential subject pronoun in Yoruba subject relatives. But first we consider the range of constituents relativizable in the language.

2. Range of Constituents Relativizable in Yoruba

Yoruba was one of the fifty languages examined by Keenan and Comrie. According to Keenan and Comrie, Yoruba relativises only subject (SU), Direct Object (DO), and Genitive (Gen) positions; the Indirect Object (IO), Oblique (Obl.), and Object of Comparison (Ocomp) are not relativizable. These three positions are therefore marked with asterisks in their paper. However, the

following sentences indicate that Yoruba relativizes all positions:

(1) ọkùnrin tí ó ra   aṣọ   ti dé
     man   that he bought cloth has come

     'the man who bought the cloth has come'

(2) aṣọ   tí ọkùnrin náà rà   dára
     cloth that man   the bought good

     'the cloth that the man bought is nice'

(3) ọmọ   tí ọkùnrin náà ra   aṣọ   fún ti dé
     child that man   the bought cloth for has come

     'the child that the man bought the cloth for has come'

(4) àdá   tí mo fi gé igi mú
     cutlass that I with cut tree sharp

     'the cutlass that I cut the tree with is sharp'

(5) iná   tí mo gbé ọbẹ̀ kà ti kú
     fire that I put soup on has died

     'the fire on which I put the soup is dead'

(6) obìnrin tí awọn olè gbé mọto rẹ̀ ti kú
     woman   that plural thief took car her has died

     'the woman that thieves stole her car has died'

(7) obìnrin tí ọkùnrin náà ba   lọ ti de
     woman   that man   the with go has arrived

     'the woman that the man went with has arrived'

(8) ọmọ   tí ọkùnrin náà ga jù   ti dé
     child that man   the tall exceed has arrived

     'the child that the man is taller than has arrived'

The examples (1) to (8) illustrate the full range of constituent types that
can be relativised: (1) is an example of a subject being relativised; (2) is
that of direct object; (3) is an indirect object; (4), (5), and (7) are ob-
liques; (6) is a genitive; (8) is object of comparison. The three positions
which have been a subject of dispute are the indirect object, oblique, and ob-
ject of comparison. Keenan and Comrie for instance do not regard the Yoruba
RC like (3), (4), (5), (7), and (8) as instances of IO, Oblique, and OComp con-

structions respectively. Therefore these positions are marked with asterisks
in their table of relativisable positions. Their reason is that the words
fún , bá , ka , and jù which precede the relativised NPs are verbs in ser-
ial verbal constructions and not prepositions. If their interpretation is cor-
rect the implication is that the NPs which are complements to these words
would be direct objects. Thus there would be no special IO, Oblique, and
OComp constructions in the language. We consider Keenan and Comrie's claim to
be valid only in the case of the OComp position (8). The word jù in (8) is
in all respects a verb. However, this is not the case with the words fún ,
fi , ka , and bá in (3), (4), (5), and (7) respectively.

It is true as some linguists have said, that these words were historically
verbs, and it is this which probably influenced Keenan and Comrie in their
analysis. However, the concern here is not and should not be with the dia-
chronic status of these words. What is at issue here is the present status of
these words, and linguists are now reanalysing these words as prepositions or
case markers. An explicit account of the present status of these words can be
found in Awobuluyi [1978] and Lawal [1986] amongst others and will not be dis-
cussed here.

3. Pronoun Retention and the Continuous Segment Principle

According to the Continuous Segment Principle quoted in section 1 above, a
given RC-forming strategy can be used to relativise only continuous segments
of the accessibility hierarchy. For example, if a language has two relativi-
sation strategies, a case coding strategy and a non-case coding strategy, and
if the case coding strategy is used to relativise only subjects, indirect ob-
jects, and Obliques, this would in effect leave a gap, i.e. the direct object
position relativised by a different strategy would destroy the continuity.

Case is coded either by means of a personal pronoun or a stranded preposi-
tion within the relative clause. Now consider the Yoruba RCs below:

(9) bàbá   [tí o ra   bata]
     father   that he bought shoes
     'father who bought shoes'

(10) bàtà [ti [baba   ra   Ø]
      shoe that father bought

      'the shoe that father bought'

(11) ọmọ   [ti [bàbá   ra   bàtà fún Ø]]
      child   that father the bought shoe for

      'the child that father bought shoes for'

(12) àdá   [ti [mo fi Ø be ìsu]]
      cutlass   that I with cut yam

      'the cutlass that I cut the yams with'

(13) iná   [ti [mo gbé ọbẹ ka Ø]]
      fire   that I   put soup on

      'that fire that I put the soup on'

(14) ọkùnrin [tí [obìnrin náà ba   Ø lo]]
      man   that woman   the with went

      'the man that the women went with'

(15) ọkùnrin [ti [awon ole   ji   mọtò rè]]
      man   that they thieves stole car his

      'the man whose car the thieves stole'

(16) ọkùnrin [ti [mo ga ju   Ø]]
      man   that I tall surpass

      'the man that I am taller than'

Starting with (9) we can see there is no missing subject in the RC. The rela-
tivized constituent which is subject is represented in the RC in the form of a
personal pronoun. In (10) on the other hand the relativized NP which is a di-
rect object is not represented in the RC; the NPrel has been deleted. The
same applies to all the other positions except the genitive position (15)
where the NPrel is pronominalised. We have decided to leave out example (16)
in this analysis because we agree with Keenan and Comrie that it is not a true
object of comparison. The NPrel here has the status of a direct object.

     Now the indication from this analysis is that Yoruba has two strategies, a
case coding strategy (pronominalisation) for subjects and genitives and a non
case coding strategy (deletion) for direct objects, Indirect Objects, and Ob-
liques. In the latter, there is no nominal element in the RC that clearly ex-

presses which NP position is being relativized.

This strategy appears to constitute a counterexample to the Hierarchy Constraint. When we subject it to the Continuous Segment Principle the RC forming strategies in Yoruba do not operate on a continuous segment. It shows a gap along the segment of the hierarchy brought on by the use of a different strategy for direct objects, indirect objects, and obliques.

The Yoruba data can also be interpreted in another way, i.e. depending on how we interpret the notion "case coding". If we take case coding to include stranded prepositions as some linguists do, then Yoruba RC strategy can be said to be case coding for all positions except the direct object. The stranded prepositions in this case are fún , fi , kà , and bá . Whichever interpretation we choose, the implications appear to be the same, which is that there is a gap along the segment. With this second interpretation the gap would be brought on by the use of a different strategy for direct objects which is second on the accessibility hierarchy.

We shall in this paper adopt the first approach, which means case coding for subjects and genitives and non-case coding for all other positions viz, direct objects, indirect objects, and obliques.

This violation of the Continuous Segment Principle in Yoruba was noticed earlier by Keenan and Comrie themselves. To explain the "violation" Keenan and Comrie [1977] claimed that the pronoun found in Yoruba subject RC is not a pronoun at all but a case of agreement. This in effect would mean that subject, direct object, indirect object, and oblique positions are all non-case coding. Only the genitive which is lowest on the Yoruba hierarchy would then be case coding, thus preserving the continuity.

4.   The Case Against the Pronoun ó as an Agreement Marker

We disagree with the claim that the pronoun in subject relatives is an agreement marker. Firstly, there is no evidence that supports such a claim. For instance, Hausa, with which Keenan and Comrie by implication compare Yoruba, is very much unlike Yoruba. Yoruba for example does not retain a full pronoun in simplex sentences with a subject NP as Hausa does. Thus the following are unacceptable in Yoruba:

(17) a. *Olú ó wá
   Olu he came

(18) a. *bàbá ó lọ
   father he went

The correct sentences are as follows:

(17) b. Olú wá
   Olu came

(18) b. bàbá lọ
   father went

Compare the above with an equivalent sentence in Hausa:

(19) a. Musa ya zo
   Musa he came

   b. *Musa zo
   Musa came

  Secondly, there is no evidence of full subject NPs accompanied by Clitic pronouns in Yoruba as justified in Givón [1976]. The pronoun present in Yoruba subject RC is a full pronoun, not a Clitic. This is illustrated with the sentence below:

(20) ó   ra   bàtà   'he/she bought shoes'
   he/she bought shoes

(21) ó   ń   sùn   'he/she is sleeping'
   he/she prog sleep

  The analysis used for Hausa and ergative languages like Tongan does not work for Yoruba.

  Another linguist who also disagrees with Keenan and Comrie is Stalhke [1976]. Stalhke is of the opinion that the element ó found in Yoruba subject RC is not an agreement morpheme as claimed by Keenan and Comrie. Stalhke argues that this ó occurs with both plural and singular subjects in focus constructions. His examples are reproduced as (22a-d) below:

(22) a. èmi ni   ó   lọ   'it is I that went'
    I   emph he/she went

   b. ìwọ ni   ó   lọ   'it is you that went'
    you emph he/she went

c. àwa ni   ó   lọ   'it is we that went'
    we   emph he/she went
d. àwọn ni ó   lọ   'it is they that went'
    they emph he/she went

Stalhke argues that if ó were an agreement morpheme, this would be diffi-
cult to explain. As he says, "It would mean claiming that the third person
singular pronoun can agree with any other person/number combination." Such a
claim, he points out, cannot be supported from any other area of Yoruba syntax,
for example, plural examples are ungrammatical with just number agreement.
Consider the examples below from Stalhke:

(22) e. *ẹ̀yin   ni   ó   lọ
         you (Pl) emph they went
     f. *àwa ni   wón lọ
         we   emph they went

The above are ungrammatical because there is no person agreement. The pro-
nouns must agree both in number and person. Stalhke's conclusion therefore is
that ó is not an agreement morpheme.

Although Stalhke's argument is against ó as an agreement marker, it also
seems to be an argument against ó as a subject pronoun. Stalhke states, "It
would be an odd subject pronoun which remained constant across this set of sen-
tences."

According to Stalhke, ó is neither a subject pronoun nor an agreement mor-
pheme. But what is it? Unfortunately, Stalhke does not seem to have an answer
That ó is a full subject pronoun is not in doubt as it occurs in simple sen-
tences like ó ra bàtà fún mi 'he bought shoes for me'. It is its status in
RC that is in doubt.

Apart from RC there is only one other construction in which a pronoun fol-
lows a full subject NP. This is focus construction. This has been used as ev-
idence in favour of analysing ó as an agreement morpheme. However, it should
not be surprising that the ó occurs also in focus constructions. Studies
have shown that many languages employ the same strategies for both focus forma-
tion and RC formation [Schachter 1973, Madugu 1982]. In all the languages ex-

amined by Schachter, the rules for forming RC and focus clauses are alike with
regard to such details as the substitution of an appropriate personal pronoun
for one of the NPs in the sentence or the insertion of an invariant marker at
the beginning of the clause and an optional deletion of the relative pronoun
in certain cases.

Yoruba is no exception in this regard. Thus, to form focus clauses in Yo-
ruba the same rules as used for relative clauses are employed, except that the
markers inserted at the beginning of the two clauses are different: the focus
marker is ni while the relative marker is tí . The following examples il-
lustrate this.

(23) a. ọ̀rẹ́   mi tí ó ra   bàtà   (Subject RC)
        friend my that he bought shoes   'my friend that bought shoes'

     b. ọ̀rẹ́   mi ni ó ra   bàtà   (Subject focus)
        friend my emph he bought shoes   'it is my friend that bought shoes'

(24) a. ọ̀rẹ́   mi tí mo ri   (Object RC)
        friend my that I saw   'my friend that I saw'

     b. ọ̀rẹ́   mi ni mo rí   (Object focus)
        friend my emph I saw   'it is my friend that I saw'

(25) a. ọ̀rẹ́   mi tí mo ra   bàtà rẹ̀ (Genitive RC)
        friend my that I bought shoes her   'my friend whose shoes I bought'

     b. ọ̀rẹ́   mi ni mo ra   bàtà rẹ̀ (Genitive focus)
        friend my emph I bought shoes her   'it is my friend whose shoes I
                                              bought'

(26) a. ọ̀rẹ́   mi tí mo ga   jù   (OComp RC)
        friend my that I tall pass   'my friend that I am taller than'

     b. ọ̀rẹ́   mi ni mo ga   jù   (OComp focus)
        friend my emph I tall pass   'it is my friend I am taller than'

Just.as in the RC, a pronoun is retained when Subject NP and Genitive NP
are being focussed, while all other positions do not retain pronouns. These
similarities between focus sentences and RC correlate with their semantics:
both constructions are said to have a common semantic property which is "fore-
grounding of one part of a sentence at the expense of the rest" [Schachter
1973]. Evidence based on a comparison with focus sentences in the language
does not therefore constitute independent evidence for the pronoun as an agree-

ment marker. It would be odd if the pronoun were to be absent in focus con-
structions since the same strategy is employed for both constructions. Before
we conclude this section we want to draw attention to another type of subject
relativization which constitutes additional evidence against the pronoun ó
as a case marker. Contrast the (a) sentences below with the (b) sentences:

(27) a. ìwọ obìnrin tí   ó   gbó   gẹ̀ẹ́sì
         you woman   that she hears English

         'you woman who understands English'

     b. ìwọ obìnrin tí   ọ   gbọ́   gẹ̀ẹ́sì
         you woman   that you hear English

         'you woman who understands English'

(28) a. àwa obìnrin tí   ó   gbọ́   gẹ̀ẹ́sì
         we women   that she hear English

         'we women who understand English'

     b. àwa obìnrin tí   a   gbọ́   gẹ̀ẹ́sì
         we women   that we hear English

         'we women who understand English'

(29) a. àwọn obìnrin tí   ó   gbọ́   gẹ̀ẹ́sì
         them women   that she hear English

         'those women who understand English'

     b. àwọn obìnrin tí   wọn   gbọ́   gẹ̀ẹ́sì
         them women   that they hear English

         'those women who understand English'

(30) a. ẹ̀yin   obìnrin tí   ó   gbọ́   gẹ̀ẹ́sì
         you (pl) women.   that she hear English

         'you (pl) women who understand English'

     b. ẹ̀yin   obìnrin tí   ẹ   gbọ   gẹ̀ẹ́sì
         you (pl) women   that you (pl) hear English

         'you (pl) women who understand English'

In the (b) sentences we have the full range of pronouns, i.e. we have 1st, 2nd,
and 3rd persons as heads of the RC, and the pronouns agree in number and per-
son with their antecedents. This is unlike the (a) sentences where the 3rd

person singular pronoun ó is used for all persons and number. The fact that the surface subject in the (a) sentences of (27-30) does not need to show agreement even strengthens the claim that the pronoun ó is not an agreement or case marker.

4. The Status of the Pronoun in Subject Relatives

To account for the pronoun ó in subject relatives it is necessary to examine the facts of the language. Perlmutter [1970] observed that for various reasons a class of surface structure constraints or input conditions must be available as one of the devices used by grammars to delimit the class of grammatical sentences, in other words, all languages have certain well-formedness conditions which determine which sentences are well formed and which are not. English for instance, is said to have such a constraint, which can be stated, "No non-imperative sentence which does not have a subject in its surface structure is well formed." This is a surface structure constraint about English.

Yoruba like other languages has surface structure constraints. One such constraint is that "no top-most sentence in Yoruba may be verbless." Another is "no tensed sentence or clause may be subjectless" [Lawal 1985]. We must note, however, that this constraint does not apply to negative sentences where the subject is a 3rd person pronoun. In Yoruba when the 3rd person subject pronoun of a sentence comes in contact with the negator kò , the pronoun gets elided. However, the sentence would still be understood as containing a subject, even though the subject no longer has an independent existence. The subject pronoun more or less fuses with the negator kò . Such sentences cannot in fact be said to be subjectless. The following is an example of this construction:

(31) kò   pa á   'he/she didn't kill it'
     he/she not kill it

The negator kò incorporates the subject. Thus the sentence cannot be interpreted as being subjectless. This phenomenon is restricted to the negator kò only. Thus (32) is ungrammatical:

(32) *ò   pa   á
       not kill it

We can now account for the pronoun in subject RC. The pronoun is there to
satisfy the above Yoruba surface structure constraint and has nothing to do
with the underlying strategy for relativization. It simply obscures the un-
derlying strategy of deletion of the coreferential NP, the same strategy as
used for direct objects.

5. Conclusions

We have shown firstly, that Yoruba relativizes all NP positions and second-
ly, that Yoruba obeys the Continuous Segment Principle. The pronoun in sub-
ject RC is a surface structure phenomenon which obscures the underlying strat-
egy for relativization.

REFERENCES

Awobuluyi, A.O. 1978. *Essentials of Yoruba Grammar.* Ibadan: Oxford Univer-
sity Press, Nigeria.

Givón, T. 1976. "Topic, pronoun, and agreement." In C. Li (ed.), *Subject
and Topic*, pp. 149-188. New York: Academic Press.

Keenan, L. and B. Comrie. 1977. "Noun phrase accessibility and universal
grammar." *Linguistic Inquiry* 8:63-99.

Lawal, S.N. 1985. "Why verbs do not gap in Yoruba." *Journal of African Lan-
guages and Linguistics* 7:155-162.

Lawal, S.N. 1986. "The classification of the Yoruba serial verbal construc-
tion." Submitted to *JOLAN*.

Madugu, I.S. 1982. "The na-na construction in Nupe." *JOLAN* 1:34-45.

Perlmutter, D. 1970. *Deep and Surface Structure Constraints in Syntax.*
New York: Holt, Rinehart and Winston.

Schachter, P. 1973. "Focus and relativization." *Language* 49:19-49.

Stalhke, H. 1974. "Pronouns and islands in Yoruba." *Studies in African Lin-
guistics* 5:171-202.

DISCOURSE DISTANCES AND
THE SWAHILI DEMONSTRATIVES*

Timothy Wilt
Michigan State University

Proximity is the key concept for understanding a speaker's
choice of Swahili demonstrative forms. The alternative
approach in Leonard [1985] that proposes a speaker's con-
centration of attention as the key concept overlooks sig-
nificant aspects of text structure. Proximity must not be
conceived of, as does Leonard, only in terms of spatial
distance between the speaker and a referent. Temporal,
narrative, and anaphoric distance must also be considered.

Leonard [1982, 1985] is the first in-depth study of the Swahili deictic
system and one of the few studies of Swahili grammar to be done from a dis-
course perspective. He argues that the choice between H, e.g. hii 'this',
H-o , e.g. hiyo 'this/that', and LE, e.g. ile 'that', demonstratives in
Swahili basically depends on the "noteworthiness of" or the "relative concen-
tration of attention on" a referent rather than, as in traditional analyses,
on the (non-)proximity of the referent to the speaker. The most noteworthy
referents, he claims, will tend to be referred to with H forms, less noteworthy
referents with H-o forms, and the least noteworthy with LE forms. Since new
items catch one's attention more than old, he then argues, one would expect
that new items will tend to be marked by H while old items will tend to be
marked by LE. He claims that a statistical analysis of four chapters randomly
selected from five novels indicates this H-for-new and LE-for-old tendency.

In this paper, I will reexamine three chapters (two different authors)

---

that Leonard [1985], hereafter referred to as L, used for his statistical anal-
ysis and argue that proximity is indeed the key criterion for determining
which demonstrative form will be used. Like L, I will limit my analysis to
that of the H and LE forms.

1. Varying Discourse Distances

    The seventh chapter of Abdulla's [1960] *Mzimu wa Watu wa Kale* (MZ) is en-
titled "Mfuko Mweusi" ("A Black Purse"). Thus, it is immediately evident what
one of the key props in the chapter will be. L's "Concentration of Attention"
(COA) hypothesis would predict that mfuko 'purse' will be most frequently re-
ferred to by the H demonstrative huu since there will be high COA on the ref-
erent. However, his "New v. Old Information" (N/O) hypothesis makes the oppo-
site prediction that mfuko will be most frequently referred to by LE since,
after the first reference, it will be "old" information. L pays little atten-
tion to the conflicting predictions of these hypotheses.

    A purely quantitative analysis of the demonstratives used with mfuko and
the chapter's other key prop karatasi 'piece of paper [found in the purse]'
reveals no significant difference in the use of the demonstratives:

	H	LE
mfuko	3	4
karatasi	3	4

Table 1: Demonstratives with key props (MZ, ch.7)

    However, closer attention to the organization of the text reveals a cate-
gorical difference between the narrator's use of H and LE that is predicted by
neither COA nor N/O. The rule is simply this: whenever mfuko or karatasi
is referred to by the narrator, he uses the LE form; whenever either of the
two is referred to by a character, the character uses the H form:

	H		LE	
	mfuko	karatasi	mfuko	karatasi
Narrator	0	0	4	4
Character speaking	3	3	0	0

Table 2: Narrator v. characters' use of demonstratives (MZ,ch.7)

Explaining this difference in terms of COA, one would have to argue that the narrator has "low concentration" on the props whereas the characters have "high concentration". This is unsatisfactory, especially in light of the title assigned by the narrator to the chapter. N/O is also unsupported since the use of the demonstrative here has nothing to do with whether or not the associated referent is mentioned for the first time.

A better explanation is that the narrator's use of LE corresponds with his lack of participation in the scene and, thus, his referring to the objects from.a narrative distance (third person voice rather than first person in which this distance would not be as consistently kept). In contrast, the characters use H to refer to objects in their immediate presence.

This distinction between the narrator's and the characters' use of demonstratives is maintained throughout the chapter. In reference to items other than the purse and the paper, characters use LE only when referring to an event or object of a past speech situation that the narrator separates from the present situation by means of a chapter division. Conversely, the narrator uses H only in (1) and in a structure which signals the immediate temporal proximity of two events (2).

(1)   upande huu na huu
      side   this and this

(2)   CLAUSE + na + huku + CLAUSE[1]
                    and here

In (1), the coupling of demonstratives conveys the same idea as the English "this (side) and that". In (2), a construction frequently used by Abdullah, huku is used to signal that the event referred to in the second clause overlaps with, or immediately follows, the event mentioned in the first clause of the sentence.

The above analysis supports the traditional one which distinguishes between

---

[1]For example, ...alisema hivyo na huku anawasha toza yake... (MZ,26) '...he said this while lighting his pipe...'

H and LE forms in terms of proximity/non-proximity (P/NP). However, it also shows that P/NP must not, in contrast to L's presentation, be considered only in relation to spatial distances but also in relation to temporal and discourse distances.

With this expanded concept of P/NP, all uses of H and LE in MZ's ninth chapter can also be accounted for. The initial distinction between the demonstratives can be made as in Table 3:

	Narrator's uses	Characters' uses
H	6	10
LE	36	9

Table 3: Apportionment of demonstratives in MZ, ch. 9

To be explained are the narrator's six uses of H and the characters' nine uses of LE.

The characters' uses of LE are readily accounted for in terms of spatial and temporal non-proximity. Eight of the nine LE forms are used to refer to a person with whom the two main characters were speaking earlier but from whom they are now separated as they talk about him. The other LE is used to refer to an action of that person (observed by the characters while with him but talked about now that they are separated from him).

Two of the narrator's six H forms occur in the na huku structure mentioned above. Two others occur in the adverbial phrase hivi sasa 'right now; immediately'; like na huku , this phrase signals close temporal proximity of two events. This leaves only the following two occurrences of H to be explained:

(3) Bali safari hii... (MZ,42)

   'But, *this time* (lit., this trip)...'

(4) Mlango wa kutokea haukuwa mdogo kama wa kwanza walioingilia. Upande huu
    ulikuwa mkubwa... (MZ,39)

   'The door for going out was not small like the first one through which
    they entered. *This side* was large.'

In both (3) and (4), the H is used to contrast one referent with another that is spatially and/or temporally removed from the present narrative situation.

Table 3 can now be revised as Table 4 to make the distinction between H and LE, as used by Abdullah, even clearer:

	Non-Proximity		Proximity	
	Narrator	Characters	Narrator	Characters
	Distanced from characters (3rd pers.)	Refer to items outside present speech situation	Indicates spatial/ temporal prox. of one referent cf. to another	Refer to items in immediate speech situation
H	0	0	6	10
LE	36	9	0	0

Table 4: Demonstrative allocation with respect to P/NP (MZ,ch. 9)

In sum, all but one[2] of the ninety-four occurrences of H and LE in MZ's chapters 7 and 9 have been accounted for in terms of spatial, temporal, and narrator vs. character P/NP. We shall now consider a different writer's use of demonstratives and see that yet another type of discourse distance may be involved in determining the choice of demonstrative forms.

2. Anaphoric Distance

L (p. 283) points out that "the supposed 'proximate' is used to refer to a location that is quite distant from the speaker" in (5), a passage from Kezilahabi's *Kichwa Maji* (KM).

(5) Upande mashariki niliweza kuona nyumba ya Baba ikitokeza juu ya miti. Nilikumbuka kwamba hapa[3] mtoto alikuwa akizaliwa. (KM,88)

'In the East I could see father's house sticking out above the trees. I remembered that *there* a child was being born.' (L's translation)

L argues that H is used to refer to the location where "the pivotal incident

---

[2] huku na huku , mentioned above.

[3] Hapa is usually translated 'here' but L translates it 'there' in this instance to reinforce his point about the spatial distance between the speaker's location and the house.

of the whole novel" is occurring and thus the speaker is giving it a high de-
gree of attention (p. 284). Similarly, he quotes a paragraph in which the
first reference to a group of cows is with LE but the subsequent two references
are with H. He concludes: "... at first the narrator views the cows with no
special interest and refers to them with LE... But when he sees them as being
directly related to a central concern, he refers to them with H" (p. 285).

However, on the page in KM following the scene concerned with the group of
cows, there is a parallel passage where the use of the demonstratives is not
consistent with the COA prediction:

(6) Juu ya jiwe nililona kibwawa kidogo cha maji. Ndani ya <u>maji haya</u>[4] nililona
     nyuki ameanguka akiogelea.... Mwanzoni nilifurahi kumwona <u>huyu mdudu</u>
     akipigania maisha.... Nilijiona mungu mdogo. Nilitazama tena pande zote,
     watoto, nyumba na miti, vyote vilikuwa chini yangu.

         Nilichukua kijiti kidogo sana ambacho kwacho nilimtoa <u>yule nyuki</u>
     majini.... Kwa <u>huyu mdudu</u> nilikuwa na nguvu ambazo haziwezi kufahamika.
     (KM,88)

     'On the rock I saw a puddle of water. In *this water* I saw a bee, which
     had fallen in, swimming.... At first I was happy to see *this insect*
     struggling for life.... I saw myself as a little god. I again looked
     around; children, houses and trees, all were beneath me.

         I took a twig which I used to take *that/the* bee from the water....
     For *this insect*, I had incomprehensible powers.'

The bee, like the group of cows, is focused on by the narrator as he develops
his philosophical musings. But, the demonstrative form changes from H to LE
to H while the concentration of attention remains constant.[5]

In all of the above-mentioned passages, however, H is used when the refer-
ent of the NP it modifies has been referred to in the preceding sentences of

---

[4]H is used with maji 'water' even though the concentration of attention
seems to be on the bee rather than the water. It might be argued that COA is
high for both. But such argumentation points to another major weakness of the
COA hypothesis: no objective criteria are given for determining the degree of
COA on a referent.

[5]N/O cannot explain the choice of demonstratives since "bee" is already
old information when H is used for the first time and H is used again after
the switch to LE.

the same paragraph. In (6) and in the passage concerning the cows, LE is used when the referent of the NP has been referred to most recently in a previous paragraph. Throughout the chapter, the demonstratives tend to be distributed in this way, as Table 5 shows.

	Anaphoric tie within paragraph	Anaphoric tie across paragraph boundaries
H	35	2
LE	2	9

Table 5: <u>Demonstrative forms related to anaphoric distance (KM,ch. 6)</u>[6]

The notion of proximity/distance is again central to understanding the use of the demonstratives here, but rather than temporal, spatial, or narrator-character distance, the relevant factor here is textual distance, the distance between an anaphor and the item to which it points.[7]

3. <u>Demonstratives and Discourse Style</u>

As well as indicating the relevance of textual proximity to demonstrative form, Table 5 shows that the large majority of anaphoric references in KM's chapter 6 are H. This is contrary to the prediction of N/O that LE will tend to be used if the referent has been previously mentioned. It also indicates that the use of demonstratives in KM is different from their use in the other novels analyzed by L for his statistical correlation between the choice of demonstrative form and whether or not the associated referent was mentioned for the first time (p. 288). In KM's chapter 6, 80% of the anaphoric references are made by H (in KM's first two chapters, the percentage is about the same); but, according to L (p. 288), in four chapters (including KM's ch. 6) from three different novels, only 38% of anaphoric references are H. Subtracting the KM entry from this total, the figure will be 11%. L did not consider the possibility of stylistic differences in the use of demonstratives, which this

---

[6]See Appendix A for a discussion of the exceptions.

[7]The concepts of textual distance and anaphoric ties are discussed in detail in Halliday and Hasan [1976].

statistical disparity indicates.

In MZ, as we have seen, the narrator's distance from the characters is more determinant of demonstrative form than anaphoric distance; in reference to the key props of MZ's chapter 7, the narrator always uses LE even though the same prop has been mentioned earlier in the paragraph, e.g. MZ, p. 25. This difference between the MZ and KM uses of the demonstratives may be due to the difference between third person (MZ) and first person (KM) narrative style; a general difference in discourse style of the two different authors; and/or the differing strength of various rules in determining which type of discourse distance (temporal, spatial, narrator vs. character, or anaphoric) is most salient.

4. Physical Distance and Demonstrative Form

L claimed that "throughout modern Swahili novels and plays we regularly encounter referents of H and LE in locations opposite to those indicated by the traditional hypothesis" (p. 283). Although this was a key argument for proposing an alternative hypothesis to account for the Swahili demonstratives, L only supported it anecdotally with two examples (both referred to in Section 2, first paragraph, above) in which factors of textual distance, though relevant, were not considered. L's quantitative statement, however, can be quantitatively tested. To do so, it is best to focus on the exophoric references of quoted speech, in which temporal and spatial factors of the speech situation are most salient, in order to avoid complicating factors such as narrative voice and anaphoric distance that affect the form of endophoric references.

We have already observed that the MZ characters used H only for referents in their immediate speech situation and LE only for those removed from their immediate speech situation (23 H, 13 LE occurrences in chs. 7 and 9). A listing of the thirty-five H/LE demonstratives used in the quoted speech of characters in KM, throughout three chapters, indicates that here, too, there is a 100% correlation between H and spatial/temporal proximity of the referent and between LE and non-proximity (see Appendix B).

5. Conclusion

The above analysis opposes L's claim that P/NP "can only account for a small part of the distribution of [Swahili demonstrative forms] in actual

texts" (p. 281). It also shows that various discourse, as well as spatial and temporal, distances play an important role in the choice of the demonstrative form. Statistical counts based on P/NP account for a much greater amount of material than L's tally based on O/N (p. 288).[8]

Some issues for further research are discourse restraints on the positioning of the demonstratives (they may occur either before or after the head noun); the discourse function of H-o ; whether or not one type of discourse distance is consistently more determinate of the demonstrative form than others; and stylistic differences between authors in their use of the demonstratives.

---

[8]This O/N tally is of questionable value anyway since it does not distinguish between endophoric and exophoric references.

APPENDIX A

Table 5 (Section 2) shows that, in KM's chapter 6, there are two exceptions each to the rule that within-paragraph anaphoric ties are made with H while cross-paragraph ties are made by LE. These exceptions are listed and briefly discussed below.

Cross-Paragraph tie but H form

(a)   ....Jina "Furaha" ambalo ni neno jingine la adhabu.
          Hii ni adhabu ambayo.... (p. 94)

      '....The name "Happy one" which is another name of punishment.
          This is a punishment which....'

(b)   Baada ya siku hizi tatu.... (p. 97)
      'After *these three days*....'

In (a), the cross-paragraph tie is between the last word of the preceding paragraph and the first one of the new paragraph. There are no LE cross-paragraph references in which the tie is as close as this (in chs. 1 and 2, as well as 6).

In (b), the H appears in the first NP of the paragraph whose aperture parallels those of the preceding two paragraphs:

¶Kwa muda wa siku tatu....   'For a period of three days....'

¶Baada ya siku tatu....   'After three days....'

¶Baada ya siku hizi tatu....   'After these three days....'

Here, the H seems to be used to signal temporal proximity as well as anaphoric relationship. The events to be reported in the third paragraph immediately follow after the three days mentioned in the first paragraph and overlap with the events mentioned in the second paragraph. There are no days intervening between the first period (of mourning) and the second (when villagers offer beer to the family of the deceased).

Within-paragraph tie but LE form

(c) Labda mwanadamu anaweza kusema kwamba kuna mwenye uwezo zaidi kuliko yule tunayemfikiria kuwa Mungu. Lakini kama ni <u>vile</u> tutakwenda nyuma mpaka wapi? (p. 89)

'Maybe even a human can say that there is one having more power than the one whom we consider to be God. But if it is *that*, how far back will we go?' (the conditional clause could be less literally translated: "But if such is the case" or "But if it is so".)

(d) ....nilitazama nyuma. Niliona /a girl in trouble/.... Nilikimbia mara moja kwenda kumwokoa huyu msichana... Nilipofika <u>pale mahali</u> sikujua la kufanya. (p. 90)

'....I looked back. I saw /a girl in trouble/.... I ran right away to go and save this girl... When I arrived at *that place*, I did not know what to do.'

In KM's ch. 6 (also, chs. 1,2), vile 'like that, so, thus' is used in reference to abstract ideas or generalizations while the H form hivi is used in reference either to a specific event/item or to an approximate measure of space or time. For example:

Kalia, why have you done this/ hivi ? (p. 90)

    Referent: attempted rape of a girl

When he finished saying this/ hivi ... (p. 90)

    Referent: speech reported by direct quote

...about/ hivi fifty paces... (p. 98)

...matters concerning girls have gotten into you such/ vile that you have forgotten the people of your house. (43)

We humans are created so/ vile that we trust in things of the future more than those of the past. (p. 93)

It may be that this semantic distinction between vile and hivi is more determinate of the form than anaphoric distance; thus, vile is used in (c).

In (d), the use of the LE form provides a from-behind perspective of the narrator's movement from one place to another. He has the reader watch him as he runs from the rock to go to the new scene of action rather than see him from in front as he comes to the action. Thus, it seems, here, that the deci-

sion to signal the physical spatial relationship of one referent to another
takes precedent over the choice to signal anaphoric proximity.

Note the difficulty of analyzing (c) and (d) in terms of COA. In (c), the
referent of vile is the startling implication of a parable from nature (see
(6)) which, when thought about, caused the narrator's head to hurt. In the
paragraph containing (d), the narrator's concentration of attention on pale
mahali 'that place' continually increases up to his arrival there.

> I heard someone crying out... I immediately stood up so that I
> could hear better where the voice came from. When I heard the
> voice a second time I looked back. I saw the girl.... I saw
> only the grass being moved.... I ran...to go and save her...
> When I arrived at *that place*... (p. 90)

APPENDIX B

   The following lists the occurrences in KM's chapters 1, 2, and 6 of the H
and LE demonstratives (with their corresponding head nouns, if explicit) which
are used in quoted speech to refer to a physical location or time period.   In
the right hand column is the English translation of the context in which the
demonstrative appears (the translation of the demonstrative is in italics).
The "Proximate/Non-proximate" columns (P, N-P) indicate the correlation between
demonstrative form and (non-)proximity.

Page	Demonstrative	P	NP	Context
7	humu	H		He comes in *here* /where the speaker is located/
	humu	H		with a stick! He thinks there's a war in *here*!
9a	hapa	H		These things did not bring me *here* /to the office in which we are speaking/ .
b	humu	H		You will be the last one to come in *here* /where we are speaking/ .
c	siku hizi	H		Where are you *these days*?
10	muda wote huu	H		For *all this time* /that we have been together in this room/ you hadn't recognized me /until I came up to introduce myself/ .
	siku hizi	H		Where are you *these days*?
	siku hizi	H		*These days* I'm working.
12	shimo hili	H		When they came up to me, I said "*This hole* /that you see here/ is so deep..."
15	siku hizi	H		You don't see people *these days*?
25	siku hizi	H		My heart is not happy *these days*.
	hapa	H		When you brought me *here* /where I am writing this letter/ ...
27	Huku	H		You would have been better off staying there. *Here*, you have brought yourself into danger.
30	mwaka huu	H		The oranges have ripened; *this year* /in which we are/ , very many!
32	Hii	H		*This* /letter I have with me/ *here*, you can read
	hapa	H		

Page	Demonstrative	P	NP	Context
33	siku hizi	H		The young people of *these days*!
	huku	H		When you were travelling to come *here* /the area in which we are speaking/ , you did not get to see her?
38	pale pale		LE	She is *there*/place away from, not visible to, speakers/ (X2)
	pale mahali		LE	A: She said /she would be at/ *that place* of yours.... She said don't take long. B: Thank you... I'll go right now.
41	siku hizi	H		...because Mama is keeping a close eye on me *these days*.
	Huku huku	H H		A: You're going which way? B: By *here*. A: I'm going *this way/here*.
42	hii barua	H		He picked up one of my books from off the table and took out a letter. "*This* is your letter" /he said/.
43	hapa nyumbani	H		Since you don't know the news of the *house here* /in which we are speaking/ , I'll tell you.
	hapa	H		Kazimoto, since you have arrived *here*, your relative has refused to eat.
45	hapa	H		You know that we were brought *here* /where we now live and are speaking/ by the government.
	hapa	H		It's possible that even now they are *here* listening to us.
85	siku hizi	H		*These days* they have ceased...
86	Wale!		LE	A: "*There*!" He showed us three cows. B: "Let's go /to where they are/ ; we'll help you.
87	upande huu	H		Stand in *this area* so that they /the cows/ won't go back there /from where we just brought them/ .
98	hapa muji ule	H 	 LE	A homestead was visible ahead of us. "If I kill him *here* /where we are now/ , *that homestead* /the people of it/ will hear," I said to myself.

REFERENCES

Abdulla, Muhammed Said. 1960. *Mzimu wa watu wa kale*. Nairobi: East African Literature Bureau.

Halliday, M.A.K. and Ruqaiya Hasan. 1976. *Cohesion in English*. London: Longman.

Kezilahabi, E. 1974. *Kichwa maji*. Dar es Salaam: East African Publishing House.

Leonard, Robert A. 1982. "The semantic system of deixis in Standard Swahili." Unpublished doctoral dissertation, Columbia University.

Leonard, Robert A. 1985. "Swahili demonstratives: evaluating the validity of competing semantic hypotheses." *Studies in African Linguistics* 16: 281-293.

RESPONSE TO WILT,
"DISCOURSE DISTANCES AND THE SWAHILI DEMONSTRATIVES"

Robert Leonard
Friends World College

1. Introduction

Wilt's analysis [in this issue] appears at first to account for signifi-
cant portions of the Swahili data. But on examination we see that each of
the rules he posits can account for only a sharply limited set of data. Each
fails to account for a wider range of data. Swahili demonstratives H and LE
have traditionally been analyzed to mean "proximity" and "non-proximity", re-
spectively. However, this analysis—both prior to Wilt and in Wilt's hands—
leaves much unaccounted for.

The "attention" hypothesis, presented in my 1985 paper[1] and disputed by
Wilt, is better able to account for the Swahili data. This hypothesis posits
that H and LE have meanings that deal with a speaker's relative concentration
of attention on a referent: H signals relatively HIGH Concentration of Atten-
tion and LE signals relatively LOW Concentration of Attention. This hypothe-
sis can not only explain all data accounted for by the traditional "proximity"
hypothesis, but can also explain data that the traditional hypothesis cannot.
For example, the proximity hypothesis can explain data in which H refers to
items that are in fact proximate to the speaker. Such items do tend to be re-
ferred to with H (and this is presumably what historically led analysts to
posit "proximity" as the invariant meaning of H).

Proximate items, however, are but a subset of items that tend to be worthy

_____

[1]Note that although Wilt lists my 1982 dissertation in his bibliography,
he seems to refer only to the short 1985 paper. I refer the reader to the
dissertation for a fuller treatment of the attention hypothesis (and the prox-
imity hypothesis). Other recent works that call into question the usefulness
of the notion "proximity" in the analysis of deictics are Hanks [1984, 1986].

of higher attention. It is a fairly common occurrence that other items, for
example ones that are thematically important, will overshadow the proximates.
In such cases H may be used to refer to something that is not proximate, or
perhaps whose position cannot even be determined. The operant variable is not
proximity but the relative amount of attention the speaker wants the hearer to
focus.

The reason proximates are so often referred to with H is because humans
are egocentric and tend to view their own environs and experiences as note-
worthy and important. Thus proximate items, which have a greater possibility
of interaction with the speaker, and a greater frequency of interaction, will
indeed tend to be referred to with H. But this higher attention is attached
to nearby items not because of their relative closeness but because of the im-
portance that often comes as a consequence of relative closeness. Yet by no
means do all proximate items accrue importance. It is crucial to realize that
most nearby items are not referred to at all; this alone would make it rather
difficult to argue that sheer proximity can cause reference by H.

The difficulties with Wilt's specific suggestions are discussed in the
following sections. It is my belief that the attention hypothesis remains su-
perior in explaining the actual usage of H and LE.

2. Wilt Neglects Data that Run Counter to his Hypotheses

2.1. Wilt's physical/temporal proximity hypothesis. Wilt's Appendix B is a
table that purports to show that there is an absolute correlation between
Proximity and H and Non-proximity and LE in quoted speech in KM chapters 1, 2,
and 6. Wilt says (p. 88) that he presents "a listing of the thirty-five H/LE
demonstratives used in the quoted speech of characters in KM, throughout three
chapters" and that it shows "a 100% correlation between H and spatial/tempo-
ral proximity of the referent and between LE and non-proximity". Although
the caption (p. 93) states that the table lists examples of H and LE from
quoted speech that "refer to a physical location or time period", Wilt lists
non-place referents as well as ones that are themselves places or times, e.g.
'this hole', 'this...letter', 'that homestead', as well as 'here', and 'these
days'. He organizes these data into a table that shows all examples of H in

the proximate column, all LE in the non-proximate.

However, Wilt neglects to include counterexamples. Following is a para-
graph containing two examples of the H demonstrative of which Wilt includes
only one in his table.

(1)   Counterexample not listed by Wilt: H-Sorcerers are distant
         (Their importance merits H.)

"Kazimoto," Kabenga alianza kunizungumzia, labda wachawi hawa wanatukalia
usiku, mimi leo mgongo unaniuma." Sikumjibu. "Inaonekana kwamba
wanakuogopa wewe," aliendelea. Sikumjibu. "Siku hizi wamekoma kufanya
matata," alisema tena. (KM 85)

'"Kazimoto," Kabenga began talking to me, "maybe H-sorcerers visit upon
us at night, today my back hurts me." I didn't reply. "It looks like
they are afraid of you," he continued. I didn't reply. "H-days they
have stopped causing trouble."'

Wilt lists 'H-days' and enters its H demonstrative in the "proximate" col-
umn as 'these days'. But inexplicably he does not list 'sorcerers' in his ta-
ble of examples.

Notice the passage makes clear that the sorcerers, referred to with H,
are not proximate in either place or time, a clear counterexample to the phy-
sical or temporal proximity hypothesis.

The sorcerers, further, are a very clear counterexample to Wilt's anaphor-
ic proximity rule that states "within-paragraph anaphoric ties are made with
H" (p. 90). By that measure, H-sorcerers are not even remotely proximate in
"anaphoric distance" for there is no prior reference in the preceding sen-
tences of the same paragraph. The passage about the sorcerers is itself the
beginning of the second paragraph of chapter 6. No mention of the sorcerers
is made even as "proximately" as in the chapter's first paragraph, nor is men-
tion made even in the preceding chapter. Yet reference is made with H.

These sorcerers are of great concern to the novel's characters throughout
the book; they are worthy of the characters' attention whether or not they are
physically near and whether their most recent prior mention is made on one
side or the other of a paragraph boundary. They are thus referred to with H.

I find two other, less striking, counterexamples that Wilt also neglected

to list. The first, mchezo huu [KM 18] 'H-play' is not occurring at the mo-
ment of speaking, but it will presently, and the speaker is urging his audi-
ence to pay full attention to the play when it does. The second, mambo haya
[KM 21] 'H-events' are not taking place at the moment of speaking, having oc-
curred the previous night. Mambo haya could also refer to 'H-matters' dis-
cussed most recently not within the same paragraph but rather in earlier par-
agraphs. It likely refers to both 'H-events' and 'H-matters', neither of
which is "proximate", yet each of which is of central concern both locally
and to the overall plotline and thematic structure of the novel.

2.2. Wilt's anaphoric proximity hypothesis. In support of this hypothesis,
Wilt attempts to demonstrate two skewings (p. 90): (1) H used for "within-
paragraph anaphoric ties", thus H within P in the table below; and (2) LE
used when the referent of the NP has been referred to most recently in a pre-
vious paragraph, thus LE across P.

    Wilt (p. 87) adduces the following skewing, based on data from a single
chapter of text:[2]

(2) Source: KM, chap. 6

	within P	across P
H	<u>35</u>	2
LE	2	<u>9</u>

    To support Wilt's claims, the data must of course skew towards the top
left cell and the bottom right cell, underlined. The skewings in this one
chapter support his hypothesis. But he neglects the other chapters discussed
in my paper, e.g. JM, chap. 1, a striking counterexample:

(3) Source: JM, chap. 1 [Kiango, S.D., 1974, *Jeraha la Moyo*. Nairobi:
              Foundation Books.]

	within P	across P
H	<u>8</u>	0
LE	38	<u>8</u>

---

[2]See REFERENCES, p. 95, for full titles of sources.

Examination of data from all four chapters discussed in my paper further does not support Wilt's hypothesis, as the combined figures show:

(4) Sources: KM, chap. 6; JM, chap. 1; MZ, chap. 9; MZ, chap. 7

	within P	across P
H	<u>63</u>	10
LE	70	<u>48</u>

So, Wilt sets out to show that anaphoric distance can explain the distribution of H and LE better than the new vs. old strategy of the attention hypothesis. He adduces data from one chapter that seem to support the anaphoric proximity hypothesis. He neglects the three other chapters that I used in the very count that he is attempting to discredit. These data do not support anaphoric distance at all. Thus, when the larger picture is examined anaphoric distance is not useful in explaining H and LE. However, all four chapters do show a correlation of H with new and LE with old [Leonard 1985:288], which supports the hypothesis that H means HIGH Concentration of Attention (COA) and LE means LOW Concentration of Attention. Once more we observe that the attention hypothesis is able to explain more data than the proximity hypothesis.

3. <u>Wilt Makes Unwarranted Claims and Assumptions</u>.

Wilt erroneously assumes that since the title of chapter 7 of MZ is *Mfuko Mweusi* 'A Black Purse', any reference to the purse other than its being "most frequently referred to by H" cannot be adequately explained by the attention hypothesis. He also erroneously assumes that what he terms my Old/New "hypothesis" would demand that the purse be frequently referred to by LE, "since, after the first reference, the purse will be 'old' information". Wilt further claims that I "pay little attention to the conflicting predictions of these hypotheses" (p. 82).

Quite the contrary. Such concerns are discussed in the text, and then condensed in a footnote:

"All other things being equal a speaker will use H for new things and LE for old. But as we saw in the discussion on proximity, other things are often not equal—other factors can outweigh the noteworthiness that stems from newness [or proximity to speaker] or make an old item [or far-away

item] worthy of attention. Examples of the former include the deliber-
ate downplaying of a new item for thematic reasons as well as the intro-
duction of new but purely background items. Examples are.... Thus we do
not expect a 100% correlation between H and new, LE and old, just as we
do not expect (and most certainly do not find) a 100% correlation between
H and proximity, LE and non-proximity. By definition, the invariant *mean-
ing* of a form [is what] correlates 100% with the form's utterance. "New"
and "old", "proximate" and "non-proximate" are but categorizations of fac-
tors that tend to influence a speaker in his choice of H (HIGH COA) vs.
LE (LOW COA)." [Leonard 1985:289]

Wilt's other assumption that the attention hypothesis would predict more
H than LE on a referent named as a chapter title is much too mechanical and
simplistic a view of the situation. To appreciate what is going on in, for
example, the particular instance of chapter 7 of MZ, we must look at the total
pattern of grammatical attention, including the choice available to the speak-
er not to direct attention at all. We must also look at how the attention-at-
attracting position of the purse accrues to referents associated with it and
not just to the referent named *purse* alone.

As stated above, HIGH COA means "the hearer should concentrate the high-
est degree of attention on the referent". The use of LOW COA is an instruc-
tion that the hearer direct a lesser amount of attention than if HIGH COA were
signaled. It should be understood that LOW COA still directs appreciably more
attention than if no attention were directed at all. It is not generally not-
ed that most items mentioned in texts receive no grammatically-directed atten-
tion. In the chapter in question (not at all unique in this regard), literal-
ly scores of referents are merely mentioned, once, by a common noun. Far few-
er are referred to by a pronominal reference such as subject marker, object
marker, etc. Far fewer still are ever referred to by demonstratives. This
general distribution is borne out by count after count performed on different
texts. That an item is mentioned by a demonstrative at all is thus, relative
to the myriad other props, quite a singling out for attention.

Further, in the narrative portions of the chapter in question we find 6 H
and 15 LE, a total of 21 demonstratives. Of this total, 4 refer specifically
to the purse, 2 refer to the place the purse is found, 2 to the sides of the
purse, 4 to the document found in the purse, and 1 to what is written on that

document. Thus 13 of 21 instances of grammatically-related attention are directed to the purse, its location, and its contents (3 of 6 H, 10 of 15 LE).
Relative to all other props mentioned in this chapter, this is a great deal of attention, quite appropriate for a central prop whose name is indeed the chapter title.

Wilt notes that in the chapter in question quoted speech is more likely to have H and narration more likely to have LE. It is true that in general, quoted speech (and transcribed conversations) will contain relatively more demonstratives and relatively more high-deictic demonstratives than will third person narration. The characters are living through the situations so the items they deal with are more important to them, and they therefore concentrate more attention than does a narrator who is not living through the situations. Although Wilt purports to have found a skewing that establishes discourse distance as a determiner of demonstrative choice, what he has done in reality is to pick up on a few more variables on the importance scale.

4. Wilt Overlooks Subtle Thematic Interplays

Wilt contrasts (p. 86) two almost adjacent passages from KM. One, adduced in Leonard [1985], shows clearly the correlation between speaker's concern and degree of COA and further shows that this correlation is more important than physical proximity.

The second passage, Wilt claims, does not bear out this correlation.
This second passage is complex and subtle in its interplay among demonstrative, lack of demonstrative, and choice of noun. The pivotal distinction here is between nyuki 'bee' and mdudu 'insect'. It would seem that Wilt considers them two words for the same thing. He thus misses the role of the demonstratives and the import of the passage.

Throughout this novel a central issue that obsesses the narrator is the power and knowledge relation between God and human, on the one hand, and human and animal on the other (the very point of the first passage in question).
In the second passage, the correlation between speaker's concern and degree of COA becomes clear when one realizes that the relation which is important to the speaker is *God : human :: human : insect* and not *human : bee*. Mdudu

'insect' is referred to exclusively with H, HIGH COA. 'H-insect' is portrayed as making the author happy in its struggling for life. Nyukl 'bee' is referred to twice with LE, LOW COA, and twice with the bare noun. The 'bee', a bare noun with no COA, is an entity that gives up hope (alikata tamaa). 'LE-bee', now an inert figure, is pulled from the water. 'To H-insect', the author realizes, "I had a strength that could not be comprehended. Perhaps H-example could help a human in understanding the mystery of God." Then doubt creeps in that there is a God—perhaps the all-powerful is just Time—and reference is made to a non-idealized 'LE-bee'. 'Insect' is a term at a higher level of abstraction and 'bee' is a particular instance of that abstraction.

This brief discussion no more than touches on the complexity of the passage. My intent simply was to outline a bit of the interplay between the grammatical forms and the thematic structure.

5. Conclusion

The preceding discussion has shown that not only does Wilt's hypothesis present considerable problems with data, it also rests on shaky conceptual grounds. I can find no justification for the claim that the different kinds of "proximity" are at all the same other than that they share the word "prox-imity". It is well-known that a single word can be used to refer to very different ideas. And indeed these are different ideas. It is unclear what motivates positing a single semantic substance that combines such disparate parameters as, on the one hand, "distance from a speaker" and, on the other hand, whether the referent "has been referred to most recently in a previous paragraph" (p. 87).

In the final analysis, proximity—even extended this way—cannot account for much of the data. As I stated in Leonard [1985], the attention hypothesis can explain all data that the proximity hypothesis can account for and it can explain data that the proximity hypothesis cannot. My conclusion then, and now, is that between the competing hypotheses, the one that better fits the linguistic facts is still the attention hypothesis.

REFERENCES

Hanks, William F. 1984. "The evidential core of deixis in Yucatec Maya." In Joseph Drogo et. al. (eds.), *Papers from the Twentieth Regional Meeting of the Chicago Linguistic Society*, pp. 154-172. Chicago: Chicago Linguistic Society.

Hanks, William F. 1986. "Markedness and category interactions in the Malagasy deictic system." CLS Working Papers.