STUDIES
in
AFRICAN LINGUISTICS

TABLE OF CONTENTS

Articles

AKAN VOWEL HARMONY:
THE WORD STRUCTURE CONDITIONS AND THE FLOATING VOWELS*

John M. Stewart
Department of African Linguistics
University of Leiden, The Netherlands

Clements, in an account of Akan vowel harmony [1981b], has a
separate autosegmental tier for the harmonic feature category
[Advanced Tongue Root]. He claims to demonstrate the superi-
ority of his autosegmental framework over the segmental
framework adopted in an earlier account of Akan vowel har-
mony by Schachter and Fromkin [1968], namely that of Stanley
[1967] and Chomsky and Halle [1968]. However, although the
segmental framework in question is indeed unsatisfactory in
certain respects as it stands, it can be readily modified to
meet Clement's objections without recourse to the autoseg-
mentalization of the harmonic feature category.

1.   Introduction

    Clements [1981b] objects to certain aspects of two linear accounts of

Akan (tongue root) advancing harmony which he examines, namely mine of 1967

and Schachter and Fromkin's (hereafter S&F) of 1968, and presents a non-

linear account of his own in which the features [+Advanced] and [-Advanced],

like the tones, are on a separate autosegmental tier, and in which some of

the [+Advanced] autosegments on that tier remain unassociated, or floating,

in surface representation.

In this paper I strongly support his initiative in attempting to apply
to vowel harmony principles that have emerged from the study of tone. I
show that his separate tier for the features of the category [Advanced] is
unjustified, but find that it does indeed appear appropriate to posit
floating segments of some kind to account for what have been called the
"zero vowels".

I begin by presenting a revised version of my 1967 account, leaving the
"zero vowels" out; this time, however, I deal not only with the Asante (As)
dialect but also with the Akuapem (Am) and Fante (Fa) dialects, and this
time I adopt the general (linear) framework of Stanley [1967] and Chomsky
and Halle [1968] as applied to Akan by S&F [1968]. I show by practical
demonstration that the S&F framework is basically satisfactory, and that it
can be readily developed in a natural way, namely by extending the scope of
Stanley's structure conditions from the morpheme to the word, to meet the
following objection by Clements [1981b:125]:

> "[S&F] adopt a rule-based model of vowel harmony which
> accounts for vowel harmony in terms of two independent types
> of statements: morpheme structure conditions determining
> co-occurrence restrictions in roots, and phonological rules
> determining the harmonic category of affixes. Within this
> framework, it is entirely accidental that the same set of
> restrictions on vowel co-occurrence should apply internally
> in roots and externally across morpheme boundaries."

A central feature of this revised account is the formulation of word
structure conditions of a particular type which I call *disharmony condi-
tions*; these capture the concept of the *harmony span* which I use in the
earlier account: the word is divided into harmony spans within which the
vowels are necessarily in harmony, and the disharmony condition for a par-
ticular dialect defines the harmony spans in that dialect by stating the
circumstances under which disharmony between two successive vowels is
possible.

After examining Clements's two main objections to my own earlier
account I turn to the "zero vowels" which I disregarded in the first
instance. I show that a more satisfactory nonlinear analysis than

Clements's is one with two tiers only, one tonal and one nontonal; by this analysis we have floating vowels in place of Clements's floating [+Advanced] autosegments.

Finally I show that once we admit floating vowels, the features of the category [Advanced] are not only not on a separate tier, but are arguably not even distinctive.

2. Proto-Akan and Akuapem

Although both Stewart [1967] and Clements [1981b] focus on the Asante dialect, I begin here not with Asante but with Akuapem as described by Christaller [1875:8-10], which, apart from one very minor complication which is peculiar to Akuapem and which I disregard in the first instance, appears to have preserved proto-Akan vowel harmony virtually unchanged.

Five vocalic binary feature categories distinguish 15 vowels, as in (1):

(1)

	ɛ	e	ɔ	o	ɩ	ῖ	i	ῖ	ɷ	õ	u	ũ	a	ã	ə
Low	−	−	(−)	(−)	(−)	(−)	(−)	(−)	(−)	(−)	(−)	(−)	+	(+)	+
High	−	−	−	−	+	+	+	(+)	+	(+)	+	(+)	(−)	−	(−)
Round	−	−	+	+	−	−	−	−	+	+	+	+	(−)	(−)	(−)
Advanced	−	+	−	+	−	−	+	+	−	−	+	+	−	(−)	+
Nasal	(−)	(−)	(−)	(−)	−	+	−	+	−	+	−	+	−	+	(−)

The categories [Round] and [Advanced] correspond for this purpose to S&F's [Back] and [Tense] respectively. A specification which is enclosed in parentheses is redundant in that it is not required for the specification of the vowel in question, but is itself provided by a segment structure condition (SgSC); the three SgSCs in (2) provide all the redundant specifications in (1):[1]

---

[1]All three conditions are given in if-then form, and it is important that the reader should not be misled by their superficial resemblance to phonological rules. Take the first condition, for instance, which says that a low vowel is nonhigh and nonround; it says also, by implication, both that a high vowel is nonlow and that a round vowel is nonlow, and

(2) a. SgSC1(10v):   b. SgSC2:   c. SgSC3:

$$\Rightarrow \begin{bmatrix} V \\ +Low \end{bmatrix} \Rightarrow \begin{bmatrix} -High \\ -Round \end{bmatrix} \qquad \begin{bmatrix} V \\ -Low \\ -High \end{bmatrix} \Rightarrow [-Nasal] \qquad \begin{bmatrix} V \\ +Low \\ +Advanced \end{bmatrix} \Rightarrow [-Nasal \quad]$$

The first condition says that a low vowel is nonhigh and nonround; the
second that a nonlow nonhigh (i.e. mid) vowel is oral; and the third that
a low advanced vowel is oral.

The specification of a low vowel for advancing is provided by the
sequence structure condition (SqSC) in (3):

(3)   SqSC1:

$$\begin{bmatrix} V \\ +Low \end{bmatrix} \Rightarrow [\alpha Advanced] \quad / \alpha \begin{bmatrix} \underline{\quad} C_0 \begin{bmatrix} V \\ +Advanced \\ -Low \end{bmatrix} \end{bmatrix}$$

This says that a low vowel is advanced if and only if it is followed by an
advanced nonlow vowel.

The simple phonological word, which will be defined presently, is
subject to the SqSC in (4):

(4)   SqSC2:

$$\begin{bmatrix} V \\ \alpha Advanced \end{bmatrix} C_0 \begin{bmatrix} V \\ -\alpha Advanced \\ +Low \end{bmatrix} \Rightarrow$$

This is a disharmony condition which says that a noninitial vowel may be in
disharmony with the preceding vowel only if it is itself low; and which
implies, of course, that any succession of vowels in which all apart from

there is thus no need to formulate these conditions separately, however
important the part they play in the provision of redundant specifications.
It will be seen that the second condition, which says that a nonlow non-
high vowel is oral, similarly implies both that a nonhigh nasal vowel is
low and that a nonlow nasal vowel is high; and that the third condition,
which says that a low advanced vowel is oral, similarly implies both that
a low nasal vowel is nonadvanced and that an advanced nasal vowel is nonlow.

the first are nonlow is either nonadvanced throughout or advanced throughout,
however many vowels there may be.[2]

    The simple phonological word consists of a single root morpheme
together with its affixes, if any, but includes no nonharmonizing affixes.
Affixes include clitics; the most important clitics are the subject pre-
fixes and the object suffixes. The nonharmonizing affixes referred to are
all clitic suffixes; examples are hɔ 'there', yi 'this', mu [m]
'inside'. Words with suffixes were in fact excluded altogether from my
earlier account [Stewart 1967], just as were words with more than one stem
morpheme; my object there was not to give anything approaching an overall
statement of the harmony but simply to establish the position of the root
of the tongue as its articulatory basis.

    Words containing the nonharmonizing clitic suffixes just mentioned
resemble words with compound stems in that where the first vowel of the
suffix or of the second part of the compound is both nonlow and advanced,
the preceding vowel is also advanced, e.g. ɔ + bɛ + ba + mu →

ɔbɛba-mu [ɔbɛbəm] 'he will come in', a + fo + tu → afo-tu

[afutu] 'advice' (from tu fo 'give advice'). In each case the problem
arises that the advanced vowel to the left of the morpheme boundary may be
in disharmony with the vowel before it (if any) even if it is itself nonlow,
and that the word then fails to conform to SqSC2. The second of the two
examples just given is a case in point; to conform, it would have to be

  *əfutu or *afoto . In principle both types of nonsimple phonological
words are outside the scope of this paper, but I shall return briefly to
these words in the final section.

---

    [2]An equivalent formulation of SqSC2 is as follows:

$$V \qquad C_0 \quad \begin{bmatrix} V \\ -\text{Low} \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} \alpha\text{Advanced} \end{bmatrix} \qquad \begin{bmatrix} \alpha\text{Advanced} \end{bmatrix}$$

This says that any noninitial nonlow vowel is in harmony with the preceding
vowel; which, of course, amounts to the same thing. See the observations in
footnote 1 on implied conditions and their redundant status.

Wherever the well-formedness condition (WFC; all SgSCs and SqSCs are WFCs, as Clements correctly points out on p. 121) in (4) would otherwise be violated (for instance, where a prefix with a nonlow nonadvanced vowel in its base form is followed by a root with a nonlow advanced vowel in its base form), the WFC is met by the application of the associated automatic phonological rule (A-rule) in (5):

(5)   SqSC2A: V → [+Advanced]

The device of the A-rule, the function of which is to show what happens where a WFC would otherwise be violated, is an original feature of the present treatment and is explained below in the section on the development of the S&F framework. The A-rule in (5) tells us that its associated WFC, namely SqSC2, is met by the replacement of nonadvanced vowels by advanced vowels wherever there would otherwise be an inadmissible combination of nonadvanced and advanced vowels.

The examples in (6) illustrate the effect of SqSC2A on the prefix ɔ- 'he', the past tense suffix -ɩ , and the perfect tense prefix a- .

(6) a. ɔ + kasa + ɩ   → ɔkasaɩ   'he spoke'

     b. ɔ + fiti + ɩ   → ofitii   'he pierced it'

     c. ɔ + bisa + ɩ   → obisaɩ   'he asked'

     d. ɔ + kəri + ɩ   → ɔkərii   'he weighed it'

     e. a + kasa   → akasa   'has spoken'

     f. a + fiti   → əfiti   'has pierced it'

     g. a + bisa   → əbisa   'has asked'

     h. a + kəri   → akəri   'has weighed it'

It will be seen that while in most cases the affix vowel displays harmony with the nearest root vowel, a prefix vowel displays disharmony where the nearest root vowel is ə . It will be recalled that SqSC2 allows dis-harmony between two vowels where the second vowel is low; as a consequence of this, there is of course no reason for SqSC2A to apply to the prefix vowel in (6d) or (6h). The disharmonic sequences in (6d) and (6h), in which the second vowel is [+Low, +Advanced] may be compared with that in

(6c), in which the second vowel is [+Low, -Advanced].

Since, in a simple phonological word, the only morpheme which can possibly have an advanced vowel in its base form by the present analysis (or by that of Stewart [1967] or by that of S&F [1968]) is the stem morpheme, the A-rule affects only affix vowels. The reason for analysing the harmonizing affixes as having nonadvanced vowels in their base forms is that, as we have just seen, the specification of these vowels for the category [Advanced], though often determined by SqSC2, is not always thus determined, and that when it is not thus determined it is always minus. As we shall see in the final section, however, it may eventually turn out to be more satisfactory to analyse nonlow prefix vowels as underlyingly advanced; we would then need (at least) two A-rules, one for the advancing of low vowels and one for the disadvancing of nonlow vowels.

As was noted earlier, the above account ignores a minor complication which is peculiar to Akuapem. This is the occurrence of the nasal vowel sequences ĩɛ̃ , ɔ̃ɛ̃ , ĩɛ̃ , ũɛ̃ , all of which violate SgSC2 which says that a mid vowel is oral, and the last two of which violate SqSC2 which says that a vowel can be in disharmony with a preceding vowel only if it is itself low; an example is the verb root ɦĩɛ̃ 'open' (cf. As ɦĩnĩ). The nasal vowel sequences in question, which contrast with the oral sequences ιɛ , ɔɛ , ie , ue , have usually been presumed to be derived from underlying ĩɛ , ɔ̃ɛ , ĩe , ũe by a phonological rule which nasalizes a mid vowel after a nasal vowel and makes it nonadvanced if it is not nonadvanced already. The phonological rule is suspect, however, as it does not operate across morpheme boundaries (see the section on the question of floating segments below), and it would appear desirable to amend SgSC2 so as to allow ɛ̃ in underlying representation and to split SqSC2 into two separate conditions, one saying that an oral vowel can be in disharmony with a preceding vowel only if it is itself low, and the other saying that a nasal vowel can be in disharmony with a preceding vowel only if it is itself nonhigh. The amended version of SgSC2 would say that a nonlow nonhigh nasal vowel is nonround and nonadvanced. An extra SqSC would be needed to state that ɛ̃ occurs only after high nasal vowels.

3.   <u>Fante</u>

　　　Apparently as a result of a merger of ə with e , the Fante dialect
has only one oral low vowel, and has consequently not ten but nine oral
vowels altogether. The two low vowels are classified as in (7a) (cf. (1)
above), and SgSCl(10v) in (2a) above is replaced by SgSCl(9v) in (7b):

(7) a.　　　　　　a　 ã　　　　b.　SgSCl(9v):

　　　Low　　　　 + (+)　　　　　$\begin{bmatrix} V \\ +\text{Low} \end{bmatrix}$

　　　High　　　 (-) -

　　　Round　　　(-) (-)　　　　⟹　$\begin{bmatrix} -\text{High} \\ -\text{Round} \\ -\text{Advanced} \end{bmatrix}$

　　　Advanced   (-) (-)

　　　Nasal　　　 - +

Whereas SgSCl(10v) says that a low vowel is nonhigh and nonround, SgSCl(9v)
says that a low vowel is nonhigh, nonround, and nonadvanced. SgSC3 in (2c)
above, which says that there is no nasal advanced low vowel, and SqSCl in
(3) above, which states the complementary distribution of nonadvanced and
advanced low vowels, are eliminated as a consequence of the loss of the
advanced low vowel.

　　　The disharmony condition SqSC2 in (4) above and its associated A-rule
SqSC2A in (5) above are both retained, but whereas in Akuapem they account
unaided for the a ~ ə alternation in prefixes, in Fante they do not
account unaided for the a ~ e alternation; for this alternation we need
in addition the A-rule in (8), which is associated with SgSCl both in its
ten-vowel variant in (2a) above and in its nine-vowel variant in (7b) above.

(8)　　　SgSClA:　　　　V → [-Low]

This A-rule says that the associated WFC, namely SgSCl(10v) or SgSCl(9v),
is met by the replacement of the feature [+Low] by the feature [-Low] wher-
ever there would otherwise be an inadmissible feature combination; the
inadmissible combinations are of course [+High, +Low] and [+Round, +Low] in
the case of SgSCl(10v), and both of these together with [+Advanced, +Low]
in the case of SgSCl(9v).

In Fante, then, where the specification [-Advanced, +Low] in prefixes is changed to [+Advanced, +Low] by the A-rule associated with the disharmony condition SqSC2 an inadmissible combination arises, which is changed in its turn to [+Advanced, -Low] by the A-rule associated with SgSC1.

It will be recalled that in Akuapem the disharmony condition is met by disharmonic sequences both of the advanced-nonadvanced type specified in (9a) and of the nonadvanced-advanced type specified in (9b):

$$(9) \quad \text{a.} \quad \begin{bmatrix} V \\ +\text{Advanced} \end{bmatrix} C_0 \begin{bmatrix} V \\ -\text{Advanced} \end{bmatrix} \quad \text{b.} \quad \begin{bmatrix} V \\ -\text{Advanced} \end{bmatrix} C_0 \begin{bmatrix} V \\ +\text{Advanced} \end{bmatrix}$$

It will be seen, however, that in Fante, as [+Advanced, +Low] vowels are disallowed by SgSC1(9v), disharmonic sequences of the nonadvanced-advanced type are impossible, and consequently all disharmonic sequences are of the advanced-nonadvanced type.

What, then, was the fate of the nonadvanced-advanced type in Fante? Compare the two Fante examples in (10) with the two Akuapem examples in (6) above which illustrate the nonadvanced-advanced type, namely (d) and (h). It will be seen that whereas kəri , with its [+Advanced, +Low] first vowel, does not trigger the advancing of prefix vowels, keri , with its [+Advanced, -Low] first vowel, does.

(10) d.   Fa   ɔ + keri + ι   →   okerii

          cf. Am   ɔ + kəri + ι   →   ɔkərii

      h.   Fa   a + keri   →   eker

          cf. Am   a + kəri   →   akəri

Now not only in these examples but in all examples of nonadvanced-advanced disharmony in Akuapem the advanced (low) vowel is the first vowel of a root and the nonadvanced vowel forms part of or constitutes a prefix, and consequently the replacement of ə by e in Fante has the effect of totally eliminating nonadvanced-advanced disharmony within the simple phonological word.

It is appropriate to mention the reason why, in Akuapem, the advanced low vowel in a nonadvanced-advanced sequence is never a prefix vowel: a

prefix with a low vowel never has an initial consonant and is never pre-
ceded by more than one other prefix, and the vowel of any preceding prefix
is totally assimilated to the low vowel which it precedes, e.g. Fa

  yɛ + a + ba → yaaba 'we have come'.

4. Asante

    In the Asante dialect the advanced low vowel ə has a more restricted
distribution than in Akuapem, the SqSC in (11) taking the place of the one
in (3):

(11)   SqSC1(As):

$$\begin{bmatrix} V \\ +Low \end{bmatrix} \Rightarrow [\alpha Advanced] \Big/ \alpha \left[\underline{\quad} C_0 \begin{bmatrix} V \\ +Advanced \\ +High \end{bmatrix} \right]$$

This says that a low vowel is advanced if and only if it is followed by an
advanced high vowel. We saw that in Akuapem, ə occurred before all
advanced nonlow vowels.

    Correspondingly, the less restrictive disharmony condition in (12a)
takes the place of the one in (4):

(12) a. SqSC2(As):
$$\begin{bmatrix} V \\ \alpha Advanced \end{bmatrix} C_0 \begin{bmatrix} V \\ -\alpha Advanced \\ [-High \quad] \end{bmatrix} \Rightarrow$$

      b. SqSC2(As)A1:
$$\begin{bmatrix} V \\ -High \end{bmatrix} \rightarrow [-Advanced] \Big/ \underline{\quad} C_0 \begin{bmatrix} V \\ +High \\ -Advanced \end{bmatrix}$$

      c. SqSC2(As)A2:   V → [+Advanced]

(12a) says that a noninitial vowel may be in disharmony with the preceding
vowel only if it is itself nonhigh. We saw that in Akuapem the dishar-
monic vowel had to be low. In addition, the two ordered A-rules in (12b-c)
take the place of the one in (5), though the one in (12c) is in fact iden-
tical to the one in (5); they say in effect that wherever possible,
SqSC2(As) is met by the replacement of e , o by ɛ , ɔ before

ɩ , ɔ , and that elsewhere it is met in the usual way by the replace-
ment of nonadvanced vowels by advanced vowels. The examples in (13)
illustrate the disharmonic mid vowels that are admissible in Asante but
not in Akuapem:

(13) a.   As   a + bɛ + tu   →   abetu   'has come and pulled it out'
        cf. Am   a + bɛ + tu   →   əbetu

      b.   As   a + kɔ + tu   →   akotu   'has gone and pulled it out'
        cf. Am   a + kɔ + tu   →   əkotu

      c.   As   ɔ + tie + ɩ   →   otiɛɩ   'he listened'
        cf. Am   ɔ + tie + ɩ   →   otiei

      d.   As   mɩ + ẅie + ɩ   →   miẅiɛɩ   'I finished'
        cf. Am   mɩ + ẅie + ɩ   →   miẅiei

In (a) and (b) the disharmonic mid vowel is advanced, and the vowel with
which it is in disharmony is a low vowel which constitutes a prefix. In
(c) and (d) the disharmonic mid vowel is nonadvanced, having become so
by SqSC2(As)A1, and both it and the vowel with which it is in disharmony
form part of the root.

    Nonlow prefix vowels appear to harmonize with mid vowels in Asante as
in Akuapem, e.g. ɔ + bɛ + tu → obetu 'he comes and pulls it out',
  yɛ + bɛ + tu → yebetu 'we come and pull it out'. Whereas in
Akuapem the harmonization is demanded by SqSC2, however, in Asante it is
not demanded by the Asante counterpart of SqSC2, namely SqSC2(As), and a
nonautomatic phonological rule on the lines of (14) seems to be needed to
account for it:

(14) $\begin{bmatrix} V \\ -Low \end{bmatrix}$ / ___]$_{px}$ C_0 $\begin{bmatrix} V \\ +Advanced \\ -Low \end{bmatrix}$ (px: prefix)

      → $\begin{bmatrix} +Advanced \end{bmatrix}$

This says that a nonlow prefix-final vowel becomes advanced before any
advanced nonlow vowel.

    It will be seen that the Asante A-rule SqSC2(As)A1 introduces alter-

nation into the root, e.g. tie ～ tiɛ 'listen'. There is one root,

namely ɲĩnsɛɲ 'become pregnant', which would appear to have had the alter-

nation in the past (*ɲĩnseɲ (unsuffixed form) ～ ɲĩnsɛɲ (suffixed form))

but to have got rid of it by generalizing the suffixed form. There are good

reasons for believing that the proto-Akan form was *ɲĩnseɲ , as will be

seen in the section on Clements's "disharmonic roots" below; and this form,

when combined with the past tense suffix, would of course have met the

structural description of the Asante A-rule in question. The generalization

of the suffixed form should perhaps be seen partly as a move in the direc-

tion of the restoration of the principle of root control, which is of

course violated by SqSC2(As)A1.

5.   Word Structure Conditions and Their Associated Automatic Rules

It will be recalled that I have made two changes in the S&F framework;

first, I have extended the scope of the SqSCs from the morpheme to the word,

and second, I have introduced a new kind of phonological rule, the automatic

rule (A-rule), the function of which is to show what happens where a well-

formedness condition (WFC) would otherwise be violated.

S&F did in fact consider the possibility of the first of these two

changes themselves; they write as follows (p. 56):

> "While it might have been possible to extend SqSC5 [their
> counterpart of my SqSC2; J.M.S] to apply to such tenseness-
> harmony constraints within the word, we have preferred to
> characterize the latter by means of a phonological rule."

The second of the two changes, namely the introduction of the A-rule,

meets a need which arises as a result of the first: to show how SqSC2 is

met in those cases in which it would be violated if each of the constituent

morphemes of the word retained its base form.

The A-rule is designed to capture Hockett's concept of *automatic

alternation*; he writes as follows [1958:279-80]:

> "Some alternations are such that if they did not take place,
> the phonemic pattern of the language would be different from
> what in fact it is."

> "Alternations of this kind are called *automatic*. One of the
> alternate shapes is the *base form*, and the other or others
> are said to replace the base form under specific conditions
> where, otherwise, there would be an arrangement of phonemes
> contrary to the phonemic pattern of the language. It is to
> be noted that the term "automatic" refers to the fact that
> the base form is replaced, but not to the particular replace-
> ment which is made. . .there is nothing about the *phonemic*
> system of the language which renders the actually-used
> device any more natural than [the theoretically possible]
> alternatives."

It will be seen that his concept of the *phonemic pattern of the language* is
already captured by the WFCs.

   Although in the present context the need for the A-rule arises in the
first instance from the extension of the scope of the SqSC from the morpheme
to the word, it is important to note that this is by no means the only type
of situation in which the A-rule is needed, and that S&F's framework is
defective without it in any case. This is evident from those situations
in which S&F posit the application of morpheme structure (MS) conditions to
the output of phonological rules (P-rules); they write as follows (p. 18):

> "Some of the MS conditions apply to strings which are the out-
> put of P-rules, as well as to those which are the input to
> this component of the grammar. To repeat these conditions as
> P-rules, or to specify the redundant features in P-rules,
> would not only be uneconomical but would obscure the general-
> ity concerning the redundancies which persist. We therefore
> include the convention that when a feature is specified as
> redundant in the MS rules, it remains redundant with the
> application of the P-rules, if the non-redundant feature is
> changed and the redundant feature is not mentioned. For
> example, at the systematic-phonemic level all [-Back] vowels
> in Akan are redundantly [-Round], and are so specified in
> the MS Segment Structure Conditions. If a P-rule changes
> the feature specification of a vowel from [+Back] to [-Back]
> and does not mention the feature [Round], this implies that
> the segment which was redundantly [+Round] becomes [-Round]
> simultaneously with the change of the backness feature."

Their example hardly fits the present analysis, which makes no use of the
distinction between [Back] and [Round], but another example can be readily
substituted. In Fante, as we have seen, [+Low] vowels are redundantly

[-High, -Round, -Advanced], and are so specified by SgSCl(9v) in (7b)
above. When the A-rule SqSC2A in (5) above changes the specification of
a low vowel to [+Advanced] (which it does without specifying a feature of
the category [Low] for that vowel), this has the consequence that the vowel
which was [+Low] becomes [-Low] simultaneously with the change from
[-Advanced] to [+Advanced]. Now whereas S&F would say that this consequence
was implied by SgSCl(9v), I maintain that it is not implied, but has to be
stated in an A-rule associated with the SgSC. As Hockett points out, there
is nothing about the phonological pattern of the language which makes the
actually used device any more natural than the theoretically possible alter-
natives; in Turkana, for instance, a is replaced in comparable circum-
stances not by e but by o [Dimmendaal 1982:23], and a different A-rule
is therefore required.

As I have already noted, S&F acknowledge their indebtedness to Stanley
[1967]; Stanley is in fact the author of the distinction between P-rules and
MS conditions as they conceive it, and also of "the convention that the out-
put of each P-rule is automatically subjected to the segment structure rules
[i.e. the SgSCs; J.M.S.]" [1967:404]. I shall now show that both of the
weaknesses in the S&F framework with which I am concerned in this section
are in fact traceable to weaknesses in Stanley [1967].

Let us look first at the difficulty that I resolve by introducing the
A-rule. Stanley writes as follows (p. 397):

> "Essentially, this proposal amounts to demanding that redun-
> dancy rules [i.e. MS conditions; J.M.S.] be clearly distin-
> guished from rules which change feature values [i.e. P-rules;
> J.M.S.]."

If, however, the MS conditions apply to the output of each P-rule in the way
he proposes then the MS conditions do change feature values. My own propo-
sal, namely that in such circumstances the feature values are changed not by
the structure conditions themselves but by associated A-rules, avoids this
difficulty.

Now let us look at the question of the scope of the structure condi-
tions. Stanley writes as follows (p. 397):

>"The fact that MS rules apply to individual morphemes, and not
>to strings of morphemes in a sentence, reflects the easily
>verified empirical fact that the constraints holding within
>single morphemes are more restrictive than the constraints
>which characterize larger units."

The error here is to compare the claims of the morpheme with those of the
sentence without considering the claims of the word.

Ultimately, then, it is this latter error of Stanley's that gives rise
to Clements's objection ([1981b:125-6]; quoted in the introduction above) to
the S&F framework. As I have shown in this section, however, the error is
by no means crucial, and we can readily adjust the Stanley/S&F framework in
a natural way without resorting to the autosegmentalization of the harmonic
feature category.

6. Clements's "Disharmonic Roots"

Clements explicitly criticizes my 1967 framework on two counts: first,
that it "provides no straightforward treatment of [the two] disharmonic
roots [pɪɲcɛ 'to come close' and ɲinsɛɲ 'to be pregnant']" (p. 132),
and second that it is characterized by "a considerable amount of indeter-
minacy with respect to the choice of where [the harmonic/prosodic features]
are to be located in linear strings of phonemes" (p. 133). I devote this
section to the first criticism and the following section to the second.

It is important to note that Clements (p. 119) distinguishes between
two types of roots which display internal disharmony: "mixed vowel roots"
such as bɪsa 'to ask' which conform to the structure conditions and are in
fact very common, and "disharmonic roots" which do not conform to the struc-
ture conditions and of which he knows no examples apart from the two just
quoted.

The first point to be made is that the present account is the first to
recognize that the Asante disharmony condition is not the same as that of
Akuapem, and that whereas in Akuapem a noninitial harmony span can generally
be initiated only by a low vowel (see SqSC2 in (4) above), in Asante a non-
initial harmony span can be initiated by any nonhigh vowel (see SqSC2(As) in

(12) above). Once this is recognized, Clements's two exceptional "dishar-
monic roots" become nonexceptional "mixed vowel roots" in Asante. This does
not dispose of the matter, however, as neither Akuapem nor Fante is entirely
free of what would still be "disharmonic roots" by Clements's analysis.

In my 1967 account I made no mention of Clements's "disharmonic roots"
as I had assumed them to be compounds, and as compounds they are entirely
regular. Clements, however, writes as follows (p. 170):

> "I know of no motivation for considering the forms cited. . .to
> be compounds, at least in the contemporary language."

My task here, therefore, is to make the motivation known.

Christaller's dictionary [1933], which is based on Akuapem, lists all
four of the items in (15a) as verb stems meaning 'to approach', and the
Fante dictionary (anonymous n.d.) lists both of the items in (15b) as verb
stems meaning 'to draw nigh, to approach':

(15) a. Am (dictionary)   pini   piŋkyɛ
                             bɛŋ   bɛŋkyɛ

      b. Fa (dictionary)   pin   pinkyɛn

The forms in the second column are exactly what one would expect if the
form in the first column was compounded with (Am) kyɛ [cɛ] or (Fa)
kyɛn [cɛn] . The difference between (Am) cɛ and (Fa) cɛn is not
explained by any regular sound correspondence; the most plausible explana-
tion appears to be that in Fante the second element in the compound has
come to be identified with the root in n-kyɛn [ɲcɛn] 'beside'.

Christaller's dictionary lists both of the items in (16a) as verb
stems meaning 'to become pregnant', and the official Fante spelling book
(anonymous 1944) lists both of the corresponding Fante items in (16b). The
Fante items are repeated in phonetic transcription in (16c):

(16) a. Am (dictionary)   yem   ninsɛŋ
      b. Fa (spelling book)   nyem   nyinsen
      c. Fa (phonetic)   ɲĩm   ɲĩnsen

The Fante phonetic form in the second column of (16c) is exactly what one would expect if the form in the first column was compounded with sen , though if it were a straightforward compound it would be phonologically
 *ɲĩn-sen and it would not take advanced vowels in prefixes as in fact it does. It is plausible to suggest that it is a compound in origin but that it has come to be treated as a single morpheme; this is something that happens sometimes even in the case of more transparent compounds such as the stem in ebufuw (not *abɔ-fuw [abufuw]) 'anger' from bɔ fuw 'get angry' (bɔ: 'breast'; fuw: 'shoot up').

It will be seen that (Fa) ɲĩnsen is not in fact a "disharmonic root" at all, and that it is plausible to reconstruct the proto-Akan form as
 *ɲĩnsɛŋ as we can explain the advanced vowel in (Am) nĩnsɛŋ only if we posit an earlier *ɲĩnseŋ . We have already explained the change of the second vowel from advanced to nonadvanced in (As) ɲĩnsɛɲ ; see the section on Asante above. Perhaps (Am) ɲĩnsɛ is to be explained as a borrowing from Asante. Compare Christaller's (Am) dictionary entry o-ʈuo 'musket, gun'; the final o represents a nominal suffix which, as S&F [1968:67] note, has a segmental realization after high vowels in Asante but has no segmental realization in any context in Akuapem or Fante. On being borrowed into Akuapem, (As) ɲĩnsɛɲ would presumably be reinterpreted as a compound parallel in structure to piɲɛ .

7.   The Representation of the Harmonic Feature

Clements writes as follows on the above topic (p. 133):

"Another problematical aspect of [Stewart's] account concerns the placement of the feature H in underlying representation. Any theory making use of prosodic features of this sort within the general framework of linear representation will be faced with a considerable amount of indeterminacy with respect to the choice of where they are to be located in linear strings of phonemes. For instance, in the case of Akan the following set of representations would all be consistent with the phonetic shape and phonological behavior of the noun root [kotoɟʷɛ] 'knee': a. Hkɔtɔɟʷɛ ; b. kHɔtɔɟʷɛ ; c. kɔHtɔɟʷɛ ; d. kɔtHɔɟʷɛ ; e. kɔtɔHɟʷɛ ; f. kɔtɔɟʷHɛ ; g. kɔtɔɟʷɛH .

>   Given this fact, the rule requiring H to occur to the right of
>   the rightmost nonlow vowel, which uniquely selects (g) as the
>   underlying representation of 'knee', is a linguistically arbi-
>   trary one."

My answer here is that how one marks the harmonic feature of a harmony
span is of no theoretical significance. What is significant is the division
of the word into harmony spans; this is captured in the present account by
the disharmony conditions in (4) and (12).

My position on the question of marking is essentially the same as that
of S&F, who write as follows (p. 13):

>   "We will mark the first vowel for this feature in the diction-
>   ary matrices, but it should be understood that this is an
>   arbitrary decision, and that we could just as easily have
>   chosen to mark the last vowel instead. It is the [structure]
>   condition itself which makes the generalization and not the
>   dictionary matrices, and therefore we need not be concerned
>   about which segment is marked."

8.   The Question of Floating Segments

One of Clements's objections to S&F's account remains unanswered. S&F
write as follows (p. 97; [U] is [u] or [ʊ] , and [I] is [i] or [ɪ]):

>   "Apparent exceptions to the claim that it is only nonlow tense
>   vowels that tense preceding vowels in grammatical morphemes
>   are provided by words such as A[m]-Fa obegua [obegwə?] /As
>   obedwa [obedɟɥə?] 'he comes to skin', okogyam [okodɟəm?]
>   'he goes to condole', etc. However, if we investigate the
>   derivation of the root morphemes in such words, we find that,
>   at some point in the derivation, there is always a [u] or
>   [i] - i.e. a nonlow tense vowel - before the low tense [ə]
>   that occurs as the first vowel in the final phonetic form.
>   Thus the underlying form of the root A[m]-Fa [gwə?] /As
>   [dɟɥə?] is /guək/ , and the [u] of the root is present
>   in the derivation until it is deleted by the [U]-deletion
>   rule. . .In the case of the root [dɟəm?] , while the under-
>   lying form is /gəb/ , the [I]-insertion rule. . .obligator-
>   ily inserts an [i] between the [g] and the [ə] , and
>   this [i] remains present in the derivation until it is
>   deleted by [the [I]-deletion rule]."

(For ə read a in every case; see Clements [1981b:116-7, 123.) Clements

has reservations about the synchronic motivation of this analysis; he writes
as follows (p. 148):

> "Most writers on the subject have assumed that [these] stems
> had their historical source in forms containing [+Advanced]
> high vowels between the initial consonant and the low vowel,
> which dropped out after having conditioned the palatalization
> (occasionally accompanied by rounding) of the consonant.
> While no synchronic alternations remain to provide a strong
> source of motivation for such an analysis in the present-day
> language (though see S&F [1968] for an attempt to support
> such an analysis with language-internal evidence), there is
> a certain amount of comparative evidence suggesting that it
> is not implausible as a diachronic analysis, at least for
> some forms."

S&F's P-rules of [U]- and [I]-deletion are certainly open to the objec-
tion that they never operate across a morpheme boundary and thus never
generate alternant forms of morphemes. Before we seek a solution, however,
let us look at another of S&F's [1968:72] P-rules, namely regressive non-
vowel nasalization, a less complex rule which is open to the same objection.

By S&F's analysis Akan has nasal consonants at the systematic-phonetic
level but not at the systematic-phonemic level, and a large proportion of the
surface nasal consonants result from the application of the regressive nonvow-
el nasalization rule, the main part of which is essentially as in (17):

(17) $\begin{bmatrix} C \\ +\text{Voiced} \end{bmatrix} \quad / \underline{\quad} \begin{bmatrix} V \\ +\text{Nasal} \end{bmatrix}$

    $\rightarrow \begin{bmatrix} +\text{Nasal} \end{bmatrix}$

This P-rule says that a voiced consonant becomes nasal before a nasal vowel;
examples are /bã/ → [mã] 'give', /dã/ → [nã] 'and'.

We can avoid this problematic P-rule if we admit underlying /m,n/
and introduce a SqSC of nasality agreement, somewhat as in (18):

(18) $\begin{bmatrix} C \\ +\text{Voiced} \end{bmatrix} \quad V$

    $\Longrightarrow \begin{bmatrix} \alpha\text{Nasal} \end{bmatrix} \begin{bmatrix} \alpha\text{Nasal} \end{bmatrix}$

This says that any CV sequence in which the C is voiced is either oral
throughout or nasal throughout.

We can avoid S&F's P-rules of $[U]$- and $[I]$-deletion in a similar way if
we are prepared to admit floating vowels comparable in status to floating
tones, distinguished from their nonfloating counterparts in terms of the
feature category [Segmental] proposed by Voorhoeve, Meeussen and De Blois
[1969]. Let us however call it [Durational] in order to avoid the apparent
contradition of talking of nonsegmental segments. The SqSC would then be
somewhat as in (19):

(19)　　　　$[C_0 \quad (V_1 \qquad) \quad V \qquad\qquad]_{syllable}$

　　　$\Rightarrow$　　　　[-Durational]　[+Durational]

This says that if a syllable contains at least one vowel, then the final or
only vowel is durational and any nonfinal vowels are nondurational; it is
assumed that another condition excludes VC-final syllables. The near-
minimal monosyllabic/disyllabic pairs of Asante verb roots in (20)
illustrate:

(20)　ɟi̥a　[ɟa]　　'accompany'　　ci=a　[cia]　　'greet'

　　　ɲĩ̥ã　[ɲã]　　'get'　　　ħĩ=ã　[ħĩã]　　'need'

　　　ɟy̥a　[ɟʷa]　　'skin'　　　dy̥i=a　[dʷʸia]　　'plant'

　　　sy̥a　[sʷʸa]　　'be small'　　ty̥i=a　[tʷʸia]　　'pay'

(i̥ , y̥ : nondurational i , u ; = : syllable boundary.)

The roots ty̥i=a , dy̥i=a , ɟy̥a are assumed to be derived from
earlier *tu=a [tua] , *du=a [dua] , *gy̥a [gʷa] ; compare the offi-
cial spellings 'tua' , 'dua' , (Fa, Am) 'gua' (As 'dwa'). The
remaining root with y̥ , namely sy̥a , is similarly spelt 'sua' ,
though its historical origin is less obvious.

It is perhaps not without interest that the "zero tone-bearing units"
which I posited in my early work on Akan tone [Stewart 1962, 1965] for what
were later to become known as floating tones were inspired by Welmers'
[1946:18-19] treatment of the floating vowels illustrated in (20) above as
"zero variants" of the vowels in question.

Clearly, if we admit floating whole vowels, there is no need for
floating [+Advanced] autosegments. This does not mean, of course, that

there is no need at all for autosegmental phonology, and I shall try pre-
sently to work out some of the implications of the floating vowels for an
Akan phonology with two tiers, one tonal and one nontonal, basically on the
lines of Clements [1981a]. First, however, I must draw attention to a cru-
cial difference between Clements's model and that of Goldsmith [1976].

Clements and Ford [1979] refer to Goldsmith's [1976] view that
"associations among [tones and tone-bearing units] are governed by a set
of principles of *well-formedness* [with] the following effect: every tone
is associated with at least one tone-bearing unit; every tone-bearing unit
is associated with at least one tone; and no association lines cross"
[1979:182]. By this view free (= unassociated) tones are of course dis-
allowed in surface representations. Clements and Ford themselves, however,
propose to admit surface free tones, claiming that they are motivated by
tonal downstep; they write as follows (pp. 204-5):

> "Throughout the synchronic phonology of Kikuyu, phonological
> and phonetic evidence converge to demonstrate that the free
> extra-low tones created by the operation of tone shift cor-
> respond precisely to the downstep elements operated upon by
> the tone rules. Throughout this set of rules, the downstep
> element functions as if it were a phonological entity on a
> par with tones. This fact follows naturally from the assump-
> tion that it *is* a tone: namely, the floating extra-low tone
> that our rules have independently generated.
>
> "From the point of view of Kikuyu phonology alone, this is
> the simplest position that we can take. Our rules have
> generated a number of extra-low tones that are subject to
> the operation of certain phrase-level rules. Subsequent to
> the operation of these rules, these free tones are inter-
> preted by the rules of pitch assignment as *operators* trig-
> gering pitch lowering. Such rules have the effect of
> lowering the register within which subsequent tones in the
> tone group are realized. No further entities such as spe-
> cial "drop tone" features or unanalysable "pitch change
> markers" need be introduced. Any rules that might be pro-
> posed to trade in free tones for such entities would be
> descriptively superfluous.
>
> "If we adopt this view, it appears possible to do away
> entirely with the theoretically suspect entity "!",
> replacing it with a known quantity. Floating tones are
> well documented in the languages of Africa, and it appears

>   likely that the downsteps identified in other languages can be
>   assigned a similar status. Thus, to take an example, it would
>   be advantageous to consider the downsteps occurring between
>   low tones in Dschang-Bamileke (see Hyman and Tadadjeu [1976])
>   as consisting of free high tones; since free high tones must
>   be generated in just the places where downstep appears phonet-
>   ically, there is no need to introduce further rules whose only
>   function is to exchange the free high tones for some type of
>   downstep entity. A free high tone is formally distinct from
>   an associated high tone, and can therefore serve as the unam-
>   biguous operator conditioning register-lowering [i.e. key
>   lowering; J.M.S.]."

Subsequent work suggests the possibility that the surface free tones which account for downstep are always low. Clements [1981a:90] sees them as low and no longer as extra-low in Kikuyu. Stewart [1981:138] shows that there is no need "to consider the downsteps occurring between low tones in Dschang-Bamileke. . .as consisting of free high tones" as Clements and Ford (passage just quoted) suggest, as even these downsteps are analysable as free low tones, and Clements [personal communication] "do[es] not know offhand of any other instances of languages in which downstep could be attributed to floating high tones."

In the light of this I suggested, in the Dutch version of this article [Stewart 1982:339), that a floating vowel was perhaps a vowel not associated with any tone just as a floating tone was a tone not associated with any vowel. (I followed S&F [1968:47] in identifying the tone-bearing units as the final vowels of syllables at the systematic-phonemic level. This raises problems, particularly as the tone-bearing units at the systematic-phonetic level are quite often sonorant consonants, but these problems need not concern us here.) I saw this as having the advantage of avoiding recourse to the feature category [Zero] ([αDurational] = [-αZero]) which I discuss there).

It now appears, however, that this must be rejected on the grounds that the category [Durational] is needed in underlying representations to distinguish between nondurational floating tones and durational floating tones. Clements [personal communication] claims that "two types of floating low tones must be recognized for Kikuyu; those that act as downstep operators, and those that don't," and Thomas L. Cook [personal communication] draws

attention to what appears to be essentially the same situation in Efik, while I myself now believe that my own analysis of the tones of the associative construction in Dschang-Bamileke [Stewart 1981] would be improved if the high-tone and low-tone associative markers were analysed as floating tones at some nonsurface stage in the derivation (as originally proposed by Tadadjeu [1974:286]) even though the floating low tone did not act as a downstep operator in this case.

It seems, then, that we should retain the category [Durational] in both tiers and revise the principles of well-formedness governing associations among tones and vowels so that they have the following effect: every durational tone is associated with at least one durational vowel; every durational vowel is associated with at least one durational tone; and no association lines cross.

So far my position has been purely defensive; I have argued merely that although the segmental framework used by S&F in their treatment of Akan vowel harmony is indeed unsatisfactory, it can be readily modified to meet Clements's objections without recourse to a separate harmonic tier. I now proceed to give three reasons why the floating [+Advanced] whole vowels posited here are in fact to be preferred to the floating [+Advanced] auto-segments posited by Clements.

First, sporadic cases of rightward vowel shift such as those illustrated in the right-hand column in (21) show that the floating [+Advanced] whole vowels sometimes become nonfloating:

(21) Am   gɥarι   [gʷarι]   Fa.Abura   gura   [gura]   'wash'
      (As   ɟɥɟarι [ɟʷarι])
      As   sɥɟaι   [sʷɣaι]   Am, Fa   sɥia   [sʷɣia]   'swear'

(Fa.Abura: the Abura subdialect of Fante.)

Second, the floating vowels account not only for the otherwise unexplained advancing of prefix vowels before nonadvanced low vowels but also for other otherwise unexplained phenomena: consonant "palatalization" before low vowels [S&F 1968:89-91], consonant rounding before nonround

vowels [S&F 1968:87-8], and, in Fante, rounding of prefix vowels before
nonround vowels [S&F 1968:102-4]. Clements says nothing about the impli-
cations of his autosegmentalization for what remains of the traditional
zero vowels.

Third, there would appear to be no objection to regarding the floating
vowels as being fully specified; as being specified, that is, for all the
vocalic feature categories without exception. An important factor favour-
ing the synchronic recovery of the specifications is the fact that, as the
pairs in (20) above illustrate, monosyllabic CV_1V sequences and disyllabic
CV_1V sequences have a number of SqSCs in common.

9.   Floating Low Vowels?

The recognition of floating advanced vowels which condition advanced
prefix vowels opens up an extremely interesting possibility: might there
not also be floating nonadvanced vowels which condition nonadvanced prefix
vowels? The symmetry commonly displayed by floating high tones and float-
ing low tones makes this a very natural question to ask; consider for
instance the Asante sentences (retranscribed from Stewart [forthcoming])
in (22):

(22) a.   cʷɛ̀ -nɔ̃̌ → cʷɛ̀nɔ̃̌   'look at him'
      b.   kòfí fìrɛ̰́ -nɔ̃̌ → kòfí fìrénɔ̃̌   'Kofi calls him'
      c.   mã̀ kòfí ḿfɩ̀rɛ̰̀ -nũ̌ → mã̀ kòfí ḿfɩ́rénɔ̃̀   'Kofi should call him'
      d.   ɔ̀- ɲ̀çʷɛ̰́ -nɔ̃̀ → ɔ̀ɲ̀çʷɛ̀nɔ̃̌   'he does not look at him'

(Acute accent: high tone; Grave accent: low tone.)

As the difference between (a) and (b) illustrates, a prepausal low-tone
object pronoun (which, as we saw earlier, constitutes a clitic suffix)
becomes high after a high tone. In (c) a floating low tone conditions a
low suffix tone after a high tone, and in (d) a floating high tone condi-
tions a high suffix tone after a low tone.

Now consider the Asante verbs in (23), which illustrates an analysis
that admits not only floating advanced nonlow vowels but also floating

nonadvanced low vowels:

(23) a.   a + ba   → aba   'has come'

      b.   a + di   → ədi   'has eaten'

      c.   a + t̬ɐɩ   → atɩ   'has picked'

      d.   a + ɲj̃ã   → əɲã   'has got'

As the difference between (a) and (b) illustrates, a nonadvanced low prefix
vowel becomes advanced before an advanced nonlow root vowel. In (d) a
floating advanced nonlow vowel conditions an advanced prefix vowel before a
nonadvanced low root vowel, and in (c) a floating nonadvanced low vowel
conditions a nonadvanced prefix vowel before a nonlow root vowel which would
be advanced but for the presence of the floating nonadvanced low vowel,
which conditions nonadvanced vowels after it as well as before it.

     It would appear that by this analysis i , u , e , o are in
complementary distribution with ɩ , ɵ , ɛ , ɔ not only in harmoniz-
ing affixes but also in roots, as roots in which the first durational vowel
is ɩ , ɵ , ɛ , ɔ are always analysable as having a floating non-
advanced low vowel before that vowel, thus: ɐɩ , ɐɵ , ɐɛ , ɐɔ ; com-
pare (c) with (b) in (23) above. Then, of course, the features [−Advanced]
and [+Advanced] are redundant and the number of contrasting oral vowels is
reduced from nine to five, and we no longer have the anomaly that whereas
the nonadvanced nonlow vowels ɩ , ɵ , ɛ , ɔ contrast with their
advanced counterparts i , u , e , o , the nonadvanced low vowel a
does not contrast with its advanced counterpart ə ; nonlow vowels are
redundantly advanced except in specified contexts, just as all along low
vowels have been redundantly nonadvanced except in specified contexts.

     Then, of course, the floating vowels are more appropriately specified
simply in terms of the feature category [Low], so that we have floating non-
low vowels and floating low vowels instead of floating advanced nonlow
vowels and floating nonadvanced low vowels. This immediately eliminates the
much-debated anomaly that the advancing of prefix vowels is conditioned not
by all advanced vowels but only by nonlow advanced vowels: we can now say
that the advancing of prefix vowels is conditioned simply by nonlow vowels.

It looks as if this approach might open up the possibility of a more satisfactory account than has yet been achieved of one aspect of Akan vowel harmony which has been excluded from the present treatment, namely harmony across boundaries other than those which occur within the simple phonological word as defined in the section on proto-Akan and Akuapem: word boundaries, boundaries between root morphemes in compounds, and boundaries preceding nonharmonizing clitic suffixes. We saw above, in the same section, the compound a + fɔ + tu → afɔ-tu [afutu] 'advice' from   tu fɔ 'give advice'; by the analysis now suggested this becomes   a + f̧ɔ + tu → af̧utu [afutu], in which what appears to be an inadmissible occurrence of a nonadvanced vowel before a nonlow advanced vowel is explained by the floating low vowel that has already been posited on other grounds. A floating low vowel then accounts for apparent exceptions to SqSC2 or SqSC2(As) in much the same way as a floating low tone accounts for downstep between adjacent high tones; downstep in this context, of course, constitutes an apparent exception to the rule that two adjacent high tones are on the same level.

It should be noted that here we are forced to regard the floating vowel as low rather than as nonadvanced, as it has to be advanced to conform to SqSC2/SqSC2(As).

It should also be noted that by this approach to harmony across word boundaries and other comparable boundaries, it may prove possible to do away entirely with the allegedly postbinary phonetic process of "vowel raising" posited by Clements [1981b:154-60], following Berry [1957], to account for these and other phenomena. Clements notes that the sentences in (24a-b) are distinct in normal speech:

(24) a. mi-i-bu bɔ̧ɔ̧ bi   'I'm breaking a stone'

      b. mi-i-bu buo bi   'I'm breaking a nest'

      c. (ɛ)bɔɔ   'stone'

      d. (e)buo   'nest'

(ɔ̧ , ɔ̧:   ɔ , ɔ affected by "vowel raising.")

By the approach I am suggesting we would have ɛbᶓɔɔ -bi → ɛbᶓɔobi

  [ɛbɷobi] 'a stone'; since, by condition SqSC2(As), any nonhigh vowel

can initiate a harmony span, that condition is satisfied by the harmoniza-

tion of the ɔ alone, and the contrast with ebuo -bi → ebuobi 'a nest'

thus survives even where the initial ɛ or e is absent.

     I am well aware that there are many questions which I have left

unanswered, such as whether we can continue to regard nonlow prefix vowels

as being nonadvanced in the base forms of the prefixes, and if not, what

the implications are. My purpose here, however, is merely to show that the

features [-Advanced] and [+Advanced], for which Clements proposes a separate

autosegmental tier, are arguably not even distinctive, and to point to

what appears to be a more promising way ahead in the attempt to capture just

what it is that tonal phenomena and vowel harmony phenomena have in common.

REFERENCES

Anonymous. 1944. *Fante Word List with Principles and Rules of Spelling*
     (Third Edition). Cape Coast: Methodist Book Depot.

Anonymous. n.d. *Fante-English Dictionary*. Cape Coast: Methodist Book
     Depot.

Berry, J. 1957. "Vowel harmony in Twi." *Bulletin of the School of
     Oriental and African Studies* 19.1:124-30.

Chomsky, N. and M. Halle. 1968. *The Sound Pattern of English*. New York:
     Harper and Row.

Christaller, J. 1875. *A Grammar of the Asante and Fante Language Called
     Tshi (Chwee, Twi)*. Basel: Basel Evangelical Missionary Society.

Christaller, J. 1933. *Dictionary of the Asante and Fante Language Called
     Tshi (Twi)* (Second Edition). Basel: Basel Evangelical Missionary
     Society.

Clements, G. 1981a. "The hierarchical representation of tone features."
     Harvard Studies in Phonology 2:50-107. Bloomington: Indiana Univer-
     sity Linguistics Club.

Clements, G. 1981b. "Akan vowel harmony: a nonlinear analysis." *Harvard Studies in Phonology* 2:108-77. Bloomington: Indiana University Linguistics Club.

Clements, G. and K. Ford. 1979. "Kikuyu tone shift and its synchronic consequences." *Linguistic Inquiry* 10.2:179-210.

Dimmendaal, G. 1982. *The Turkana Language.* Ph.D. dissertation, University of Leiden.

Goldsmith, J. 1976. *Autosegmental Phonology.* Bloomington: Indiana University Linguistics Club.

Hockett, C. 1958. *A Course in Modern Linguistics.* New York: Macmillan.

Hyman, L. and M. Tadadjeu. 1976. "Floating tones in Mbam-Nkam." In Larry M. Hyman (ed.), *Studies in Bantu Tonology*, Southern California Occasional Papers in Linguistics 3:57-111. Los Angeles: University of Southern California.

Schachter, P. and V. Fromkin. 1968. *A Phonology of Akan: Akuapem, Asante, and Fante.* UCLA Working Papers in Phonetics 9. Los Angeles: University of California.

Stanley, R. 1967. "Redundancy rules in phonology." *Language* 43:393-436.

Stewart, J. 1962. *An Analysis of the Structure of the Fante Verb, with Special Reference to Tone and Glottalization.* Ph.D. dissertation, University of London.

Stewart, J. [1965.] "The typology of the Twi tone system." *Bulletin of the Institute of African Studies* [Legon, Ghana] 1:1-27.

Stewart, J. 1967. "Tongue root position in Akan vowel harmony." *Phonetica* 16:185-204.

Stewart, J. 1981. "Key lowering (downstep/downglide) in Dschang." *Journal of African Languages and Linguistics* 3:113-38.

Stewart, J. 1982. "Vokaalharmonie in het Akan: de woordstructuurcondities." *Glot* 5:321-41.

Stewart, J. Forthcoming. "The Asante Twi tone shift." In I. R. Dihoff (ed.), *Current Approaches to African Linguistics* 1. Dordrecht: Foris.

Tadadjeu, M. 1974. "Floating tones, shifting rules, and downstep in Dschang-Bamileke." In W. R. Leben (ed.), *Papers from the Fifth Annual Conference on African Linguistics*, pp. 283-90. *Studies in African Linguistics*, supplement 5. Los Angeles: Department of Linguistics, University of California.

Voorhoeve, J., A. Meeussen, and K. de Blois. 1969. "New proposals for the
     description of tone sequences in the Igbo completive phrase."
     Journal of West African Languages   6:79-84.

Welmers, W. 1946. *A Descriptive Grammar of Fanti*. Supplement to *Language*
     22.3: Language Dissertation no. 39. Baltimore: Linguistic Society
     of America.

DISCOURSE STRATEGIES IN PULAAR: THE USE OF FOCUS

Sonja Fagerberg
University of Wisconsin--Madison

This paper uses the concept of discourse cohesion as a basis
for looking at certain sentence emphasis patterns in Pulaar
(Fula). In particular, four patterns of focus are examined,
first syntactically and then pragmatically. Focus in Pulaar
can be divided according to clefting and pseudo-clefting pat-
terns for subject and non-subject focus. These four patterns
are discussed in depth. Then, I suggest that the idea of
discourse connectedness provides a model for explaining the
choice of a focus pattern in a given environment. Examples
of this are drawn from a transcription of the story of
Gelaajo Ham Bodeejo, performed by Mammadu Ñon Giise, a griot
from northern Senegal.

1. Introduction

    Linguistic literature in recent years has been full of references to
the pragmatic concept of sentence emphasis. This paper looks at the feature
of [+focus] in Pulaar[1] as one special type of sentence emphasis in that lan-
guage. Drawing from theoretical work in pragmatics, functional sentence
perspective, and discourse analysis, I will move from a discussion of the
shape of non-verb focus in Pulaar to a discussion of the use of focus in a
Pulaar text.[2] Language is viewed as a text-forming, cohesive phenomenon in
which certain features, emphasis being one of them, must be studied in a
context or environment. I will discuss first the syntactic rules (insertion

---

    [1]Pulaar is a language of the West Atlantic branch of Niger-Congo. It
is spoken throughout the West African sahel. It is known as Fulfulde in all
dialects spoken from Mali eastward. (In the literature it is most commonly
referred to as "Fula", "Fulani", or "Peul".) The particular dialect under
consideration here is that of the Haalpulaar'en (or "Toucouleur") in
northern Senegal.

    [2]This article is based upon research carried out in Senegal from 1976-
1981 for a doctoral dissertation. The text examples which are used in this
paper come from a transcription of the narrative of *Gelaajo Ham Bodeejo*.
This text, of roughly 500 lines, is included in the appendix of the disser-
tation. (See bibliography.)

rules, deletion rules, movement rules, copying rules, etc.) which mark
the feature of focus in Pulaar. And then I will discuss some of the
semantic implications of and pragmatic reasons for the use of emphasis,
and for the choice of one emphatic pattern over another in a given con-
text. Throughout I am indebted to the distinction which Givón [1979:32]
draws between "categorial" rules and "strategy" rules. According to
him:

> "The human communicator is not a deterministic user of an
> autonomous, subconscious grammar as Chomsky would have us
> believe. Rather, he makes *communicative choices*. He uses
> rules of grammar for a communicative effect."

The following discussion should bring out both the categorial rules of
focus in Pulaar, and some of the strategy "rules" (or reasons) for its
use. We will look at both the shape or form of non-verb focus, as well
as the possible rules governing the choice to use or not to use focus.
And when it is used, we will examine the factors determining the choice
of one focus pattern from four possibilities.

2.   The Shape of Focus

    Semantically, I am taking "focus" to mean a sentence which has a *marked
assertion* which contrasts with its *presupposition*. In Pulaar, focus can
first be divided into two types: verb focus and non-verb focus. Verb focus
is achieved through the use of a special set of suffixes replacing the
tense-aspect-mood markers, and will not be considered here. Non-verb focus
is achieved through the addition of the focusing particle ko [3] to a

---

[3]The particle ko carries a very heavy load in this particular dia-
lect. (It is interesting to note that it has dropped many or most of these
functions in the eastern dialects, including its function as a marker of
emphasis or focus.) Among its most important functions are: 1) as the
copula joining two nominals in a relationship of equivalency, identity,
location, or role, 2) as the "presentative" marker ('it is...') followed
by a nominal, 3) as the relative pronominal equivalent to the English
'what', 4) as the interrogative question word equivalent to "what?",
5) as the complementizer before an S-complement. It can also take several
idiomatic roles and meanings.

neutral sentence[4] with the incumbent morpho-syntactic changes which accompany ko insertion. The use of the emphatic particle ko activates three additional focus features, which may be optional or obligatory. These include: 1) a change in the neutral SVO word order (optional), 2) the use of the relative (dependent or consecutive) verb form (obligatory), and 3) a non-clitic pronoun or NP from the independent series of pronominals (either emphatic, deictic, or referential) when the NP is in focus (obligatory).

There are four patterns of ko insertion, two being used for subject focus and two for non-subject focus. As we shall see, these result in two cleft patterns and two pseudo-cleft patterns, distinguished in Pulaar by the position of ko and/or the focused NP. These four patterns are variants of a neutral sentence. If we take the following neutral SVO sentence:

(1) a. a   walii   mo   'you helped him'
         you   have-helped him
         S   V   O
         clitic

it can be permuted in the following four ways:

PATTERN 1 - <u>Subject focus through clefting</u>

(1) b. ko   aan   walii   mo   'it is you who helped him'
         it-is you   help   him
         FOCUS S   V   O
               emphatic relative

This sentence pattern could be formulated as follows. (The parentheses indicate something which is syntactically, though not pragmatically, optional. The forms which the pronominal and verb may take are listed below each symbol. The X represented all non-subjects.)

---

[4]The following discussion is based upon the assumption that there is a distinction between a neutral sentence and a sentence with a morpho-syntactic marker for emphasis. I take the following two features as indicative of a neutral sentence pattern in Pulaar: 1) that it take the word order (Ad)SVO(Ad), and 2) that at least the first verb takes as its tense-aspect-mood marker or suffix an independent morpheme, rather than a consecutive or relative form.

PATTERN 1:   (ko   )   S   V   (X)
                             emphatic   relative
                             deictic
                             referential
                             nominal

PATTERN 2 - <u>Subject focus through pseudo-clefting</u>

(1) c.   mo   walli   mo   ko   aan   'the one who helped him was you'
          he-who   help   him   it-is   you
          S COPY   V   O   FOCUS   S
          REL PRO   relative   emphatic

Expressed in a formula, we get:

PATTERN 2:   S   V   (X)   ko   S
                    relative copy   relative   emphatic
                    pronoun   deictic
                                                             referential
                                                             nominal

PATTERN 3 - <u>Non-subject focus through clefting</u>

(1) d.   ko   kaŋko   mballu   -ɗaa   'it is he whom you helped'
          it-is   him   help   you
          FOCUS   O   V   S
                   emphatic   relative   postposed clitic

The formula for this pattern is a bit more complex because the order of sub-
ject and verb depends upon the person, a regular feature in the relative
verb forms. Basically, first and second person subject clitics follow the
verb, while the third person clitics precede. This fact is indicated by the
symbol $\frac{S-V}{V-S}$.

PATTERN 3:   (ko   )   X $\frac{S-V}{V-S}$
                             emphatic   relative
                             deictic
                             referential
                             nominal

PATTERN 4 - <u>Non-subject focus through pseudo-clefting</u>

(1) e.   mballu   -ɗaa   ko   kaŋko   'whom you helped was him'
          help   you   it-is   him
          V   S   FOCUS   O
          relative   postposed clitic   emphatic

Or we can formulate this pattern as:

```
            S-V
PATTERN 4:  V-S          ko  X
                relative     emphatic
                             deictic
                             referential
                             nominal
```

These four patterns illustrate the four primary syntactic forms which focus can take. In order to correctly interpret the above patterns, notice the following features. First, the concept of non-verb focus hinges upon the introduction of the particle ko into the neutral sentence. However, ko in S-initial position is optional syntactically and may be deleted. In a focused sentence, ko deletion has no impact upon the syntax of the sentence itself. The syntactic adjustments related to the use of ko as a particle of emphasis remain (i.e. a focused pronominal cannot be a clitic and the verb must take the relative form). Whether or not it is deleted seems to be a communicative choice, not a syntactic one. Statistically, ko is deleted in a minority of cases, and usually in a context where several focused sentences follow one upon the other, the first one using ko while the others delete it.

However, ko deletion does not apply to anything but S-initial position. Within the sentence, ko must appear. For example:

(2) a. ko aan walli mo   'it is you who helped him'

    b. ∅ aan walli mo   'it is you who helped him'

    c. mballu-ɗaa ko kaŋko   'who you helped was him'

    d. *mballu-ɗaa ∅ kaŋko

The final example in the series is incorrect because ko in that position cannot be deleted. To generalize, one could say that ko is *optional in cleft sentences, but obligatory in pseudo-clefts.*

Secondly, for the non-verb focus patterns, the relevant NP's can be divided into subject (marked by S) and non-subject (marked by X). That is we can talk about patterns of subject focus and patterns of non-subject focus. And any NP in the sentence may be focused upon in both a cleft

(S-initial) and pseudo-cleft (S-final) pattern.

Third, the word order of the subject and verb in a focus sentence depends upon person. This is marked by the patterns $\frac{S-V}{V-S}$. Whenever the relative verb form is used, it is marked by a combination of verbal suffixes and the position (and form) of the subject clitic. In a relative or dependent verb form, the first and second person subject clitics follow the verb. However, the third person clitics, as well as the first person plural exclusive pronoun, precede it.

Fourth, a clitic pronoun can never be the NP in focus. The NP in focus must be either a full nominal, or one of the three independent forms--the emphatic, deictic, or referential pronouns. These pronominal forms can function similarly to full nominals in a variety of syntactic environments in Pulaar, focusing being one of them.

Finally, notice that pattern 2 depends upon a pronominal copy which holds the place of the focused subject, now shifted to S-final position. The copy is a relative pronoun. This particular pattern is used very rarely, perhaps because shifting the focused NP out of S-initial position into S-final position is not congruent with the function of focus. Focusing and S-initial position seem to be features which support each other. When it is a S-final non-subject which is focused upon, it can be left in it normal S-final position, or it can be frontshifted to S-initial position. But to consciously shift an S-initial subject into S-final position when it is the NP in focus, as a pragmatically ambiguous operation, requires an additional syntactic rule of pronominal copying. The place of the backshifted subject must be held by a pronominal copy of the subject, which now appears in S-final (or pseudo-cleft) position. As was already mentioned, this strategy for focusing is very rarely used in spontaneous speech.

The features relevant to non-verb focus in Pulaar can be summarized in Table 1. Patterns 1 and 2 have in common the fact that they both focus upon subjects, while 3 and 4 focus upon non-subjects. However, patterns 1 and 3 are similar in that they are both clefting patterns, whereas 2 and 4 are pseudo-clefts.

Table 1. Non-verb focus in Pulaar

	Pattern 1	Pattern 2	Pattern 3	Pattern 4
ko in S-initial position	X		X	
clefting pattern	X		X	
ko in S-final position		X		X
pseudo-clefting pattern		X		X
optional ko deletion	X		X	
subject focused	X	X		
non-subject focused			X	X
relative verb form used	X	X	X	X
non-clitic pronominal in focus position	X	X	X	X
word order changed		X	X	
frontshifting change			X	
backshifting change		X		
pronominal place-holder		X		
asserted NP is S-initial	X		X	
asserted NP is S-final		X		X

The feature of [+focus] as discussed above has been captured by the *rule of* ko *insertion*. That rule states that any NP of a sentence may be brought into focus through the addition of the focusing particle ko , plus some combination of the following three "rules":

(i)   a movement rule: It is common to find the focused non-subject pro-
moted to S-initial position (PATTERN 3). It is
less common but still possible to find a focused
subject demoted to S-final position (PATTERN 2.)
In that case, a special pronominal copying rule
also applies.

(ii)   a deletion rule: Whenever ko occurs in S-initial position it may
                          optionally be deleted. Rules for the application
                          of this feature are probably pragmatic, rather
                          than syntactic. This applies to PATTERN 1 and
                          PATTERN 3. Any application of rule (1) or rule
                          (3) that applied before ko was deleted is main-
                          tained.

(iii) an agreement rule: The use of ko calls for two types of agreement
                          rules: a) the use of the relative verb form, and
                          b) the use of a non-clitic pronoun in focused
                          position (either a full nominal or an independent
                          pronoun).

The application of these rules is summarized in the following table.

Table 2. Rule application

| | Movement rule | Deletion rule | Agreement rules | |
			verb /	pronoun
Pattern 1: (subject) (cleft)		X	X	X
Pattern 2: (subject) (pseudo-cleft)	X (plus pronoun copy)		X	X
Pattern 3: (non-subject) (cleft)	X	X	X	X
Pattern 4: (non-subject) (pseudo-cleft)			X	X

The differing degrees of syntactic complexity between the four patterns
become evident in this last chart. But it also should be clear that each
pattern depends upon the manipulation of the same basic features. Only
pattern 2, a statistically uncommon pattern, makes use of an additional
syntactic operation.

3.   The Pragmatic Features Governing the Choice of [+focus]

     The above discussion has been primarily interested in the syntactic
features of focus. According to the definition quoted earlier by Givón

[1979], these were largely *categorial rules*--the 100% rules applied for grammatical "correctness". However, the question of *strategy rules* also came up. For example, the question of optional ko deletion was left to possible pragmatic factors.

In this section, I will take up the question of pragmatic features by trying to determine what differentiates between the two possible patterns of non-subject focus.[5] In other words, when would one choose pattern 3, and when pattern 4 in a discourse situation? What I should like to determine is the distributional range of the two following sentences:

PATTERN 3:   ko kaŋko mballu-ɗaa   'it is *he* whom you helped'

and:

PATTERN 4:   mballu-ɗaa ko kaŋko   'whom you helped was *he*'

Both have focus marked by ko . Both focus upon the non-subject--in this case the object mo (him/her) represented here by the emphatic pronoun kaŋko . But I would like to suggest that their distribution within a text would not be the same.

Too much of the literature on focus (and topic) has been confused by attempts at a static, "semantic" definition of the focused NP as the "new" or "unknown" element being asserted in an uncontextualized sentence. This has led to definitions of the following sort:

> "While *theme* is the given point of departure, *focus* is new
> information in the sense that it is textually and situation-
> ally nonderivable, nonanaphoric, although not necessarily
> factually new information" (Justus [1976:219].)

or:

---

[5]Theoretically, one should be able to draw the same distributions between patterns 1 and 2, differentiating the patterns of subject focus. However, after transcribing several hundred pages of various texts, I did not find a sufficient number of sentences using pattern 2 (which "demotes" the focused NP to S-final position) to draw any conclusions. And the text of *Gelaajo Ham Boɗeejo*, which provides the basis of evidence for this paper, does not use this pattern even once.

> "I use 'topic' to refer to a major consituent...that is
> usually in sentence-initial position, and that expresses
> known, given information...In contrast, I use 'focus' to
> refer to a major constituent carrying new information..."
> (McIntosh [n.d.:19]).

There seems to be some confusion over the idea of focus as being "new" and
therefore "nonderivable". As we shall see shortly, focus in Pulaar may be
derivable from either what precedes or what follows, depending upon whether
it is a cleft or a pseudo-cleft pattern.

If we use the new/old information model, we immediately run into diffi-
culties on the level of discourse. A focus sentence is not an entity in
isolation with something in it presupposed and something else asserted, as
so many of these definitions limited to the intrinsic value of the focused
NP would have us assume. A focus sentence is one which is *marked* as having
a *special discourse function*. The reason for choosing to use focus cer-
tainly has more to do with the needs of the discourse and the communicative
choices open to the speaker than this static portrayal of its component
parts would have us believe. After all, sentence emphasis is a preeminently
contextualized phenomenon. It occurs within a larger context to meet the
needs of that context--or more precisely, the needs of communicating within
a speech context.

It seems to me that the pragmatic value of a focus sentence is less
adequately defined by emphasizing the supposed nature (new/old, thematic/
rhematic, known/unknown) of the focused NP itself, than by visualing the
environment in which the two types of focus patterns (cleft or pseudo-cleft)
can occur, and marking the *connections* between them. Hetzron [1975:348]
comes to the point with this simple definition:

> "With Jesperson [1924:145] I feel that the "new information"
> is not always contained in the predicate (this term being
> used here in a sense equivalent to 'rheme'), but it is
> always inherent in the *connexion* of the two elements - the
> fact that these two elements are put together. At issue is
> not the novelty of the element, but what the speaker intends
> to build up in the discourse."

In this section, we are talking less about categorial rules and more about

strategy rules for effective communication. That is, rules or generaliza-
tions which can help to predict and understand communication choices when a
language is actually in use. And generally speaking, with strategy rules we
are probably always talking about "degrees of adequacy", rather than any
grammatical absolutes measured as "correct" or "incorrect". The first
definitions offered above I find inadequate to a number of situations when
reading a Pulaar text. Whereas I find a much higher degree of adequacy--a
higher degree of predictability and a less complicated model for describing--
in the following definition offered by Hetzron [1975:364]:

> "Both constructions, cleft and cataphora [or pseudo-cleft]
> are instances of focusing (*mise en relief*) which elevate
> the communicational importance of an element above the
> level of the rest of the sentence. Yet the motivation for
> such focusing may be varied. When an element is focused
> because it fills a gap in previous knowledge, it is brought
> *forward* in a cleft construction or another type of emphatic
> construction. When the focusing is necessary for paving
> the way for later use of the same element in the discourse
> or for a pragmatic reaction, the cataphoric construction
> that moves the focused element to the *end* of the sentence
> is created."

That is to say, there is a relationship implied and established by a focus
construction. That relationship may either go backwards or forwards in the
discourse. It is that *relationship* which is the essence of a focus con-
struction. It is that relationship which determines the choice, in English,
between an *it*-cleft and a *WH*-pseudo-cleft (see Prince [1978]). And as we
shall see, in Pulaar it is that relationship which determines the choice
between pattern 3 and pattern 4.

In fact, I would like to present a simple hypothesis about the use of
these two patterns which is related to the *position* of the asserted NP. It
is simply this: that in Pulaar an asserted NP is brought forward in the
sentence when it is "connected" to something which has preceded it in the
discourse, whereas the asserted element is postposed if it is connected to
something which is still to come. *Frontshifting* or *backshifting* are simply
factors in the larger *backward* or *forward connection*.

This very simple model for explaining the choice may, of course, be

elaborated upon with more "semantic" explanations. Obviously, when the con-
nection runs between the focus sentence and something which follows in the
discourse, the focus sentence *in its entirety* is probably introducing some-
thing new which is now picked up upon in the following discourse. Thus the
definitions "new", or "theme", or even "introduction" could be tacked onto
this focused sentence. And when the focus sentence comes in final position
after a flow of discourse, connecting the focused piece with what has pre-
ceded, the entire focus sentence could be called "old" or "given", or even
a "summary" of what came before. But these definitions must consider the
total focus sentence and its environment, not simply the focused NP in an
isolated sentence. Only then can these kinds of definitions and explana-
tions of the focus function have a place.

   Throughout the Pulaar narratives which I have transcribed to date, the
forwards or backwards connection is generally marked by either: a) direct
repetition of a lexical item, b) duplication of a given semantic range by
two lexically distinct items, or c) pronominalization. These clear and
easily observed syntactic features therefore become the primary means for
identifying the direction of the connection; and thereby for explaining
the choice between a cleft and a pseudo-cleft pattern. (Other features
may of course exist to mark connection. However, the three markers listed
above account for all the data collected so far.) Taking the text of
Gelaajo Ham Bodeejo[6] as a basis for testing the above hypothesis, I found
that there was 100% correlation between the placement of ko (and/or
the focused NP) and the direction of a clear connection with the other
elements in the discourse for the thirty examples of non-subject focus
occurring in the text. That is, as a model it explained and predicted
the phenomenon every time.

   In the following example from the story of Gelaajo, the focus sentence

---

   [6]This text, recited by Mammadu Ñon Giise, a griot from northern Sene-
gal (Hoore Fonde), is included at the end of my doctoral dissertation. It
is roughly 500 lines in length, and includes numerous examples of focus in
operation.

(using the particle ko [7]) comes second. This sentence is an example of
frontshifted non-subject focus (pattern 3). The placement of ko within
the focused sentence is to the left. And the connection between the
focused sentence and the larger discourse is also to the left, marked in
this case by repetition.

(3) Sabu   o wi'ii   ko   o hirataa   mi laarat golle makko.
     because he said   that   he isn't-jealous I look-at deeds his
                 PERF. COMP.   NEG.IMPERF.   IMPERF.

     Ko   silsil golle   makko ɣeewan   -mi.
     it-is reality of-deeds his   look-for I
     FOCUS OBJECT   RELATIVE SUBJECT
                                    IMPERFECT

     'Because he said that he isn't/wouldn't be jealous, I am going to check
     on his actions. It is the reality of his actions (not just his words)
     which I am going to examine.'

The connection between these two sentences is indicated through repetition.
Notice the direct repetition of golle makko 'his deeds, actions', which
establishes a clear link between the sentences. Secondly, the semantically
similar idea of 'to look at, examine' is given in two lexically distinct
verbs, ɣeewa and laarat . The connection between the first non-focused
sentence and the second sentence with the focused element in initial posi-
tion is made twice; once through word-for-word repetition and once through
the repetition of a semantic idea.

In the following example which illustrates the same set of relation-
ships, all of the elements occur within one sentence which has two clauses.
The second clause is a focus construction. Using pattern 3 again, it fronts
an S-final non-subject to S-initial position.

---

[7]Notice that the first ko in (3) is the complementizer, not the
focus particle.

(4) Faɗo pii- maa-mi adan, ko   ŋgoo piino-ɗaa Gelaajo.
     shoe hit   you I first it-is that hit   you Gelaajo
         RELA. O   S   FOCUS I.O. RELA. S   D.O.
         PERF.   PERF.

   '(The) shoe (which) I hit you (with) first, it is *that* (with which)
   you had hit Gelaajo.'

This time, the connection is made both through repetition and pronominali-
zation. The verb fi'a (realized in both cases in the form pii-)
meaning 'to hit' is repeated. Secondly, the nominal faɗo 'shoe' from
the first clause is pronominally copied by the concordant class pronoun
ŋgoo in the deictic form. The focused or asserted NP is brought forward,
and the connection is drawn between the focus sentence and what has pre-
ceded, as my hypothesis would have predicted.

      In contrast, postponing the focused element is due to a connection
between the focus sentence and something which follows in the discourse.
Once again, there are several clear examples in the text of *Gelaajo Ham
Boɗeejo* which use direct repetition and pronominalization to mark this
feature.

(5) O yontaa   ko   to debbo jeewo to. O raaŋani   suudu
     he is-due-at it-is at wife first the he draw-towards room
     S   FOCUS COMPLEMENT

     debbo   jeewo o.
     of-wife first the

     'Where he is due is at his first wife's (room). He drew near his
     first wife's room.'

This passage introduces completely new material within the text where it
occurs. Up to this point, the narrator has been concerned with the exploits
of Gelaajo in a far-off village. Suddenly, he now introduces Gelaajo in
his own home. Semantically, the focused sentence provides a certain amount
of new material in a new setting, which makes what follows comprehensible.
It introduces something new through a focus construction with the ko
clause in S-final position. The repeated element shared by these two
sentences is debbo jeewo o 'the first wife'. The focused non-subject NP
remains to the right in the sentence, and it is connected by this repetition

to a sentence on the right.

One final example of this rightward connection appears in the text sentence:

(6) Ŋgarnoo-mi ko   yoptaade.   Mi yoptiima   paɗe ɗe.
    came   I it-is to-get-revenge I got-revenge shoes the
    RELA.   S FOCUS COMPLEMENT   S PERF.
    PERF.

    'What I came (for) is revenge. I have avenged (the slap across the
    face with) the shoes.'

Here we find the idea of yoptoo 'to get revenge, to avenge' repeated,
once in the infinitive form and once in the perfect. Focus is to the right
within the sentence, as is the discourse connection.

Marking the focus on the right in the above example connects it to
what follows (on the right as well). That describes the inter-sentence
relationships which we are observing here. Semantically, we can perhaps
say now that the focused sentence is being used to introduce new materials
into the narrative. The idea of revenge, which is the focused element, is
mentioned here for the first time. However, while this might explain the
"why" of focus, it does not explain why this particular pattern of focus
is used, rather than another. The explanation involving discourse con-
nectedness (*cohesion*) does explain the choice of pseudo-clefting over
clefting.

4. Conclusions

This paper has considered some of the basic features of focus, as one
type of sentence emphasis, in Pulaar. It began with a look at four of the
primary sentence patterns which can be identified syntactically as focus
patterns. Each involved the introduction of the focusing particle ko .
It became evident that the position of ko (or the focused clause in the
case of ko deletion) is relevant to the interpretation of the patterns as
cleft or pseudo-cleft. Both types of pattern can be applied to any NP in
the sentence, but the NP's can be grouped as either subjects or non-subjects.

Secondly, I suggested an hypothesis which might account for the distri-

butional differences between the two common forms of non-subject focus
patterns, one a cleft pattern and the other a pseudo-cleft. The hypothesis
is simple: that a focused sentence is one which expresses *connections in
discourse*, and that the type of focus sentence chosen depends upon the
direction of the connection and the sentence-internal *position* of the ko
clause. A focused NP in S-initial position is mirrored by a connection
branching to the left. And a focused element in final position signals
connections to the right. The semantic interpretations of these syntactic
features may include readings such as "new", "old", "given", etc. But the
model of connection is much simpler to use. Secondly, these semantic
definitions must apply to the entire focus sentence and its role in the
discourse, not simply to the component parts of the focus construction.
Finally, these definitions may be able to explain the feature of [+focus]
in the sentence, but they are not adequate for explaining the choice
between a clefting pattern (patterns 1 and 3) and a pseudo-clefting pattern
(patterns 2 and 4).

REFERENCES

Akmajian, A. 1970. "On deriving cleft sentences from pseudo-cleft
      sentences." *Linguistic Inquiry* 1:149-168.

Arnott, D. 1970. *The Nominal and Verbal Systems of Fula*. Oxford:
      Clarendon Press.

Chafe, W. 1976. "Givenness, contrastiveness, definiteness, subjects,
      and topics." In Charles Li (ed.), *Subject and Topic*, pp. 27-55.
      New York: Academic Press.

Dijk, T. V. 1977. *Text and Context: Explorations in the Semantics and
      Pragmatics of Discourse*. New York: Longman.

Fagerberg, S. 1982. "Syntactic expansions in text: beyond SVO in Pulaar
      oral narrative performance." Ph.D. dissertation, Madison, Wisconsin.

Givón, T. 1979. *On Understanding Grammar*. New York: Academic Press.

Grimes, J. 1978. *Papers on Discourse*. Dallas: Summer Institute of Linguistics.

Halliday, M.A.K. and R. Hasan. 1976. *Cohesion in English*. London: Longmans.

Hetzron, R. 1975. "The presentative movement or why the ideal word order is V.S.O.P." In Charles Li (ed.), *Word Order and Word Order Change*, pp. 115-147. Austin: University of Texas Press.

Jesperson, O. 1924. *Philosophy of Grammar*. London: Allen & Unwin.

Justus, C. 1976. "Relativization and topicalization in Hittite." In Charles Li (ed.), *Subject and Topic*, pp. 215-245. New York: Academic Press.

McIntosh, M. Forthcoming. "Topic and focus in Fulfulde."

McIntosh, M. (n.d.). "Fulfulde as a VSO language." Unpublished paper.

Prince, E. 1978. "A comparison of *WH*-clefts and *it*-clefts in discourse." *Language* 54.4:883-906

Sylla, Y. 1979. "Grammatical relations and Fula syntax." Ph.D. dissertation, UCLA.

THE SYNCHRONIC BEHAVIOR OF BASIC COLOR TERMS IN TSWANA
AND ITS DIACHRONIC IMPLICATIONS*

Ronald P. Schaefer
University of Benin

The synchronic distributional pattern of potential basic color
terms in one dialect of Tswana is examined in a wide range of
construction types. From this pattern the non-basic status of
the term lèphùtsí emerges, as well as a constraint requiring
the exclusion of animals from the semantic extension of basic
terms designating hue. Accepting lèphùtsí as non-basic, how-
ever, leaves a pattern of semantic reference violating a widely
assumed universal constraint governing historical stages in the
evolution of color names. To resolve this dilemma, a compara-
tive analysis of color term reference in the Sotho languages is
undertaken. Based on this analysis, the semantic reference for
one basic color term in Tswana is hypothesized to have under-
gone a historical change, whereby the universal constraints on
color naming give way to the constraint governing basic terms
for hue.

1.   Introduction

The semantic categorization of the color domain in Tswana, a Southeastern

Bantu language, appears on initial examination to challenge a theory of the

historical evolution of color naming which has received wide recognition

(Berlin and Kay [1969], Kay [1975], and Kay and McDaniel [1978]). More

precisely, a system of color names obtained over a period of months from a

Tswana informant appears inconsistent with predictions from this theory. To
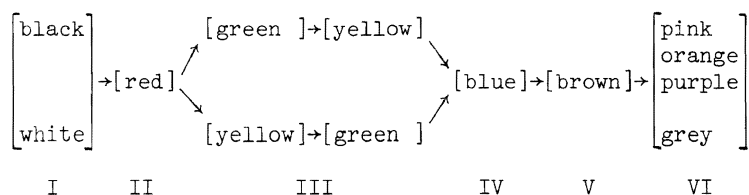
establish this point, the synchronic distributional behavior of potential
basic color terms will be examined in a wide range of construction types.
On the basis of this behavior, we will argue that the inconsistency is due
to a conflict between language specific and language general constraints.

2.   Review of Basic Color Terms

Over the past decade, the theoretical foundation for a series of
interlocking studies in linguistics, anthropology, psychology, and neuro-
physiology has been Berlin and Kay [1969]. Their investigation, based on
experimental evidence and dictionary analyses of 98 languages, argued that
the categorization of color by natural language is far from arbitrary,
contrary to Gleason [1955] and Ullman [1962]. In each language they
examined, no more than eleven, and no fewer than two, basic color terms
could be identified, with each term across languages referring to one of
eleven focal areas in the color spectrum. Concurrent with these synchronic
claims, they postulated a relatively fixed universal sequence of stages
through which a system of basic color terms would historically evolve.
This sequence, originally interpreted as the successive encoding of the
eleven foci, is shown in Table 1.

Table 1.   Evolutionary sequence on color terms from Berlin and Kay [1969]

$$\begin{bmatrix} \text{black} \\ \\ \\ \text{white} \end{bmatrix} \to [\text{red}] \begin{matrix} \nearrow \\ \searrow \end{matrix} \begin{matrix} [\text{green }]\to[\text{yellow}] \\ \\ [\text{yellow}]\to[\text{green }] \end{matrix} \begin{matrix} \searrow \\ \\ \nearrow \end{matrix} [\text{blue}]\to[\text{brown}]\to \begin{bmatrix} \text{pink} \\ \text{orange} \\ \text{purple} \\ \\ \text{grey} \end{bmatrix}$$

$$\quad\text{I}\qquad\text{II}\qquad\qquad\text{III}\qquad\qquad\text{IV}\qquad\text{V}\qquad\text{VI}$$

Subsequent research has altered the details of this sequence as well
as sharpened our understanding of the nature and source of the categories
underlying basic color terms. In the revised color term sequence, shown in
Table 2, the stages are viewed as a progressive differentiation of the

dimensions hue and brightness (Kay and McDaniels [1978]).[1]

Table 2. Revised evolutionary sequence of color terms from Kay [1975] and Berlin and Berlin [1975]

$$
\begin{bmatrix} \text{MACRO-BLACK} \\ \\ \\ \text{MACRO-WHITE} \end{bmatrix} \rightarrow [\text{ROW}] \begin{matrix} [\text{GRUE}]\rightarrow[\text{yellow}] \\ \nearrow \\ \searrow \\ [\text{yellow}]\rightarrow[\text{GRUE}] \end{matrix} \quad [\text{green \& blue}]\rightarrow[\text{brown}]\rightarrow \begin{bmatrix} \text{pink} \\ \text{orange} \\ \text{purple} \end{bmatrix}
$$

---------------------------------------[grey]

```
  I        II         III              IV          V      VI
```
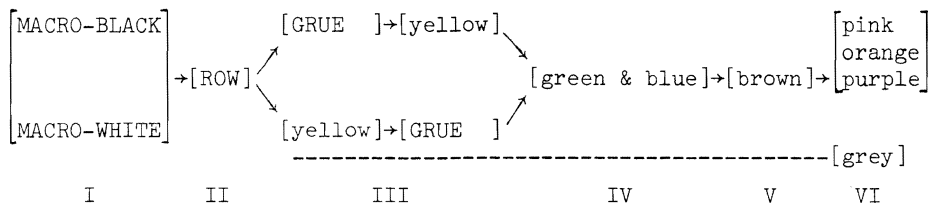
For example, the semantic categories in the first stage combine the brightness and hue dimensions such that foci identified by light/warm colors (MACRO-WHITE) and foci identified by dark/cool colors (MACRO-BLACK) are contrasted. In the second stage, the light/warm category is differentiated as a light (white) and a warm (red) category. Hue and brightness, along with saturation, also serve as the traditional dimensions for analyzing color categorization. Of these, the role of saturation, referring to chromatic purity, is little understood, and, even today, only the dimensions hue and brightness have been linked to specific neurophysiological components (Bornstein [1973a,b] and McDaniel [1974]).

The revision of the color naming sequence has been strengthened by studies of synchronic variation (Kay [1975] and Berlin and Berlin [1975]). Though all speakers of a language may not manifest a uniform stage in their color naming behavior, they do apparently manifest adjacent stages. The general finding is that development of new basic terms from non-basic terms, and distribution of basic and non-basic terms across dialects, conforms to the stages delineated in the revised Berlin and Kay sequence. Basic color term systems existing in different dialects of a language should thus mani-

---

[1]The color foci for the categories identified with capital letters are: BLACK, GREEN, and BLUE for MACRO-BLACK; RED, YELLOW, and WHITE for MACRO-WHITE; RED and YELLOW for ROW; and GREEN and BLUE for GRUE.

fest a single stage. For example, if a system representative of a dialect
contains a basic term for BLUE and GREEN, it should also contain a basic term
for YELLOW. Subsequent discussion herein will take up this prediction as it
applies to the data from Tswana.

Based on recent studies, it is apparent that there may be psycho-
physiologically natural points in the color spectrum which ground the
semantic categorization of color. Despite this grounding, the linguistic
task of identifying basic color terms has proven problematic for a number
of investigators, Wescott's [1970] analysis of Bini, a Kwa language of
Africa, being a case in point. Wescott attempted to identify basic color
terms by using the Berlin and Kay [1969] criteria, but they proved indeter-
minate in assigning Bini to one of the original diachronic stages.[2]

To strengthen the means for determining basic color terms, Westcott
[1970] posited additional criteria.[3] What is significant about these
criteria is that they flesh out the general criterion of "similar distri-
bution" advanced by Berlin and Kay [1969] and call attention to the value
of extensive examination of color term distribution. Accordingly, the
criteria of Berlin and Kay, combined with those from Wescott, were used to
identify a set of potential basic color terms in the Tswana data under con-
sideration.

Two dominant properties of the distributional behavior of these terms will
receive attention. The first involves the single term lèphùtsí , which

---

[2]The criteria of Berlin and Kay [1969] specify that a basic color term
should be the following: it should be monolexemic, its meaning should not be
included in that of any other term, it should not be restricted to a narrow
class of objects, it should be psychologically salient, it should have the
same distribution as other basic terms, it should not name an object having
that color, and it should not be a load word.

[3]The supplemental criteria proposed by Wescott [1970] include: poly-
lexicity (the number of words for each color term), polytypy (the number of
parts of speech represented by each color term), polymorphy (the number of
allomorphs for each color term), onymicity (the extent a color term is used
as the name of a person or place), and metaphoricity (the extent a color term
is used in figurative expressions).

deviates from the distributional pattern established by the other potential basic color terms. The second involves a pattern of complementary distribution within this set of terms.

3. Basic Color Terms in Tswana

3.1. <u>Distribution of lèphùtsí</u>. In Berlin and Kay [1969] one of the principle criteria defining basic terms is their similar, if not identical, pattern of distribution. The behavior of the Tswana color term lèphùtsí does not meet this criterion. In an extremely wide range of construction types, the distributional behavior of lèphùtsí, meaning YELLOW, is irregular. An initial argument supporting this conclusion is the abstract noun construction. In order to refer to the abstract nature of a color, the Class 7 singular marker, bò-, is prefixed to the relevant color term root. As the construction types in Tables 3 and 4 reveal, this marker is not prefixed to lèphùtsí.

Table 3. Class 7 singular constructions

1. bò-ǹtšhó   ké-m̀málá   ká-sètšwáná
   black   be color   in Tswana
   'black   is a color in Tswana'

2. bò-šwèú   ké-m̀málá   kà-sètšwánà
   'white   is a color in Tswana'

3. bò-húbídú   ké-m̀málá   ká-sètšwánà
   'red   is a color in Tswana'

4. bò-tàlá   ké-m̀málá   ká-sètšwánà
   'green   is a color in Tswana'

5. lèphùtsí   ké-m̀málá   ká-sètšwánà
   'yellow   is a color in Tswana'

6. bò-pùtšwà   ké-m̀málá   ká-sètšwánà
   'blue   is a color in Tswana'

7. bò-sètlhá   ké-m̀málá   ká-sètšwánà
   'grey   is a color in Tswana'

Table 4. <u>Class 7 singular constructions</u>

1. bò-ǹtšhò   ké-m̀málá   mó-sèkéténg̀ sélè
    black   be color   in skirt   that
    'black   is a color in that skirt'

2. bò-šwèú   ké-m̀málá   mó-sèkéténg̀ sélè
    'white   is a color in that skirt'

3. bò-húbídú   ké-m̀málá   mó-sèkéténg̀ sélè
    'red   is a color in that skirt'

4. bò-tàlá   ké-m̀málá   mó-sèkéténg̀ sélè
    'green   is a color in that skirt'

5. lèphùtsí   ké-m̀málá   mó-sèkéténg̀ sélè
    'yellow   is a color in that skirt'

6. bò-pùtšwà   ké-m̀málá   mó-sèkéténg̀ sélè
    'blue   is a color in that skirt'

7. bò-sètlhá   ké-m̀málá   mó-sèkéténg̀ sélè
    'grey   is a color in that skirt'

The irregularity of lèphùtsí is demonstrated further by its dual
meaning. According to Berlin and Kay's [1969] criteria, a basic color term
should not name a color and an object possessing that same color, e.g. gold
as the name of an object and a color in English. Consider then the behavior
of color terms in Table 5, where a set of simile constructions employing each
term is introduced. lèphùtsí occurs both as a term for color, YELLOW, and
as a term for an object manifesting that color, a pumpkin. None of the other
potential basic terms exhibits this behavior.

A third argument showing the irregular behavior of lèphùtsí is found
in constructions where color terms are adjoined to the different class pre-
fixes marking the descriptive copulative. Tables 6, 7, 8, 9, 10, and 11
offer a representative sampling of descriptive copulative constructions.
The irregular patterning in these constructions is evident: lèphùtsí does
not adjoin with the first person singular marker kèmò- , the Class 1
singular marker mò- , the Class 1 plural marker bábá- , the Class 4
singular marker sésé- , and the Class 5 plural marker dídí- . More
specifically we notice that the second syllable of two syllable markers is
deleted while single syllable markers are deleted altogether.

Table 5. <u>Simile constructions</u>

1. é-ǹtšhò   jááká bò-síxò
   it be black   like the night
   'it is black like the night'

2. é-tšhwèú   jááká lè-rù
   'it is white like the cloud'

3. é-khúbídú   jááká mà-dí
   'it is red   like blood'

4. é-tàlá   jááká lè-tlhàrè
   'it is green like a leaf'

5. é-léphùtsí   jááká lèphùtsí
   'it is yellow like a pumpkin'

6. é-pùtšwà   jááká lè-wápì
   'it is blue   like the sky'

7. é-tshètlhá   jááká mò-lɔ̀rà
   'it is grey   like ash'

Table 6. <u>Descriptive copulative constructions with the</u>
          <u>first person singular marker</u>

1. kè-mò-ǹtšhò   xò-xó-fétà
   I be black   you be surpassed
   'I am blacker than you'

2. kè-mò-šwèú   xò-xó-fétà
   'I am whiter   than you'

3. kè-mò-húbídú   xò-xó-fétà
   'I am redder   than you'

4. kè-mò-tàlá   xò-xó-fétà
   'I am greener than you'

5. kè-lèphùtsí   xò-xó-fétà
   'I am yellower than you'

6. kè-mò-pùtšwà   xò-xó-fétà
   'I am bluer   than you'

7. kè-mò-sètlhá   xò-xó-fétà
   'I am greyer   than you'

Table 7.   Descriptive copulative constructions with the
           Class 1 singular marker and present tense

1.   mò-ngwè lé mó-ngwé ó-mó-ǹtšhò
     one   and one   be black
     'everyone is black'
2.   mò-ngwè lé mó-ngwé ó-mó-šwèú
     'everyone is white'
3.   mò-ngwè lé mó-ngwé ó-mó-húbídù
     'everyone is red'
4.   mò-ngwè lé mó-ngwé ó-mó-tàlá
     'everyone is green'
5.   mò-ngwè lé mó-ngwé ó-léphùtsí
     'everyone is yellow'
6.   mò-ngwè lé mó-ngwé ó-mó-pùtšwà
     'everyone is blue'
7.   mò-ngwè lé mó-ngwé ó-mó-sètlhá
     'everyone is grey'

Table 8.   Descriptive copulative constructions with the
           Class 1 singular marker and past tense

1.   mò-ngwè lé mó-ngwé ó-nè á-lé mó-ǹtšhò
     one   and one   be   past black
     'everyone was black'
2.   mò-ngwè lé mó-ngwé ó-nè á-lé mó-šwèú
     'everyone was white'
3.   mò-ngwè lé mó-ngwé ó-nè á-lé mó-húbídù
     'everyone was red'
4.   mò-ngwè lé mó-ngwé ó-nè á-lé mó-tàlá
     'everyone was green'
5.   mò-ngwè lé mó-ngwé ó-nè á-lé léphùtsí
     'everyone was yellow'
6.   mò-ngwè lé mó-ngwé ó-nè á-lé mó-pùtšwà
     'everyone was blue'
7.   mò-ngwè lé mó-ngwé ó-nè á-lé mó-sètlhá
     'everyone was grey'

Table 9. <u>Descriptive copulative constructions with the</u>
<u>Class 1 plural marker</u>

1. bó-tlhé bá-bá-ǹtšhò
    all of them be black
    'all of them are black'

2. bó-tlhé bá-bá-šwèú
    'all of them are white'

3. bó-tlhé bá-bá-húbídù
    'all of them are red'

4. bó-tlhé bá-bá-tàlá
    'all of them are green'

5. bó-tlhé bá-léphùtsí
    'all of them are yellow'

6. bó-tlhé bá-bá-pùtšwà
    'all of them are blue'

7. bó-tlhé bá-bá-sètlhá
    'all of them are grey'

Table 10. <u>Descriptive copulative constructions with the</u>
<u>Class 4 singular marker</u>

1. sélé sésé-ǹtšhò
    that yonder be black
    'that is black'

2. sélé sésé-šwèú
    'that is white'

3. sélé sésè-húbídù
    'that is red'

4. sélé sésé-tàlá
    'that is green'

5. sélé sé-léphùtsí
    'that is yellow'

6. sélé sésé-pùtšwà
    'that is blue'

7. sélé sésé-sètlhá
    'that is grey'

Table 11. <u>Descriptive copulative constructions with</u>
           <u>Class 5 plural markers</u>

1. dì-lɔ̀ tsélé   dí-dí-ǹtšhò
    things those yonder be black
    'those things are black'

2. dì-lɔ̀ tsélé   dí-dí-tšhwèú
    'those things are white'

3. dì-lɔ̀ tsélé   dí-dí-khúbídù
    'those things are red'

4. dì-lɔ̀ tsélé   dí-dí-tàlá
    'those things are green'

5. dì-lɔ̀ tsélé   dí-léphùtsí
    'those things are yellow'

6. dì-lɔ̀ tsélé   dí-dí-pùtšwà
    'those things are blue'

7. dì-lɔ̀ tsélé   dí-dí-tshètlhá
    'those things are grey'

A more subtle argument substantiating the distributional irregularity of
lèphùtsí is found in descriptive copulative as well as identificative
copulative constructions. In both of these copulative construction types,
the Class 5 singular prefix assumes the surface form én- with monosyllabic
root forms and é- with multisyllabic roots. Irrespective of surface form,
the Class 5 prefix, when adjoined to root initial segments with a low
strength value, conditions the application of a phonological strengthening
process.[4] The weak initial segments of the color terms -šwèú (š > tšh),
  -húbídù (h > kh), and -sètlhá (s > tsh), in Tables 12 and 13, manifest
strengthening when the Class 5 prefix is adjoined.[5] In contrast, the weak
initial segment of lèphùtsí , /l/ , does not manifest strengthening. If
it did, the resulting form would be tèphùtsí . What might motivate this
pattern of irregular behavior?

---

[4]See Cole [1955] and Schaefer [1980, 1982] for discussion of the
strengthening process.

[5]At some earlier diachronic stage, BLACK was probably encoded by -šò ,
becoming -ǹtšhò under strengthening and reanalyzed as a single lexical unit.

Table 12.   Descriptive copulative constructions with the
            Class 5 singular marker

1. à-é-ǹtšhò
    Q - it be black
    'is it black?'
2. à-é-tšhwèú
    'is it white?'
3. à-é-khúbídù
    'is it red?'
4. à-é-tàlá
    'is it green?'
5. à-é-léphùtsí
    'is it yellow?'
6. à-é-pùtšwà
    'is it blue?'
7. à-é-tshètlhá
    'is it grey?'

Table 13.   Identificative copulative constructions with the
            Class 5 singular marker

1. ké-é-ǹtšhò
    it be one black
    'it is a black one'
2. ké-é-tšhwèú
    'it is a white one'
3. ké-é-khúbídù
    'it is a red one'
4. ké-é-tàlá
    'it is a green one'
5. ké-é-léphùtsì
    'it is a yellow one'
6. ké-é-pùtšwà
    'it is a blue one'
7. ké-é-tshètlhá
    'it is a grey one'

Partial motivation for the failure of lèphùtsí to manifest the
strengthening process may be due to its morphological structure. In its
non-color usage, the lexical unit lèphùtsí is composed of the Class 3
singular prefix lè- adjoined to the root -phùtsí . Class 3 in Tswana is
a miscellaneous class, consisting of names for parts of the body, some animals,
plants and collective nouns. The Class 3 singular prefix lè- thus has a
morphological status equivalent to the various prefixes not manifested in the
previous construction types. It is perhaps then the morphological heritage
of lèphùtsí that accounts in part for its irregular behavior.

The composite morphological structure of lèphùtsí , the color term, is
revealed in some construction types. Intensifier constructions, requiring
reduplication of a color term root, are presented in Tables 14 and 15. These
constructions indicate that only a portion of the lexical form meaning YELLOW,
-phùtsí, is reduplicated. A similar partial segmentation under reduplication
is not manifested by any of the remaining color terms. Based on these intensi-
fier constructions, it would appear that the irregular distributional behavior
of lèphùtsí is a reflection of its morphological structure prior to becoming
a term for color.

Distributional irregularity within the system of potential basic color
terms is not entirely confined to lèphùtsí . The term for BLUE, -pùtšwà ,
exhibits a degree of irregular patterning, though it is not as consistent or
widespread as the pattern we have already witnessed. In Tswana, a color term
can be used in forming personal names by prefixing to it the Class 1 singular
marker mò- . Using the resulting stem as a base, one can also express the
location or place of this person's clan by prefixing to it the locative
marker xá- . A list of constructions expressing personal names and the
locations of individuals with these names is shown in Table 16. As can be
seen, the prefix mò- in the name and location constructions is deleted
before lèphùtsí . Deletion, surprisingly, also occurs before -pùtšwà .

One last argument supporting the irregular behavior of lèphùtsí , and
to some extent of -pùtšwà , can be discerned in figurative expressions.
Figurative expressions provided by my informant for each of the basic color

terms are presented in Tables 17, 18, and 19.

    Table 14. <u>Descriptive copulative intensifier constructions with</u>
              <u>the Class 9 marker</u>

1. xóxó-ǹtšhò-ǹtšhò
    it be black black
    'it is completely black'

2. xóxó-šwèú-šwèú
    'it is completely white'

3. xóxó-húbídù-húbídù
    'it is completely red'

4. xóxó-tàlá-tàlá
    'it is completely green'

5. xó-léphùtsí-phùtsí
    'it is completely yellow'

6. xóxó-pùtšwà-pùtšwà
    'it is completely blue'

7. xóxó-sètlhá-sètlhá
    'it is completely grey'

    Table 15. <u>Descriptive copulative intensifier constructions with</u>
              <u>the Class 5 marker</u>

1. é-ńtsé   ǹtšhò-ǹtšhò
    it become perfect black black
    'it has become exceedingly black'

2. é-ńtsé   tšhwèú-tšhwèú
    'it has become exceedingly white'

3. é-ńtsé   khúbídù-khúbídù
    'it has become exceedingly red'

4. é-ńtsé   tàlá-tàlá
    'it has become exceedingly green'

5. é-ńtsé   lèphùtsí-phùtsí
    'it has become exceedingly yellow'

6. é-ńtsé   pùtšwà-pùtšwà
    'it has become exceedingly blue'

7. é-ńtsé   tshètlhá-tshètlhá
    'it has become exceedingly grey'

Table 16.   Name and location constructions with the
            Class 1 singular marker

1. mò-ǹtšhò   xá-mó-ǹtšhò
    Mr. Black   place of the clan of Mr. Black

2. mò-šwèú   xá-mó-šwèú
    Mr. White   place of the clan of Mr. White

3. mò-húbɩ́dù   xá-mò-húbɩ́dù
    Mr. Red   place of the clan of Mr. Red

4. mò-tàlá   xá-mó-tàlá
    Mr. Green   place of the clan of Mr. Green

5. lèphùtsɩ́   xá-lèphùtsɩ́
    Mr. Yellow   place of the clan of Mr. Yellow

6. pùtšwà   xá-pùtšwà
    Mr. Blue   place of the clan of Mr. Blue

7. mò-sètlhá   xá-mó-sètlhá
    Mr. Grey   place of the clan of Mr. Grey

Table 17.   Figurative expressions involving the Tswana color terms

1. xó-nè xó-lè lèfìfɩ̀ xó-ré tšhò
    it be past dark   it said black
    'it was dark, it said black'

2. ó-nè   á-ì-tshásɩ́tsé   lè-tsòkú á-ré šwèú
    he past he himself put on ochre   he said white
    'he had put on ochre, he said white'

3. ó-nè   á-ì-tshásɩ́tsé   lè-tsòkú á-ré húbè
    'he had put on ochre, he said red'

4. nóxá é-nè   é-répálétsé   mó-sè-tlhàréñ é-ré tálà
    snake be past it stretched out in a tree   it said green
    'a snake was stretched out in a tree, it said green'

5. - - -

6. mà-rù á-nè á-à-póxɩ̀lè   lè-wápɩ́ lé-ńtsé lé-ré pùtšwà
    clouds past they dispersed the sky it past it say blue
    'the clouds had dispersed, the sky was saying blue'

7. ó-nè   á-ì-téɩ́lwé   ké-lè-fátshé á-ré sètlhá
    he past he himself beaten up by world   he said grey
    'he had been beaten up by the world, he said grey'

Table 18. <u>Figurative expressions involving the Tswana color terms</u>

1. kè-mó-bò-ǹtšhòng̀
    I inside blackness
    'I am lost'

2. ké-mò-thò yóó-pèló   tšhwèú
    person   who be hearted white
    'he is a person who is white hearted'

3. kè-ḿ-mátlà   ká-mà-tlhɔ́ á-mà-húbídù
    I am look for with eye   which be red
    'I am looking for him with the red eye'

4. - - -

5. - - -

6. - - -

7. ké-mò-thò yóó-pèlò   tshètlhá
    'he is a person who is grey hearted'

Table 19. <u>Figurative expressions involving the Tswana color terms</u>

1. á   lè-káù lélé-tšhwànà
    what youth who be black
    'what a black youth'

2. á   lè-káù lélé-šwánà
    'what a white youth'

3. á   mò-sé-tsànà yóó-mà-rámá ámà-húbítswànà
    what a girl   who be cheek that be red
    'what a girl who is red cheeked'

4. á   mà-hùlɔ̀ ámà-t̪álàná
    what pasture which be green
    'what a green pasture'

5. - - -

6. - - -

7. - - -

In the first type of figurative construction, lèphùtsí is the only color term for which an expression is lacking. In the second type, a figurative expression could not be formed with -tàlá , -pùtšwà , and lèphùtsí . In the final type of expression, -sètlhá and lèphùtsí do not participate. According to Wescott's [1970] criteria, (see also Derrig [1978]) basicness in the color domain is indexed by occurrence in figurative expressions and by the range of this occurrence. That is, the longer a term has referred to color, the greater the likelihood that it will be used in figurative expressions. Since lèphùtsí does not enter into any figurative expressions, its status as a basic term is definitely suspect. To a lesser degree, -pùtšwà is also suspect. Its behavior pattern, however, may reflect the fact that, next to lèphùtsí , it is the most recent addition to the set of Tswana basic terms.

3.2. <u>Complementary distribution of basic terms</u>. A second pattern character-izing color terms within the basic set is one of complementary distribution. At the core of this pattern is a co-occurrence constraint which differentially affects potential basic color terms in Tswana. This pattern argues that at the semantic level two sub-classes of basic color terms are recognized in Tswana. One class is composed of terms designating color on the brightness dimension:
-ǹtšhò , -šwèú , and -sètlhá . The other, in contrast, is composed of terms designating color on the hue dimension: -húbídù , -tàlá ,
-pùtšwà , and lèphùtsí .

As a background for viewing these two classes of basic terms, the following may be helpful. Tswana has a large number of terms which ascribe a color, a combination of colors, or a combination of color and other prominent marking, e.g. horns, to cattle and other animals. A representative sampling of these terms is listed in Table 20. As the glosses indicate, not every term can be ascribed to cattle, some can only be ascribed to sheep. More thorough discussion of these terms is precluded by limitations on the length of the present paper. It is the existence of these terms, however, to which we wish to call attention at this time.

Table 20. <u>A sample of terms which ascribe color to animals (male and female)</u>
<u>in Tswana</u>

1. bòbírwà bòbírwánà   'black, only of sheep'
2. bòdílúdì bòdílótsànà   'black and white spotted, zebra'
3. bòfàtshwà bòfàtshwànà   'black and white in a large pattern'
4. bòfìfàdù bòfìfàtšwànà   'black'
5. bòfìtšhwà bòfìtšhwànà   'very dark brown, of cattle and horses'
6. bògwàrìpà bògwàrìpànà   'red and white, running into one another'
7. bògwὲbà bògwὲbànà   'red and white in small spots'
8. bòhúnóù bòhúnwánà   'red, bay colored, reddish brown'
9. bòkhùkhwà bòkhùkhwànà   'brown with yellow at extremities, or yellow with brown at extremities, of goats and dogs'
10. bònàlá bònàánà   'red and white in a large pattern'
11. bòngòlò bòngòlwànà   'black with white or yellow underparts, of goats and dogs'
12. bònkgwὲ bògwànà   'red, black with white along spine and underparts'
13. bòntlhwá bòtlhwànà   'dark brown, chocolate colored, of sheep'
14. bòràmàgà bòràmàgànà   'large spotted brown and white, maybe black'
15. bòròkwà bòròkwànà   'brown, dark red, yellowish red'
16. bòsàṁpà bòsàṁpánà   'white and red striped or streaked'
17. bòšùmú bòšùmúnyánà   'red, black with white extending to face, and possible throat and belly'
18. bòtlhàbà bòtlhàbánà   'brown turning to yellow at extremities, or yellow turning to brown at extremities, cattle'
19. bòtùbà bòtùbànà   'coffee colored, fawn, yellow'
20. bòwὲbù bòkwὲbù   'grey *roan*, red *roan*'

Returning to the main issue, an initial argument substantiating a pattern of complementary distribution within the set of basic terms rests on the Class 7 prefix. As shown earlier, one refers to the abstract quality of a color, e.g., blackness, by prefixing the Class 7 singular prefix bò- to a color term. To the resulting stem, the form -ànà can be suffixed. This marker, shown in Table 21 constructions, denotes either femaleness of

the cattle species or diminutiveness (endearment), when adjoined to a
potential color term. (Maleness and non-diminutiveness are unmarked.)

Table 21. <u>Construction involving the Class 7 and the
            sex/diminutivity marker</u>

1. bò-ǹtšhò   bò-šwànà   'blackness'

2. bò-šwèú   bò-šwàánà   'whiteness'

3. bò-húbídù   bò-húbìtswánà   'redness'

4. bò-tàlá   bò-tàlánà   'greenness'

5. lèphùtsí   'yellowness'

6. bò-pùtšwà   bò-pùtšwànà   'blueness'

7. bò-sètlhá   bò-sètlhánà   'greyness'

Not all potential basic color terms can adjoin with the suffix -ànà
to convey both of these meanings. Notice first that it does not adjoin with
lèphùtsí to convey either the diminutive or female meaning. More signifi-
cant is a constraint extending beyond lèphùtsí , namely that potential
basic terms designating hue are never ascribed to female cattle. The color
terms -húbídù , -tàlá , and -pùtšwà , when adjoined to the suffix
-ànà , can only refer to diminutiveness (or endearment). The referential
scope of the terms -ǹtšhò , -šwèú , and -sètlhá stands in contrast.
The latter terms, all designating colors on the brightness dimension, can
adjoin with -ànà and refer either to diminutiveness or femaleness. In
other words, both meanings of the suffix -ànà can be conveyed with basic
terms designating color on the brightness dimension. The unmarked forms in
Table 21 (those in the left hand column) abide by the same constraint, i.e.
terms designating hue (including lèphùtsí) cannot be ascribed to male
cattle, while terms designating brightness can.

A second argument substantiating complementary distribution within the
set of basic terms is based on the root -phólòxólò . This root form,
meaning ANIMAL, co-occurs with only a subset of the basic color terms, as
shown in Tables 22 and 23.

Table 22. <u>Constructions involving the root form meaning ANIMAL</u>

1. é-nè é-lé phólòxóló é-ǹtšhò
    it be past animal   black
    'it was a black animal'

2. é-nè é-lé phólòxóló é-tšhwèú
    'it was a white animal'

3. - - -

4. - - -

5. - - -

6. - - -

7. é-nè é-lé phólòxóló é-tshètlhá
    'it was a grey animal'

Table 23. <u>Constructions involving the root form meaning ANIMAL</u>

1. fá-xòngwé dì-phólòxóló tsédí-ǹtšhò   xà-dí-bónáxálè
    sometimes animals   which be black not they seeable
    'sometimes animals which are black are not seeable'

2. fá-xòngwé dì-phólòxóló tsédí-tšhwèú   xà-dí-bónáxálè
    'sometimes animals which are white are not seeable'

3. - - -

4. - - -

5. - - -

6. - - -

7. fá-xòngwé dì-phólòxóló tsédí-tshètlhá xà-dí-bónáxálè
    'sometimes animals which are grey are not seeable'

Examination of these tables indicates that -phólòxóló only occurs with
 -ǹtšhò , -šwèú , and -sètlhá . The form meaning ANIMAL, therefore,
does not co-occur with basic terms designating hue. We have thus bolstered
our argument that there is a systematic constraint, specified by the domain
ANIMAL, which governs the semantic extension of potential basic color terms.

    Although not obviously tied to the domain of cattle, a third argument
supports the pattern of complementary distribution within the set of
potential basic color terms. This argument pivots on the behavior of the

derived verb suffix -hàlà (-fàlà in some dialects). As Cole [1955]
states, -hàlà regularly is suffixed to roots denoting qualities, such as
color terms, to express an inchoative meaning, i.e. the coming into
existence of a state. The form -hàlà is also conjoined with the causative
marker, giving the form -hàtsà , and suffixed to basic terms. The resulting
form refers to causing a state to come into existence. It is the complemen-
tary distribution of basic color terms in expressions of inchoative and
causative meaning to which we now turn.

Inchoative constructions involving potential basic color terms will be
discussed first. In addition to -hàlà , the so-called deficient verb
 -ńná can also express the inchoative meaning. Tables 24 through 29
offer sample constructions expressing the inchoative meaning. As we can see
in Table 24, the marker -ńná co-occurs with each potential basic term,
irrespective of whether it designates color on the brightness or hue
dimension. However, the marker -hàlà , in Table 25, does not behave in
a similar fashion, co-occurring only with the brightness terms -ǹtšhò ,
 -šwèú , and -sètlhá . The variety of constructions expressing an
inchoative meaning in the tables above substantiate this pattern
of complementary distribution, those with future and past tense in Tables
26 and 27, and those with perfective aspect in Tables 28 and 29. The dis-
tribution of the marker -hàlà is thus restricted compared to the marker
 -ńná , for it fails to co-occur with basic terms designating color on
the hue dimension.

Causative constructions involving the potential basic color terms dis-
close an identical pattern of complementary distribution. Just as construc-
tions with the marker -ńná express the inchoative meaning alongside those
with -hàlà , so constructions with the marker -dìrà express a causative
meaning alongside those with the marker -hàtsà , the causative form of
 -hàlà . And as -hàlà did not co-occur with color terms designating hue,
 -hàtsà does not co-occur with terms designating hue. Sample causative
constructions are shown in Tables 31 through 33.

Table 24.   Inchoative constructions with the deficient verb -ńná

1. é-ńná   ǹtšhò
    it becomes black
    'it is becoming black'
2. é-ńná   tšhwèú
    'it is becoming white'
3. é-ńná   khúbídù
    'it is becoming red'
4. é-ńná   tàlá
    'it is becoming green'
5. é-ńná   léphùtsí
    'it is becoming yellow'
6. é-ńná   pùtšwà
    'it is becoming blue'
7. é-ńná   tshètlhá
    'it is becoming grey'

Table 25.   Inchoative constructions with -hàlà and -ńná in
            present tense

1. é-á-ǹtšhò-hàlà
    it black become
    'it is becoming black'
2. é-á-šwèù-hàlà
    'it is becoming white'
3. é-ńná khúbídù
    'it is becoming red'
4. é-ńná tàlá
    'it is becoming green'
5. é-ńná léphùtsí
    'it is becoming yellow'
6. é-ńná pùtšwà
    'it is becoming blue'
7. é-á-sètlhà-hàlà
    'it is becoming grey'

Table 26. <u>Inchoative constructions with -hàlà and -ńná in</u>
           <u>future tense</u>

1. é-tlá-à   -ǹtšhò -hàlà
    it future black become
    'it will become black'

2. é-tlá-à   -šwɛ̀ù -hàlà
    'it will become white'

3. é-tlá-à-ńná   khúbídù
    it future become   red
    'it will become red'

4. é-tlá-à-ńná   tàlá
    'it will become green'

5. é-tlá-à-ńná   léphùtsí
    'it will become yellow'

6. é-tlá-à-ńná   pùtšwà
    'it will become blue'

7. é-tlá-à-sɛ̀tlhà-hàlà
    'it will become grey'

Table 27. <u>Inchoative constructions with -hàlà and -ńná in</u>
           <u>past tense</u>

1. é-nè   é-ńtšhó -hàlà
    it past be black become
    'it was becoming black'

2. é-nè   é-šwɛ́ú -hàlà
    'it was becoming white'

3. é-nè   é-ǹná   khúbídù
    'it was becoming red'

4. é-nè   é-ǹná   tàlá
    'it was becoming green'

5. é-nè   é-ǹná   léphùtsí
    'it was becoming yellow'

6. é-nè   é-ǹná   pùtšwà
    'it was becoming blue'

7. é-nè   é-sɛ́tlhá -hàlà
    'it was becoming grey'

Table 28.   <u>Inchoative constructions with -hàlà and -ńná in</u>
            <u>perfective aspect</u>

1. é-ńtšhò -hètsè
    it black become perfect
    'it has become black'

2. é-šwὲù -hètsè
    'it has become white'

3. é-ńtsé   khúbídù
    it become perfect red
    'it has become red'

4. é-ńtsé   tàlá
    'it has become green'

5. é-ńtsé   léphùtsí
    'it has become yellow'

6. é-ńtsé   pùtšwà
    'it has become blue'

7. é-sὲtlhà -hètsè
    'it has become grey'

Table 29.   <u>Inchoative constructions with the deficient verb -ńná</u>
            <u>and perfective aspect</u>

1. é-ńtsé   ńtšhò
    it become perfect black
    'it has become black'

2. é-ńtsé   tšhwὲú
    'it has become white'

3. é-ńtsé   khúbídù
    'it has become red'

4. é-ńtsé   tàlá
    'it has become green'

5. é-ńtsé   léphùtsí
    'it has become yellow'

6. é-ńtsé   pùtšwà
    'it has become blue'

7. é-ńtsé   tshὲtlhá
    'it has become grey'

Table 30. Causative constructions with -dìrà and -hàtsà

1. ó-ńtšhó-hàtsà   pìtsá
    she black become cause the pot
    'she caused the pot to become black'

2. ó-šwéú-hàtsà   pìtsá
    'she caused the pot to become white'

3. ó-dírá   pìtsá khúbídù
    she cause pot   red
    'she causes the pot to become red'

4. ó-dírá   pìtsá tàlá
    'she causes the pot to become green'

5. ó-dírá   pìtsá léphùtsí
    'she causes the pot to become yellow'

6. ó-dírá   pìtsá pùtšwà
    'she causes the pot to become blue'

7. ó-sétlhá-hàtsà   pìtsá
    'she causes the pot to become grey'

Table 31. Causative constructions with -dìrà in perfective
           aspect

1. ó-dìr-ílé   pìtsá ǹtšhò
    she make perfect pot   black
    'she made the pot black'

2. ó-dìr-ílé   pìtsá tšhwèú
    'she made the pot white'

3. ó-dìr-ílé   pìtsá khúbídù
    'she made the pot red'

4. ó-dìr-ílé   pìtsá tàlá
    'she made the pot green'

5. ó-dìr-ílé   pìtsá léphùtsí
    'she made the pot yellow'

6. ó-dìr-ílé   pìtsá pùtšwà
    'she made the pot blue'

7. ó-dìr-ílé   pìtsá tshètlhá
    'she made the pot grey'

Table 32. <u>Causative constructions with -hàlà in perfective</u>
          <u>aspect</u>

1. ó-ǹtšhó-hádítsé   pìtsá
    she black become cause perfect iron pot
    'she blackened the iron pot'

2. ó-šwèú-hádítsé   ǹkxhó
    'she whitened the clay pot'

3. - - -

4. - - -

5. - - -

6. - - -

7. ó-sètlhá-hádítsé   lètsóxòlá xáxwέ ká-mólɔ̀rà
    she grey become cause perfect hand her   with ash
    'she greyed her hand with ash'

Table 33. <u>Causative constructions with -hàlà in perfective</u>
          <u>aspect</u>

1. ó-è   -ǹtšhó -hádìtsè
    he it black become cause perfect
    'he caused it to become black'

2. ó-è-šwέú-hádìtsè
    'he caused it to become white'

3. - - -

4. - - -

5. - - -

6. - - -

7. ó-è-sέtlhá-hádìtsè
    'he caused it to become grey'

    The consistent and widespread pattern of complementary distribution
just examined appears to reflect a semantic constraint operating within the
set of color terms in Tswana. Underlying this constraint on basic terms is
the recognition of two sub-classes, classes which appear semantically natural
in that each is defined by one of the dimensions giving rise to the categori-
zation of color, hue, and brightness.

3.3. <u>Consequences of the distribution pattern</u>. The two patterns character-
izing the distribution of potential basic color terms lead to the following
consequences. Relative to the behavior of other terms, the irregular
behavior of lèphùtsí argues that it is not a basic color term. The fact
that it is object-derived strengthens this argument. Next to lèphùtsí ,
the most recent addition to the set of basic color terms seems to be
-pùtšwà , whose distribution deviates in small measure from that of the
other terms. The fact that the behavior of -pùtšwà conforms to the
overall pattern of the other terms establishes, nonetheless, that it is a
basic term. Finally, since the terms -ǹtšhò , -šwèú , and -sèt!há
have the least restricted distribution, being ascribable to animals as well
as objects, perhaps only these terms should be considered basic.

   Careful examination shows all but one of these consequences to be
theoretically problematic. Rejecting all but -ǹtšhò , -šwèú , and
-sèt!há as basic terms seems inadequate given the general character and
use of the entire set of potential terms. In some ways, rejecting all but
these terms would be comparable to rejecting *yellow*, *blue*, and *green* as
basic color terms in English because each does not combine with the morpho-
logical marker *-en* , i.e. **yellowen*, **bluen*, and **greenen*. These Tswana
terms do highlight a significant fact, namely a systematic point of contact
and overlap between the set of basic color terms and the set of auxiliary
color terms referring to animals.[6]

-----

   [6]The extent to which similar patterns of overlap exist in other dialects
of Tswana and other languages in cattle herding economies would seem to deserve
attention. An intriguing issue raised by the overlap of terms in the basic and
auxiliary sets is why terms designating brightness, rather than hue, overlap.
Further investigation might explicate the motivation for this particular type
of overlap. The set of auxiliary color terms in Tswana raises another intri-
guing issue. Berlin and Kay [1969] correlate the general expansion of color
term systems with a cultural variable, technological advance. The pattern of
color term expansion they envisage, though, is one where a semantic category
emerges from a set of object categories having no previous reference to color.
The situation in Tswana stands in contrast. In Tswana, an elaborate system of
categories designating both the hue and brightness of color in the animal
domain is available at the semantic level. In some languages, therefore, a
color category may not emerge as basic at the semantic level so much as

To return to the set of consequences, rejecting lèphùtsí as a basic
color term poses a serious theoretical issue. Rejecting lèphùtsí results
in a diachronic stage not recognized by the revised, or the original,
evolutionary sequence of color terms shown at the outset. There is no
allowance for a stage where BLUE and GREEN are lexicalized but YELLOW is
not. Of course, one could claim that the dialect under examination stands
as an exception to the postulated universal diachronic sequence. This
explanation seems too hasty at present though. As an alternative, one
might claim that -pùtšwà , as well as lèphùtsí , is not a basic term.
The dialect of Tswana examined herein would then be at Stage III in the
revised evolutionary sequence. The regular distributional behavior of
-pùtšwà argues against such a conclusion however. Still a third alterna-
tive rests with a comparative and historical analysis of Tswana color terms.
Information gleaned from comparative and historical sources, indicating the
semantic extension and semantic reference of forms cognate with the Tswana
color terms, may provide a situation under which we can mollify the effect
of rejecting lèphùtsí as a basic color term.

4. Comparative Analysis of Color Terms

The evidence available from historical sources on first glance seems to
impede rather than facilitate understanding of color term development in
Tswana. Especially affected are the terms -pùtšwà and -sètlhá . The
Rev. T. Brown's Setswana-English:English-Setswana Dictionary, the only
extant dictionary of Tswana, was published in 1925 in its revised form.
In the main, the semantic record provided by Brown's dictionary agrees with
the system of semantic reference outlined by the preceding examples, except
for the semantic reference of -sètlhá . Close inspection of the entries
in this dictionary, moreover, indicates that the inclusion of animals or
non-animals in the semantic extension of color term is not consistently noted.

---

transfer from a non-basic to the basic color domain. It is perhaps differ-
ences between the processes of emergence and transfer that will lead to a
clearer understanding of the situation in Tswana.

According to the color term data discussed thus far, the term -sètlhá
refers to GREY and the term lèphùtsí refers to YELLOW. The Brown dictionary,
in contrast, lists -sètlhá as referring to YELLOW. No examples illustrating
this usage are provided. A second contrast involves the form -pùtšwà , and
its morphological variant -pùdùdù , which are listed as referring to DARK
GREY, BLUE and DARK BLUE, GREY, respectively.[7] In the data examined thus far,
 -pùtšwà has referred only to BLUE. We will take up these problems of
reference shortly.

    Additional information bearing on the referential scope of the potential
color terms is available in dictionaries of languages genetically related to
Tswana, Northern Sotho and Southern Sotho.[8] Overall, the names these diction-
aries provide for the primary colors conform with the pattern established by
Brown and my informant, except for the naming of YELLOW, BLUE, and GREY. In
the Southern Sotho-English Dictionary edited by Paroz [1961] and the Compara-
tive Northern Sotho Dictionary:Northern Sotho-Afrikaans/English [1977], the
term -tàlá refers to GREEN or BLUE, i.e. Berlin and Kay's GRUE. This fact
distinguishes Northern and Southern Sotho from Tswana, for Brown's dictionary
and my informant indicate that -tàlá refers only to GREEN.[9] A difference in
referential scope also characterizes the term -pùtšwà . Brown glosses it as
BLUE or BLUE GREY, and my informant indicated that it referred to BLUE. In
Northern and Southern Sotho, -pùtšwà is glossed as GREY, referring to
colored objects or animals. No mention is made of the BLUISH GREY found in
the lexical entries in Brown.

---

    [7]A process of haplology, combined with the palatalization process dis-
cussed in Cole [1955], d > ts, could specify the relatedness of the forms
 -pùtšwà and -pùdùdù . The entries for these terms also suggest that they
may not have been ascribed to cattle or animals. The entry for -pùdùdù , at
least in Brown [1924], makes no reference to cattle. Under -pùdùdù , the
sample illustrations are non-animal: blue sky and blue clothes. A separate
term, -kwèbú , is listed as referring to the BLUE of an ox.

    [8]See Doke [1954] for discussion of the Sotho family of languages.

    [9]See Squires [1942] for evidence that the term -tàlá may have
referred to BLUE and GREEN in some northern Tswana dialects in more recent
times.

For present purposes, the most conspicuous similarity between the dictionary entries for Northern Sotho, Southern Sotho, and Tswana involves the form -sɛ̀tlhá . In both the Northern and Southern Sotho dictionaries a form -sɛ̀hlá , cognate with the Tswana form -sɛ̀tlhá , refers to YELLOW. Examples illustrating this color in the Southern Sotho dictionary include a chestnut horse and a pale, light-skinned man or animal. Presumably, -sɛ̀hlá was attributed to animals, whereas Brown's dictionary for Tswana fails to provide a similar level of detail for -sɛ̀tlhá . Nonetheless, an intriguing hypothesis, derived from these examples, is that the historical antecedent of -sɛ̀hlá/-sɛ̀tlhá referred to a faint YELLOW, a desaturated YELLOW, which allowed for the "light-skinned" and "pale man" examples. But before going further, let us attempt to clarify the discussion thus far by listing in Table 34 the color terms obtained from the various dictionaries and from my informant, along with the semantic reference of each term.

Table 34.   Color terms obtained from various dictionaries of the Sotho languages and from my informant

	Tswana (Brown)	N. Sotho	S. Sotho	Tswana (Informant)	Referent
1.	-ntšho	-tsho	-tsho	-ntšho	BLACK
2.	-šwɛu	-šwɛu	-šwɛu	-šwɛu	WHITE
3.	-hubidu	-hubedu	-hubidu	-hubidu	RED
4.	-tala	-tala	-tala	-tala	GREEN
5.	-sɛtlha	-sɛhla	-sɛhla	lephutsi	YELLOW
6.	-pucwa, -pududu	-tala	-tala	-putšwa	BLUE
7.	-pududu	-putšwa	-putšwa, -pududu	-sɛtlha	GREY

A further check on the referential scope of color terms in the Sotho languages is available in Louw [1957]. Louw compared the color terms ascribed to cattle in the three Sotho languages. Based on his analysis of one speaker from each language, -pùtšwà/-pùdùdù and -sɛ̀tlhá/-sɛ̀hlá refer to DARK GREY

and YELLOW, respectively. The Tswana terms -sèt|hǎ and -pùtšwà most likely
were then ascribable to cattle at one point in their history.

     The semantic information derived from these sources has clear implications
for inferring the development of the referential scope of color terms in the
Tswana dialect under consideration. In each of the non-Tswana dictionaries
consulted, the term for YELLOW was cognate with the Tswana form -sèt|hǎ .
The saturation level of this YELLOW color, judging by the Southern Sotho
record, appeared to be low, encompassing YELLOW, as well as a light colored,
pale YELLOW. When my informant was confronted with the information that
 -sèt|hǎ referred to YELLOW, she indicated that perhaps the GREY color of a
lion (a pale hue) was the intended meaning.

     It seems reasonable to conclude from these different sources that at some
earlier date, -sèt|hǎ referred to a range of color encompassing desaturated
YELLOW and LIGHT GREY. Such a conclusion seems in harmony with Christman's
[1971] assertion that as hue becomes desaturated, it is perceived in terms of
the brightness dimension, e.g. GREY.

     A comparable conclusion seems warranted for -pùtšwà . The dictionaries
assign to -pùtšwà the meaning BLUE and DARK GREY in Tswana and GREY in
Northern and Southern Sotho. For some earlier diachronic stage, it seems
reasonable to postulate that the historical antecedent of -pùtšwà referred
to a range of color which encompassed desaturated BLUE and DARK GREY.

     Assuming the correctness of these postulates regarding the referential
scope of -pùtšwà and -sèt|hǎ , it appears that both referred at some
earlier diachronic stage to a highly desaturated range on the hue dimension
which naturally intermixed with the brightness dimension. Is there evidence
from outside linguistics proper which could further substantiate these pos-
tulates? Seemingly supportive evidence is available in Bornstein [1973a,b].

5.   Additional Support for Comparative Analysis

     Bornstein [1973a,b] has offered a theoretical explanation for cultural
differences in optical illusion susceptibility and color naming that is at
base psychophysiological. Previous explanations of behavior in these two

domains have assumed that across populations physiological structures involved in perceptual processing are identical. Bornstein, in contrast, calls attention to potential physiological differences in the pigmentation density of the ocular structures which process incoming perceptual information. In particular he points out a correlation between patterns of increased yellow pigmentation and reduced susceptibility to optical illusions.[10] For instance, populations in Africa, which exhibit high retinal pigmentation, have shown less susceptibility to the Muller-Lyer optical illusion than populations in Europe, which exhibit less retinal pigmentation [Berry 1971]. In the domain of color naming, the pigmentation of the ocular structure evinces a comparable effect.

Bornstein [1973a,b] examined primary color terms in over 145 different languages for what he termed "semantic confusions." A semantic confusion occurred when a single color term referred to more than one of the primary colors, i.e. BLACK, WHITE, RED, GREEN, YELLOW, and BLUE. His major finding was the high frequency with which BLUE was lexicalized with BLACK, GREEN, or both of these colors. A correlated set of confusions, though less pervasive, involved WHITE and YELLOW (GREY unfortunately is not considered a primary color).

To explain this "BLUE" confusion pattern, Bornstein [1973a,b] pointed out the phenomenal similarity between the color naming practices defining the pattern of BLUE confusion across languages and the color naming practices of individuals with weak blue-yellow vision. These individuals are characterized by a highly dense yellow pigmentation of the ocular structures. Probing further, Bornstein located the pattern of BLUE confusion primarily in areas of relatively intense sunlight populated by dark skinned peoples. In this highly sunlit environment, the adaptive function of increased pigment density is to absorb potentially harmful short wavelength radiation. More important, the

---

[10]See Segall [1979] and Jahoda [1971, 1975] for some counter-arguments to the Bornstein hypothesis. It is important to note that these arguments are not conclusive and frequently involve a within culture experimental design, such as comparing white and black skinned individuals living in a highly technological society.

effect on color perception is to attenuate incoming visual stimulation of
BLUE light and, so it appears, YELLOW light as well.

     The referential scope of color terms among a population characterized by
high pigment density, such as Tswana speakers, might then be expected to mani-
fest a phenomenal similarity with the referential scope of color terms among a
population characterized by weak blue-yellow vision. A quotation from Pickford
[1951:103], where extensive investigation of color vision is reported, reveal-
ing a close affinity between the naming practices of blue-yellow weak individ-
uals and Tswana speakers should not be surprising:

> "With the yellow-blue blind and even the yellow-blue weak
> either or both blue and yellow are diminished in saturation
> compared with the normal. Dark grey tends to invade blue and
> light grey and white tends to invade yellow."

The phenomenal similarity between the color naming condition outlined by
Pickford and the postulated scope of reference of the terms -pùtšwà (DARK
GREY and BLUE) and sètlhá (LIGHT GREY and YELLOW) cannot be easily dismissed.
The referential scope postulated for these terms may thus rest on a natural
psychophysiological property of the visual system of Tswana speakers.

6.   Resolution of -sètlhá Issue

     A major issue still remains. That is, was -sètlhá , or its historical
antecedent, a member of the set of basic terms or the set of auxiliary terms,
whose extension is confined to animals, when it did refer to YELLOW. If
  -sètlhá was a basic term referring only to YELLOW, why does it now, as a
basic term, refer to GREY? On the other hand, if -sètlhá , on becoming a
basic term did not refer to YELLOW, but instead to GREY, then the revised
evolutionary sequence of basic color terms is not satisfied (Kay [1975], Kay
and McDaniel [1978]).

     In order to choose between these alternatives, it seems advantageous to
employ the construct "overburdening" discussed by Kristol [1980]. Two crucial
aspects of semantic overburdening are pertinent to our discussion. First, the
referential scope of a color term is subject to language general and language

specific constraints. And second, the natural consequence of violating such a
constraint is a reduction in referential scope.

Recall now from the data examined at the outset, the constraint governing
the inclusion of animals and cattle in the semantic extension of some basic
color terms. Specifically, terms designating a color on the hue dimension
could not be ascribed to animals. Let us add to this the comparative data
obtained from the Sotho languages, from which we postulated the semantic char-
acter of -sètlhá at some earlier diachronic stage. We found, to repeat,
that -sètlhá referred to both the hue and brightness dimensions and included
animals within its extension during that stage. This semantic condition, how-
ever, violates the constraint noted above, i.e. basic color terms designating
hue in Tswana are not ascribable to animals. In other words, the semantic
condition of -sètlhá is overburdened. To amend this condition, the refer-
ential scope of -sètlhá could be restricted to YELLOW and its extension
could exclude animals, or its referential scope could be restricted to GREY
and its extension allowed to include animals. Of these two semantic changes,
the latter seems the path of least effort and the one which maintains the
greater portion of -sètlhá's former semantic integrity, i.e. requiring only
a change in referential scope rather than a change in reference and extension.
This change would necessitate, however, the introduction of a new term for
YELLOW, such as lèphùtsí , and set the stage for what appeared an exception
to the postulated universal sequence of color term development.

7.   Summary

To review, the distributional behavior of potential basic color terms in
one dialect of Tswana was examined in a wide range of construction types.
This behavior led to the rejection of one of the color terms, lèphùtsí ,
as a basic term and to the recognition of a language specific constraint
requiring the exclusion of animals from the extension of basic terms desig-
nating hue.

Rejecting lèphùtsí proved problematic, since the remaining terms were
inconsistent with universal constraints on the historical stages basic color

terms should instantiate. After a comparison of color term semantics in the
Sotho languages, a historical state of reference and extension for the Tswana
term -sὲtlhá was postulated. Assuming the existence of this earlier seman-
tic state made the resulting color terms consistent with universal constraints,
although -sὲtlhá's meaning then violated the Tswana specific constraint
governing the extension of hue terms. Apparently as a result of violating
this language specific constraint, a change in the semantic character of
 -sὲtlhá was necessitated, leading to a narrowing of its referential scope
and the introduction of a new color term, lὲphὺtsí .

REFERENCES

Anonymous, n.d. *Comprehensive Northern Sotho Dictionary: Northern Sotho,
        Afrikaans/English*. Johannesburg: J. L. Van Shaik LTD.

Berlin, B. and E. A. Berlin. 1975. "Aguarana color categories." *American
        Anthropologist* 2:61-87.

Berlin, B. and P. Kay. 1969. *Basic Color Terms*. Berkeley: University of
        California Press.

Berry, J. W. 1971. "Muller-Lyer susceptibility: culture, ecology or race?"
        International Journal of Psychology 6:193-197.

Bornstein, M. H. 1973a. "Color vision and color naming: a psychophysio-
        logical hypothesis of cultural differences." *Psychological Bulletin*
        80:257-285.

Bornstein, M. H. 1973b. "The psychophysiological component of cultural
        differences in color naming and illusion susceptibility." *Behavior
        Science Notes* 8:41-101.

Brown, Rev. J. T. 1925. *Setswana-English:English-Setswana Dictionary*,
        third edition. Johannesburg: Pula Press.

Christman, R. J. 1971. *Sensory Experience*. Toronto: Intext Educational
        Publishers.

Cole, D. T. 1955. *An Introduction to Tswana Grammar*. Capetown: Longman.

Derrig, S. 1978. "Metaphor in the color lexicon." In D. Farkas, W. M. Jacobsen and K. W. Todrys (eds.), *Papers from the Parasession on the Lexicon,* pp. 85-96. Chicago: Chicago Linguistic Society.

Doke, C. M. 1954. *The Southern Bantu Languages.* London: Oxford University Press.

Jahoda, G. 1971. "Retinal pigmentation, illusion susceptibility and space perception." *International Journal of Psychology* 6:199-208.

Jahoda, G. 1975. "Retinal pigmentation and space perception: a failure to replicate." *Journal of Social Psychology* 97:133-134.

Kay, P. 1975. "Synchronic variability and diachronic change in basic color terms." *Language and Society* 4:257-270.

Kay, P. and C. McDaniels. 1978. "The linguistic significance of the meanings of basic color terms." *Language* 54:610-646.

Kristol, E. 1980. "Color systems in southern Italy." *Language* 56:132-145.

Louw, J. A. 1957. "The nomenclature of cattle in the southeastern Bantu languages." Pretoria: Communications of the University of South Africa C,2.

McDaniel, C. K. 1974. "Basic color terms: Their neurophysiological bases." Paper presented at the American Anthropological Association, Mexico City.

Paroz, R. A. 1961. *Southern-Sotho English Dictionary.* Morija: Morija Printing Works.

Pickford, R. W. 1951. *Individual Differences in Color Vision.* London: Routledge and Kegan Paul.

Schaefer, R. P. 1980. "On the motivation and structure of a strengthening process in Tswana." In P. Hamel and R. Schaefer (eds.), *Kansas Working Papers in Linguistics* 5.2. Lawrence: University of Kansas.

Schaefer, R. P. 1982. "A strength hierarchy for a morphophonemic process in Tswana." *Studies in African Linguistics* 13:147-176.

Segall, M. H. 1979. *Cross-Cultural Psychology: Human Behavior in Global Perspective.* Monterey, CA: Brooks/Cole.

Squires, B. T. 1942. "Colour vision and colour discrimination amongst the Bechuana." *Transactions of the Royal Society of South Africa* 29:29-34.

Ullman, S. 1962. *Semantics*. Oxford: Blackwell.

Wescott, R. W. 1970. "Bini color terms." *Anthropological Linguistics*
     12:349-360.

Witkowski, S. R. and C. H. Brown. 1977. "An explanation of color nomencla-
     ture universals." *American Antrhopologist* 79:50-57.

GENETIC RELATIONSHIP AND THE CASE OF MA'A (MBUGU)*

Sarah G. Thomason
University of Pittsburgh

This paper addresses the general question of genetic vs. non-
genetic language development, in the context of a structural
and historical discussion of Ma'a (Mbugu), a language with
Cushitic basic vocabulary that is spoken in Tanzania. The
grammatical structure of Ma'a is compared to characteristic
Cushitic and Bantu structures. The conclusion that emerges
from this comparison is that Ma'a probably does not have
enough Cushitic grammar to qualify as a Cushitic language
in the full genetic sense; and if it does not, its origin
must be nongenetic. The final section of the paper seeks
to determine the particular route of nongenetic development
that Ma'a has followed, using the direct evidence of pub-
lished comments about its speakers' history and the indirect
evidence of comparison with other languages whose origin is
nongenetic or, like Ma'a, on the borderline between genetic
and nongenetic.

1.   Introduction

     Ma'a (Mbugu)[1] is a favorite battleground for proponents and opponents of

hypotheses of language mixture, since it is said to have Bantu grammar but

not Bantu vocabulary. It is best known to non-Africanists from Goodman's

description, and probably most people who know about the language would

agree with his conclusion that "the development which Mbugu has undergone

defies easy categorization; it remains a unique linguistic specimen" [1971:

253]. In this paper I will argue that Ma'a is indeed a "mixed language"

---

[1]Elderkin comments that the self-name *Ma'a* is preferable to the more
familiar *Mbugu* as a designation for this language, since the name *Mbugu* is
also used for a Bantu language spoken in the same region [1976:280].

which probably should not be classified genetically into any language family, and that its history can be partly inferred from its structure and from published information about the history and cultural traits of its speakers. I will also show that this type of linguistic mixture is not unique, but is found in a few other languages whose history is better attested.

Greenberg was apparently the first, and for some time the only, linguist to classify Ma'a genetically as a Cushitic language. In making this classification, he used his method of mass lexical comparison [Greenberg 1955, 1963], so his criterion for the grouping was strictly lexical. Since 1971, more extensive lexical studies have shown conclusively that the basic vocabulary of Ma'a is primarily of Cushitic origin (see especially Ehret [1980]), and as a result of these studies the genetic grouping with Cushitic now seems to be widely accepted. Welmers, for instance, observes that "the development of Ma'a, though certainly unusual, is within the familiar framework of continuous language history with extensive external influence", and that "the continuous or genetic history of Ma'a is Cushitic" ([1973:8]; see also Ehret [1974, 1976, 1980]).

Opposing views can be found even in the relatively recent literature, however. Dolgopolskij groups Ma'a with Bantu because of its Bantu morphology ([1973]; cited by Zaborski [1976:83]), and the older "mixed language" claim is favored by Whiteley [1960a]. Elderkin also seems to prefer the last treatment, and to conclude from the case of Ma'a that genetic classification of languages is in principle unscientific: "classification of languages rests on the selection of one part of a language to typify that language, and this selection is arbitrary" [1976:296-7]. He goes on to say that Ma'a is a crucial example because "no theory of linguistic relationships should have an exception" (p. 297).[2] Elderkin is quite right to suspect that something is wrong with a genetic linguistics that takes just one part of a language as

---

[2]This is a rather common view. Note, in the same volume, Grover Hudson's remark that genetic classification is 'exhaustive, since no languages can be left out' [1976:237]. The same position is reflected, I think, in the strenuous efforts that have been made by some creolists to justify genetic classifications for Caribbean creoles (see e.g. Meillet [1921:82] and Hall [1958:370f.]).

diagnostic for purposes of classification. I could, in theory, graft Russian
lexical morphemes onto my English grammar, but I would not then be speaking
Russian; and I would no longer be speaking English, either.

   The flaw in Elderkin's reasoning is his assumption that genetic classi-
fication of languages must be exhaustive. The idea that some languages do
not fit into the standard genetic picture dates from the nineteenth century,
with Schuchardt's work on pidgin and creole languages. It remains contro-
versial even for these classic examples of "mixed languages",[3] but the
controversy arises, in my opinion, from a failure to take seriously enough
the first principle of genetic relationship--namely, that a daughter language
is a changed later form of its single parent language. On this view, we must
surely assume that any given daughter language in a family tree arises
through an unbroken series of generation-to-generation transmissions of an
entire language, that is, a complex set of interrelated lexical, phonological,
morphosyntactic, and semantic structures. Changes, both externally and
internally motivated, accumulate gradually enough that (as can be shown for
Indoeuropean, for instance) systematic reflexes of proto-language structures
can be found in all linguistic subsystems of a daughter language even after
five or six thousand years. But if the chain of transmission is broken at
any point, then the resulting language no longer belongs in any family tree,
because it is not a changed later form of any single parent language: it
does not meet the conditions for genetic classification.[4] Such a break in
transmission will always be reflected, as I have argued elsewhere [Thomason

---

   [3]See Thomason [1980] for a discussion of this controversy.

   [4]The notion of a "break in transmission" is necessarily vague, because
there are borderline cases where transmission is neither clearly normal and
continuous nor clearly abnormal and discontinuous. One example is Afrikaans,
which is viewed by some linguists as a creole and by others as a direct out-
growth of Dutch; other examples are cases where two or more languages in an
intimate contact situation are so closely related that the source of many
structures cannot be determined. One such case is found in those areas of
England where Old English and Old Norse were spoken, at a time when they
still shared many lexical and grammatical features. Other possible examples
of this type might be certain pidginized Bantu languages that arose in exclu-
sively Bantu-speaking areas. These and other borderline cases are described
in Thomason & Kaufman, Forthcoming.

1980], in a mismatch between the vocabulary and the grammar of the resulting
language; it will not be possible to show that both have arisen from the same
parent language.

The most obvious candidates for languages with nongenetic origins are
pidgins, which do not arise through any sort of transmission. Instead, they
are created as new languages in multilingual contact settings[5] in which a
single lexicon is adopted--usually, though not always, taken almost entirely
from a single language--and a new grammar evolves through intergroup communi-
cation, generally without a single-language target. The next most obvious
nongenetic languages are creoles like those of the Caribbean, which developed
rapidly among linguistically diverse groups of slaves who adopted the lexicon
of the slavemasters and constructed a new grammar that apparently did not, at
least in the beginning, involve any serious attempt to learn a single lan-
guage's grammar.[6] These languages arose, therefore, outside of normal
transmission processes. In most pidgins and in the Caribbean creoles, the
vocabulary is taken from a single language, and the grammar is not derived
from that language or from any other single language. The nongenetic histor-
ical development, though not evident in the single-source vocabulary of such
a language, is reflected clearly in the grammatical structures: the least

---

[5]Or, much more rarely, in bilingual contact situations. Whinnom (1971)
has argued that, for social reasons having to do with the availability of the
target language, no pidgin can develop in a bilingual setting. But his model
does not take into account the possibility that speakers of the vocabulary-
base language might deliberately withhold access to their language. Such
deliberate withholding is attested for the 17th-century Delaware-based Amer-
indian pidgin and for Mobilian Jargon, and it probably also accounts in part
for the emergence of Tây Bôi between French and Vietnamese speakers in
Vietnam.

[6]The Caribbean creoles are considered by some creolists, e.g. Alleyne,
Bickerton, and Thomason & Kaufman, to have developed in a process of abrupt
creolization--that is, without going through a fully crystallized pidgin
stage. Most of them have remained in contact with the vocabulary-base
language under social circumstances that encourage convergence toward that
language, i.e. decreolization. As a result, these creoles may safely be
assumed to be more like the European vocabulary-base language now than they
were when they first crystallized as creole languages. Caribbean creoles
like Saramaccan which have *not* remained in contact with the vocabulary-base
language show more African, and fewer European, structural features.

decreolized Caribbean creoles and European-vocabulary pidgins like Tok Pisin
(Neomelanesian) show few or no universally marked features characteristic of
European languages, but they do have marked features characteristic of the
relevant substrate languages (African, Melanesian).

The lesson to be learned from these cases is that lexical correspondences,
no matter how numerous and systematic they are, cannot stand alone as suffi-
cient evidence of normal transmission, and hence of genetic classifiability.
From a retrospective viewpoint, in order to rule out the possibility of
nongenetic development, we must show systematic correspondences in grammati-
cal as well as in lexical structures, and cognation in grammatical as well as
in lexical morphemes.[7]

If we look at Ma'a from this perspective, we will not focus on the lexi-
con, because the Cushitic origin of the basic vocabulary is no longer in
doubt. Ma'a thus belongs either in the Cushitic group or in no genetic
group. And we will not focus primarily on the mere fact that Ma'a has some
Bantu grammatical features, because most languages that acquire foreign gram-
matical features do so without losing their genetic continuity. Instead, the
crucial question has to do with Cushitic grammar: does Ma'a have enough of
it to qualify as a Cushitic language in the full genetic sense?

This question is addressed in Section 2 below, in a systematic compari-
son of Ma'a structures with characteristic Cushitic and Bantu grammatical
structures. The method of comparison is primarily typological. The reason
for this typological emphasis is that even "hard-to-borrow" features like
inflectional affixes might show regular phonological correspondence with
comparable affixes in one language, but close functional and/or positional
correspondence with affixes in another, as a result of interference. So, for
instance, the so-called "second genitive" case in Russian, a partitive con-

---

[7]The insistence on grammatical correspondences in languages that are
claimed to be genetically related is of course not new; many historical
linguists have emphasized the importance of such correspondences at least
since Gyarmathi's time (late 18th century). But this aspect of genetic
linguistics has sometimes been neglected because vocabulary is easier to
elicit, easier to compare, and certainly easier to quantify than
grammar.

struction, has as its marker a native Russian suffix -u , but the grammati-
cal distinction between partitive and non-partitive genitives entered Russian
through the influence of neighboring Finnic languages. Examples of this sort
are common, and they show the need for extreme caution in interpreting his-
torically the products of the most intensive contact situations, like the one
in which modern Ma'a arose: solid evidence of a particular origin must be
sought in structural features in which the potential source languages dis-
agree. Where the sources agree typologically, a definite origin can be
assigned only to morphemes that agree in form *and function* with one source
or another.

In the case of Ma'a, we will see that in general its structures are
similar to Cushitic structures only where Cushitic and Bantu are typologi-
cally similar. Where Cushitic and Bantu differ, Ma'a usually agrees with
Bantu. Specifically, Ma'a corresponds to Cushitic in a few phonological
units, syntactic structures, and derivational processes, and in one feature
of the inflectional morphology. Otherwise Ma'a matches Bantu closely, and
most strikingly in the inflectional morphology, where it has a complete and
productive set of Bantu inflectional structures. Overall, few productive
nonlexical structures in Ma'a can be shown to be of definite Cushitic origin,
whereas, by contrast, many can be shown to derive from Bantu. Usually, as
several authors have observed, the Bantu structures can be traced to the
Bantu language Pare, a southern dialect of Asu, whose speakers have been in
intimate contact with the Ma'a people since about the seventeenth century
[Ehret & Nurse 1981:141-2]. The other Bantu source is Shambaa, whose
speakers are now neighbors of the remaining Ma'a speakers in the northeastern
corner of Tanzania.

2.   Ma'a, Cushitic, and Bantu Structures

In the discussion that follows, I am basing my statements about Ma'a
primarily on the five most useful published sources available to me: Ehret
[1980], Copland [1933-34], Green [1963], Tucker & Bryan [1974], and Elderkin
[1976]. The first four sources make use of primary data; Elderkin's analysis
is based on secondary sources. All five sources together provide only a

fragmentary sketch of Ma'a grammar, but additional information has been pro-
vided by Christopher Ehret [personal communication, 1982].[8] For Cushitic
grammar I am relying heavily on the very useful sketches in Bender [1976].[9]
I will emphasize Iraqw and Dahalo, the two Southern Cushitic languages for
which I have descriptions, since Southern Cushitic is the branch that Ma'a
matches lexically.

2.1. Phonology. As far as phonology is concerned, the Ma'a inventory con-
sists mostly of phonemes that are common in both Cushitic and Bantu (and,
for that matter, in other language groups around the world). Given the fact
that the mix of structures in Ma'a must reflect extensive contribution from
both sources, the presence of such common phonemes in Ma'a cannot be
ascribed definitely to either source. In particular, though these phonemes
correspond regularly to identical or similar phonemes in other Southern
Cushitic languages, their phonetic representations in Ma'a might just as well
be due to the fact that Bantu languages also have such sounds as to direct
inheritance from Cushitic. Evidence for phonological inheritance from
Cushitic, or for interference from Bantu, must therefore rest on the presence
or absence in Ma'a of phonemes that occur in only one of the two groups, and

---

[8]The only major sources that are not available to me are Meinhof [1906]
and the source called FILE by Tucker & Bryan [1974]. But since the sources
I do have make frequent references to these, it is unlikely that information
crucial to my argument is missing. The Ma'a data that Ehret used in his 1980
study came from his own field work in 1967 and 1973 and from Bernd Heine's
field work [Ehret 1980:11f.].

[9]This book contains sketches of the following Cushitic languages: Iraqw
and Dahalo [Elderkin 1976]; Beja [R. Hudson 1976]; Highland East Cushitic
[G. Hudson 1976]; Werizoid [Black 1976]; Afar [Bliese 1976]; Oromo (Galla)
[Gragg 1976]; and Dasenech [Sasse 1976]. Other sources on Cushitic languages
that I have used are Welmers [1952] and [1973] on Saho; Bender et al. [1976]
on Hadiyya (Highland East Cushitic) and Oromo; Whiteley [1960b] on Iraqw;
R. Hudson [1974] on Beja; and Tucker & Bryan [1966] on Cushitic languages
in general (especially Galla, Somali, Awiya, Bilin, and Beja).

the strongest evidence will be phonemes that are universally marked.[10]   Ma'a
does have a few phonemes that provide such evidence, as we will see on
examining the phonemic inventory in Table 1 (data from [Ehret 1980:113]).

<div align="center">

Table 1. <u>Ma'a phonemes</u>

</div>

<u>Consonants</u>					<u>Vowels</u>	
p	t	č	k	ʔ	i	u
b	d	ɟ	g		e	o
ᵐb	ⁿd	ⁿɟ	ᵑg		a	
f	s	ɬ	š	x	h	
v	z		ɣ			
m	n	nʸ	ŋ			
r	l					
w		y				

<div align="center">

<u>Tones</u>

HIGH   LOW

</div>

     The Ma'a inventory has one phoneme, the voiceless lateral fricative
/ɬ/ , which is rather highly marked in universal terms and which is clearly
a Southern Cushitic (SC) inheritance. Lateral obstruents occur in other SC
languages, but such phonemes apparently do not occur in other branches of
Cushitic, and they do not occur widely in Bantu. So in this respect, at
least, Cushitic words are pronounced in Ma'a with a characteristic SC sound.
Ma'a has two other phonemes which, though not universally marked, are more
common in Cushitic than in Bantu: /ʔ/ and /x/ . Of these /ʔ/ is far
more common than /x/ as a phoneme in Cushitic. Other than these, no Ma'a
phonemes look like promising candidates for specifically Cushitic inheri-
tances.
     One subphonemic phonetic feature of Ma'a seems likely to be due to Bantu
influence. The voiced stop phonemes /b d ɟ/ have implosive pronunciation

---

[10]By "universally marked" I mean, here, a phoneme or phoneme type that
is uncommon in languages of the world--so uncommon that it can probably be
considered relatively unlikely to arise spontaneously, and thus more likely
to be present as a result of inheritance or convergence.

[Ehret 1980:130], and even /g/ is pronounced with weak implosion [Ehret, p.c. 1982]. Now, in many Cushitic languages implosives (usually only /ɗ/) and ejectives occur as phonemes. Dahalo, for instance, has several glottalized phonemes of each type, and Ehret [1980] reconstructs implosives for Proto-SC. The fact that the implosive pronunciation is not distinctive in Ma'a would therefore suggest that Ma'a matches Bantu in this respect, since Bantu languages often have implosives as allophones of pulmonic voiced stop phonemes. But Ehret comments that the implosive feature is not "attributable to Bantu and, where it occurs in East African Bantu languages, it can be laid to pre-Bantu habits of articulation" [p.c. 1982]. Nevertheless, the loss of glottalization as a distinctive feature in Ma'a still requires an explanation, and of course it is quite possible for a Bantu language to have acquired allophonic implosives from earlier interference and then later to have influenced Ma'a to lose glottalized stops as a distinctive phoneme type. Since both Pare and Shambaa have allophonic implosive pronunciation of voiced stops [David Odden, p.c. 1982], this seems the most likely source of the Ma'a implosives.

Several features of Ma'a phonology can definitely be ascribed to Bantu influence. The most striking one is the highly marked series of four prenasalized voiced stops, /ᵐb ⁿd ⁿɟ ᵑg/ , which entered the language first in Bantu loanwords [Ehret 1980:113].[11] The other non-SC phonemes in Ma'a are not universally marked, but the Bantu influence is clear. Ma'a acquired the phonemes /v ɟ ɣ/ in Bantu loanwords, and Bantu /ɟ/ and /ɣ/ caused the original Ma'a allophones [ɟ] and [ɣ] to assume phonemic status. Moreover, the *absence* in Ma'a of several marked phoneme types is noteworthy, because Bantu also lacks them. Since all of these occur in other SC languages and are reconstructed by Ehret for Proto-SC, their absence in the Ma'a cognates

---

[11]Although Ehret [1980] reconstructs a series of prenasalized voiceless stops for Proto-SC, the prenasalized voiced stops in Ma'a do not correspond to these. The only SC language with prenasalized voiceless stops is Dahalo (though Ma'a has corresponding nasal + stop clusters in some environments), so the reconstructed set seems rather dubious for Proto-SC, especially since Dahalo, like Ma'a, has undergone heavy interference from neighboring languages--with, however, very different results.

strongly suggests elimination under Bantu influence. Ma'a has lost pharyngeal
fricatives, which are common in Cushitic in general and occur in Iraqw and
Dahalo in particular. The labialized dorsal stops that are characteristic of
Southern, Northern, and Central Cushitic do not occur in Ma'a. The Proto-SC
ejectives and retroflex stops are gone. Besides these marked phoneme types,
Ma'a has lost the contrast between long and short vowels that characterizes
most Cushitic languages, including Proto-SC. I cannot judge the probability
of Bantu influence on this loss, since, though many Bantu languages lack
phonemic vowel length [Welmers 1973:25], I have no specific information about
this feature in Pare or Shambaa. As for prosodic features, Ma'a has two
phonemic tones and predictable stress on the first stem syllable [Ehret, p.c.
1982]. According to Ehret, Bantu loanwords contributed to the development of
tone in Ma'a; he does not reconstruct lexical tone phonemes for Proto-SC, and
he considers Iraqw to have neither phonemic tone nor phonemic stress [p.c.
1982]. Dahalo, like Ma'a, has lexical tone distinctions. In general,
Cushitic languages have phonemic stress, while Bantu languages have phonemic
tones.[12] Finally, at least one phonotactic feature of Ma'a may be due to
Bantu influence. Quite recently, to judge by the chronology indicated by
other sound changes, Ma'a lost all word-final and verb-stem-final consonants;
this change, according to Ehret, may have been motivated by "the Bantu pat-
tern in which no word could end in a consonant" [1980:110].

In sum, the Ma'a phonemic inventory has two phonemes that are character-
istic of Cushitic but not of Bantu, namely /ɬ/ and /ʔ/ . But in most
phonological features in which Bantu and Cushitic differ, Ma'a matches Bantu
rather than Cushitic. These features include the presence of prenasalized
voiced stops and of phonemic tones, and the absence of pharyngeal fricatives,
labialized dorsal stops, ejective and retroflex stops, and final consonants.
Ma'a also differs from typical Cushitic structure in its lack of distinctive

----

[12]Welmers [1973:78] observes that, though Cushitic languages are not in
general tonal, some are analyzed as having phonemic tones. But, he says,
Saho pitches are predictable if stress is treated as phonemic, and he
believes that this is likely to be true for at least some other Cushitic
languages too.

vowel length, but I do not know the status of vowel length in the relevant
Bantu languages. In addition to these general structural divergences from
Cushitic, Ma'a has acquired the phonemes /v ɟ ɣ/ under Bantu influence.

2.2 <u>Morphology</u>. When we turn to the morphology, we find a sharp distinction
between inflectional and derivational patterns as far as their historical
sources are concerned. The inflectional system of Ma'a is almost entirely
of Bantu origin, but the derivational affixes seem to be about evenly
divided between Bantu and Cushitic suffixes. Typologically, however, Bantu
and Cushitic agree in the particular kinds of derivational processes attested
in Ma'a. So, in the morphology as in the phonology, where the two groups
differ typologically Ma'a almost always resembles Bantu rather than Cushitic.

2.2.1. <u>Nominal inflection</u>. In the nominal subsystem the most important
grammatical categories are noun classification, number, pronominal posses-
sion, and adjectival agreement. Almost everyone who has written about Ma'a
has emphasized the presence of Bantu noun-class prefixes, both on nouns and
as agreement markers on adjectives, verbs, and certain particles. These
prefixes, usually from Pare but occasionally (judging by the phonological
shape) from Shambaa, fall into the standard Bantu classes 1/2, 3/4, 5/6, 7/8,
9/10, 11, 13, 14, 15, and 17 [Ehret 1980:131]. Most writers have also
remarked on the inconsistent use of the prefixes. Bryan reports, for
instance, that one informant gave mu-haraza as the citation form for
'river', but later used the unprefixed form haraza in conversation [Tucker
& Bryan 1974:192]. Moreover, some adjectives are invariable, i.e. they do
not agree with the nouns they modify , and some nouns are also invariable.[13]

---

[13]Although the type of data is limited--only words of Cushitic origin,
and (presumably) only citation forms given--a count of the noun classes repre-
sented in Ehret's published Ma'a data [1980] gives a rough idea of the level
of attestation for the various classes: 134 nouns have no prefix (but some
of these would be prefixless in the singular in Bantu too); 239 nouns have
class prefixes, including m(u)- (class 1 or 3; 60 nouns), mi- (cl. 4; 1
noun), i- (cl. 5; 72), ma- (cl. 6; 17), ki- (cl. 7; 35), N- (cl. 9/10;
17), lu- (cl. 11; 20), ka- (cl. 13; 4), (v)u- (cl. 14; 10), and ku-
(cl. 15; 3). This last class contains all verb infinitives, so the actual
number of Ma'a nouns in ku- is no doubt very large, even though Ehret's

But none of the Ma'a sources highlights the point of major significance for
the genetic question. The crucial fact about the Ma'a system of noun classi-
fication is not that it resembles Bantu and has morphemes of Bantu origin,
but rather that it is so *unlike* Cushitic that it cannot possibly be viewed as
a continuation, or even as a partial remodelling, of an earlier Cushitic
classificatory system. Typologically, Bantu and Cushitic languages differ
in every respect in the ways in which they classify nouns.

    First of all, the semantic content of the Bantu noun classes is quite
varied. Typical semantic features used for classifying nouns are human;
trees and other plants; animals; long, thin objects; paired objects; items of
material culture; diminutives; augmentatives; abstract nouns; and the verbal
infinitive. Cushitic and other Afroasiatic languages, by contrast, have just
two noun classes, based on the semantic feature of biological gender: mascu-
line and feminine.[14] Second, the formal expression of noun classification in
Bantu is entirely prefixal, with paired singular/plural prefixes on nouns and
agreement prefixes (sometimes different from the noun prefixes) on adjectives
and other modifiers, on verbs, and on some particles, e.g. the associative
particle used in genitive constructions. Cushitic languages do not have a
uniform set of gender affixes on nouns themselves, though in some of the
languages the gender of at least some nouns is reflected in the form of the
noun. Examples are found in Saho, in which stressed nouns and unstressed
nouns ending in a consonant are regularly masculine, while unstressed nouns
ending in a vowel are feminine [Welmers 1973:222]; Dasenech, in which dis-
tinct masculine and feminine singulative suffixes are added to collective
nouns [Sasse 1976:203]; and Afar, which has a masculine vocative suffix
opposed to a feminine vocative suffix [Bliese 1976:150]. In some Cushitic
languages modifiers agree with head nouns in gender, usually by the presence

---

data did not contain many. Ehret observes that prefixless nouns belong in
class 9/10 for purposes of concord [1980:131].

    [14]As in most languages with noun classification systems, Bantu and
Cushitic noun classes contain many nouns which do *not* meet the semantic cri-
terion, but which are classified arbitrarily or merely by their phonetic
shape. For instance, not all feminine Cushitic nouns refer to female
creatures, and Bantu languages have inanimates in the "animals" class.

or absence of a feminine suffix. Many of the languages distinguish masculine and feminine third person singular verb forms in combined tense-aspect/ subject suffixes, and most have distinct masculine and feminine third singular pronouns. Cushitic languages clearly do not all treat gender in the same way, and, according to Greenberg [1963:45], some of the western languages lack grammatical gender entirely. Nevertheless, all but one of the languages for which I have data have a masculine/feminine distinction at least in third singular free pronouns.[15] Cushitic languages that distinguish gender else-where as well typically have suffixes that vary according to gender. As far as the Southern Cushitic languages are concerned, Elderkin mentions a mascu-line/feminine distinction in Iraqw only in the free pronouns, while Dahalo is said to have natural gender which is marked at least in free pronouns, adjec-tives (by suffixes), and demonstratives. Proto-SC nouns, according to Ehret, were marked for gender by suffixes attached to the noun stem [1980:48].

Ma'a noun classification follows the Bantu pattern faithfully. It has prefixes of Bantu origin attached to Cushitic noun stems that are "associated with the same classes as the semantically comparable Bantu stem would be" [Elderkin 1976:289]. Most adjectives take the appropriate noun-class pre-fixes for the nouns they modify, and verbs take concordial prefixes to agree with full-noun subjects and objects. In genitive constructions, the 'of' particle -a is combined with the noun-class prefix appropriate for the head (possessed) noun, as in afá yá mohέ 'a person's goat' (lit. 'goat class=9- of person') vs. afá ɉá mohέ 'a person's goats' (lit. 'goat class=10-of person) [Tucker & Bryan 1974:200]. In all, as noted above, fifteen noun classes are attested in noun and concordial prefixes. Gender markers of Southern Cushitic origin do occur in Ma'a; compare, for instance, iŋi 'brother' and iŋinta 'sister' or i'alú 'sheep' and i'alé 'ram'. But, though common, they are not productive [Ehret, p.c. 1982]. The language's current pattern of noun classification is Bantu, and it has replaced an older

---

[15]The exception is Dasenech, which does distinguish masculine and femi-nine gender in nouns, adjectives, and verbs. So, all the Cushitic languages described in my sources have grammatical gender.

Cushitic pattern, which is now attested only in relic word pairs. Even the
personal pronouns, which are of Cushitic origin, lack the usual Cushitic
masculine/feminine gender distinction.[16]

The number category also reveals a deep typological division between
Bantu and Cushitic, and here again Ma'a is Bantu in type. As mentioned above,
number is marked in Bantu nouns by the noun-class prefixes, which occur in
paired singular/plural sets. It is therefore an obligatory category, both in
the noun itself and in other words that agree with the noun. In Cushitic, by
contrast, the category of number is not obligatory, at least for some nouns.
When plural is marked, a wide variety of markers, apparently with lexically
governed distribution, is used in many of the languages, including Iraqw and
Dahalo. For a given language, these markers may include several suffixes, an
infix, accentual alternations, reduplication of the final consonant, and a
change in vowel pattern. More striking still is the marking of number in
some nouns by adding a singulative affix instead of a plural one. Compare,

---

[16]Ehret [p.c. 1982] observes that "there is no specific evidence for
attributing the loss of gender in Ma'a to Bantu influence," and that "in the
pronouns the masculine forms were generalized as might be expected with nor-
mal processes of language change." It is true, of course, that a Bantu-style
system of noun classification is not inherently incompatible with a Cushitic
system based semantically on biological gender. A language *could* have both.
But though exceptions certainly exist, most languages in the world do not mix
biological-gender (or animacy) classification with other kinds of noun clas-
sification. Bantu, in any case, does not. So the circumstantial evidence
for the loss of Cushitic noun classes in Ma'a *because of* the rise of Bantu
noun classes is very strong, particularly in light of the fact that Ma'a
agrees typologically with Bantu, and differs from Cushitic, in so many other
respects: if, as Ehret believes, Ma'a was once an ordinary SC language, it
has shifted typologically toward Bantu in all its grammatical systems. And
if it has shifted toward Bantu in other subsystems, why not assume that Bantu
interference was the causal factor in this instance as well as in the more
obvious instances (like the agreement patterns or the phonemic tones)?
Similarly, "normal processes of language change" may not demand explanations
as dramatically as apparently abnormal ones do, but that does not mean that
we should not seek explanations for them. Since even the most natural changes
often fail to occur, it is never inappropriate to ask why a particular change
happened when it did; and if a reasonable explanation is available, it should
not be rejected merely because similar changes have occurred under different
antecedent conditions.

for instance, Dahalo ʔúšò 'male elephant' : plural ʔùšàšè and kì:dzò
'old man' : plural kí:dzò:mà , but á:dʒù 'lung fish' : singulative
á:dʒùmè [Elderkin 1976:292].

Ma'a marks number by the paired Bantu noun-class prefixes, both in nouns
and in other words that agree with nouns. Examples (from Green [1963]) are
mu-'o m-gititu 'small mouth' (-'o 'mouth' and -gititu 'small') : plural
mi-'o mi-gititu . Ehret [p.c. 1982] cites two Ma'a nouns of Cushitic origin
that have typical SC number marking: nihi 'animal' : plural nihena , and
tambala 'snail' : plural tamba (the singular of 'snail' has a suffix -a ;
stem-final -l has been lost in the suffixless plural by a regular sound
change). Ma'a has few, if any, other traces of Cushitic number marking [Ehret
1980:48], though my sources list several nouns with a Cushitic suffix -no
that indicates mass quantity [Ehret, p.c. 1982] and apparently functions
sometimes as a quasi-plural: iɬare (Bryan) or ɬare (Green) 'cloud' :
plural ɬareno (Green), maɬareno (Green, Bryan), or maɬare (Bryan);
ŋgile 'bee' : ŋgileno 'swarm of bees' (Green); kunge 'Kweme nut' :
plural kungeno (Green); and 'i'alú 'a sheep' : plural ale:no (Meinhof
[1906]) or ma'alú (Bryan) (Green [1963:185]; Tucker & Bryan [1974:207]).
Even here it is noteworthy that the variant forms for 'clouds' and 'sheep',
with only the Bantu plural prefix ma- , occur in the most recent source;[17]
Ehret [p.c. 1982] confirms that the -no suffix is losing ground in the
language. (The hybrid form maɬareno 'clouds', with both the Bantu plural
prefix and the Cushitic plural suffix, suggests that one mechanism for the
replacement of Cushitic patterns by Bantu ones may have involved double-
marking of nouns at one stage.)

Case inflections do not occur in Ma'a or, in general, in Bantu. It is
hard to say whether or not this lack puts Ma'a in significant contrast with
Cushitic, however. Most Cushitic languages seem to have at least a two-case
distinction, between a subject case and an accusative or "absolute" case; but

---

[17]It should also be noted, however, that Bryan collected her Ma'a data
in 1959, and Green does not say when he collected his, so that these two
sources might represent contemporaneous usages. The Meinhof data, of course,
is the oldest of the three.

I have no information about the category of case in Southern Cushitic. In
one construction involving case Ma'a does differ significantly from the Bantu
languages of the area. Shambaa, like Swahili, uses a locative suffix -nì
beside a locative construction with the associative particle -a (see below)
to express location. For instance, 'on the mountain' is mwima-nì or
zuu ya mwima (literally 'aboveness of the mountain'). By contrast, Ma'a
uses a preposition in locative expressions, as in aná Ìongorí 'on the
mountain'.[18] I have little information about locatives in Cushitic, so I
cannot tell whether the Ma'a construction agrees with those in Cushitic
languages or not (though most branches in the group have postpositions rather
than prepositions).

　　Pronominal possession is marked on nouns in most Cushitic languages by
a set of suffixed possessive pronominals which is etymologically related
entirely or in part to the set of independent pronouns. In Iraqw these
suffixes are added to the noun after a suffixed class marker (a syntactic
class marker that determines the person and number of verb concord). In
Bantu, pronominal possessors are full pronouns which, like other adjectives,
follow the noun and take noun-class agreement prefixes. Ma'a pronominal
possessors are of Cushitic origin and are apparently suffixed, as in
Cushitic. Goodman states that they do not take concordial agreement [1971:
245], but Tucker & Bryan observe that this is true only when the possessives
are used attributively. They do take the appropriate Bantu concord prefixes
when used predicatively, as the following examples show: yá iŋí ni mu-ɣɔ́
'this child is mine' (DEM child BE class=1-my) and yá va-iŋí ni va-kánù
'these children are ours' (DEM class=2-child BE class=2-our). Compare
attributive use in mu-harega go 'my arm' (class=3-arm my), i-ɬέ ɣɔ́
'my name' (class=5-name my), and ki-kire go 'my stool' (class=7-stool my)
[Tucker & Bryan 1974:202]. The Ma'a attributive possessive construction thus
lacks the concord prefix and connective we would expect in a Bantu language,
as e.g. in Swahili wa-toto wa-a-ŋgu 'my children' (class=2-child class=2-

---

　　[18]This information about Ma'a locatives, and the Ma'a and Shambaa
examples, were provided by Christopher Ehret [p.c. 1982].

CONNECTIVE-my) vs. vĭ-su vi-a-ŋgu 'my knives' [Gleason 1955:48]. The
attributive construction in Ma'a is similar to analogous constructions in
Cushitic and is, in fact, the only inflectional pattern in Ma'a whose origin
is clearly Cushitic. Moreover, when attached to kin terms pronominal posses-
sors beginning in a vowel are frequently preceded by a Cushitic connective
morpheme -r- [Ehret, p.c. 1982], so that even the morphophonemic behavior
of the possessives is Cushitic.

2.2.2. Verbal inflection. Verbs in both genetic groups, Cushitic and Bantu,
inflect for the same general categories: tense/aspect, person, number, and
noun class. However, the two groups differ sharply in the actual verb con-
struction. Typologically, Cushitic verbs tend to be flexional, with partially
or wholly unsegmentable tense-aspect/subject affixes, while Bantu is agglu-
tinative. Thus, the Iraqw second person plural marker is composed of the
second singular suffix -t plus -a (plural + present/future) or -e
(plural + past); similarly, the Lowland East Cushitic language Oromo (Galla)
has a third singular masculine imperfective suffix -a opposed to a third
singular masculine perfective suffix -e .[19] Most Cushitic languages, like
Iraqw and Oromo, have suffixed tense-aspect/subject markers. In Bantu, by
contrast, tense/aspect and subject agreement are expressed by separate
affixes--subject prefixes, tense/aspect prefixes and often suffixes too.
Another difference is that object marking is rare in Cushitic verb morphol-
ogy, and it does not occur, apparently, in Southern Cushitic,[20] but Bantu
languages have full sets of object prefixes as well as subject prefixes.

     The Bantu third person agreement prefixes occur in paired singular/
plural sets according to the class of the noun referent; these are prefixed
to nouns and adjectives. First and second person prefixes do not inflect for

---

[19]Bender et al. use the traditional terms 'imperfect' and 'perfect' to
refer to these two aspects in Oromo, and probably they correspond in the
typical Afroasiatic way to Elderkin's tense-labeled categories in Irawq.

[20]*Pace* Elderkin [1976:294f.], who speaks of a particle which either pre-
cedes the verb or is suffixed to it. I take it that this is not a true
suffix when it follows the verb, but rather still a particle. An example is
ʔákùwàté ìùβò ~ ìùβókùwà 'I hit(past) (someone) for you'.

noun class. Cushitic verbs typically show a masculine/feminine distinction
in the third singular only, not in the plural, and even this minimal gender
distinction seems to be lacking in Iraqw and Dahalo. Number agreement with
the subject is marked either flexionally, as in Dahalo third singular -í
vs. third plural -e: , or agglutinatively, as in Iraqw third singular -/í/
vs. third plural -/i̧ + r/ .

Tense/aspect in Cushitic generally follows the familiar Afroasiatic
pattern in which the major division is between the imperfective and perfec-
tive aspects. Subcategories of tense and, especially, aspectual functions
are expressed by derivational processes like suffixation and initial-syllable
reduplication. Bantu tense/aspect systems also distinguish completive from
incompletive aspect [Welmers 1973:350,384], and they have in addition a
paradigmatically related set of prefixes expressing features like simple pre-
sent, present continuous, future, immediate past, remote past, stative (or
perfect), and mood (e.g. conditional). Tense is in general a more important
inflectional category in Bantu than in Cushitic.

Cushitic has a number of other characteristic morphological features,
mostly derivational ones, in its verb system. Among these are prefixed
reduplication as an intensive or frequentative formation, verb negation by
means of a suffix or a prefix, and passive and causative formation by suffix-
ation. Bantu verbs often share the last two features, e.g. negation in
Kinyarwanda by means of a prefix, as in /nhi-ba-geénd-â/ 'they are not
going' (NEG-they-go-ASPECT) [Kimenyi 1978:313], and passive formation by
means of a suffix -wa , as in Venda funa 'love' : fun(i)wa 'be loved'.

In verb morphology, Ma'a patterns with Bantu wherever Bantu and Cushitic
differ. Its verb morphology is agglutinative, and the inflectional morphemes
are arranged in the common prefixal Bantu pattern:

(NEG +) Subject + Tense + (Object +) ROOT (+ extension)

(In Copland's text [1933-4:243,245], however, there is no object prefix if a
full-noun object is present in the sentence.) Examples are ve-ne-tu-ifi
'they (the Masai) will destroy us' (class=2-FUT-us-destroy) and
te-tu-ta-zaxo 'we will not hold' (NEG-we-FUT-hold). This ordering

contrasts sharply with the predominantly suffixing Cushitic patterns. More-
over, the inflectional affix morphemes themselves are all Bantu in origin,
and most of them can be identified with Pare and/or Shambaa affixes. Ma'a
even has some nonautomatic morphophonemic alternations characteristic of
Bantu verb morphology, most notably the distinction in the negative between
first person singular si- , e.g. si-'ánthù 'I do not cook', and the other
persons with a separate negative prefix, e.g. tɛ-tú-'ánthù 'we do not cook'
(NEG-we-cook; the present tense marker is zero) [Tucker & Bryan 1974:204];
compare Swahili si-anguki 'I do not fall' and ha-tu-anguki 'we do not
fall' [Loogman 1965:200]. The class 1 (personal singular) verb agreement
marker also varies as in Bantu. Ma'a has a- (or e- , in an alternation
also found in Asu) as subject, as in Mwa-gilu a-ka-ba' 'Elder said'
(class=1-elder class=1-PAST-say) [Copland 1933-4:244], and m- as object,
as in n-áà-m-ma 'I have hit him' (I-PERFECT-class=1-beat) [Tucker & Bryan
1974:201]. Compare Swahili a-me-m-piga 'he has beaten him' (class=1-
PERFECT-class=1-beat) [Gleason 1955:26].

The inflectional morphemes of Bantu origin in Ma'a verbs include subject/
object markers for all three persons, both singular and plural, with a variety
of noun-class prefixes attested for the third person; tense prefixes ta-
(present or future), ne- (future), aa- (perfect), ka- (past); a tense/
aspect suffix -ye (past NEG and conditional); a conditional prefix ku-
(in the tense/aspect position); and a negative prefix si-/te- .

2.2.3 Derivational morphology. As mentioned above, derivational affixes in
Ma'a seem to be divided about evenly between suffixes of Bantu origin and
suffixes of Cushitic origin. But although the *morphemes* come from both
sources, the productive derivational *patterns* in Ma'a are all, as far as I
can tell, patterns in which Cushitic and Bantu agree, at least in ordering:
all the processes involve suffixation. Cushitic derivational processes that
are not found in Bantu, most notably prefixed reduplication and infixation,
are apparently not productive in Ma'a, though lexicalized frequentative
reduplicatives are rather common in Ehret's Ma'a data [1980], e.g. fufu
'to catch breath, rest' (from Proto-SC *fook'- , with prefixed reduplica-

tion and loss of the stem-final consonant).

    The most common causative formative in modern Ma'a, according to Ehret,
is probably the Cushitic suffix -ti [1980:63], e.g. -gugúlu 'run' :
-gugulúti 'drive away'. The productivity of this suffix is indicated by
(among other things) the fact that it has been added to stems since the recent
and probably Bantu-influenced sound change that eliminated final consonants.
Another Cushitic verb extension, -'V , is also still quite productive, and
three others, -mu , -u , and -au , were productive until fairly
recently [Ehret, p.c. 1982]. But a number of Bantu verb extensions are also
quite productive in Ma'a and are used with Cushitic as well as with Bantu
verbs. For instance, in Ehret's Ma'a data [1980] I find seventeen Cushitic
verbs with the causative suffix -ija , ten with the stative -Vka , seven
with the passive -wa , four each with the reciprocal -ana and the inten-
sifier -ya , and a few other suffixes in one or two verb forms each. (Note
that the entries in this data list are generally citation forms only, not
sets of inflected and/or derived forms; still, the occurrence of the Bantu
suffixes in the list gives some indication of their penetration into the
Cushitic vocabulary.) Finally, Ehret [p.c. 1982] observes that "the Cushitic-
derived amplificative of both nouns and verbs, -ša , is a very productive
suffix" with "no direct parallel in neighboring Bantu languages." He also
points to an apparently still productive adjective-forming suffix -'i of SC
origin. Ma'a thus has a number of Cushitic derivational affixes beside Bantu
affixes, but typologically there is no contrast between the two groups in
this grammatical subsystem.

2.3 Syntax. As in the derivational morphology, the syntactic structures of
Ma'a are divided between patterns of Cushitic origin and patterns of Bantu
origin. But unlike the derivational patterns, two of the relevant syntactic
differences between Cushitic and Bantu are typologically significant. First
of all, Cushitic languages have dominant SOV word order. Most of them also
have other word-order features often associated with SOV languages (cf.
Greenberg [1966], e.g. postpositions and Adjective-Noun word order; but
Noun-Adjective order is dominant in Iraqw and Dahalo, and in some Lowland

East Cushitic languages as well. Ma'a has as dominant word order patterns SVO and Noun-Adjective, and these are typically Bantu.[21] Ma'a also has prepositions rather than postpositions, e.g. locatives he 'to' and na 'from', both from Proto-SC verb roots [Ehret, p.c. 1982]. Prepositions are expected in an SVO language, but these particular ones cannot be attributed directly to Bantu influence, since, according to Ehret, in Asu and Shambaa "verbs of movement do not generally require the insertion of a directional marker" [p.c. 1982].

Bantu and Cushitic genitive constructions also differ, and Ma'a uses both types, but typologically the two are not far apart. Bantu uses an associative particle -a with appropriate noun-class prefixes. In Cushitic a possessor noun sometimes takes a case suffix or a subordinative particle, but sometimes the possessed and possessor nouns are simply juxtaposed without special marking of the construction. Iraqw and another SC language, Burunge, use the latter Cushitic construction, and so does Ma'a; but Ma'a also has the Bantu construction with prefixed -a between the two nouns. Ehret [p.c. 1982] remarks that the Cushitic usage seems almost as common as the Bantu pattern in Ma'a.

Another construction type in which Ma'a uses both Cushitic and Bantu patterns is the copula. The Bantu morpheme ni is used in Ma'a as the copula and to introduce the agent of a passivized verb [Goodman 1971:248], but, according to Ehret [p.c. 1982), forms of SC verbs for 'to be' are more common, and the copula relationship is obligatorily marked in Ma'a--unlike Asu and Shambaa, which frequently omit the marker. In this syntactic feature, however, as in the genitive construction, the typological difference between Bantu and Cushitic is not very great, so the mixture of construction types in Ma'a does not in any case seem likely to cause serious communication difficulties for speakers using different patterns.

The same point can be made for one of the two syntactic features that

---

[21]Ehret [p.c. 1982] believes that "earlier SOV order can be internally reconstructed for pre-Ma'a," and that the shift to SVO order is probably recent, dating from a period when Bantu influence had become significant.

have come to my attention in which Ma'a clearly matches Cushitic but not
Bantu. Although most (but not all) Bantu languages have a class of words
that can be identified as adjectives, the class is quite small in most of the
languages [Welmers 1973:271]. But in Ma'a, as in other SC languages, this
class is a large and important one [Ehret, p.c. 1982].

     The other exclusively Cushitic feature of Ma'a syntax is the means of
expressing possession, a "normal transitive verb -|o 'to have', fully
conjugatable and occurring in the normal syntactically verbal contexts"
[Ehret, p.c. 1982]. Ehret points out that this usage matches the pattern
in SC languages but contrasts sharply with the patterns in all the nearby
Bantu languages. These languages, he says, "combine person markers with a
connective (a-na 'he-with') for expressing having in present action con-
texts and use a form 'to be with' for other tense/aspects." Like the SVO/SOV
word order distinction, this difference in expressions of possession is typo-
logically significant; but, while Ma'a agrees with Bantu in word order, it
agrees with Cushitic in this feature.

2.4 Lexicon. Although the basic vocabulary of Ma'a is, as already mentioned,
primarily of Cushitic origin, and although much cultural vocabulary is also
Cushitic, the language has a very large number of Bantu words as well--at
least 50% of the vocabulary, according to Ehret [p.c. 1982] (the other infor-
mation in this section is from the same source). The earliest layer of Bantu
words pre-dates the Bantu-influenced grammatical changes and shows features
like Cushitic suffixes that are no longer productive in Ma'a. A later influx
of Asu words constitutes the largest set of Bantu words, and more recently
these have been supplemented, and sometimes supplanted, by Shambaa words.
Included in the vocabulary of Bantu origin are many verbs and some body parts
as well as cultural words.

     Ehret notes, however, that in spite of the large proportion of Bantu
words the lexical semantics of the SC portion of the vocabulary is still
Cushitic, at least in some lexical fields. In particular, Ma'a uses "a five-
part color division--black, white, red, yellow, green--with each expressed by
a simple adjective," while Bantu languages typically "express only black,

white, and red with simplex terms."

Table 2 summarizes the features discussed in Section 2. Examination of this table shows that Ma'a agrees with Bantu rather than Cushitic in far more features in which the typological disagreement between Bantu and Cushitic is clearly significant.

3.   How Did Ma'a Get Mixed?

The comparative discussion above substantiates the usual claim that Ma'a grammar--and most clearly the inflectional morphology, which is the grammatical subsystem usually assumed to be least susceptible to foreign interference--is mostly of Bantu origin. More to the point for the genetic question, it shows that Ma'a has few definite and productive Cushitic grammatical features, and some of the ones it does have (like the obligatory copula) do not differ much from Bantu typologically. If we set the paucity of systematic nonlexical Cushitic features against the presence of so many Bantu categories, grammatical morphemes, and even allomorphy, we must consider the possibility that Ma'a is not a changed later form of a single Cushitic parent language. Before we look at the available information about the history of the Ma'a people, we should ask what conclusions can be drawn from the linguistic evidence alone. That is, since genetic relationship of languages must in principle be established solely on the basis of the linguistic data, what are the possible routes of nongenetic development? And, does the structure of Ma'a fit any of these possibilities?

According to the model of language contact phenomena developed by Terrence Kaufman and me [1975 and Forthcoming], there are three basic lines of historical development that may culminate in a language native to a speech community whose basic vocabulary is demonstrably not from the same source as its grammar. First, speakers of a language A may shift to another language B under social conditions so pathological that only the vocabulary of B is successfully acquired. The languages most likely to have arisen through such a process are the Caribbean creoles, which emerged when enslaved Africans were forced (because they were put into linguistically diverse groups) to shift away from their various native languages, but without the opportunity

Table 2. Summary of typological agreements between Ma'a and Bantu or Cushitic

Cushitic		Ma'a		Bantu
/ɬ/	=	/ɬ/		-----
/ʔ x/	=	/ʔ x/		(/ʔ x/)
phonemic glottalics		allophonic implosion	=	allophonic implosion (Pare/Shambaa)
[ʝ ɣ]		/v ɟ ɣ/	=	/v ɟ ɣ/
-----		/ᵐb ⁿd ⁿɟ ᵑg/	=	prenasalized voiced stop phonemes
/ʕ ɦ/		-----	=	-----
labialized dorsal phonemes		-----	=	-----
ejective stop phonemes		-----	=	-----
retroflex stop phonemes		-----	=	-----
phonemic vowel length		-----	?=	-----(?)
C#		no C#	=	no C#
fem:masc; especially suffixes		Bantu N classes; prefixes	=	Bantu N classes; prefixes
optional sg:plu; singv.		oblig. sg:plu; no singv.	=	oblig. sg:plu; via prefix sets
postpositions	≠	LOC via prepositions	≠	LOC via -ni or particle -a
N-possessive pronoun	=	N possessive pronoun		N -a possessive pronoun
flexional; suffixes		agglutinative; mostly prefixes	=	agglutinative; mostly prefixes

Table 2. (continued)

Cushitic	Ma'a		Bantu
SOV	SVO	=	SVO
GEN = N N(-gen.)	(=) GEN = $\left\{ \begin{array}{l} \text{N N} \\ \text{N -a N} \end{array} \right\}$	(=)	GEN = N -a N
obligatory copula	(=) obligatory copula		optional copula
many adjectives	(=) many adjectives		few adjectives
trans. verb 'have'	= trans. verb 'have'		no trans. verb 'have'
vocabulary	= ± 50%; most basic vocabulary		vocabulary
5-part basic color system	= 5-part basic colors		3-part basic color term system

Note: (=) marks agreements in which the typological difference between Bantu and Cushitic is minor.

and/or the motivation to learn the slavemasters' language as a whole. Second,
a well-established pidgin may get nativized (i.e. become a creole). An
example is Tok Pisin (Neomelanesian), which has an English lexicon but mark-
edly non-English grammar, and which is now learned increasingly as a first
language by children in New Guinea. And third, speakers of a language A may
maintain their own language for a long period of time in the face of inter-
ference from B so great that only the vocabulary of A is successfully main-
tained, in some cases only as a special alternative (secret) vocabulary. In
such a case, we believe, all speakers of A will be bilingual, at least at one
stage in the language's history. Anglo-Romani is an example: English
gypsies all speak English, but they maintain Romani vocabulary as a secret
code. That is, their version of the language is Romani only in its lexicon;
the phonology, morphology, and syntax are all English, so that for them
Romani is, in effect, a lexical substitution code. (The Romani spoken else-
where, e.g. in Russia, is normally transmitted Romani--an Indic language--in
grammar as well as in vocabulary.) All three of these processes result in
languages that have arisen outside of normal transmission; their origins are
therefore nongenetic.

   If we oversimplify the Ma'a case for the moment, and suppose that all
its vocabulary is Cushitic and all its grammar Bantu, we can imagine five
possible origins for the language from these three lines of development,
depending on whether language A is assumed to be Bantu or Cushitic. But of
these five possibilities four can be eliminated as being in fact impossible
or implausible.

   If Ma'a were a case of unsuccessful shift from A to B, where A is
Cushitic and B is Bantu, then we would expect Ma'a vocabulary to be basically
Bantu: as Kaufman and I have argued, lexicon is acquired first in any pro-
cess of language shift. But Bantu lexical morphemes in Ma'a are not common
in the basic vocabulary--not, for instance, in a hundred-word Swadesh list--
so that Ma'a could not reasonably be supposed to be the product of an imper-
fect shift from a Cushitic to a Bantu language.

   Similarly, Ma'a could not be the product of massive interference from a
Cushitic language B in a Bantu language A, because in this case we would

again expect Ma'a basic vocabulary to be primarily of Bantu origin. English
is often said, for instance, to have a vocabulary that is mostly of French
and/or Latin origin, but the basic vocabulary of English actually has only a
scattering (ca. 7%) of Romance loanwords.

A more interesting argument might be made for interpreting Ma'a as a
result of imperfect shift from a Bantu language A to a Cushitic language B.
In this case we would expect, as we indeed find, Cushitic basic vocabulary.
Problems arise, however, when we consider the grammar, in particular the
morphology. For the most part, Ma'a inflectional morphology can be traced to
a single Bantu language, Pare. The inflectional morphology is somewhat
simpler than that of Pare, for instance in the optional use of some noun-
class prefixes, but typologically it does not differ from ordinary Bantu.
Its noun classification system is not significantly reduced, and both noun
and verb inflections even include some Bantu allomorphy, as described above.

Now, this picture corresponds to no known shift situation. The closest
analogue is perhaps the case of the Caribbean creoles, whose speakers origi-
nally acquired vocabulary from one of several B languages during a process of
shift. In those instances the shift itself is explained by the desperate
need of Africans from diverse linguistic backgrounds to communicate with one
another. Without such a pathological social situation, it is hardly con-
ceivable that any group would attempt a shift to a language that was avail-
able as a model only in the most superficial way.

Moreover, the most famous characteristic of Caribbean creole grammar--
the near-total absence of inflectional morphology--is explained by the diver-
sity of the languages that entered into the formation of the creoles. Afri-
can *pidgins* which have arisen among speakers of exclusively Bantu languages
show morphological reduction far more sweeping than the minor sorts of simpli-
fication observed in Ma'a (see e.g. Polomé [1971:58], on Katanga Swahili, and
Nida & Fehderau [1970] on Kituba). The inevitable conclusion, even if we did
not know that Pare existed, is that the elaborate Ma'a morphology must come
from a single source, or perhaps from one primary source with later influence
from a second of the same group--but not from several sources that were all
influential at the same time. So, if Ma'a represented a case of unsuccessful

shift, then we must assume that a group of speakers who shared a single Bantu
language (Pare) shifted to a Cushitic language whose availability as a model
was as severely restricted as that of the European languages in Caribbean
creole formation. The linguistic facts of Ma'a would fit such a hypothesis,
but the necessary social assumption is so improbable that it constitutes
strong evidence against the hypothesis.

The same basic argument applies to the hypothesis that Ma'a represents a
nativized pidgin. In all pidgins, the morphological structures of the parti-
cipating languages are significantly simplified, even when--as with the Bantu-
based African pidgins--all the languages involved in their formation share
general principles of morphological organization to a very large extent. Given
this salient characteristic of pidgins, an assumption that Ma'a had its origin
in a pidgin which arose from contact between one Bantu language and one
Cushitic language is untenable. (Goodman [1971:253] gives a similar argu-
ment.) Another difficulty with such a hypothesis, though maybe not an
insurmountable one in itself, is the relative rarity of pidgins in two-
language contact situations. Usually more than two languages are involved.

The fifth and last possible line of nongenetic development is the one
that most closely fits Ma'a: Ma'a arose as a product of massive interference
from a Bantu language B (Pare) in a Cushitic language A. The circumstances
that must be assumed to support this claim are by no means unique so far as
the linguistic outcome is concerned, and the hypothesized social situation,
though rare, is neither unknown in general nor dubious in this case, given
the recorded history and cultural traits of the Ma'a people.

The external history of Ma'a has been described from oral traditions by
Kimambo [1969] and Feierman [1974] in their studies of the Pare and the
Shambaa people, respectively. According to these traditions, the Ma'a people
came to the South Pare mountains about three hundred years ago, settling in
Vudee in the northwestern region of South Pare. Some time after that, appar-
ently in "an attempt to resist encroachment on their ways of living"
[Kimambo 1969:62], a large group of Ma'a moved southward to settle in the
Usambara Mountains. The Ma'a clans that remained in South Pare shifted to
Pare, but the main Ma'a group in Usambara did not shift either to Pare or to

Shambaa. One small offshoot of the Ma'a had moved to Usambara before the main migration; these people, the Nango, shifted to the Shambaa language. As Feierman [1974:77] puts it, citing several oral traditions, "the Nango chose to live the Shambaa way," but the Mbugu (Ma'a) did not.

This picture of the Ma'a as resisters of cultural assimilation is supported by comments by Copland and Green. Copland emphasizes that "the Mbugu were not disposed to sink their individuality" as some clans had done [1933-4:242], and Green remarks that "the Wambugu are a reserved and uncommunicative people regarding their past history and present customs... They keep themselves apart from the indigenous Wasambaa [Shambaa]" [1963: 175]. Green also refers to their "present independent attitude," to their "isolation from the Wasambaa," and to their "extreme conservatism" as seen in their resistance to cultural influence from Europeans as well as from the Shambaa (p. 177). This desire for isolation was no doubt reinforced by the efforts of some Shambaa chiefs, late in the nineteenth century, to enrich themselves by selling Ma'a (among other people) into slavery--efforts which forced the Ma'a to barricade themselves behind palisades "where they could farm in peace and not be enslaved" [Feierman 1974:172].

Nevertheless, the wish to preserve their cultural autonomy clearly did not keep the Ma'a from having regular contacts with their Bantu-speaking neighbors. Kimambo and Feierman report such contacts with the Pare and the Shambaa, and Ehret [p.c. 1982] points out that even now the Ma'a are neigh- bors of Asu groups who live on the edges of the Usambaras. The intimacy of these contacts is evident from the fact that all the Ma'a today are apparently fluent in both Pare and Shambaa [Ehret, p.c. 1982]. (But it is not clear that this was true fifteen years ago, because Feierman had to use a translator when he collected Ma'a oral traditions in 1968, though he was fluent in Shambaa [1974:8).

But, though continuing regular and close contact with the Asu and the Shambaa is certain, this is not the most striking sociolinguistic feature of Ma'a history. Probably the most important Bantuizing influence in the development of Ma'a, and the one that accounts for the unusual nature and degree of linguistic mixture in the modern language, is the ethnic link

between the Ma'a speakers in the Usambaras and the Pare-speaking Ma'a clans
who remained in Vudee in the South Pares. Long after the main body of the
tribe had moved to the Usambaras, tradition demanded that they return to
Vudee for the annual initiation rites. Copland, Green, Kimambo, and Feierman
all report independently that, as Kimambo puts it [1969:62], "the original
residence in Vudee remained the shrine of all the Mbugu," and that the Ma'a
of Pare and Shambaai were regularly reunited there to initiate their young
men. Feierman believes that this practice may have helped to create the
impression among the Shambaa that the Ma'a of the Usambaras were not fully
cooperative, and therefore not trustworthy, neighbors [1974:81]. Copland
was told that "the last pilgrimage from Usambara took place in 1921" [1933-4:
242], which suggests that the practice was on the decline by 1933. This is
reinforced by Green, who comments that the Ma'a of the Usambaras used always
to go to Vudee for the rites, but that "lately the rite has been performed at
Shume in the Usambaras. They still occasionally go to Vudei, four or five
hundred strong when the mood takes them" [1963:175-6].

In other words, recent use by the Ma'a of Bantu languages has never been
confined to trade and other limited communicative functions with Bantu neigh-
bors, from whom the Ma'a have deliberately kept their cultural distance. The
remaining Ma'a speakers[22] have maintained contact with their Bantu-speaking
kinsfolk in the South Pares, and in earlier years, at least, this contact
necessitated regular unrestricted communication during the annual visits to
Vudee. The clans that had remained in Pare--by implication, those Ma'a who
did not resist cultural assimilation so strongly--had shifted to the Pare
language and forgotten the Ma'a language, so that the Ma'a of the Usambaras
would have to know Pare in order to talk to them. Meanwhile, however, the
Ma'a of Pare may have kept the memory of at least some Ma'a vocabulary for
use in the rituals, much as the Copts of Egypt still use Coptic in their
religious ceremonies even though they all speak Arabic natively. That is, in

---

[22]Ehret [1980:11] says there are several thousand of them today. White-
ley says that there were about 11,000 Ma'a in the Usambaras at the time of
the 1948 census [1960:96].

the context of ethnic reunion for ritual purposes, the most salient part of
the ethnic-heritage language--the original Cushitic vocabulary--may well have
played an important role, even though the language for ordinary communication
between the two Ma'a groups must have been Pare. In any case, as far as the
pilgrim Ma'a were concerned, this same salient part of their language was
eventually the major relic of their Cushitic heritage, stubbornly maintained
in spite of the pressure to shift to Bantu in their Bantu-surrounded home-
land and the more intimate pressure to adopt Pare so that they could talk to
their kin in Pare.

The degree of resistance to total cultural assimilation that is well
attested in Ma'a is probably unusual, but other situations can be found that
are comparable, though not identical, both in their cultural and in their
linguistic results. One is the case of the English gypsies, who speak English
but also maintain Romani vocabulary, the Indic lexicon of their original
ethnic language, for use (with English grammar) as a secret code. But, while
most English gypsies speak English most of the time, Ma'a is still (as far as
I can tell from sources that are largely silent on this point) the ordinary
every-day language of its speakers. Another difference between Ma'a and
Anglo-Romani is that, as we have seen, Ma'a still has a few productive
Cushitic grammatical features, while Anglo-Romani has no Romani grammar at
all. So Anglo-Romani is an even more extreme example of grammatical replace-
ment than Ma'a.

Another language that can usefully be compared with Ma'a is the one
spoken by members of the scattered ethnic Greek communities in Asia Minor.
When they were studied by Dawkins early in this century, these communities
had been under constant cultural pressure from surrounding Turkish speakers
for hundreds of years. Many Greeks had adopted the Turkish language and
culture (including, for instance, the Moslem religion), but others had
retained both their language and their religion, along with other cultural
traditions. However, through centuries of bilingualism the Greek spoken in
Asia Minor became heavily Turkicized, though not to the extent that Ma'a has
become Bantuized: numerous Greek grammatical features, including the bulk of
the inflectional systems, remained in Asia Minor Greek [Dawkins 1916].

A third example is the spectacular case reported by Menovščikov [1969],
in which the Aleut spoken on Mednyj, one of the Commander Islands, has had
its entire elaborate finite verb morphology replaced by that of Russian,
though the original elaborate noun morphology and non-finite verb morphology
are retained. The historical interpretation of this case is difficult,
partly because both Aleut and Russian speakers on Mednyj were apparently
bilingual to some extent in each other's language during the relevant period
(see Thomason [1981] for discussion), but it is clear that Aleut was partially
maintained in the face of very strong pressure from Russian.

Of these three cases, the one that most closely resembles Ma'a in its
developmental characteristics is Asia Minor Greek. To the extent that Anglo-
Romani is spoken natively at all, it has become a native language (in some
gypsy families) through expansion in function of the secret-code jargon con-
sisting of Romani words and English grammar. This has taken place since the
shift to English by English gypsies, so that it is probably best viewed as a
re-emergence of first-language learning of Romani vocabulary--that is, after
a break in transmission, and after a period in which the first-learned
language in the community was English. The case of Mednyj Aleut is more
complicated, but here too it seems fairly likely that the transmission pro-
cess was abnormal, because second-language learners, mainly Russian husbands
of Aleut wives, must have participated in the transmission of the language to
children born into the culturally and linguistically diverse community.

But Asia Minor Greek, like Ma'a, developed in communities which, though
largely bilingual, were relatively homogeneous culturally. I see no room for
doubt in either case about the existence of cultural and ethnic continuity
from a period pre-dating Turkish and Bantu influence, respectively. This con-
tinuity includes normal transmission of the community's language, in that
(unlike Anglo-Romani and Mednyj Aleut) there is no evidence that would force,
or permit, us to infer a break in the transmission of an entire language at
any period in the language's history. As far as we can tell, the ancestors of
the current Ma'a speakers did not shift to Bantu while keeping their original
Cushitic lexicon, and they did not experience the disruption of mixed Bantu-
Ma'a households that might have prevented children from learning Ma'a as a

whole language. So to this extent, at least, Welmers is right in claiming
that the development of the language is "within the familiar framework of
continuous language history" [1973:8].

Nevertheless, in one respect Ma'a resembles Mednyj Aleut and especially
Anglo-Romani rather than Asia Minor Greek, and this is the crucial point for
the question of genetic relationship: it is not possible to show for Ma'a,
any more than for Anglo-Romani or Mednyj Aleut, systematic form/function
correspondences in all grammatical subsystems. The history of Ma'a is similar
in kind to the history of Asia Minor Greek, but the amount of foreign inter-
ference is far greater in the Ma'a case. Therefore, although the Bantu inter-
ference features presumably accumulated gradually in Ma'a, so that the
difference in the language of any two adjacent generations was minor, the
net effect is a language whose grammatical morphemes are almost entirely of
Bantu origin (if we assume, as many linguists do, that derivational affixes
belong to the lexicon), and whose definitely Cushitic phonological and syn-
tactic features are also outweighed by the Bantu features.

If we consider only the linguistic structures of modern Ma'a, I do not
believe that a convincing case can be made for treating the language as a
changed later form of Proto-Southern Cushitic: there are too many Bantu sub-
stitutions in the grammar, and too few remaining systematic Cushitic gramma-
tical features. If this is true, then Ma'a cannot be said to be genetically
related to Cushitic languages, unless we reformulate the notion of genetic
relationship to fix on the vocabulary as the sole criterion for establishing
relationship. But then genetic relationship would cease to be a historical
concept and become merely a synchronic taxonomic one. Probably the most
significant consequence of such an approach would be that comparative recon-
struction could no longer be claimed to be a guess at the structure of a real
language that was spoken in the past, and that consequence would surely be
unacceptable to most historical linguists. Ma'a morphological data certainly
cannot be treated, for purposes of reconstruction, as gradually modified
Proto-SC, and Ma'a phonology is also of limited value for reconstructing
Proto-SC phonology, given the highly irregular effects of contact-induced

changes in phonological shapes of words.[23]

Just as we could infer from the structures of the least decreolized Caribbean creoles that they arose through imperfect shift or as nativized pidgins, we can infer from the structure of Ma'a, as I argued at the beginning of this section, that it arose in a long-term situation of language (and culture) maintenance under conditions of intense cultural pressure from Bantu. We can therefore establish that Ma'a arose from a Cushitic language. But since Ma'a as a whole is no longer an appropriate object for comparative reconstruction with Cushitic, the integrity of genetic relationship as a claim about gradual linguistic divergence over time is best preserved by putting Ma'a, along with languages like Anglo-Romani, Taki-Taki, and Tok Pisin, outside the genetic model.

---

[23]This does not, of course, mean that Ma'a is irrelevant for the reconstruction of Proto-SC. But its usefulness is similar to the usefulness of borrowed elements in neighboring languages, e.g. Finnish borrowings from early Germanic: the Ma'a data must, in most instances (/ɬ/ is an exception), be considered in light of Bantu phonology as well as Cushitic in order to arrive at a reasonable interpretation of the relevance to Proto-SC reconstruction.

REFERENCES

Bender, M. L. (ed.). 1976. *The Non-Semitic Languages of Ethiopia*. East
     Lansing, MI: African Studies Centre.

Bender, M. L., E. Mulugeta and D. L. Stinson. 1976. "Two Cushitic lan-
     guages." In M. L. Bender, J. D. Bowen, R. L. Cooper and C. A.
     Ferguson (eds.), *Language in Ethiopia*, pp. 130-54. London: Oxford
     University Press.

Black, P. 1976. "Werizoid." In Bender [1976], pp. 222-31.

Bliese, L. 1976. "Afar." In Bender [1976], pp. 133-65.

Copland, B. D. 1933-34. "A note on the origin of the Mbugu with a text."
     Zeitschrift für Eingeborenen-Sprachen 24:241-5.

Dawkins, R. M. 1916. *Modern Greek in Asia Minor*. Cambridge: University
     Press.

Dolgopolskij, A. B. 1973. *Sravnitel'no-istoričeskaja Fonetika Kušitskix
     Jazykov*. Moscow: Nauka.

Ehret, C. 1974. *Ethiopians and East Africans: The Problem of Contacts*.
     Nairobi: East African Publishing House.

Ehret, C. 1976. "Cushitic prehistory." In Bender [1976], pp. 85-96.

Ehret, C. 1980. *The Historical Reconstruction of Southern Cushitic Pho-
     nology and Vocabulary* (Kölner Beiträge zur Afrikanistik, 5). Ber-
     lin: Reimer.

Ehret, C. and D. Nurse. 1981. "The Taita Cushites." *Sprache and Ge-
     schichte in Afrika* 3:125-68.

Elderkin, E. D. 1976. "Southern Cushitic." In Bender [1976], pp. 278-
     97.

Feierman, S. 1974. *The Shambaa Kingdom: A History*. Madison, WI: Uni-
     versity of Wisconson Press.

Gleason, H. A., Jr. 1955. *Workbook in Descriptive Linguistics*. New York:
     Holt, Rinehart & Winston.

Goodman, M. 1971. "The strange case of Mbugu." In D. Hymes (ed.),
     Pidginization and Creolization of Languages, pp. 243-54. Cam-
     bridge: University Press.

Gragg, G. 1976. "Oromo of Wellega." In Bender [1976], pp. 166-95.

Green, E. C. 1963. "The Wambugu of Usambara (with notes on Kimbugu)." *Tanganyika Notes and Records* 60-61:175-89.

Greenberg, J. H. 1955. *Studies in African Linguistic Classification*. New Haven.

Greenberg, J. H. 1963. *The Languages of Africa*. Bloomington: Indiana University Press.

Greenberg, J. H. 1966. "Some universals of grammar with particular reference to the order of meaningful elements." In J. H. Greenberg (ed.), *Universals of Language*, pp. 73-113. Cambridge, MA: MIT Press.

Hall, R. A., Jr. 1958. "Creolized languages and genetic relationships." *Word* 14:367-73.

Hudson, G. 1976. "Highland East Cushitic." In Bender [1976], pp. 232-77.

Hudson, R. A. 1974. "A structural sketch of Beja." *African Language Studies* 15:111-42.

Hudson, R. A. 1976. "Beja." In Bender [1976], pp. 97-132.

Kimambo, I. N. 1969. *A Political History of the Pare of Tanzania, c. 1500-1900*. Nairobi: East African Publishing House.

Kimenyi, A. 1978. "Grammatical tone neutralization in Kinyarwanda." *Studies in African Linguistics* 9:303-17.

Loogman, A. 1965. *Swahili Grammar and Syntax*. Pittsburgh: Duquesne University Press.

Meillet, A. 1921 (1958). "Le problème de la parenté des langues." Reprinted in his *Linguistique Historique et Linguistique Générale*, pp. 76-101. Paris: Champion.

Meinhof, C. 1906. "Linguistische Studien in Ostafrika. X. Mbugu." *Mittheilungen des Seminars für Orientalischen Sprachen* 9(3):293-323.

Menovščikov, G. A. 1969. "O nekotoryx social'nyx aspektax èvoljucii jazyka." In *Voprosy Social'noj Lingvistiki*, pp. 110-34. Leningrad: Nauka.

Nida, E. and H. W. Fehderau. 1970. "Indigenous pidgins and koinés." *IJAL* 36:146-55.

Polomé, E. 1971. "The Katanga (Lubumbashi Swahili Creole." In D. Hymes (ed.), *Pidginization and Creolization of Languages*, pp. 57-59. Cambridge: University Press.

Sasse, H.-J. 1976. "Dasenech." In Bender [1976], pp. 196-221.

Thomason, S. G. 1980. "Continuity of transmission and genetic relationship." In E. C. Traugott, et al. (eds.), *Papers from the Fourth International Conference on Historical Linguistics*, pp. 27-35. Amsterdam: Benjamins.

Thomason, S. G. 1981. "Are there linguistic prerequisites for contact-induced language change?" Paper presented at the Tenth Annual University of Wisconsin-Milwaukee Linguistics Symposium: Language Contact.

Thomason, S. G. and T. Kaufman. 1975. "Toward an adequate definition of creolization." Paper presented at the 1975 International Conference on Pidgins and Creoles, Honolulu.

Thomason, S. G. and T. Kaufman. Forthcoming. *Language Contact, Creolization, and Genetic Linguistics*.

Tucker, A. N. and M. A. Bryan. 1966. *Linguistic Analyses: The Non-Bantu Languages of North-Eastern Africa. Handbook of African Languages*. London: Oxford University Press.

Tucker, A. N. and M. A. Bryan. 1974. "The 'Mbugu' anomoly." *BSOAS* 37:188-207.

Welmers, W. E. 1952. "Notes on the structure of Saho." *Word* 8:145-63; 236-51.

Welmers, W. E. 1973. *African Language Structures*. Berkeley: University of California Press.

Whinnom, K. 1971. "Linguistic hybridization and the 'special case' of pidgins and creoles." In D. Hymes (ed.), *Pidginization and creolization of Languages*, pp. 91-115. Cambridge: University Press.

Whiteley, W. H. 1960a. "Linguistic hybrids." *African Studies* 19(2): 95-7.

Whiteley, W. H. 1960b. "The verbal radical in Iraqw." *African Language Studies* 1:79-95.

Zaborski, A. 1976. "Cushitic overview." In Bender [1976], pp. 67-84.

Department of Linguistics

and

African Studies Center   announces
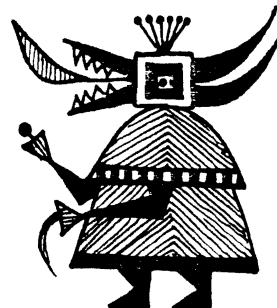
The 15th

Conference on

AFRICAN LINGUISTICS

March 29–31, 1984

at

U.C.L.A.

Please fill out and send to: The 15th Annual Conference on African Linguistics
                              Department of Linguistics, UCLA, Los Angeles, CA 90024

Send further announcements____   I plan to attend_____
I would like to present a paper tentatively entitled_____
    (Abstracts of between 500-800 words due on January 1, 1984)

I would like to participate in a "Working Group" on _____
    (e.g. tone, serial verbs, language policy, Bantu, Chadic, etc.)

Name _____   Address_____