Swahili and Sabaki
A Linguistic History

Derek Nurse and
Thomas J. Hinnebusch

Edited by Thomas J. Hinnebusch
With a special addendum by Gérard Philippson

UNIVERSITY OF CALIFORNIA PRESS
Berkeley • Los Angeles • London

UNIVERSITY OF CALIFORNIA PUBLICATIONS IN LINGUISTICS

Editorial Board: Leanne Hinton, Pamela Munro,
Johanna Nichols,William Shipley, Sandra Thompson

Volume 121

UNIVERSITY OF CALIFORNIA PRESS
BERKELEY AND LOS ANGELES, CALIFORNIA

UNIVERSITY OF CALIFORNIA PRESS, LTD.
LONDON, ENGLAND

Library of Congress Cataloging-in-Publication Data

Nurse, Derek.
Swahili and Sabaki: a linguistic history / Derek Nurse and Thomas J. Hinnebusch; edited by Thomas J. Hinnebusch; with a special addendum by Gérard Philippson.
p. cm. — (University of California publications in linguistics: v. 121)
Includes bibliographic references and index.
ISBN 0-520-09775-0 (alk. paper)
1. Swahili language—History. 2. Bantu languages—History. I. Hinnebusch, Thomas J. II. Philippson, Gérard. III. Title. IV. Series.
PL8701.N87 1993
496'.392—dc20 93-4560
CIP

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984. ♾

2

A COMPARATIVE HISTORICAL PHONOLOGY OF SABAKI

§1.0. Introduction

The objective of this chapter is twofold. First, we outline the basic features of the consonant and vowel systems as we have reconstructed them for Proto-Sabaki (PSA) and, where necessary, set out our justification for the reconstruction. In some cases this will be a fairly straightforward, noncontroversial task, obvious to Bantuists who are familiar with Meinhof's work (1899-1932) and Guthrie's Common Bantu (CB) (1967-71). In other cases, the proposed reconstruction requires outlining the evidence and reasoning behind our choice, and sometimes detailing the tentative aspects of the analysis. The data upon which this analysis is based are largely represented in the PSA reconstructed lexis in Appendix 2.

Second, we reconstruct the historical sound changes which are necessary to understand the evolution of the individual Sabaki phonological systems, specifically how they came to develop the systems they exhibit today. The goal is not, however, to specify all the details of the present synchronic systems, but only those areas, as we understand them, that shed light on Sabaki's development into a rich and diverse dialect cluster, as well as those features of the system that reveal the group's internal phonological relationships. In so doing we develop a body of systematic evidence useful for subgrouping.

In working toward these goals, we deal with each set of reconstructed proto segments, beginning with stops, followed by continuants, nasals, nasal-consonant sequences, vowels, and other aspects of the system. A section is devoted to each set of segments, with subsections for each of the two goals: how Guthrie's CB is to be interpreted vis-à-vis the Sabaki dialects, by focusing on the justification for our reconstruction, followed by a subsection on post-PSA developments.

§1.1. Methodological comments. Part of Meinhof's and Guthrie's approach to CB involved stating the relationship between a reconstructed form and its reflex in extant languages. Thus in Guthrie's *Comparative Bantu* (1967-71) we find tables of correspondence relationships for many individual Bantu languages, e.g., Swahili CB *p : p, *t : t, *k : k, etc. This is partially the approach taken here, as can be seen in the various tables throughout this chapter. As important as this information is, it is one-dimensional and

provides neither an understanding of how the various components of such statements relate internally and systematically within one language nor an interrelated systematic view across languages. Thus, we are also working toward several other goals: (a) to show how correspondence relationships, as phonological processes, in one language or dialect relate to those in another within Sabaki; (b) to distinguish pre-PSA from PSA events; and (c) to identify major post-PSA events. The correspondence relationships that we find in the Sabaki languages have various sources. Some are inherited from Northeast Coast Bantu (NEC) and East Highlands Bantu (see Chap. 1, §2.3 - 3.1); others stem from a common PSA period when the proto community was still a compact unity, or at least confined within a geographical space which allowed for intimate contact and mutual linguistic influence among its various subcomponents. Other phonological events occurred after the fragmentation and dispersal of the proto community. These also are of interest, because they contribute to defining subgroups within Sabaki.

Working out the chronology of these events is an important component of this chapter. By working out the details, we have a means to help us decide at what point phonological events occurred. If a shared process can be shown to be earlier than another we can then make distinctions between retention and innovation, important for subgrouping.

In reconstructing the synchronic Sabaki situation, we want to attempt an understanding both of the diachronic-synchronic relationships themselves—that is, what processes underlie such correspondences—and of how the various processes interrelate through space and time. This is not always an easy task, and at best what we can hope to accomplish at times is a statement of a hypothesis which can form the basis for future discussion, study, and investigation. We begin with correspondences, that is, where Guthrie left off, e.g., CB *p : Swahili /p/ : Pokomo /ɸ/ : Giryama /h/ : Digo /β/ : Vumba /β/ : Makunduchi /v/ : Ngazija /β/ : Nzuani /v/, etc.

There are two points to be made about this list of correspondences. First, Swahili is conservative in retaining /p/. We consider this apparent, given our confidence in the reconstruction of Bantu *p, a labial stop, and given what we know about the distribution of the reflexes of *p both within and outside Sabaki. Second, where change is attested, we see that the reflexes of *p are all weakened variants. Accepting this as a phonetic/phonological fact, we can go a step further and make the claim that this shared feature is not an accident. It is an attested feature of a set of geographically contiguous languages which shares a range of other linguistic features. Within this context, then, we can confidently assume that these languages share this feature because they are related, and that lenition itself is a common shared historical event. Whether lenition is a feature that spread to Sabaki from language groups outside Sabaki, or is due to an internal innovation, is not at issue at this point.

Furthermore, the reflexes, as diverse as they may seem at first glance, can be shown to be related in terms of a series of connected shifts (i.e., a seriation). Thus in the closely related Comorian dialects it is easy to see the connection between /β/ and /v/, where one is the source of the other. But in Mijikenda, another set of mutually intelligible dialects, it is

less easy to see the connection between Giryama /h/ and Digo /β/, unless we view the total array of reflexes in Sabaki as the result of a connected series of articulatory shifts. When we observe /ɸ/, the reflex in Pokomo, we have the link for a series of changes which begin to make some sense of both the phonetics and the distribution of the reflexes. Thus, we can hypothesize that the change *p > [+lenition] is a stage we could claim for much of Sabaki. If we assume that one result was a voiceless continuant, identical or similar to /ɸ/, the next set of reflexes is predictable in terms of /ɸ/. This segment was then further weakened, losing features which identified it as both a labial and an obstruent in becoming Giryama /h/, a result consistent with lenition. Digo, along with several other Sabaki dialects, moved in a slightly different direction, voicing /ɸ/ to yield /β/. We can thus posit for Mijikenda, Pokomo, and Comorian an intermediate innovation, and a two-part seriation, *p > ɸ > h and *p > ɸ > β > v (see *p-lenition, §4.1.1.1).

Such a seriation is analytically much more satisfying than the simple listing of correspondences. It provides an underlying uniform account for surface diversity, shows the potential historical relatedness of otherwise unrelated facts, and constitutes evidence for subgrouping where otherwise none would exist. This analysis of *p-lenition assumes that the seriations *p > ɸ > h and *p > ɸ > β > v make more sense phonetically than for *p to directly and independently become /h/ or /β/ or /v/. Those familiar with dialect geography will recognize the application of seriation here (Anttila 1972:294ff.; also see "telescoping" in Hyman 1975:173-175). As applied here, it attempts to relate apparently autonomous events, based on what we know of change in general, of phonetically motivated change in particular, and of the possible kinds of change that specific segments can undergo. Seriation plays a role in demonstrating that different dialects potentially share parts (intermediate stages) of specific processes with other dialects, which otherwise would not be apparent by only looking at autonomous correspondences. Thus a subset of Sabaki shares the intermediate shift of *p > /ɸ /. Seriation allows us the possibility of looking at a series of events in specific dialects as potential manifestations of the same shared phonetic event.

While seriation is one dimension of the methodology applied in this chapter, stratigraphy (Anttila 1972) is another. This is concerned with the spatial distribution of correspondences, and what might be inferred from such information about both earlier proto forms and their chronological development into the modern forms. Reflexes with the widest distribution are often the oldest, and the most localized the most recent; centrally located versus peripherally located reflexes can also provide help in developing the chronology of a series. For example, the co-occurrence of nasal-devoicing and post-consonantal aspiration ($\underset{\circ}{N}\underset{\circ}{C}^h$ < *$N\underset{\circ}{C}$) in northern and southern peripheral areas of NEC points to an earlier state of affairs underlying the evolution of aspirated voiceless stops ($\underset{\circ}{C}^h$ < $\underset{\circ}{N}\underset{\circ}{C}^h$ < *$N\underset{\circ}{C}$) in the more centrally located Swahili and Mijikenda languages. Seriation predicts the same thing: *$N\underset{\circ}{C}$ > $\underset{\circ}{N}\underset{\circ}{C}^h$ > $\underset{\circ}{C}^h$ is easier to understand phonetically than the development *$N\underset{\circ}{C}$ > $\underset{\circ}{C}^h$. In this example, phonetics (seriation) and distribution (stratigraphy) coincide, allowing a strong claim that Swahili and Mijikenda went through an intermediate stage; and on that presumption we can claim historical interrelatedness, if not by virtue of

sharing the middle stage precisely as stated, at least as an areal drift-like event leading to aspirated voiceless stops. In contrast, elsewhere in Bantu, we do not find this interesting congruence. This adds further weight to the argument that events in Swahili and Mijikenda are related to events in other geographically contiguous areas. (See Möhlig 1981, etc. for another view of stratigraphy.)

While seriation and stratigraphy yield insights on the nature of proto forms, intermediate stages, and the evolutionary development of modern forms, another dimension is gained by considering the intrinsic ordering relationships of sets of processes. In the case of a number of changes, e.g., *A > Z and *B > Y, stratigraphy and rule-ordering together can help in establishing a chronology. Where both give congruent results, we have independent confirmation of a seriation of changes and stages of change. For example, Bantu *p, where not protected by a preceding nasal, undergoes lenition in a large number of East Highlands Bantu languages (/ɸ/, /h/, /β/, Ø, etc. have all been attested); but *p before the Bantu high vowels *į and *ų spirantizes. Typical examples, which mirror the situation in many Sabaki and NEC languages, would include the following somewhat hypothetical, but representative, forms: *mpepo 'wind, spirit' > mpeho, *mulapį 'glutton' > mulafi. The only possible ordering here, without fudging the rules and using exception features and similar formal devices, is (a) Spirantization (*p > f/___į), followed by (b) Lenition (*p > ɸ/[-stop]___) to correctly derive the proper reflexes. This is illustrated in the derivation on the left; the other ordering gives the incorrect output (Rule a = Spirantization, Rule b = Lenition, Output includes the operation of other rules):

*N-pepo	*mu-lapį		*N-pepo	*mu-lapį	
-----	f	Rule a	h	h	Rule b
h	-----	Rule b	-----	-----	Rule a
mpeho	mulafi	Output	mpeho	*mulahi	Output

This derivational order very likely matches the historical chronology. Other evidence comes from stratigraphy: *p-spirantization is an earlier process than *p-lenition, as indicated by the fact that *p-lenition is not as widely distributed; furthermore, *p-lenition is not uniform in its results (*p : ɸ : β : v : h : Ø) within Sabaki. Moreover, stratigraphy suggests that spirantization is more likely an inherited feature within Sabaki (however, see §5.1.1-5.1.5), whereas the PSA continuum was already starting to diversify when *p-lenition either spread areally into PSA or innovated within the Sabaki geographical area. Of course, not all the processes of change discussed in this chapter lend themselves to this neat sort of congruence.

§1.1.1. Assumptions. The application of seriation, stratigraphy, and rule-ordering in this chapter assumes what we outlined earlier in Chapter 1 (§5). We are dealing with a historical proto language or finely knit proto dialect cluster whose speakers interacted in various ways and in various degrees over a period of time with each other and with non-Sabaki Bantu and non-Bantu groups. We recognize several sources and dynamics of change, including, but not exclusively, perseverance and retention of older forms, conver-

gence (independent development of similar features), shared innovation, and intimate borrowing of lexis, phonology, and other linguistic structures within and from outside the group. Thus part of what we try to do, where possible, is to discriminate between various sources of change and make claims for sound changes accordingly. This is important for classifying subgroups which share a common proto period, or period of mutual interaction that has led to shared development. On this basis, we only want to identify groups and subgroups that share specific innovations or common areal features. Convergence, and the retention of common features, are less helpful, and do not lead to historically verifiable subgroups. It has usually been easy to distinguish between retention and innovation in the course of this analysis, especially where we can rely on either Guthrie's or Meinhof's analysis of (Common) Bantu, or are confident in our own reconstruction of pre-Sabaki proto stages. If we are sure of our reconstruction of *A for a particular set of reflexes, then if L(anguage)1, L2, and L3 share the correspondence *A : Z, whereas L4 and L5 still attest a direct reflex of *A (*A : A), and we are furthermore confident of the direction of change, then we can assume that L1, L2, and L3 form a subgroup by sharing the innovation of *A > Z.

Where identical developments can be demonstrated for a close-knit group of dialects, it is assumed that they are due to common innovation when they are, for the most part, restricted to the group. Innovations that are attested in a great many units in the group, as well as outside it, may well be evidence of a larger proto group.

§1.1.2. Methodological difficulties. It has been a working assumption of this chapter that where two languages share identical correspondences, they do so for historical reasons. This is an easy assumption to make because we are dealing with a contiguous group of languages. An example is provided again by *p-lenition. Some Mijikenda languages, e.g., Digo and Duruma, share the correspondence *p : β, and do so presumably because they share a common proto period and thus a common development of *p ultimately to /β/. A few Swahili dialects (Chifundi, Vumba, and a variant of Makunduchi) attest the same correspondence, *p : β. As a working hypothesis, we assume that this is evidence for a common Mijikenda/Southern Swahili proto period. This has been the simplest procedure to follow in similar cases, but we recognize that this assumption may not always be justified, nor supported by other evidence. In this instance, because on the one hand we are dealing with Mijikenda and on the other with Swahili, both of which include members other than Digo/Duruma or Vumba/Chifundi/Makunduchi respectively, it may well be that the correspondences are not genetically shared once all the facts are considered. They may be due to independent innovation, or an areal effect spread by fairly recent borrowing, in this case from Mijikenda into Swahili (see §15.14). Here we have no simple heuristic device to decide between potential sources of identity: independent innovation, genetics, areal spread, borrowing, or convergence. In similar cases throughout this chapter, in the absence of clear evidence or arguments, we push the methodological assumption to the extreme and posit genetic sources for such identity. Our intention here is to develop a body

of evidence within a consistently applied methodology. We leave it to other scholars to adduce other hypotheses and evidence (for further specific discussion of Vumba, see Chap. 5, §13.5 and §13.6).

Another problem concerns seriation, and how it has been applied here. We have usually assumed that the relationship between a series of changes is a simple matter of articulatory change. This is the simplest methodological approach to the data. So we have presented *p-lenition within Sabaki as a series of successive lenitions involving the change of single articulatory features. Other analyses are possible. That the voiceless labial stop is acoustically less distinct than other voiceless stops allows for nonserialized changes that may be perceptually or acoustically motivated (see Maddieson 1984:36-37). Thus /β/ in Digo may have no serialized connection to /ɸ/ in Pokomo. We cannot always be sure of the exact genesis or evolution of change in specific cases. Even so, subgrouping evidence can often still be adduced. In our example, we can still distinguish a group of languages which share a connected drift-like event (Sapir 1921:150, 171) from a group which does not. In this sense, lenition in Pokomo, Mijikenda, and Comorian is a commonly shared event, whereas specific stages in a seriation model might not be.

Just as seriation does not always yield clean results, there are shortcomings connected with stratigraphy. It is clear, for example, that the changes of the stops to fricatives before the high vowels must be an early change in Bantu, and that the lenition of *t (e.g., *muti 'tree' > muri (Vu); see §4.1.2) in parts of Sabaki is more recent. The former is uniformly attested over a wide area and in a large number of languages, while *t-lenition is mostly restricted to some Sabaki dialects and only sporadically attested in other groups. Stratigraphy cannot fail to give results in such obvious cases. However, there are instances where distribution cannot be so easily quantified. What is statistically significant is not an easily answered question, so stratigraphy, at best, is a gross sort of indication of chronology, but in some cases it is all we can rely on to make some assessment of chronology.

Finally, rule-ordering works well where we have bleeding and feeding relationships between rules; e.g., *p-spirantization bleeding *p-lenition. Otherwise it is often not possible to determine ordering. For instance, *p-lenition in Sabaki is obviously independent of the harmony rule which predicts the height of suffixal vowels in Bantu verbal inflectional morphology (/-Vl-, -Vk-, -Vš-, etc., where V is determined by the V of the root or stem, e.g., Sw -nunu-li-a 'buy for' and -som-e-a 'read for').

There are other problems of a general nature which are faced by comparative/ historical linguists who work with unwritten languages. It is not always possible to know when we are dealing with a shared innovation from a common proto period, or with a feature that is due to convergence, borrowing, or shared retention, or in the case of Sabaki to distinguish between pre, and post, PSA events, even by appealing to our methodological heuristics. Many Bantu languages share similar phonological and morphological systems, and while many different types of change are theoretically possible, in practice the types of documented change are limited. Thus the chances are reasonable that two not particularly closely related Bantu languages might share the same sort of innovation, e.g., the dissimi-

lation we find in Ngazija (Comorian), where *NÇVNÇV > NÇVØÇV (e.g., *nkondo 'war' > /nkodo/ [nkoɗo]) is similar to a process found in Bantu languages in southwestern Africa (see Kwanyama Law, §8.1.3). As uncommon as this change happens to be in Bantu, as compared to Meinhof's Law (§7.1.1), the change is naturally motivated, and due to the commonly shared Bantu morpheme structure, namely *NÇVNÇV sequences. Thus we do not want to assume genetic association between Comorian and southwestern Bantu, other than common Bantu-ness, on the basis of one change. In this case we cannot appeal to stratigraphy, given the great distance separating the two areas and, more important, the lack of other shared innovations specific to the languages in question. Finally, single shared innovations are a poor basis for subgrouping; a better measure is shared sets of phonological or morphological innovations.

§2.0. The Sabaki Languages

The dialect clusters of Sabaki (see Chap. 1, §2.2) considered specifically in this chapter are Elwana (El), Pokomo (Po), Mijikenda (MK), and Swahili (Sw)—with emphasis on Unguja (Ung), Comorian (Com), and Mwani (Mn). The different dialect clusters or dialects within these are often abbreviated in the rest of the book thus:

UP	Upper Pokomo
LP	Lower Pokomo
ND or ND_1	Northern dialects of Swahili (Mwiini, Tikuu, Siu, Pate, Amu, Malindi/Mambrui, and the Mombasa area dialects)
ND_2	Mwiini, Tikuu, Siu, Pate, and Amu
ND_3	Tikuu, Siu, Pate, and Amu
SD	Southern dialects of Swahili
NMK	Northern Mijikenda
SMK	Southern Mijikenda

For now these labels are mainly used as terms of reference, but it will become increasingly clear that they also have historical and genetic significance.

Other abbreviations used in tables for individual languages/dialects are:

Am	Amu	Maf	Mafia	Ng	Ngazija
Chi	Chifundi	Mak	Makunduchi	Nz	Nzuani
Ch	Chonyi	Ma	Maore	Pa	Pate
Di	Digo	Mt	Mtang'ata	Pe	Pemba
Du	Duruma	Mv	Mvita	Ra	Rabai
Gi	Giryama	Mh	Mwali	Si	Siu
Ji	Jibana	Mn	Mwani	Ti	Tikuu
Jo	Jomvu	Mw	Mwiini	Tu	Tumbatu
Ka	Kauma	Nga	Ngare	Vu	Vumba

§3.0. An Overview of the Phonological System

We have reconstructed the following consonant and vowel system for PSA. The system itself is largely nondistinct, with the exception of some details, from the system that would be reconstructed for Proto-Northeast Coast Bantu (PNEC). In the following chart, *W is a labial approximant; segments in parentheses are attested in very few cases in the set of reconstructed lexis (App. 2); problematic reconstructions are indicated with a question-mark. Nasal-consonant + stop (*NC) combinations are considered to be prenasalized stops if voiced. Some *NC derive from m+W, n+l, etc., sequences. Vowels may be long.

Consonants

p	t	c	k
W (ɓ)	l (ɗ)	j	g
f	s	š	
v	z		
mp	nt	nc (= ɲc)	nk (= ŋk)
mb	nd	nj (= ɲj)	ng (= ŋg)
m	n	ny (= ɲ)	(ŋ ?)
w	(r ?)	y	

Vowels

į	ų
i	u
e	o
a	

Consonant-glide (CG) sequences are well represented. Nearly a full set of consonants followed by *w and *y was reconstructed (e.g., *pw, *tw, *py, *fw, etc.), apart from those created by suffixing the Passive *-w- or the Causative *-y-. Various nasal + consonant sequences (e.g., *mf, *mv, *ns, etc.) and nasal + consonant + glide clusters (e.g., *mpw, *ntw, *nsw, etc.) also exist.

Though all extant Sabaki languages, with the exception of Elwana, have five vowels, a PSA seven-vowel system has been reconstructed: *į, *i, *e, *a, *o, *u, and *ų (§5.4, §10 and Chap. 5, §12.2). Long vowels are also reconstructible (see §10). For lack of data and analysis, we have been forced to largely ignore tone. The little work that has been done on tone in Sabaki languages (e.g., Tucker and Bryan 1970, Kisseberth 1984, Philippson 1986, 1988) would indicate a proto language with a highly reduced system (but see Philippson's discussion below in §16).

A good deal of the PSA phonological system is inherited virtually unchanged from more remote proto periods, such as the Proto-East Highland group (Heine 1973) and as yet unidentified intervening groups. The PSA nasals, prenasalized consonants (*mp, *mb, etc.), and the glides *w and *y are unchanged from earlier proto periods. Other consonant sets are inherited, as we assume is the case for the voiceless stop series *p, *t, and *k. The voiced series *W, *l, and *g, a reconstruction which diverges from the more familiar formulations of *b, *d, and *g (Guthrie's CB, Meeussen 1967) and *β, *l, and *ɣ (Meinhof 1932), will require discussion for justification (§4.2). The PSA palatals *c, and *j, though appearing to be identical to CB *c and *j, and thus inherited, are also the focus of special discussion (§4.1.4, 4.2.4) for reasons having to do with the distribution of the

reflexes of CB *c and CB *j in Sabaki and surrounding languages. The fricatives *f, *v, *s, *z, and *š result from Bantu Spirantization, a process whose effects are widely attested in eastern Bantu. Here, again, appearances are deceiving, and evidence is considered that Sabaki Spirantization is due to a complex of inherited and internal Sabaki events (§5). PSA's seven-vowel system is an inherited feature, as are the long vowels. In the following, each section is normally organized to discuss (a) the development of PSA from an earlier proto system, (b) the PSA system, its explication and rationale, and (c) the development of this system within Sabaki.

In this chapter, PSA forms are marked with an asterisk, e.g., *p. Guthrie's forms are also asterisked, but are preceded by the code CB (Common Bantu), e.g., CB *p.

§4.0. Proto-Sabaki *p, *t, *c, *k; *W, *l, *j, *g

§4.1. Voiceless *p, *t, *c, *k. In non-prenasalized position, we have reconstructed [-continuant, -nasal] obstruents. Examples are given for stem-initial and intervocalic positions:

*p	*-pola 'become cool'	*-lapa 'swear'
*t	*-tokota 'boil'	*-ɨkuta' be satisfied'
*k	*-kali 'sharp, etc.'	*-fɨka 'arrive'
*c	*-canga 'assemble'	*jɨjɨco 'eye'

For the specific reflexes of these items, see App. 2.

This analysis is supported by both internal and comparative evidence. Although Comorian, Pokomo, and Mijikenda show forms of lenition affecting *p, *t, and *k in intervocalic environments, these are all demonstrably late PSA developments (discussed in §4.1.1.1). In most cases the nonstopped variants alternate synchronically with stopped variants in postnasal position in individual Sabaki dialects. However, we reconstruct for PSA a system in which there is no alternation, and argue that various lenition rules created the synchronic alternation that we see today, roughly around the time of the break-up of the original proto community. To hypothesize any other system involves much more complicated assumptions and rule sets than are warranted by existing evidence. Although synchronic grammars have gone through restructuring where the lenis variants are now phonemic, and the stops are predictable, there is no evidence for reconstructing such a system for PSA. This is obvious in the case of *t and *k, which have undergone lenition in only part of the dialect continuum, and where obstruents are well attested comparatively. For example, *t-lenition (e.g., *t > r ~ h, etc.) occurs in Mijikenda, Pokomo, Comorian, and a few Swahili dialects, but not in the majority of Swahili dialects; given its restricted distribution within Sabaki, and in NEC generally, it would appear to be a late PSA innovation. The lenition of *k is even more restricted, found only consistently in Comorian. It is clearly not a PSA development, although it is ultimately due to a generalization of the historical Sabaki lenition. On the other hand, *p-lenition is widespread throughout Sabaki, with the exception of Elwana, Mwani, and most Swahili dialects. It is also found widely in

NEC and East Highlands Bantu, so there is the very real possibility that PSA inherited a lenis voiceless labial. In the following subsections we examine the evidence for reconstructing *p, *t, and *k for PSA.

§4.1.1. Proto-Sabaki *p. PSA *p has weakened in Sabaki, with the exception of Elwana, Mwani, and most Swahili dialects, which are conservative in their preservation of stops. In postnasal position, and in some dialects following *i̧, we find a stop, e.g., Ngazija /-piha/ 'cook' (< *-i̧pika). The following reflexes are the norm in intervocalic environments (see Chart 1; other examples are given in App. 2):

CB *p : PSA *p :

p	El, Mw, ND, SD, Mn	ɸ	Po
β	Di, Du, Vu, Chi, Com (not Nz)	v	Nz
v ~ β	Mak[1]	h	Gi, Ch

*-pangul- 'wipe' -pangula (El), -ɸangula (UP), -ɸanguya (LP), -paanguɬa (Mw), -βangua (Ng), -vangua (Nz), -hangusa (Gi), -vangusa ~ -βangusa (Mak)

*-tapik- 'vomit' -t̪apeka ~ -t̪apika (El), -r̪aɸika (UP); -haɸika (LP), -hahika (Gi), -βaβika (Du), -tapika (Am,Mv,Ung), -raβika (Vu), -tapika ~ -tavika (Mak), -raβiha (Ng), -raviha (Nz)

*-peepel- 'wave' -ɸeɸela (UP); -ɸeɸea (LP), °-pepeɬa (¬ ee) (Mw), -pepea (Am, Mv,Ung), -pepa 'sway, etc.' (Am,Ung), -βeβera (Vu), -pepeya (Pe). (CB *-pèèp-, *-pèp- 'blow as wind'). Also:

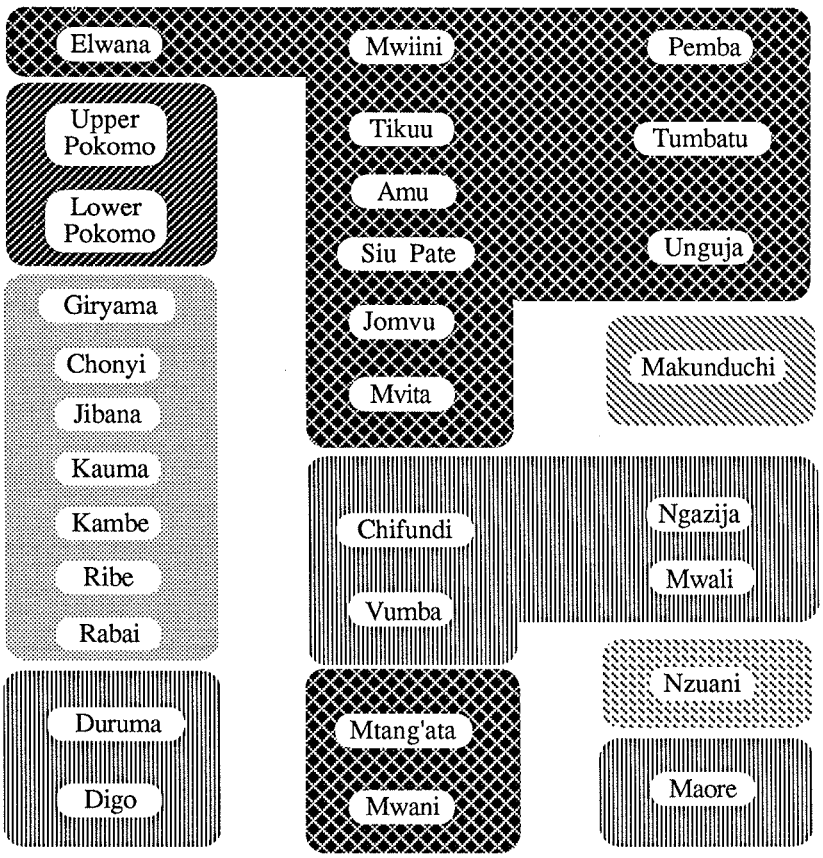
*kipeepelo 7/8 'waving, swaying object'
kipepelo 'butterfly' (El), kiɸeɸelo 'butterfly, bellows' (UP), kiɸeɸeo (LP), kiheho 'fan' (Gi), kipepeo 'fan, butterfly' (Am,Ung), °špepo (¬ ee) (Mw)

*lupeepo/mpeepo 11/10 'wind, cold, spirit'
bépɔ ~ béèpɔ 'cold' (El), mpeɸo 'cold, spirit' (LP), pʰeho 'cold, wind' (Gi,Du), °lpepo 11 (¬ ee) (Mw), °mpepo (¬ ee) 'spirit' (Mw), upepo/pʰepo 11 (Mv,Ung), uβeβo 11 (Vu), °upepo/ nyipepo 11 (°p ¬ β) (Ng), mbeβo 'heaven, fan' (Ng), mpevo 'wind' (Nz), peβo (Ma), mepo (Mn)

[1]Data in Chum and Lambert (1962/3a, b, and c) and Whiteley (1959) list for one Makunduchi dialect a voiced bilabial fricative as a reflex of *p. However, during recent field work (TJH, 1988) with several speakers of Makunduchi, including Haji Chum, Lambert's consultant, a voiced labiodental fricative, [v], was consistently elicited instead of [β]. Further field work would be desirable in order to say more about these reports, specifically to check the status of [β] and determine whether we are dealing with dialect differences or a documented generational shift of bilabial to labiodental articulation. Interestingly, Chum (pers. comm.) has reported that Lambert collaborated with him via the mail, and added that he did not do field work in Zanzibar.

Chart 1

Sabaki *p



Key

*p : p

*p : ɸ

*p : h

*p : v

*p : β

*p : β ~ v

Chap. 2, §4.1.1

TJH

Similar weakening of Bantu *p is widely attested in East Highlands Bantu. North of a line running west from the Rufiji Delta in Tanzania to at least as far as Lake Tanganyika (see Chart 25, or Guthrie 1967:70, topogram 5), most if not all the Bantu language groups attest similar weakening. This extensive distribution might indicate that lenition in Sabaki is a relatively early innovation compared to other shifts of a more restricted distribution.

However, the Sabaki situation is quite distinct. Adjacent non-Sabaki languages are quite uniform in attesting /h : *p/. Sabaki languages, in contrast, are not uniform. This lack of congruence and uniformity with the rest of East Highlands, plus the fact that /p/, a stop, is still attested, clearly as a relic, within part of the Sabaki continuum (Elwana, Mwani, and most Swahili dialects), points to later innovation for Sabaki. Setting up a lenis proto labial consonant for Sabaki involves other difficulties. To hypothesize something like *ɸ would force us to set up a rule for Elwana, Mwani, and Swahili that strengthens *ɸ, i.e., CB *p > PSA *ɸ > /p/ (Elwana, Swahili, Mwani), an unlikely, though not impossible scenario. However, there is little support for this, and we know of no cases where a voiceless bilabial continuant becomes stopped in non-postnasal environments. Furthermore, the distribution of /p/ within East African Bantu languages suggests that it is a relic. It is the urban Swahili dialects, sociologically prestigious, and Elwana, Mwani, and Nyilamba (F31), the latter three geographically isolated, where the bilabial voiceless stop is still attested. Also, if we look at the nature of *p-lenition in non-Sabaki languages, we find that there is a regular distribution of languages with /h/ < *p, forming a continuous band. Languages to the north of this area (in the so-called Bantu borderlands), however, show considerable variation in the form of their lenis segments: Central Kenya and the northeastern Lacustrine languages (Ganda-Soga, Luhya), parallel with Sabaki, attest a wide range of lenis reflexes for *p, including ɸ, h, Ø, w, β, and v, suggesting that *p-lenition was innovated elsewhere and is the result of relatively recent northerly diffusion into these borderlands after contemporary subgroups differentiated. All these facts support our contention that *p-lenition is a late PSA event, though not necessarily unconnected with lenition in the rest of East Highlands.

§4.1.1.1. *p-lenition. The distribution of the reflexes of PSA *p suggests the following two-part seriation:

*p	>	ɸ	>	h	>	Ø	Stage 1
			>	β	>	v	Stage 2

Whether in fact each of the dialects underwent precisely this ordered shift is unprovable, but our understanding of phonological change in general leads us to suggest this scenario. Pokomo, and at least two Mijikenda dialects, Giryama and Chonyi, share Stage 1; another part of Mijikenda (e.g., Digo and Duruma), three Swahili dialects (Vumba, Chifundi, and Makunduchi), and Comorian participate in Stage 2. The diagram assumes that *ɸ is an intermediate innovation shared by all the dialects with the exception of Elwana, Mwani, and most Swahili. The seriation allows us to see the potential relationship between the various aspects of lenition in Sabaki. What the dialects share is a disposition to weaken

*p. The seriation makes that explicit and suggests that *p, a bilabial stop, did not simply independently become /h/, or /ɸ/, or /β/ or /v/. Phonetically, it is easier to understand what has happened if we see that an /h/ (as in Giryama, Chonyi, etc.), bleached of anything that might suggest it was once a bilabial stop, can most readily be derived intermediately from a non-stop, namely, *ɸ (as in Pokomo), either articulatorily or acoustically. The development of /β/ (as in Digo, Duruma, Comorian, etc.) stems most naturally from *ɸ , a simple voicing change. That it is still attested in Pokomo lends support to this analysis. Finally, in Comorian a feature change of bilabial to labiodental (*β > /v/, Nz) completes the chain of events. The same change is apparently attested for Makunduchi; however, see fn. 1.

Although *ɸ, a voiceless bilabial fricative, is a convenient source for what has happened (Hinnebusch 1973; Hinnebusch, Nurse, and Mould 1981), /ɸ/, as such, may never have been part of the phonological inventories of pre-Digo, pre-Comorian, or pre-Swahili. This is especially clear in Swahili, where the dialect variants are only /p ~ β ~ v/, never /ɸ/. Simply the phonetic "weakness" of /p/ could be the actual source (see Maddieson 1984:37). Nevertheless, we retain this analysis, understanding that the late PSA innovation we are claiming here is an icon that stands for the shared Sabaki propensity for lenition. It is this tendency, whatever its source, that sets up the conditions for the individual Sabaki dialects to innovate their synchronic systems (see Heine and Möhlig 1980:38ff. for a similar use of seriation relative to these changes). We summarize this development in the following terms:

Stage 1:	PSA *p (pre-PSA inherited)
Stage 2:	Innovation of Sabaki *p-lenition: *p > *ɸ (late PSA: Com, Po, MK; a few SD?[2])
Stage 3a:	*ɸ > /h/ (e.g., Gi, Ch)
Stage 3b:	*ɸ > /β/ (Di, Du, Com; also Vu, Chi, Mak?)
Stage 4:	*β > /v/ (Nz, Mak?)[3]

There are several points before concluding. That Elwana and Mwani did not participate in the shift is consistent with their geographical position at the periphery of the group. We can speculate that most dialects of Swahili did not share the change because of sociolinguistic variables that served consciously or unconsciously to keep the urbanized Swahili population separate from their rural relatives. Whatever the ultimate source of lenition in Sabaki dialects, it is clear from both the variety of change here and its distribution within the Sabaki continuum that *p-lenition is a late arrival on the scene; the speakers of SD, ND, Elwana, and Mwani were likely already in place, or sufficiently isolated geographically or sociologically, to not share in the shift. However, those that share lenis reflexes—Comorian, Mijikenda, Pokomo—do so because of shared history. They were most likely in

[2]The question-mark here indicates that lenition in the few Swahili dialects where it occurs may be due to nongenetic sources; see Chap. 5, §13.6.

[3]Question-marks are indicated for Makunduchi in both Stages 3b and 4 because we are not certain of the status of /β/ in this dialect.

close contact, possibly sharing adjacent territories, which would explain the diffusion of lenition. Following this, the linguistic or geographic break-up of this group must have followed fairly quickly to explain the diversity after Stage 2.

§4.1.1.2. *p-lenition and chronology. See §5.2.1 for a possible intersection of *p-lenition and spirantization.

§4.1.2. Proto-Sabaki *t. CB *t in non-postnasal position has the following reflexes in Sabaki (Chart 2):

CB *t : PSA *t :			
t	Mw, ND,[4] SD, Mn	t̪	El[5]
č	Si, Pa, Ti	r̪	Upper UP
h	Lower UP, LP, MK	r	Vu, Chi
r ~ tr	Com[6]		

*muti 'tree' mót̪ɪ (El), mur̥i (UP), muhi (LP,Gi,Ra,Di,Du), muti (Mw,Mn), nči (Ti,Si,Pa), mti (Am,Mv,Ung), nti (Nga,Jo), mri ~ muri (Chi), mɟiti (Tu), mɟiti/miɟiti ~ uɟiti/miɟiti (Mak), nti/miti (Pe), mri (Vu), mri [mdri]/miri (Ng), muri (Mh), mwiri (Nz,Ma)

*luti 11/10 'stick, shaft' lut̪i (El); yuhi (LP); luti (Mw); uti/nyuti (Am,Ung); uri/ndri (Ng); uri 'magic, charm' (Nz)

*-tapik- 'vomit' -t̪apeka ~ -t̪apika (El), -r̪aɸika (UP), -haɸika (LP), -hahika (Gi), -tapika (Am,Mv,Ung), -raβika (Vu), -tapika ~ -tavika (Mak), -raβiha (Ng), -raviha (Nz)

*-tafun- 'gnaw' -r̥afuna (UP), -hafuna (LP,Gi), -tafuna (Mw,Am,Mv,Ung,Mn), -čafuna (Si,Ti), -trafunya (Ng)

Though most Sabaki dialects attest reflexes that derive from weakening (or less commonly affrication), the evidence indicates that PSA inherited a stop. All Sabaki's NEC neighbors attest a stop, and most East Highlands languages in the northeast also directly attest a stop (exceptions are the Chaga group and Nyaturu (Rimi). Further, the variation and seriation of the lenis forms within Sabaki point to an original stop, which is still attested, of course, in Elwana, Mwani, and most Swahili dialects. Thus, the lenition seen in the Sabaki group is a strictly local phenomenon, and must be a late PSA development.

---

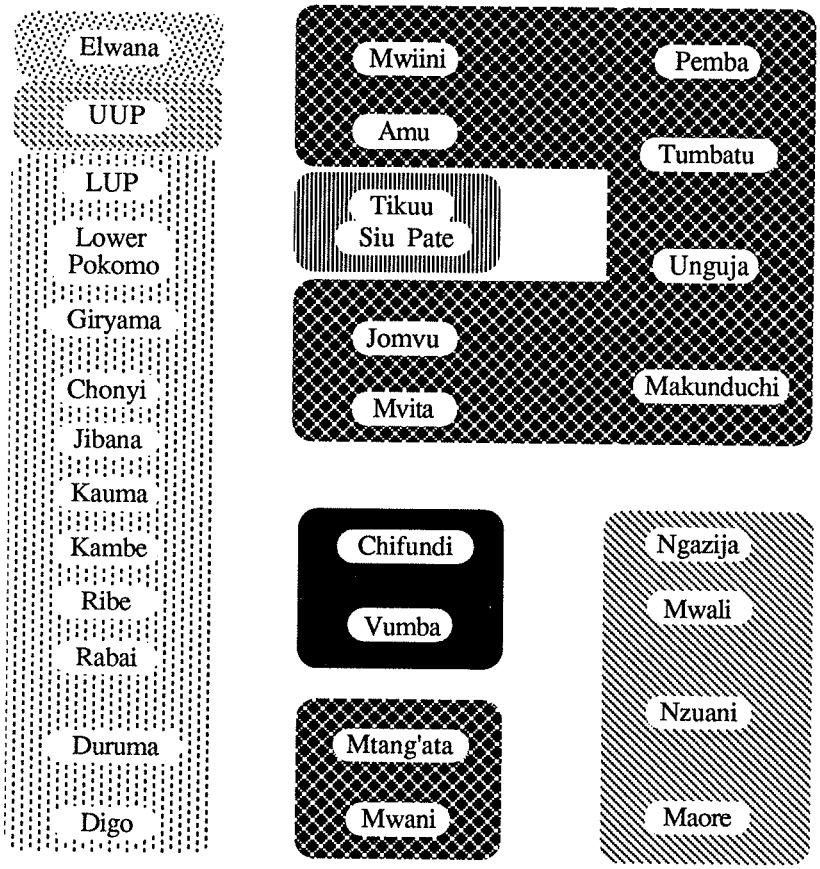
[4]In Mvita, /t/ is sometimes heard with a retroflexed off-glided pronunciation [tɻ] (TJH, field work 1988).

[5]We follow Maddieson and Sands (pers. comm.) and mark all t's as dental. However, we note that Möhlig (1986a) records both alveolar and dental t's in his phonological inventory for Elwana, but only gives examples of alveolar /t/ in his 100-word list, e.g., -tóòna (El), cf. -r̥una (UP), -huna (LP) 'cut'; -katɛ (El) as in mwékátɛ 'daytime', cf. mučikar̥e (UP); máfóta (El) 'fat, oil'; -tátʊ (El) 'three'; mòtɪ (El) 'tree'.

[6]Comorian /tr/ is phonetically [tɻ].

Chart 2

Sabaki *t



Key

*t : t

*t : č

*t : h

*t : r ~ tr

*t : t̪

*t : r̥

*t : r

§4.1.2.1. *t-lenition. PSA *t demonstrates a seriation similar to the one we have seen above for *p (also see Heine and Möhlig 1980:46):

*t	>	(t̪ʳ)	> r̥	> h	Stage 1
				> r	Stage 2

This seriation is motivated by the same considerations that prompted the seriation for *p; each intermediate stage is presently attested in the dialect continuum. Moreover, phonetically it makes more sense to suppose that each point in the series arose from the previous stage rather than to hypothesize, for example, that a voiceless *t shifted directly to a voiced /r/, or that *t weakened directly to /h/.

There is another set of facts that needs to be noted. The same dialects that share *t-lenition also attest forms of *p-lenition: Pokomo, Mijikenda, Comorian, Vumba, and Chifundi. The parallelism does not stop here. For example, it is Pokomo which shares earlier stages of both lenitions, though not completely. Only Upper Pokomo still attests the earlier voiceless /r̥/, the rest having /h/, while all Pokomo attests /ɸ/. This cannot be an accident. Lenition of *t in Sabaki is clearly linked with *p-lenition. For the most part, the same set of languages shares both *t- and *p-lenition (Mijikenda, etc.), and it is the same set which is conservative (Mwani, Elwana, ND, and most of SD—the exceptions being Vumba and Chifundi). We hypothesize the following stages of change:

PSA	*t	Pre-PSA inherited
Stage 1:	*t > *r̥	Late PSA: MK, Po, Com; /r̥/ attested in upper UP
Stage 2a:	*r̥ > /h/	MK, LP, part of UP adjacent to LP
Stage 2b:	*r̥ > /r/	Vu, Chi, Com

An alternative analysis is suggested by some Comorian data. PSA *t in intervocalic environments in all Comorian dialects has, for the most part, shifted to /r/ [r], e.g., /nkara/ (Ng, Nz) 'headpad' (*nkata). In Class 5 nouns, however, *t is preserved as /tr/ (realized as a stop with a voiceless liquid off-glide, namely, [t̪ʳ]) as in traho/maraho 5/6 'buttock' (*i̪tako) (see Sabaki Strengthening, §6.0.), in contrast to the dental stop /t̪/ which is found predominantly in borrowed vocabulary, e.g., Ngazija /t̪afau̪t̪i/ 'different', mai̪t̪i 'corpse', both from Arabic. The off-glided /tr/, varying sometimes with /r/ crossdialectally, is also found in some non-Class 5 contexts, but not in a consistent, predictable way, as seen in the following examples (those marked SB are from Silbertin-blanc 1980:46):

itranda 7 (Ng), šitranda 7 (Nz) 'bed' (*kitanda, CB *-tànda) (SB)
ilatru 7 (Ng), šilaru 7 (Nz) 'shoe' (*kilatu) (SB)
mafura (Ng), matra (Nz, Ma) 'oil, fat' (*mafuta)[7]
-ra mahio (Ng) 'shout', -ta ~ -tra mare (Ng) 'spit', -tra (Nz) (PSA *-t- 'hit, throw')

[7]We assume that /matra/ in Nzuani and Maore derives from *mafutra via the followings steps: *mafutra > mafu̥tra > maftra > /matra/. We have not found other examples of the loss of [fu]; it may correlate to loss of /si/ (see Chap 5, §9.5).

There is no immediate explanation for this, although one possibility is that we are looking at a stage of Comorian in which lenition either is still active, though incomplete, or became inactive before it affected all lexical items. In either case, [t̪ʳ] represents an intermediate stage of change, and where it exists synchronically (but only predictably in Class 5 after *i̪; see §6), it is a relic of that stage of lenition. Thus

Stage 1:	*t > t̪ʳ	Context-free Com shift
Stage 2:	*t̪ʳ > r̪ > r	Mostly regular intervocalically; there is some variability, but not in Class 5, nor after *i̪

In some cases, /tr/ in nouns (e.g., /ilatru/ 'shoe') may be explained on analogy with Class 5 nouns where it is preserved and regularly alternates with /r/ in the plural, but this would not explain the presence of /tr/ in verbs (e.g., /-tra/ 'hit'). At this point we prefer to consider such cases relics, which for whatever reason did not undergo the next stage of lenition. This analysis of Comorian suggests that *t-lenition in other Sabaki dialects may also have attested a similar intermediate change of *t to /tr/, a variable observed in Mvita (TJH, field work 1988). If so, the following would better represent the evolution of Sabaki *t-lenition:

PSA *t		Pre-PSA inherited
Stage 1a:	*t > t̪ʳ	Late PSA: Com, Po, MK; Vu, Chi; /tr/ attested in Com
Stage 1b:	*t̪ʳ > /r̪/	Intermediate in Po dialects, MK; still attested in some UP dialects
Stage 2a:	*(t)r̪ > /h/	MK, LP and adjacent UP
Stage 2b:	*(t)r̪ > r	Com; Vu, Chi

Our own field notes for Comorian (TJH, 1983) also show another sort of variation, namely, [t̪] ~ [t̪ʳ] ~ [r], either elicited variably or seen in comparison with Sacleux (Sx: Chamanga and Gueunier 1979):

t̪asi ~ tʳasi/maṯasi (Ng) 'morning'
t̪avu (~ tʳavu ~ ntʳavu, Sx) 9/10 'mud'
it̪unguu ~ itʳunguu 7 (Ng); širungu 7 (Nz) 'onion' (*kitungulu)

The precise direction of the variation, whether [t̪ʳ] is replacing [t̪] or vice versa, is not certain, but we can offer the following hypothesis. This particular sort of variation does not involve the Class 5/6 alternation ([t̪ʳ ~ r]), nor do the many recent borrowings from Arabic, French, ND, etc., show variation; they typically attest /t̪/, and do not vary with /t/ [t̪ʳ]. Only inherited Bantu lexis appears to show such variation. This fact, and the Ngazija ~ Nzuani dialectal alternation (/it̪unguu ~ itrunguu ~ širungu/), suggest that /t/ [t̪ʳ] is shifting variably to /t̪/. This is also somewhat predictable by morphophonological structure: /t/ [t̪ʳ] alternating with /r/ in the Class 5/6 paradigm is allowed, but nonalternating /t/ [t̪ʳ] is considered aberrant, and is beginning to fall in with borrowed lexis with /t̪/. The matter requires more field verification.

As a final note we point out that Mijikenda, particularly Digo, has a nasalization process associated with /h/, regardless of source, whereby adjacent vowels are nasalized, e.g., *muti : /muhi/ [mũhĩ] 'tree', /-ba(h)a/ [bãhã] 'big' (a loan word).

§4.1.2.2. ***t-palatalization (ND).** Three ND dialects (Tikuu, Siu, and Pate) attest an unrelated shift, wherein *t becomes /č/:

*t > č	Pate, Siu, Tikuu	
-čeča	'quarrel'	(*-teta, CB *-tét-; cf. -teta Ung)
-čunda	'pick fruit, flowers'	(*-tunda; cf. -tunda Ung)
nčočo	'child'	(*mutoto; cf. mtoto Ung)

This change requires little comment except to point out that it is an example of a chain-shift, motivated by an earlier shift in which PSA *c in ND becomes /t̪/ (§4.1.4.1. below). The Tikuu, Siu, and Pate shift of *t > /č/ thus recreates the earlier phonemic pattern (Nurse 1985b). The Swahili SD also participate in a similar affrication process wherein *t/___ ia becomes /č/: *t > č /___ia; thus -ča 'fear' (*-tia; cf. -tia Mw, -ria Nz).

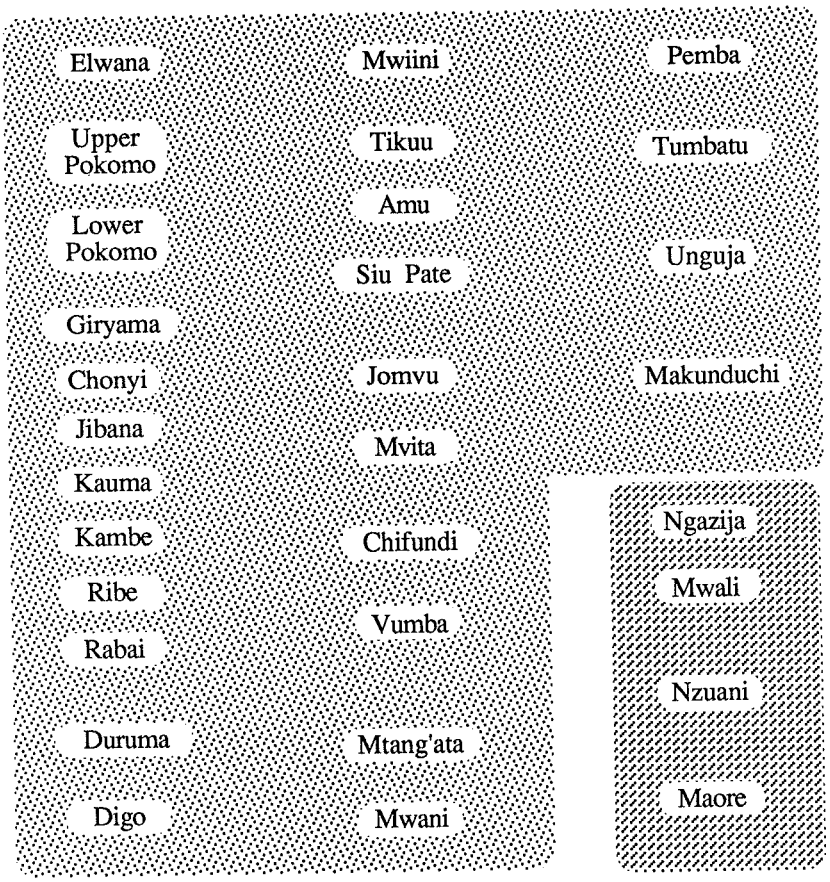
§4.1.2.3. ***t-lenition and chronology.** Stratigraphy indicates that *t-lenition is later than *p-lenition. *p-lenition is a widespread Bantu shift; *t-lenition is highly restricted. Our working hypothesis, therefore, has *p-lenition occurring earlier than *t-lenition. There is, unfortunately, little independent internal Sabaki evidence to support this, beyond the following: while there is interaction between *p-lenition and Spirantization (see §5.2.1), indicating that *p-lenition may well be as early as Spirantization, there is no such evidence for *t-lenition, suggesting that *t-lenition is later than Spirantization. Furthermore, *t-lenition does not bleed other rules affecting *t, e.g., Spirantization or Class 5 Strengthening (§6). Finally, there is internal evidence that indicates that at least in Comorian *t-lenition is later than *p-lenition: *p-lenition applies uniformly (*p > β > v), *t-lenition applies variably (*t > tr ~ r).

§4.1.3. **Proto-Sabaki *k.** For most of Sabaki, *k is attested as /k/; the exception is Comorian, where it has shifted to /h/. There are also context-specific changes affecting *k, as we will outline below. CB *k in intervocalic and some stem-initial positions has the following reflexes in Sabaki (Chart 3):

CB *k :	PSA *k :	
	k	El, Po, MK, Mw, Sw, Mn
	h	Com
*-anika	'set out to dry'	-anika (LP,Gi,Ung,Am,Mn), -aanika (Mw), -anyiha (Nz), -aniha (Ng,Ma)
*-ceka	'laugh'	-čɛka (El), -tseka (LP,Gi,Du), -t̪eka (Mw,Am,Mv), -čeka (Ung), -tseha (Ng,Nz,Ma), -seka (Mn)
*-kam-	'milk'	-kama (Gi,Du,Mw,Am,Mv,Ung), -hama (Ng,Nz)

Chart 3

Sabaki *k



Key

*k : k

*k : h

Our assumption that PSA inherited a voiceless velar stop from a pre-PSA stage, and thus our reconstruction of *k for PSA, is amply supported and noncontroversial.

§4.1.3.1. *k-lenition: Comorian. Lenition of *k > /h/ is limited to Comorian. Nevertheless, it is reasonable to claim that it is an extension of *p- and *t-lenition that Comorian, and other Sabaki dialects, have undergone. It occurs under the same conditions and constraints.

Outside Comorian there are a few idiosyncratic cases of *k-lenition, but it is uncertain whether these changes were motivated by the general sort of lenition affecting *p and *t, or were simply independently motivated. In each instance where this has happened it is restricted to specific environments. For example, the Mwiini infinitival marker *ku- has become /x-/ (/x-pika/ 'to cook', /x-taaɬa/ 'to take', /x-sooma/ 'to read', etc.); this is part of a language-specific process involving, as well, the PSA prefixes *ki-/*vi̧- (Class 7/8) which shift respectively to /ši-, s-/ (via či-/zi-) before voiceless obstruents, e.g., /š-kombe, s-kombe/ (Whiteley 1965, Abasheikh 1978).

Another case is the PSA primary negative prefix *nka-, which changes to /ha-/ in Swahili via several intermediate stages: *nka > n̥kʰa- > kʰa- > ha- (see §8.2. on nasal-devoicing and aspiration). Another example is the ND and Comorian verbal innovation *ni ku- 'habitual' which becomes /hu-/ through much the same intermediate stages, namely, *ni ku- > n(i)ku- > nku- > n̥kʰu- > kʰu- > hu-; see Chap 4, §3.3.1.1).

Finally, in Tikuu there is a small set of items where *k has weakened (e.g., /hondre/ 'field' (*nkonde), /hundre/ 'beans' (*nkunde), /hon̥doo/ 'sheep' (Sw kondoo), /-ehundru/ 'red' (*-enkundu). These are most likely due to the same sort of sporadic shift of *nk > h just described for Swahili (however, see Bakari 1985:229 for a different analysis). This particular sort of lenition has not been observed in Comorian, so it is unlikely that these are loans from Comorian.

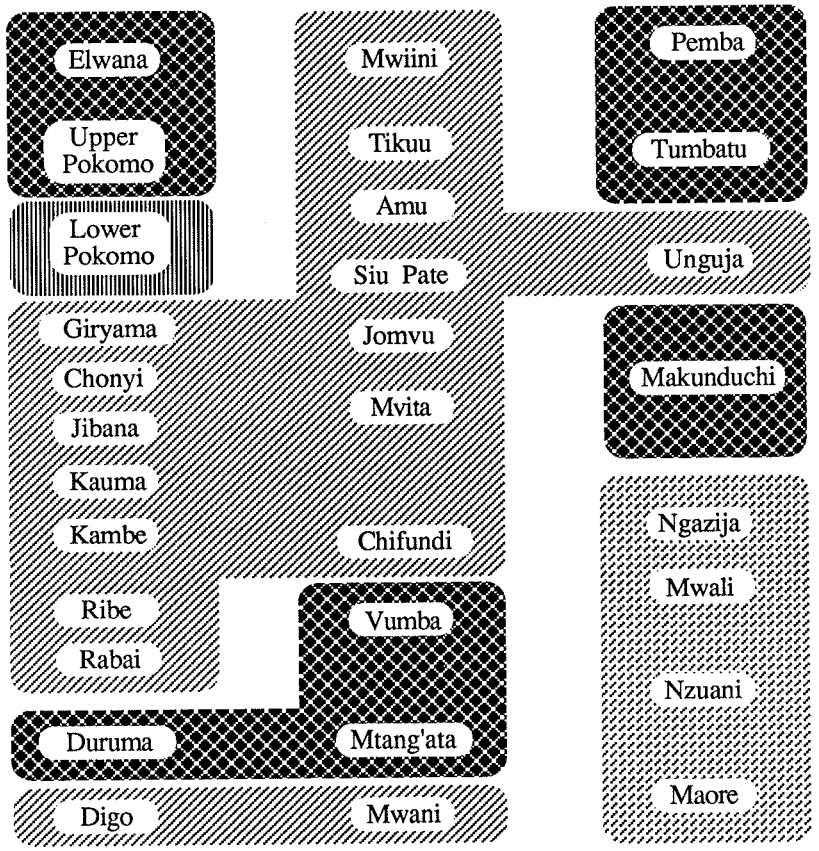
§4.1.3.2. *k-palatalization. This is a label we use to refer to the pre-PSA innovation where CB *k before *i + V is realized as PSA *kyV and a subsequent series of changes involving *ky. It is this palatalized velar which is the source for /ky/, for those /č/ which are not derived from CB *c, and for those /š/ whose source cannot be traced to CB *k/___ i̧ (PSA *š, see §5.1.4). It is *ky, furthermore, which leads to subsequent Sabaki palatalizing changes affecting *k before *i and *e. Correspondences for PSA *ky and *ki in potential palatalizing environments follow:

1. CB *ki/___V : PSA *ky (Chart 4)

ky	El, UP, Du, Vu, Pe, Mt, Mak, Tu
ky ~ č	LP
č	MK except Du; Ung, ND, Mw, Mn
š	Com

Chart 4

Sabaki *ky



Key

*ky : ky

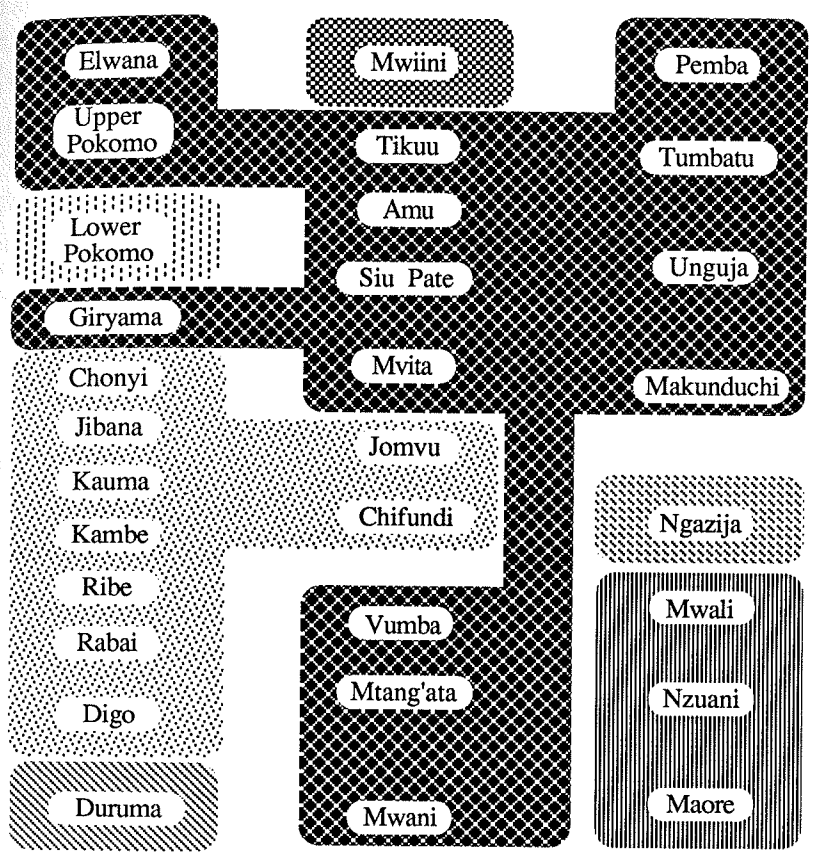
*ky : ky ~ č

*ky : č

*ky : š

Chart 5

Sabaki *ki



Key

*ki : ki

*ki : či ~ ši

*ki : ki ~ či

*ki : ky

*ki : či

*ki : (h)i ~ š

*ki : š

*-kya 'dawn'	-ča (Gi,Ch,Mw,Mv,Ung,Chi,Mn), -kya (Du,Pe,Vu), -ša (Ng,Ma)
*kyula 'frog'	kyola (El), kyula (UP,Du), čuya (LP), kyula (Du), čula (Gi), čulwa (Ch), čuuɬa (Mw), čwačwa ~ čuačua (Am,Ti,Pa,Si), čura (Mv,Ung), kyura (Vu), kyura ~ šura (Pe), šwa-toto (Ma)
*kyakulya 'food'	kyakuɟa (UP), čakuɟa (LP), čakurya (Gi,Mn), kyakurya (Du), čakuǰa (Mw), čakula (Am,SD), kyakulya (Mak,Vu,Pe), šahula (Ng,Nz,Ma)
*kyala 'finger'	kyáàla (El), kyala (Du), čala (Gi,Maf,Mn), čaaɬa (Mw), šaya (Ng), ša (Nz)
*kyelupe 'white'	kyauɸe (LP), čeupe (Ung)

2. CB *ki/___C : PSA *ki (Chart 5)

ki	El, Gi, UP, ND except Mw, SD, Mn
ki ~ či	LP
ky	Du
či	MK (not Gi), Jo
či- ~ ki-	Chi[8]
ci- ~ š(i)-	Mw (see 3. immediately below)
(h)i	Ng (but /ši/ in, e.g., *mukila 'tail')
ši	Nz, Ma, Mh

*kicwa 'head'	kečwa ~ kíčwa (El), čitswa ~ kitswa (LP), kitswa (Gi), čitswa (Ch,Ra,Di), kyitswa (Du), čiṯa (Mw), kiṯwa (Ti,Am,Si,Pa, Mv), čiṯwa ~ kiṯwa (Chi), kičwa (UP,Ung,Vu,Mt,Pe,Tu,Mak, Ung,Maf), hitswa (Ng), šitswa (Mh,Nz,Ma), čiṯwa (Chi), kiswa (Mn)
*kikapu 'basket'	čikaɸu (LP), kyikaBu (Du), čikaBu (Di), škapu 7/8 (Mw), kikapu (Am,Ung), °ikapu/zi- (°k/p ¬ h/B) 7/8 'type of basket from Madagascar' (Ng)
*kintu 'thing'	kintu (LP), kit^{h}u (Gi), kyit^{h}u (Du), čiint^{h}u (Mw), k^{h}iču (Ti,Pa, Si), k^{h}itu (Am), kit^{h}u ~ k^{h}itu (Mv,Ung), k^{h}itu (Nga), čitu (Jo), kitu (Pe), hindru (Ng), šintru (Nz), šitru (Ma), kinu (Mn)

[8]Lambert (1958b:22) says: "The prefix of Class 7 is almost invariably 'chi', but in Funzi 'ki' is occasionally heard as an alternative in a few words."

*mukila 'tail'	mukila ~ mukela (El), mukila ~ muki(y)a (UP), mči(y)a (LP), mukira (Gi), mukyira (Du), mučira (Ch,Ra,Ka,Ji,Ka), mčira (Di), mkiɬa (Mw), nkia (Ti,Si,Pa,Pe), °nkia (°k ¬ č) (Jo), mkia (Am,Mv,Ung,Vu,Mt,Tu,Mak), mčia (Chi), mšia 'tail, tail bone' (Ng), mšia (Nz); nkira (Mn)

3. PSA *ki/___Ç :

š	Mw
*kikapu 'basket'	škapu
*kipande 'piece'	špande

4. PSA *k/___e :

č	Gi, Ch, Di, Jo, Chi
ky	Du
š	Com
*-kema 'cry out'	-čema (Gi) 'shout', -kema (Ung), -šemeza (Ng), -šemeledza (Nz), but nkeme 'noise' (Nz)
*-kenga 'deceive'	-čenga (Jo), but -kenga (Gi)
*-kešya 'spend night'	-česa (Gi,Di,Chi), -šesa (Ma), -šedza (Nz), kešo (Ung) 'tomorrow'
*-pokela 'receive'	-počea (Jo), -Bočera (Di), but -hokera (Gi)
*įkende 'testicle'	čende (Gi,Ch), kyende (Du) but kenze 5 (Nz (Sx))
*įkenda 'nine'	čenda (Gi,Ch,Di), šendra (Nz)
*įvuke 'vapor'	vuče (Gi), vuše (Ng) 'smell'
*muke 'wife'	muče (Gi,Ch,Chi), mukyetu (Du), mše (Ng)

There are several related phenomena, all involving palatalization, apparent in this array of data: (a) the shift of PSA *ky to /č/; (b) the development of Comorian /š/; (c) the generalization of *ky-palatalization to Digo, Chonyi, and Mwiini; (d) the Mwiini development of a morphophoneme /š/; and (e) the palatalization of *k before *e.

We are assuming that present-day Sabaki dialects which attest /č/ before vowel-stems derive this reflex from *ky, e.g., /č-akula/ (Ung) : CB *k___i+V). This is the obvious source of /č/ (distinct from /č/ < CB *c) in the Swahili dialects, as predicted by phonetic seriation, a hypothesis which is supported by the fact that some Swahili dialects still attest /ky/, e.g., /kyakulya/ 'food' (Mak, Vu, Pe). A generalized form of this change also operated in most Mijikenda dialects, Mwiini, and Comorian, where *k not only became /č/ (/š/ in Comorian) before *y but before *i as well, a type of rule simplification wherein the conditioning environment is any [+high, +front] segment. At some point a further simpli-

fication occurred and the change affected *k before any [+front] segment, namely, /y, i, e/; this stage is most clearly attested in Comorian. So far, then, we distinguish the following:

CB *ki : PSA *ky /___ V	PSA (/ky/ still attested in El, Po, Du, SD (not Ung)
*k > č/___ y	Stage 1: post-PSA: all of MK, Ung, ND, Chi, Mn, Com (if Com /š/ < č < *k/___y; see below)
*k > č/___ i	Stage 2: post-PSA: MK (not Gi), Mw, Jo, Chi, presumably also Com (see below)
*k > č/___ e	Stage 3: post-PSA: MK (including Gi but see below), Jo, Chi, presumably also Com (see below)

Stage 1 is widely shared by most Mijikenda dialects, ND, Unguja, Mwani, and Comorian (by virtue of sharing a form of palatalization, namely, /š/ versus /č/) (see Chart 32). Stage 2 is more restricted, having affected Mwiini, Jomvu, Chifundi, and most of Mijikenda (not Giryama), and Comorian. Stage 3 clearly affected Comorian, but it is not clear that the palatalization we see in Mijikenda is due to Stage 3 (i.e., an extension of Stages 1 and 2) or another palatalization process. Mijikenda is not uniform in how it is affected by palatalization. Giryama variably palatalizes *k and *g before *e (muče 'wife' < *muke, -jeza 'try' < *-gez-, but -kenga 'deceive' < *-keng-, -ogerera 'swim' < -ogelel-), but not before *i (e.g., -gita 'cook', giza 'darkness', ngira 'path', ki- Cl. 7), whereas the rest of Mijikenda for the most part regularly shows palatalized reflexes of both *k and *g before *i and *e (e.g., -ojerera 'swim' (Du, Di, Ra, Ch), njira 'path' (Du, Di, Ch, Ra, etc.), či- Class 7) (see end of §4.2.3.1 for further discussion). Thus only part of the cluster is congruent with Stage 3; Giryama is out of step in terms of both the seriation and stratigraphy of the process. While clearly the palatalization we find in Giryama before *e is ultimately due to Stage 1, the facts seem to suggest that Stage 3 for Mijikenda is independent of the operation of Stage 3 in Comorian.

Does palatalization provide any subgrouping evidence? Most of Mijikenda, ND, and Comorian fall together in sharing Stage 1; Unguja fits as well but this may be another ND influence that we see elsewhere, given that the rest of SD still attests an unchanged reflex of *ky. That Mwiini falls in with Mijikenda in sharing Stage 2 is probably due to convergence since in the rest of ND we have only variable attestations of *k-palatalization. In ND_3 all the dialects except Pate exhibit variable synchronic palatalization in Class 7: /kisu hiki ~ hiči/ 'this knife' (Am, Si, Ti), /kisu kilee ~ čilee/ 'that knife' (Am, Si, Ti), /kisu kimevunḍika ~ čimevunḍika/ 'this knife is broken' (Si, Ti) (Nurse 1982b:84-85). Similar variable palatalization has also been noted elsewhere among Mombasa and Zanzibar speakers of Swahili. All such cases, including Mwiini, can be traced to the fact that *ky has become /č/ in these languages; given this precedent, and the fact that palatalization is such a natural phonetically motivated change, it is not surprising to find either the synchronic variation or the de-facto palatalization in Mwiini. Stage 1 sets things off, and the proto dialect cluster was probably more homogeneous when it got started, and thus it seems reasonable to think that it is a genetically shared event, but the dialects attesting Stage 2 were

likely already in place and distinct from the others (e.g., Mwiini from ND, Duruma and Giryama from the rest of Mijikenda) when it began. Comorian (see below for details) departs from the others in attesting /š/ rather than /č/, but this could be a late adjustment after Comorian separated from the others. Whether Comorian was still part of the mainland continuum when Stage 2 innovated is not known, but for subgrouping purposes we have included Comorian within the isogloss. The palatalization we see in Jomvu and Chifundi may well be traced to Mijikenda influence.

Some further language-specific peculiarities have to be noted. In Mwiini, only Class 7 *ki ~ *ky- shifts to /č/; *ki in word-internal positions is unaffected (e.g., /čisu/ 'knife' < *kifyu, versus /mkiɬa/ 'tail' < *mukila, but /-čiimbiɬa/ 'run' < *-kimb-). Lower Pokomo has regular lexis with /či/ from *ki, but both /ki-/ and /či-/ in Class 7 C-initial stems. Older Lower Pokomo speakers still use Class 7 /či-/; younger speakers use Class 7 /ki-/, which is a systematic borrowing of the Swahili Class 7 shape (H. Ipu, pers. comm.). Stage 3 in Mijikenda has not affected all lexis, e.g., /-čenga/ (Ch) versus /-kenga/ (Gi) 'deceive' < *-kenga. However, Duruma is conservative in all environments, preserving /ky-/. Lambert (1958a:21) notes that in Jomvu, palatalization does not happen when /k/ occurs before an inflexional morpheme, such as the present negative /-i/, the subjunctive /-e/, the prepositional /-il-/, etc. Thus, /-ṯukia/ 'hate', /kʰaṯaki/ 'he does not want', /-čokeša/ 'make tired', /nifike/ 'let me arrive'.

In Comorian *k has shifted to /š/ (or /h/) in all three environments, before /y, i, e/, e.g., *-kya : /-ša/ 'dawn', *kicwa 'head' : /šitswa/ (Nz), *muke 'wife' : /mše/. In Ngazija, *ki- Class 7 is realized as /š(i)-/ before V-stems, e.g., /šahula/ 'food'; before C-stems it is realized as /hi-/ (monosyllabic stems), and /i-/ (polysyllabic stems), e.g., /hitswa/ 'head', /iloo/ 'fishhook'. The shift of *k in the Class 7 prefix to /h/ could well be part of Comorian *k-lenition (*k > h), but the alternation within Ngazija, and the evidence in Nzuani (*ki- 7 > /ši-/), indicates otherwise. Whether these events in Comorian are independent innovations or part of the general Sabaki series of shifts is not clear. However, given the fact that Comorian takes part in other Sabaki shifts (e.g., the various consonant lenitions) we hypothesize that the Comorian situation results from an intermediate shared stage with Sabaki, and that Comorian innovated by shifting an intermediate *č to /š/:

*k > č/___y,i,e	Stage 3 inherited post-PSA shift
*č > š	Stage 4: Comorian[9]
*š > h (Class 7 C-stems)	Stage 5a: Ngazija
*h > (Ø) (Class 7)	Stage 5b: Ngazija polysyllabic stems

A minor adjustment in the formulation of the rule for Stage 3 is glide absorption, whereby *y deletes during the palatalization of *k. Also, Stage 4 in Comorian must be ordered after the change that shifts PSA *c (: CB *c) to /ts/ (see §4.1.4.3 below).

[9]Alternatively, PSA *k > š/___ y,i,e independently.

There is an unconnected change in some lexical items in some SD where Swahili *č irregularly shifts to /š/, e.g., *ncawa 'louse' : /šawa/ (Pe). The sources of /š/ in this and Comorian are different (see §4.1.4.2).

A piece of evidence possibly supporting the shift of a Comorian *č > š is an interesting alternation in Mwiini. In Mwiini, Class 7 prefix *ki- is realized alternatively as /či-/ before voiced consonant-initial stems, or /š-/ before stems with initial voiceless consonants, e.g., /p, t, t̪, k, f/, but not before /s, š/, and not in monosyllabic stems, e.g., /čisima/ 'well', /čišiindo/ 'noise', /čit̪a/ 'head'. A few Class 7 nouns showing the alternation follow:

čiβli	'shadow'	špete	'ring'
čiguwo	'cloth, rag'	štuunguɬu	'onion'
čiǰana	'small child'	št̪ana	'comb'
čiwaɬa	'scar'	šfinikio	'cover'

Thus, Mwiini has a synchronic rule in which an underlying /či-/ [+Cl. 7 nominal prefix] becomes /š-/. This is part of a more general rule in which the nominal prefixes /ku-/ Cl. 15, and /zi-/ Cl. 8 shift to /x-/ and /s-/ respectively in the same environment (see Whiteley 1965:68); we will call this Stage 6 for purposes of exposition only:

*č > š/__...[-voice] (Cl. 7) Stage 6: post-PSA: Mwiini

What set up this change, we argue, would have been an initial vowel-devoicing process affecting the vowels of nominal prefixes which have a voiceless underlying consonant followed by a stem-initial voiceless consonant. With vowel-deletion, a dissimilation process creates more natural fricative-stop clusters from stop-stop clusters. Thus, /čikapu > čik̥apu > škapu/ 'basket'. The phonological processes for fricative-initial stems are somewhat different; stems with initial fricatives are only affected if the potential final result does not mask the marking function of the prefix or the phonological structure of the stem. Therefore, /čifinikio > čif̥inikio > šfinikio/ 'cover' is possible, but not *šsima (< /čisima/ 'well') or *šsepe (< /čišepe/ 'old cloth'), because these violate conditions on preferred consonant clusters or morpheme structure. On the other hand, the conditions that shift /zi-/, the Class 8 plural prefix, to /s-/ cannot be strictly phonological, since the conditions governing vowel-devoicing do not hold. Vowel-loss here occurs on analogy with the vowel-loss affecting the Class 7 and Class 15 prefixes; this then creates the conditions leading to the assimilation of /z-/ to the following voiceless consonant.

Because of the rather special circumstances under which *č becomes /š/ in Mwiini, it is unlikely that there is any connection between this change and the general situation in Comorian. Furthermore, Mwiini does not participate in Stage 3, whereas Comorian has shifted *k > *č > /š/ before *y, i, and e. It does show, of course, the ability for /č/ to become /š/.

§4.1.3.3. *k-lenition, *k-palatalization, and chronology. We have discussed the following rules so far:

1. PSA *p-lenition
2. PSA *t-lenition
3. Comorian *k-lenition
4. *k-palatalization before V-stems
5. *k-palatalization before C-stems
6. *k-palatalization before y,i,e
7. Comorian *k > č > š/___ y,i,e
8. Mwiini *k > č > š/___ i+[-voice]
9. Comorian ši- > (h)i- (Ng)

The first three lenition changes can only be ordered relative to each other on the basis of stratigraphy, ranging from *p-lenition, the most widely attested, to *k-lenition, the most restricted. We can order Rule 3, *k-lenition, and Rules 4 through 7, the various palatalization rules, relative to each other on the basis of internal reconstruction. The palatalization rules—whether or not Comorian shared the earlier stages with the other Sabaki dialects—must be earlier than *k-lenition (*k > h), otherwise lenition would have affected *k in palatalizing environments. This is illustrated in Nzuani and Ngazija, where this order correctly derives the desired reflexes:

*kyakulya	*kicwa	*kicwa	
----	kitswa	kitswa	*c > ts (see §4.1.4.1 below)
čakula	čitswa	čitswa	(6) *k > č/___y,i,e
šakula	šitswa	šitswa	(7) *č > š/___y,i,e
šahula	----	----	(3) *k > /h/
----	----	hitswa	(9) ši- > (h)i-
šahula	šitswa	hitswa	Output
'food'	'head'	'head'	
(Ng, Nz, Ma)	(Nz, Ma)	(Ng)	

This ordering is consistent, as well, with what a stratigraphy would give us. The palatalizing processes are more widely distributed than *k-lenition, which is only locally attested; thus the late ordering of *k-lenition. With a different order, that is, *k-lenition ordered earlier than Rules 6 and 7, wrong reflexes for two of the three forms would result:

*kyakulya	*kicwa	*kicwa	
----	kitswa	kitswa	*c > ts (see §4.1.4.1 below)
hyahula	hitswa	hitswa	(3) *k-lenition
----	----	----	(6) and (7) not applicable
*hyakula	*hitswa	hitswa	Output

This order works for the Ngazija term for 'head' only because Ngazija has two rules which converge: (a) an earlier *k-lenition rule, and (b) a fairly recent one, still reconstructible internally and presumably synchronically operative, where /š/ yields /h/.

The order of applicability for the various palatalizing processes follows that suggested by phonetic seriation. Rule 5, while it has to be later than 4, could be ordered after any of the subsequent processes, since it is independent of them and they of it. Thus Class 7 /ki-/

to /či/ in Digo, Chonyi, and Mwiini could be as recent a change as the Ngazija shift of Class 7 /ši-/ to /(h)i-/.

Finally, there are cases where palatalization has not occurred, although the apparent surface conditioning environments are present, e.g., /kionda/ (Am) 'sore', /kiatu/ (Ung) 'shoe', /kiazi/ (Ung) 'potato'. These, and others, all have reconstructed *l-initial stems (*kilonda, *kilatu, *kilazi, etc.). The consonant at one point was deleted by a sound change. Here palatalization did not apply because it is an earlier rule than the *l-deletion (see §4.2.2.2 for further discussion). We know from written evidence that this /l/, at least in ND, was deleted only in the last two or three centuries. For a synchronic analysis of palatalization, see Kutik (1983). That palatalization is an older rule and no longer productive can also be seen by comparing recent loan-words with sequences potentially eligible for palatalization. Thus /kiasi/ 'amount', which is a recent loan word from Arabic, shows no palatalization, and must have been borrowed after palatalization ceased to be active.

§4.1.3.4. The morphological limits of palatalization. In Swahili, at least, palatalization has certain morphological limits. There are other contexts in which it might be expected to apply. Thus after /-ki-/, the Class 7 object marker, or /-ki-/ 'imperfective, if', there are many sequences in which the context for palatalization exists: /tu-li-ki-ona/ 'we saw it' (Class 7), or /tu-ki-enda/ 'if we go'. In these contexts, palatalization does not occur nor, as far as we can see from older Swahili texts, has it ever occurred. Thus we have to conclude that palatalization, at least in Swahili, is not a purely phonological rule. We do not have adequate data to enable us to say whether this is true of other Sabaki languages.

§4.1.4. Proto-Sabaki *c. The modern reflexes of CB *c in Sabaki are for the most part affricates, although not all preserve the palatality of the reconstructed segment. Reconstructing *c is sufficiently supported for Sabaki in contrast to most Bantu languages. In fact, it was the Sabaki languages, specifically SD /č/, that provided Guthrie with a major block of evidence in reconstructing *c for CB. The range of reflexes with examples is:

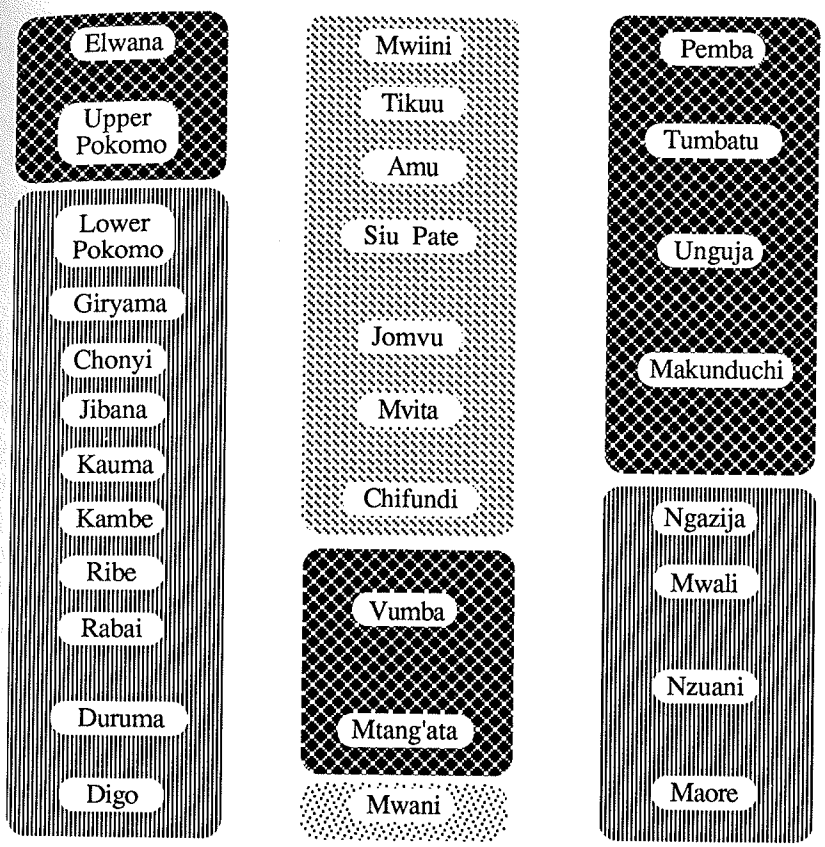
CB *c : PSA *c (Chart 6)

ts [t̪s]	Com, LP, MK
č	El, UP, SD
t̪	ND, Mw, Chi
s	Mn

*-Wici 'raw'	°-besi (El), -witsi (LP), -itsi (Gi,Du,Ng,Nz,Ma), -wit̪i (Mw, Am), -vit̪i (Ti), -bit̪i (Mv), -biči (Ung), -wisi (Mn)
*i̪-ji̪co 'eye'	ɪzíǐčɔ (El), ǰičo (UP), dzitso (LP,Gi,Ch,Du,Di,Ng,Nz,Ma), dzitso (Ng), it̪o (Mw,Am,Si,Pa), yit̪o ~ žit̪o (Ti), ɟit̪o (Mv, Chi), ɟičo (Ung,Vu,Mt), ɟičo ~ zičo (Pe), ɟičo ~ dičo (Mak), dičo (Tu), liso (Maf), riso (Mn); also see *kicwa App. 2.

Chart 6

Sabaki *c



Key

*c : č

*c : ts

*c : t̪

*c : s

We have reconstructed a voiceless palatal, *c, for PSA. Although we have chosen not to mark the segment in our reconstructed lexis for [+ or - delayed release], i.e., whether the segment was an affricate or stop, it is a fact that most of the reflexes of *c are affricates, /č/ and /ts/, or the stop /t̪/, which may be derived from the dental affricate */ts/ [t̪s]; thus there is good support for assuming that Sabaki *c was phonetically an affricate. The only exception is Mwani, at the extreme southern end of the Sabaki continuum, which attests the continuant /s/. This may have come from an areal influence of P20 languages; these mainly have /h/ or Ø, as reflexes of *c, which derive from an intermediate *s (*c > s > h > Ø).

The Sabaki languages are unique in NEC in attesting palatal (or dental) [-continuant] segments corresponding to Guthrie's CB *c. Surrounding languages attest either non-palatal voiceless or voiced continuants, usually /s/ and /z/, and reflexes derivable from these, e.g., /h/ < /s/ (e.g., Zone G10, 60) or /θ/ < /s/ (G23, E50). Most of Bantu, in fact, attests /s/ (see Guthrie 1967:76, topogram 11). The question now to be addressed is whether PSA *c is an innovation or an inherited segment from an earlier proto period? If we simply take Guthrie's CB reconstruction at face value, then we can argue that PSA preserved an earlier proto palatal, while surrounding languages along with the rest of Bantu have innovated by mostly shifting CB *c > s. Guthrie's reconstruction, however, should not be accepted without question. Guthrie (1967:62) speculated that CB *c was possibly the affricate [tš], a likely source for both series of commonly occurring reflexes: /š - s/ (Bantu) and /č - ts - t̪/ (mostly only Sabaki). On the other hand, Mann (1973:32-33) argued that *s was a more likely reconstruction, since the noncontinuant palatals were so restricted in distribution; only Sabaki supports Guthrie's reconstruction of a [-continuant] segment. Mann's analysis has its problems as well; it is never explained, for example, what triggered the shift of a nonpalatal /s/ to a palatal /š/. There is other evidence to consider as well and for this we must go beyond NEC.

Guthrie's hypothesis supports viewing Sabaki *c as inherited from PB. Mann's arguments support our contention that it is a PSA innovation, and therefore part of the body of evidence that supports the reality of Sabaki as an NEC subgroup. What further independent evidence exists to support this? Given the current view of the overall genetic classification of Bantu (see Watters 1989; Hinnebusch 1989; and Williamson 1989a and b), NEC is a subgroup of East Highlands Bantu, which in turn is a major sub-branch of Narrow Bantu; most languages of Guthrie's Zones A, B, and some of C relate to East Highlands Bantu at an even remoter node. At each point in the downward-branching tree, including where Zones A, B, and C link up with the rest of Narrow Bantu, East Highlands Bantu, and intervening nodes, up to the point where Sabaki is reached, we have little evidence (except in Sabaki, and perhaps Guthrie's Group P30) for reconstructing a palatal noncontinuant. We do have compelling evidence for reconstructing a continuant, either *s or *š, for most pre-PSA proto subgroups above the PSA node. The reconstruction of a continuant, moreover specified as palatal, would allow us to explain why a scattered number of languages throughout Bantu attest /š/ (see Guthrie 1967:76, topogram 11). A proto *š would, as well, explain the preservation of palatality in PSA. If so, then PSA would

have to have innovated a rule of palatal hardening in which CB *š shifted to PSA *c, whereas the rest of Bantu depalatalized CB *š to *s, except for a few relic areas where a voiceless palatal continuant was preserved. Such a hypothesis has to be considered in relation to another widespread Bantu process, Spirantization (§5.1). For the moment it is enough to remark that, at least for PSA, the hypothesized palatal hardening must have preceded Spirantization, otherwise palatal continuants which arise through Spirantization, as in Swahili, would also fit the structural description of the rule and become noncontinuants.

Of course this hypothesis (namely, CB *š rather than Guthrie's *c, and Sabaki Palatal Hardening) is at best tentative. Were we to unequivocally support it, we would have a clear innovation supporting the integrity of Sabaki as a subgroup. Otherwise, if we were to assume Guthrie's reconstruction, Sabaki could only be defined as a group by events that occur outside the group where those languages innovate a rule whereby CB *c > /s/ or /š/. In either case, we find Sabaki defined and forming a unit: in the former analysis, a genetic unit defined by an innovation; and in the latter, an areal group that did not share in innovations which occurred in neighboring groups. For our exposition in this volume, we take the somewhat conservative position and work from Guthrie's reconstruction of *c.

§4.1.4.1. **Proto-Sabaki *c-dentalization**. Given the total array and distribution of reflexes of *c in Sabaki, we see two possible scenarios. In the first, *c is depalatalized, first becoming a dental affricate, then a dental stop: *c > č > ts > t̪:

*c			PSA: Retained as palatal affricate /č/ in El, SD
*č	>	ts [t̪s]	Stage 1: Post PSA innovation (MK, Po, Com, ND)
*ts	>	t̪	Stage 2: Post PSA innovation (Mw, ND, Mv, Chi)

In the second, *c is depalatalized/dentalized independent of Stage 1, thus:

*č	>	ts [t̪s]	Stage 1: Post PSA innovation (MK, LP, Com)
*č	>	t̪	Stage 2: Post PSA innovation (Mw, ND, Mv, Chi)

In the first hypothesis, Mijikenda, Lower Pokomo, Comorian, and ND all shift to the dental affricate /ts/ [t̪s] as a first step, with ND subsequently moving further to the dental stop; whereas in the second, ND and Lower Pokomo-Comorian behave independently in their treatment of *č (for this treatment of ND, see Nurse 1985b). In both scenarios the fronting is unconditional. The overall effect of both seriations is dentalization, shared by most of Sabaki. It is once again Elwana, Upper Pokomo, and SD which are conservative in retaining the palatal, while the others innovate. As for the motivation for both changes, see Nurse (1985b), who argues that dentalization is an areal feature whose source is non-Bantu, probably Dahalo (Southern Cushitic) or possibly Aweera (Eastern Cushitic).[10]

[10]There are reasons to prefer one hypothesis over the other. The first claims that dentalization is an event shared by MK, LP, Com, Mwiini, and ND, whereas the second views dentalization in MK, LP, and Com as separate from the process in ND. For methodological consistency we prefer the first hypothesis, because it is in line with both the methodological bias of this chapter (see the discussion of seriation in §1.1), and

Because we are assuming that /s/ in Mwani probably resulted from the influence of an areal change affecting surrounding languages in Zone P, where CB *c : /s/, it does not figure as part of this scenario except at the highest level, where we assume it shared an original palatal affricate with the rest of Sabaki.

§4.1.4.2. *c > š. In some SD dialects *c has de-affricated to /š/. At best this is a sporadic change, not attested in all possible lexical items nor affecting all dialects equally. The reflex shows up most commonly in the rural Zanzibar dialects and in Pemba, but Sacleux (1939/41) gives others for the "Mrima" coast, and they exist in Unguja:

*-cenga 'sift, cut'	:	-šenga 'sift' (Pe), but -čenga 'cut, lop' (Ung)
*ncawa 'louse'	:	šawa (Pe), but čʰawa (Ung)
*ncuku 'cupping horn'	:	šuku (Pe), but čʰuku (Ung)
*ncoloko 'type pea'	:	šooko (Tu), but čʰooko (Ung)
*ncale 'arrow'	:	mšale (Ung)
*ɟcoka 'axe'	:	šoka (Ung, Po, El)
*ɟcavu 'cheek'	:	šavu ~ čavu (Ung) (CB *-càgụ)

This process is distinct from that in which Comorian *č has shifted to /š/ (§4.1.3). In Comorian it is generally only the palatal affricate which derives from PSA *k/___ y, i, e that shifts to /š/ whereas PSA *c in Comorian has regularly become /ts/.[11] There are probably several different sources for Swahili /š/ in these and similar examples. They may be loans.[12] Alternatively, if the rural dialects are conservative, as seen in their retention of certain PSA features (e.g., *fy and *sw), then /š/ (: CB *c) could represent evidence for a PSA stage with *š (see discussion in §4.1.4). There is also some evidence that /š/ derives from *c following Class 5 *ɟ-. This is seen most clearly in ND, e.g., *ɟ-caka 'bush' > išaka (Ti), šaka (Am, Mv), and as exemplified by place names in the Lamu Archipelago such as *i-shaka-ni* or *Shanga* (*-canga 'sand'), and the name of an area of Muqdisho,

---

the assumption that identical processes are due to common history (§1.1.1 and §1.1.2). For further discussion of these competing analyses and their implications for subgrouping, see Chap 5, §4.1.

[11]An exception is Comorian /šenga/ 'grain' < *ncenga < CB *-cèng-. This may be a borrowing from one of the rural Swahili dialects, but it could as well be the result of the Comorian rule that deaffricated *č > š) (< *k/___ y,i,e) which applied in this case because the *c-dentalization process (*c > ts), which is an earlier rule, had not applied to this item. This item is indirect support for our analysis in §3.1.3.2 that *k before *y, *i, *e in Comorian first became č before becoming /š/.

[12]Virtually all the items with this "aberrant" /š/ in SD also occur in that form in Shambala, e.g., Standard Swahili /čavu ~ šavu/ 'cheek', /čawa/ 'louse, flea', /-čenga/ 'cut', /-čoma/ 'pierce', /-čunga/ 'look after', and SD/Shambala /šawa/, /šavu/, /-šenga/, /-soma/, /-šunga/. Although there are no explicit historical records of Shambala presence on the coast, the Shambala community was the major presence just inland of SD communities in northern Tanzania for at least several centuries before the 19th. Further, Shambala is the major member of the Seuta subgroup, and Seuta (Shambala *Sheuta*) is so-called precisely because all its members all claim descent from an ancestor of the same name. Several Swahili groups along the Mrima coast (e.g., the Mtang'ata) also claim descent from this same Seuta. That would suggest that these Swahili groups at one time absorbed people whose first language was a Seuta language. The somewhat irregular pattern of distribution of these /š/ items in SD would be plausibly explained by differential absorption from Shambala over several centuries.

Shanga-ni. Such Class 5 examples (e.g., šoka 'axe') as we have listed above may well be ND loans. Here we have the same problem. Not all eligible items are affected, and those vary from dialect to dialect today.

Whatever the source of such cases of /š/, they are independent of the *č-dentalization seriation, since we are aware of no evidence in Sabaki which would show a voiceless alveopalatal fricative [š] deriving from either a dental affricate [ts] or a dental stop [t̪]. Further, given their sporadic, almost patternless distribution, we have no evidence for subgrouping.

§4.1.4.3. *c-dentalization and chronology. In only two cases are we able to say much about the chronology of this process. Rules such as *p/t-lenition do not structurally interfere with the operation of *c-dentalization, so rule-ordering is of no help in deciding which might be the earlier process. Stratigraphy does not yield results here either, since within Sabaki *c-dentalization and *p/t-lenition cover approximately the same geographical area, and approximately the same dialects. The notable difference in their respective distributions is that while *c-dentalization affected ND, lenition did not.

We can glean information about chronology in respect to the rules which have created palatal affricates, namely, *k-palatalization (§4.1.3.2) and the local ND shift in which *t became /č/ (§4.1.2.2). These must be later changes which were innovated after *č-dentalization. Otherwise, *č-dentalization would have applied to the output of these rules:

*ky-	*ki-	*c	*c	*t	
--	--	ts	ts	--	(1) *č > ts Stage 1
--	--	t̪	--	--	(2) *ts > t̪ Stage 2 (ND)
č	--	--	--	--	(3) *k > č/___y
--	č	--	--	--	(4) *k > č/___i
--	--	--	--	č	(5) *t > č (Ti, Si, Pa)
č	č	t̪	ts	č	
Sabaki	SMK	ND	MK	Ti, Si, Pa	

If Rules 3, 4, or 5 are ordered earlier, *c-dentalization will apply and give an incorrect result:

*ky-	*ki-	*c	*c	*t	
č	--	--	--	--	(3) *k > č/___y
--	č	--	--	--	(4) *k > č/___
--	--	--	--	č	(5) *t > č (Ti, Si, Pa)
ts	ts	ts	ts	ts	(1) *č > ts Stage 1
--	--	t̪	--	t̪	(2) *ts > t̪ Stage 2 (ND)
*ts	*ts	t̪	ts	*t̪	Output
Sabaki	SMK	ND	MK	Ti, Si, Pa	

Here only the ND and Mijikenda items are correctly derived from *c.

The correct derivation and ordering are also partially congruent with stratigraphical results: *c-dentalization (both Stages 1 and 2) is more widely attested than Rules 4 and 5 and would thus have to be considered an earlier change, using this metric alone. As for *c-dentalization and Rule 3, both changes have roughly the same distribution in terms of the number of dialects affected, so we can tell nothing from stratigraphy here and have to rely solely on rule-ordering for results.

§4.2. The PSA voiced consonants *W, *l, *j, *g. For PSA, in all non-postnasal stem positions, we reconstruct a voiced labial approximant, a voiced lateral, a voiced palatal, and a voiced velar stop, as exemplified in the following word-initial and intervocalic environments:

*-Wici	'raw, unripe'	*WucaWi	'witchcraft'
*-luma	'bite'	*i-jala	'rubbish heap'
*-jenga	'build'	*muji	'village'
*-gula	'buy'	*i-gogo	'log'

*W = a labial approximant

In contradistinction to the set of voiceless nonnasal obstruents (*p, *t, *c, and *k), the set of voiced nonnasal reconstructed segments is somewhat unusual in that its members do not form a neat natural class, other than the fact that they all share the feature [+voice]:

	*W	*l	*j	*g
cont.	+	+	-	-
voice	+	+	+	+
approx.	+	+	-	-

The series consists of two stops, *j, and *g, and two nonstops, *W and *l. It differs from Guthrie's CB reconstruction of stops (*b, *d, *j, and *g) for the same series (with its presupposition that [±voice] is distinctive for the series), but does approach more closely Meinhof's system (*β, *l, and *ɣ), which differs insofar as the feature [±continuant] is distinctive and voicing can be redundantly specified accordingly. Guthrie's reconstruction of CB is appealing because of its symmetry: a series of voiceless and voiced stops. Meinhof's is equally so in terms of distinctive feature specification. In contrast, the lack of symmetry in the proposed PSA reconstruction creates some doubt. Nevertheless, there is compelling evidence that such a system, as reconstructed, is reasonable and natural. There are eastern Bantu languages, as seen in the following table, which have similar synchronic systems, i.e., with a voiced labial approximant or continuant, a voiced lateral continuant, a voiced palatal, and a voiced velar stop in the environment of a following [-round] vowel

Thus, PSA's asymmetrical system for the voiced series with its many parallels in modern Bantu languages is well motivated on comparative external evidence. There are, of course, languages with a symmetrical system. For example, Manda (N11) and Tumbuka (N21) have nonstop series, respectively /ʋ, l, and ɣ/ and /β, r, and ɣ/.

Guthrie CB:[13]	***p**	***t**	***c**	***k**	***b**	***d**	***j**	***g**
PSA	*p	*t	*c	*k	*W	*l	*j	*g
Gogo (G11)	h	t	h	k	w	l	z	g
Nhwele (G32)	h	t	s	k	w	l	z	g
Pogolo (G51)	p	t	s	k	w	l	z	g
Sango (G61)	p	t^h	h	x	β	ř	s	g
Hehe (G62)	p	t	h	k	ʋ	l	s	g
Kinga (G65)	p	t	h	x	β	l	ts	k
Yao (P21)	p	t	s	k	β	l	s	g
Pimbwe (M11)	p	t	s	k	β	l	ʔ	g
Safwa (M25)	p	t	s	k ~ x	β ~ w	l	z	g
Nyakyusa (M31)	p	t	s	k	β	ř	ʄ	g
Manda (N11)	p	t	h	k	ʋ	l	ʄ	ɣ
Maŋanja (N31c)	p	t	s	k	Ø	l	z	g

The internal evidence that supports the reconstruction is not as straightforward as we would like. Given the variation of the attested reflexes for *W and *g in Sabaki some further consideration must be given to these reconstructions, especially to *W and its feature specification. The reconstructions of *l and *j are less controversial, and the reflex evidence clearly supports them. In each case, however, we need to address the question of whether these are inherited from a pre-PSA stage or are PSA innovations.

§4.2.1. Sabaki *W. There is a great deal of variety in the reflexes of the Sabaki labial approximant *W (: CB *b). Some of this diversity, however, is due to internal dialect variation. Further difficulty is created by conflicting reports given by various scholars who have recorded these languages. In the following discussion we will attempt to clarify the situation. Initially we will only consider *W in nonround vowel environments:

PSA *W/___[-round][14] (Chart 7)

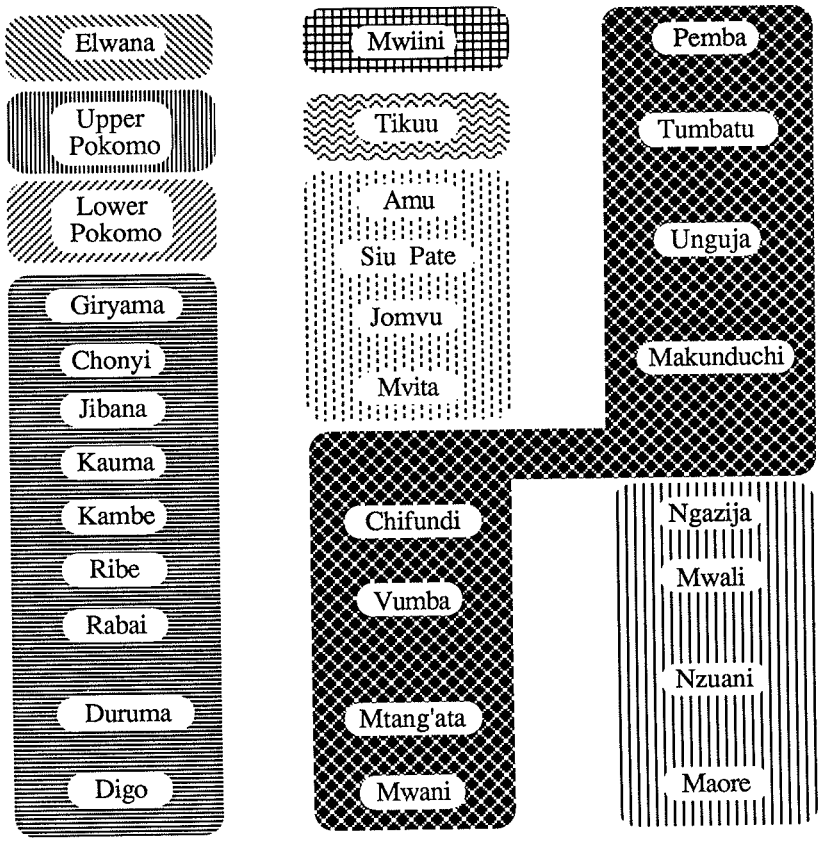
β ~ w	El
v	Ti, UP-Milalulu
ʋ ~ β	UP
ʋ ~ w	LP

[13]Data based on Guthrie (1971); for Nhwele we follow Nurse (1979a) who gives /w/ : CB *b in contrast to Guthrie's Ø; /ʋ/ is a labiodental approximant.

[14]There is considerable variation in the reports on the reflexes of PSA *W in the literature. For Mwiini, Whiteley (1965) records /β/, Abasheikh (1978) records /w/, and M. Bakari (pers. comm. 1986) reports /ʋ/, a labiodental approximant. Reports also vary for different dialects of Pokomo: Ehret (pers. comm.) for UP (Milalulu) records /v/, Nurse (field notes) for UP (Ndera) gives /β/, Hinnebusch (1973; field notes) for UP (Masalani) records /ʋ ~ β/, and for LP (Garsen) /ʋ ~ w/, Nurse (field notes) gives /w/ for LP (Ngao). Reports from Ipu (field notes), Philippson (field notes), Townsend (field notes), and Van Otterloo (1980) show a diverse situation for Elwana as well. The display above represents what is found before the vowel /a/; it does not reflect the correspondence relationship in all environments nor all allophonic variation.

Chart 7

Sabaki *W



Key

*W : w

*W : υ

*W : w ~ Ø

*W : Ø

*W : β ~ υ

*W : β ~ υ ~ v

*W : v

*W : υ ~ w

*W : w ([υ], [β])

ʊ	ND (Bakari 1985:238)
w	Ung, Pe, Vu, Mak, Mn (?)
Ø (Ø ~ [h])	MK ([h] insertion is haphazard; see fn. 15)
w ([ʊ], [β])	Mw
w ~ Ø	Com[15]

*-WaWa 'itch'	-wawa (LP,Ung,Ng,Nz,Mn), -ʊaʊa (Am), -vava (Ti)
*-Wi(Wi) 'bad'	-bɛ́ɛ̀bɛ (El), -ii (Gi,Du), -vi (Ti), -ʊi (Si,Pa,Am,Mv), -wi (Pe,Vu), -i (Ng)
*ncaWa 'louse'	jawa (El), nčaβa (UP), ntsawa (LP), ts^{h}aha (Gi,Ch,Du), ts^{h}aa (Di), t̪^{h}aʊa (Am,Pa,Si,Mv), t̪ava (Ti), nt̪awa (Mw), č^{h}awa (Ung), sawa (Mn) (see App. 3)
*-laWa 'go away'	-yawa (LP,Tu), -dawa 'start' (LP), -laa (Gi), -ɬaʊa (Mw), -awa ~ -lawa (archaic) (Ung), -yawa ~ -lawa (Mak), -awa ~ -yawa (Pe), -awa (Vu), -lawa (Ng,Nz,Mn)

For other examples, see *mwiWa 'thorn', *njiWa 'dove sp.', *kiWanja 'open area', *-leWa 'be drunk', *muWa 'sugarcane', *WucaWi 'witchcraft', *maWele 'millet', *-Wili 'two' and *-Wika 'crow' (App. 2).

Each of these reflexes, except of course Ø in Mijikenda, could be a reasonable candidate for the reconstructed segment. What is certain and well-supported is that PSA had a segment that was, at least, [+continuant, +labial, +voice]. This is the minimal feature specification for the Sabaki proto segment. What is less clear is whether the proto segment was an approximant, i.e., *w, or nonapproximant, i.e., *β, and if the former, whether it had bilabial or labiodental articulation. The evidence is somewhat ambiguous in regard to both questions.

The comparative evidence from non-Sabaki NEC suggests that pre-PSA had *w, a labial approximant, rather than either *β, a fricative, or *b, a stop. Sabaki's nearest relatives, Seuta and Ruvu, have /w/, a bilabial approximant, and thus support the reconstruction of an approximant for NEC. Some East Highlands languages, such as most of Lacustrine Bantu, and Sukuma/Nyamwezi, support *β, but these have no direct bearing on Sabaki except at rather more remote proto periods. There is almost no support in East Highlands Bantu for reconstructing a labial stop (see Guthrie 1967:71, topogram 6). Where stops are attested, they are found in specific environments and appear to be local secondary developments.

Some Southern Highlands (SH) languages, however, have /β/ as well as /v/, or /ʊ/, a labiodental approximant, somewhat comparable to what we find in the northern Sabaki languages. Because SH presumably has a relatively closer genetic affinity to NEC than other Tanzanian Bantu languages (see Hinnebusch, Nurse, and Mould 1981), this is rele-

[15]Comorian also attests /h/ where PSA *W has been deleted. Such /h/-insertion is probably a late secondary development and not a reflex of *W per se.

vant evidence in helping to decide whether pre-PNEC had *β, *ʋ , or *w. But because the two sources we rely on here for SH data differ in their reports, it is not completely clear what we want to reconstruct for SH. Guthrie (1971) records /ʋ/, a labiodental approximant, for Hehe, and /β/ for the rest of SH. Nurse (1979a), on the other hand, records /v/ for Hehe, Sangu, and Bena, while the rest of SH has /β/. Guthrie's report suggests *ʋ vs. *β; Nurse's suggests *v versus *β. We could compromise by choosing *β, or decide on a reconstructed labiodental segment, either *ʋ or *v. There is no principled reason for deciding on either, but here the important point to be made is that the data support a labiodental articulation at some point in the derivation of these languages, and possibly a reconstructed labiodental segment. Thus, if pre-NEC had a labiodental, it is possible PNEC had a labiodental, albeit at this juncture a continuant, namely, *ʋ, which was preserved and bequeathed to PSA and is directly attested in some of the modern Sabaki languages. While *β, or less likely *v (which also have their modern counterparts in SH and northern Sabaki), are possible, we rule them out in favor of, at least, a labial approximant *W. This is far more frequently attested as either [w] or [ʋ] in the modern languages under consideration. We also rule out *β and *v for PSA, because where [β] and [v] exist in Sabaki they would appear to be secondary developments, given their environments and distribution.

Whether in fact PSA *W had bilabial or labiodental articulation is difficult to decide. The modern reflexes can be derived equally plausibly from either. From *ʋ the presently attested labiodental approximants in ND can be directly derived, as well as the labiodental fricative in Tikuu and Milalulu Pokomo via a simple strengthening process (*ʋ > v). On the other hand, a rule changing the feature specification from labiodental to bilabial would give the bilabial approximants in SD, Comorian, part of Pokomo, and Mijikenda (which undergoes the further shift of /w/ to Ø). Derivations from *w would be equally plausible via a feature change which would yield /ʋ/ for northern Sabaki, and in turn /v/ by strengthening. The labiodental could also be regarded as the intermediate source of /β/ where it has been recorded in northern Sabaki. An intermediate labiodental approximant is also the likely source of the contradictory transcriptions for Mwiini and those found elsewhere.

Mwiini is reported by Whiteley (1965) to have a bilabial fricative, but it is clear from Abasheikh (1978) and Bakari (1985; pers. comm.), and Kisseberth (pers. comm. 1986) that the segment in question is an approximant, and not a fricative. Moreover, as confirmed by Bakari (pers. comm.) it has labiodental articulation. The apparently mistaken recording of /β/ for /ʋ/ is understandable if the recorder was paying more attention to the acoustic aspects of the sound than to the role played by the articulators. In our own experience with recording ND and Pokomo, it was easy to miss both the labiodental articulation and the *varying* degrees of approximation that exist.

Although it seems to us that where [β] exists in Sabaki it is due to secondary developments, as well as to phonetic variation arising from a proto labiodental approximant, there is the possibility that [β] is a relic, a retention from an earlier proto period. Meinhof reconstructed PB with *β, so the possibility exists that /β/, where it is found (except where

derived from *p), is a holdover from an earlier PSA *β. We take the opposite view and argue that /β/ is the result of post-PSA events.

There are several arguments in support of this position. First, and most important, the evidence for reconstructing *β for Sabaki and Ruvu, as outlined above, is not very convincing. To assume *β for PSA means assuming it for the rest of NEC, and this is not well supported. Second, the preponderance of languages and dialects within Sabaki with labiodental articulation (/ʋ/ in ND, /v/ in Tikuu and Milalulu Pokomo) points away from *β. Third, where /β/ does exist in modern languages, in both Elwana and Pokomo and in the Southern Highlands languages, it always appears to occur where there are closely associated dialects which have either a labiodental approximant or labiodental fricative. For example, in Upper Pokomo the more northerly dialects have /v/ (e.g., Milalulu), and the more southerly ones have /β/ (e.g., Ndera). Farther south in the Kinakomba Upper Pokomo dialect, both [β] and [ʋ] were recorded (TJH, field notes 1972-3). The greatest number of items, however, were recorded with [ʋ]. In some cases the same item fluctuated in pronunciation between [ʋ] and [β], with various degrees of approximation of the articulators, and varying, as well, between bilabial and labiodental articulation. And, finally, Lower Pokomo has /ʋ ~ w/. Since in other respects, the Upper Pokomo dialects are conservative, the chain of forms from north to south (v – β – β ~ ʋ – ʋ ~ w) calls for an original dental, rather than bilabial, source. This association is easily explained by positing an intermediate *ʋ which can yield /ʋ/, /v/, /β/, or /w/. Fourth, attestations of [β] in Elwana (as in our reconstructed PSA lexis) tend to be found in specific environments, i.e., before or after a front vowel or front (palatal) glide, e.g., [βebe ~ βiβi] 'bad', [iβingu] 'cloud', [-βina] 'dance', [-iβa] 'steal', [mukeβa] 'deserted person', [-βiba] 'thatch'. Also, where Ipu (field notes) and Van Otterloo (1980) agree in recording [β], it also tends to be found in the same environment: *mavi : [maβi] 'excrement', *vyala : [βyala] 'give birth', *mwivi : [mwiβiani] 'thief'. The apparent operative rule here is one which takes a labial segment, either *v (from Bantu Spirantization) or *ʋ, to [β]. Furthermore, there is evidence that the rule has been generalized to include other segments. In Elwana, some items with reflexes deriving from *p, which otherwise would generally be realized as a voiceless stop, are found with [β] before a front vowel or glide: *pa : [βeya] 'give'; *fyagila (CB *-piagid-) : [βyela] 'sweep'; *-ogelela : [-oβela] (Van Otterloo 1980) : [-opewela] (Ipu) 'swim'. That reflexes of *p are involved seems good evidence that the appearance of [β] in Elwana is a recent phenomenon. Though it is not clear why labials should shift to [β] before a front vowel or glide, the facts seem clear enough to rule out the possibility that the bilabial fricatives in northern Sabaki are relics from an earlier proto period, even though in some cases we are not sure of the quality of the transcriptions we have.

Thus, we have two conflicting hypotheses, with a slightly different set of rules, that can explain the distribution of modern Sabaki reflexes: (1) where PSA *W = [ʋ], a labiodental approximant, and (2) where PSA *W = [w], a bilabial approximant. If case 1, we would then have the following series of shifts:

*ʊ > w
> v > β (/ʊ/ is retained in ND; SD, etc. innovate bilabial articulation)

In case 2, we would have:

*w > ʊ > v > β (/w/ is retained in SD; ND, etc. innovate labiodental articulation)

At this time we cannot choose between these alternatives. Nevertheless, the implications for PSA are not trivial. With *ʊ in the first case we can assume either that PSA innovated the labiodental articulation from a bilabial approximant, which it presumably inherited from NEC, or that the labiodental was inherited from some deeper node. On the other hand, if we assume the *w in case 2 for PSA, then the labiodental articulation is a post-PSA innovation affecting only part of the Sabaki dialect cluster. One argument in favor of *w is the fact that Seuta and Ruvu, except possibly Gogo, all have /w/. Furthermore, to assume that Sabaki is a NEC subgroup and that *ʊ is a PSA innovation means that we have the following chain of changes:

PNEC *w > PSA *ʊ > w (SD, Com)
> ʊ (ND, etc.)

Although this hypothesis requires a shift from bilabial articulation (PNEC) to labiodental articulation (PSA) back to bilabial articulation (SD, Com), it will be our working hypothesis for outlining the internal Sabaki changes in the next section. However, in recognition of the tentative nature of this analysis, we use the neutral symbol *W here and in Appendix 2. Finally, given the distribution of /β/ in East Highlands Bantu, we assume that *β may well be the ultimate source of PSA *W. The same data also support our contention that PSA developed from a pre-PSA series *β, *l, and *g.

§4.2.1.1. Changes affecting PSA *W. The labial approximant behaves differently depending on whether the following vowel is [+round], or [-round]. Before [-round] vowels, *a, *e, *i, we can set up the following seriation:

*W > ʊ > w > Ø (1)
> v > β (2)

(1) *W > ʊ PSA innovation (/ʊ/ still attested UP, ND, Mw, (El?))
*ʊ > w Stage 1: Post-PSA (LP, SD, Com, MK?, Mn?)
*ʊ > Ø Stage 2a: Post-PSA (MK?)
*w > Ø Stage 2b: Post-PSA (Com); before *i, and *e

(2) *ʊ > v ~ β Stage 3a: Post-PSA (Upper UP)
*ʊ > v Stage 3b: Post-PSA (Ti)

This seriation is attractive because again it is the northern area of the Sabaki continuum which retains the assumed earlier dental articulation of the approximant. Some dialects, however, further shift *ʊ to /v/, maximizing the dentality of the segment. This also correlates quite well with what we saw developing with *c-dentalization, an innovation

which appropriated dentality in the fronting of *c. With Nurse's work (1985b) that argues for a non-Bantu external source of dentality, we might want to maintain, instead, that the labiodental articulation of *W is a Northern Sabaki innovation, rather than retention. We of course cannot be sure whether Mijikenda also shared [+dental] with ND prior to the deletion of Mijikenda's reflex of *W, but that it does share dentality with ND in other cases suggests that it did.

§4.2.1.2. ***W/___[+round]**. Before [+round] vowels, *W is generally deleted throughout Sabaki. There are exceptions, and some indeterminacy. In Tikuu, *W has regularly become /v/, e.g., /vač^{h}u/ < *Wantu; but before [+round] vowels we have a mixed situation: (a) before *o, *W has been retained as /v/, e.g., /nguvo/ 'clothing' < *nguWo, /kiovo/ 'hook' < *kiloWo; (b) before *u, deletion is the rule within stems, and for Class 14 *Wu- on C-stems, e.g., /u̪avi/ 'witchcraft' < *WucaWi 14, /-eua/ 'inspect' < *-laWula; (c) for Class 14 *Wu- on V-stems and M-stems we find, e.g., /uvongo 'earth', /vuongo/ 'clay' < *Wulongo; /vuso/ 'face' < *Wuso, /uvuča/ 'bow' < *Wuta. The final set suggests that Tikuu had regular deletion of *W before *u except in the case of Class 14 V- and M-stems, which have subsequently undergone, or are undergoing, paradigmatic leveling, a morphological change, to fit the normal pattern seen in Class 14 C-stems

Mwiini shows a similar pattern: (a) *W-deletion before *u in all cases, e.g., /uso/ 'face' < *Wuso, /-umba/ 'create' < *-Wumba; (b) retention before *o, e.g., /-aʊo/ 'their' < *-aWo, /-ʊoɬa/ 'rot, be rotten' < *-Wola, /-ʊona/ 'see' < *-Wona. An interesting exception is /čiloho/ 'hook' < *kiloWo. Some examples illustrating the Sabaki patterning here are:

*WucaWi 'witchcraft'	učawi (El,Ung), učaβi (UP), utsai (Gi), u̪aʊi (Am,Mv), u̪avi (Ti), učwai (Ng), usawi (Mn)
*Wulili 'bedstead'	uriri (Gi), ulili (Am,Mv,Ung,Nz)
*Wulongo 'soil'	ulongo (Gi,Du), uɬoongo (Mw), uʊongo ~ uongo (Am), uvongo 'earth' (Ti), vuongo 'clay' (Ti), udongo (Ung,Mv)
*Wuso 'face'	oso (El), ʊuso (ND obs), uso (LP,Gi,Du,Am,Mw,Mv, Ung,Ng,Nz,Mn), vuso (Ti)
*Wuta 'bow'	u̪a (El), uha (LP,Gi,Du), uta (Am,Mv,Ung,Ng), uvuča (Ti)
*nguWo 'clothing'	guo (El), nguo (LP,Gi,Du,Am,Si,Mv,Ung,Ng,Nz,Mn), nguvo (Ti), nguʊo (Mw)
*kiloWo 'fishhook'	kilowa (El), kiloa (LP), kiloo (Gi), kyiloo (Du), čiɬoho (Mw), kioo (Am,Si,Pa,Mv,Ung), kiovo (Ti), iloo (Ng); cf. uloo 'gaff' (Ung)

The loss we see in these examples is quite natural and based on the feature configuration of *W, a [+labial] approximant, followed by a [+round] (= [+labial]) vowel. The approximant assimilates the features of the following labial vowel and consequently deletes,

thus *WucaWi > uucawi > Øucawi 'witchcraft'. It is probable that the deletion was an automatic consequence of the assimilation, which never shows up on the surface.

A similar motivation is responsible for the apparent deletion of *W following round vowels; the approximant is assimilated to the labiality of the preceding round vowel and deletes, e.g., *ɻtuWi 'coconut juice' : /tui/ (Am, Ung), *ncuWi 'leopard' : /čʰui/ (SD), /ʈʰui/ (Am), *-zoWela 'be used to' : /-zoea/ (Ung, Am, LP), /-zoera/ (Gi), /-zea/ (Ng), etc. However, there are exceptions: *luWingu 'sky, heaven' : /uwingu/ (Ung) : /uʊingu/ (Am), /yuʊingu/ (LP), *nguluWe 'pig' : /nguyuwe/ (Po), /nguuʊe/ (Am), /nguuwe/ (Mak), /nguruwe/ (Ung), /ngulu(w)e/ (Mn). It should be said, however, that the presence or absence of *W in these cases may be due to inconsistent application of an orthographical convention, or improper synchronic phonemic analysis. We suspect that in the underlying synchronic grammars of these languages we should posit /w/. This has some comparative support in Tikuu, which attests /v/ where we expect to have /w/ in the other dialects: /ičuvi/ 'coconut juice', /-ðovea/ 'be used to', /ʈʰuvi/ 'leopard', etc. This data, at least, points to a late stage when Sabaki, including Tikuu, attested an approximant in post round vowel environments. On the other hand, in some cases in some languages [w] may simply be a transitional glide which is sometimes given orthographic representation, e.g., Swahili *nguwo* 'clothing' instead of *nguo*; other examples are frequently seen in various redactions of Swahili verse. These might well represent retentions from an earlier proto level. It is difficult to say, and thus we will not consider the diachronic implications of this aspect of *W-loss.

In the two cases of *W-loss so far described, feature-identity underlies segment-loss. This, however, will not explain why *W in Mijikenda dropped out in all vowel environments, e.g., /atʰu/ 'people' < *Wantu, nor why in Comorian it was deleted before the front vowels, */i, e/, e.g., *-Wik- 'crow' : /-iha/, *-Wek- 'put' : /-(y)eha/). Deletion in Mijikenda could be interpreted as a simplification and extension of the rule which deleted *W before */u, o/, thus:

*W > Ø /___V [+round]	Stage 1: Sabaki (see above for exceptions in Ti, Mw)
*W > Ø/___V	Stage 2: Mijikenda

This is the analysis in Hinnebusch (1973) and Hinnebusch, Nurse, and Mould (1981:44ff.), and is the one adopted here. Comorian represents a similar but different sort of simplification, where *W is deleted before all vowels except the [+low] vowel /a/. Comorian first deleted *W before [+round] vowels, a shared event with the rest of Sabaki, and then extended the rule to matching [-low] vowels, only retaining the approximant before the [+low] vowel. Certain constraints operated here; for example, if *W is the Class 11 prefix vowel, it will not delete: /w-embe, ny-embe/ 'thin, sharp knife' (*lwembe), /w-ingu/ 'heavens' (*lu-Wingu) (all examples from Nzuani in Ottenheimer 1986).

In response to loss, some Sabaki dialects have introduced other segments to break up the resulting VV sequences, thus *kiloWo 'fishhook' > /čiɬoho/ (Mw), *-Weka 'put' >

/-yeha/ (Ng), *-Wila 'boil up' > /-(h)ia/ (Ng), *ncaWa 'louse' > /ts^{h}aha/ (Gi, Du, Ch). It is not certain without further synchronic analysis how systematic this is.

§4.2.1.3. *W rules and chronology. Most of the rules affecting *W are local internal Sabaki processes, and therefore chronologically late. Only two have affected Sabaki generally. The total set that we need to consider, in the order determined by stratigraphy is:

***W-loss**

1. *W > Ø/___ [+round] (most Sabaki dialects; exceptions in Ti and Mw)
2. *W > Ø (MK)
3. *W > Ø/___ [+front] (Com)

***W-hardening**

4. *W > v (Ti, part of UP)
5. *W > [β]/___ [+front, +high] (El, part of UP)

Rule 1, *W-loss before [+round] vowels, is not just a Sabaki rule, but is shared with most NEC and eastern Bantu languages (see Nurse 1979a, App. 2, passim). Stratigraphy would indicate early innovation for this process in East Highlands Bantu, and thus presumably it might not be a PSA rule at all, but inherited from a pre-PSA level. However, were it not for the fact that Tikuu and Mwiini preserve a reflex of *W, we would have argued that PSA inherited Ø in [+round] environments. So at least for Proto-Swahili we have to posit *W before *u and *o. And to preserve the notion of a uniform proto language for PSA, we assume *W in such environments for all Sabaki. The reality of the situation may be something else. We are dealing with a change that combines genetic and areal elements. First of all, the Tikuu data argue against an inherited Ø before round vowels from a pre-PSA stage, but the uniformity of the correspondences throughout eastern Bantu indicates an early innovation, and early spread of the change to the Sabaki continuum, which must have already been somewhat diverse to get such differential behavior within Sabaki. When the change did spread to Sabaki, Tikuu and Mwiini, on the far northern coast, were already sufficiently distinct to behave differently in this regard. Other widespread eastern African Bantu changes present a similar picture (see Spirantization, §5.2).

On the other hand, given the naturalness of the underlying process leading to *W-loss before [+round] vowels, the uniformity of attestation over a wide geographical area could well be due to a number of separate independent innovations, but without evidence supporting such a hypothesis we will not consider it further. In any case, *W-loss before [+round] vowels is an early Sabaki phenomenon, and has to be ordered first in this subset of processes involving *W-loss.

Earlier we hypothesized that Rules 2 and 3 are extensions (via rule simplification) of *W-loss before round vowels. The interesting question before us now is whether 2 and 3 are independent innovations or shared Comorian-Mijikenda events, namely:

Stage 1: *W > Ø/___u,o (post or early post-PSA, including Com, MK)
Stage 2: *W > Ø/___i,e (Com, MK)
Stage 3: *W > Ø/___a (MK)

This is attractive because the results of stratigraphy and phonological seriation match perfectly, and it sets up a further piece of evidence supporting a link between Comorian and a specific subpart of Sabaki and is in line with other shared innovations linking Comorian and Mijikenda, e.g., lenition of *p, *t (Charts 1 and 2). However, it is not a strong piece of evidence, because these are innovations involving loss and could therefore represent independent convergence.

Rule 4, *W-hardening, only affects one ND dialect, Tikuu, and the northern Upper Pokomo dialects. Since there is considerable territory, and a number of dialects which do not share this change, separating Tikuu from northern Upper Pokomo, it is unlikely that the two events are connected other than the fact that they share a common intermediate stage, the dental articulation of *W [ʋ]. On the other hand, cultural influence on Upper Pokomo, emanating from the Lamu area, was strong, and we may argue for a historical connection (Arabic /tena/ 'again' is /čena/ in Upper Pokomo and Tikuu).

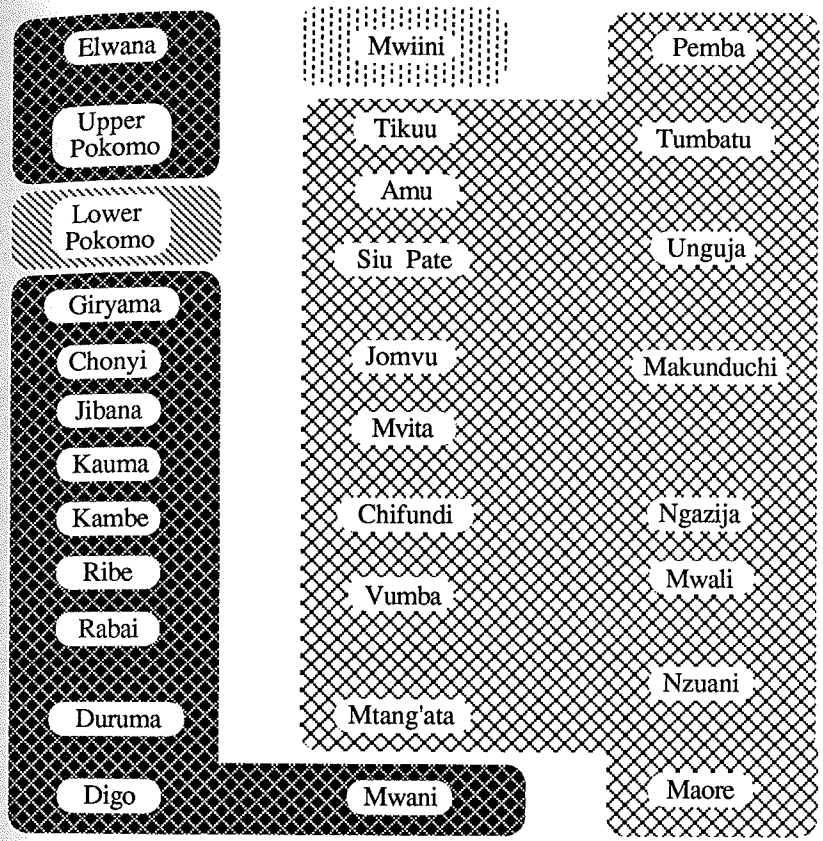
Rule 5 is specific to several Upper Pokomo dialects and is a late development, also dependent on the dental articulation of *W [ʋ].

Finally, we made a decision early in the discussion of *W which has subgrouping implications. We first of all argued for a Sabaki-wide dental articulation of *W, namely, [ʋ], and a subsequent de-dentalization, namely, ʋ > w. The subgrouping effect is to contrast SD and Comorian against ND, Mw, Upper Pokomo, and probably Elwana, with Lower Pokomo apparently falling in with the de-dentalizing dialects, though in the latter there is variation (/ʋ ~ w/). We unfortunately don't know how Mijikenda should fit in here, since it has deleted its reflexes of *W. We observed earlier that because Mijikenda may share *c-dentalization with ND, it is likely that it might have shared dentality of *W with ND; even so, we are reluctant to subgroup on the basis of this. Though our analysis of *W in this section is speculative, the classification based on it roughly correlates with the behavior of Sabaki *c, which divides the continuum into major north/south groups (see §4.1.4.3 above).

§4.2.2. PSA *l. It is clear that PSA inherited *l from an earlier period. Most NEC languages still attest *l (Nurse 1979a). Where *l has been deleted, its reflex can frequently be reconstructed internally for most Sabaki dialects, except possibly Lower Pokomo, which regularly attests an *l-weakening rule (*l > y/[-stop]___), which presumably also underlies the loss of *l in some other Sabaki dialects in specific contexts. A statement of correspondence relationships without stating environments is complex, especially where loss is involved; the following is only a simplification which will be discussed in more detail in the following paragraphs.

Chart 8

Sabaki *l



Key

*l : l ~ (r)

*l : ɬ ~ l

*l : y ~ Ø

*l : l ~ Ø ~ (y)

*CB *d : PSA *l (Chart 8)

ɬ ~ l	Mw
l ~ (r = [r] or /r/)	El, UP, MK, Mn
l ~ y	Lower UP
y ~ Ø	LP
l ~ Ø ~ (y)	Sw, Com

*-lil- 'cry'	-lila (El,Ng,Ma), -iya (LP), -rira (Gi,Du,Mn), -la (Mw), -lia (Am,Mv,Ung,Nz)
*-leW- 'be drunk'	-rea (Gi,Du), -ɬeewa (Mw), -yewa (Am), -ewa (Pa,Si), -eva (Ti), -lea ~ -rea (Jo), -lewa (Ung,Ng,Nz), -rewa (Mn)
*-leet- 'bring'	-yeha (LP), -reha (Gi,Du), -ɬeeta (Mw), -yeta (Am), -eča (Ti), -leta (Mv,Jo,Ung), -rera (Vu), -yera (Ng), -erea (Ng)
*-lum- 'bite'	-lúma (El), -luma (UP,Gi,Ch,Ra,Du,Mn), -luma (ND obs), -uma (LP,Am,Pa,Si,Ti,Jo,Mv,Ung,Vu,Chi,Mt,Ng), -ɬuma (Mw)
*-lol- 'look at'	-lola (El,Gi,Du,Ng,Mn), -yoa (LP), -ɬooɬa (Mw), -olewa 'be seen' (Am), -owa (Mv), -ola (Vu,Mt), -yoya ~ -lola (Mak), -ola ~ -oa (Pe), -lolea 'think about' (Ng)
*-laal- 'sleep'	-yaa (LP,Chi,Vu,Mt), -lala (Gi,Du,Am,Si,Pa,Mv,Jo,Ung,Pe, Ng,Nz,Ma,Mn), -rara (Di), -ɬaaɬa (Mw), -yala (Ti), -yaya (Tu)

§4.2.2.1. Changes affecting PSA *l. Two types of change affect *l in Sabaki. Mostly we find innovation leading to loss of the reflex; the other type involves feature changes. We will discuss cases of loss first. The Sabaki dialects which regularly show evidence of *l-loss are Lower Pokomo, ND (but not Mwiini), SD, and Comorian. The others conserve a consonantal reflex of *l, namely, Mwiini, Upper Pokomo, Elwana, Mijikenda, and Mwani. Various stages of *l-loss can be distinguished, depending on the quality of the following vowel and whether it is stressed or not. We find the following regularities in those dialects attesting loss (examples can be found in App. 2):

*l/___u [±stress]	>	Ø	Sw, Com, LP	(*mulungu, *mujukulu)
*l/___o [-stress]	>	Ø	Sw, Com	(*kivyalo, *nkolo)
*l/___o [+stress]	>	Ø	Sw, Com	(*mulongo, *-polozya)
*l/___a [-stress]	>	Ø	Sw, Com	(*-ikala, *-kali)
*l/___a [+stress]	>	Ø	Sw, (Com)	(*-kalanga, *mulala)
*l/___e [±stress]	>	Ø	ND, LP	(*nkelele, *-kale)
*l/___i [±stress]	>	Ø	LP	(*-lila)

Lower Pokomo and adjacent Upper Pokomo dialects attest a stage which we assume underlies all examples of *l-loss in Sabaki. The *l in all environments shifts to /y/ and then deletes, but only before the front vowels */i, e/. There is ample evidence that a similar shift

of *l > y underlies loss where we find it in the other dialects. For example, in Comorian, *l after *a and before an unstressed *a is realized as /y/ or Ø, e.g., *njala 'hunger' : /ndzaya ~ ndzaa/ (Ng), /ndza/ (Nz, Ma), *-katala 'refuse' : /-haraya/ (Ng), /-hara/ (Nz). Other evidence comes from some SD, where /y/ or Ø is the reflex of *l before *a and *o:

*-lola 'look at'	-yoya (Tu,Ma), -ola (Vu,Mt)
*-laWa 'go away'	-yawa ~ -lawa (Ma,Tu), -yawa ~ -awa (Mt), -awa (Vu)
*-laala 'sleep'	-yaya (Tu), -yaa (Mt)
*-lola 'marry'	-lola ~ -yoa (Ma), -yoya (Pe), -oa (Vu)

Amu, Siu, Pate, and Tikuu also attest /y/ from *l in certain items, usually before /e/ and sometimes before /a/, e.g.:

*-leWa 'be drunk'	-yewa ~ -ewa (Am)
*-lemela 'be heavy'	-yemea (Am)
*-lala 'sleep'	-yaa (Am), -yala (Ti), -yaða 'put to sleep' (Pa,Si,Ti)

On the basis of this evidence, then, we will assume that an intermediate shift to *y underlies all cases of Sabaki *l-loss.[16]

*l-loss is most regularly attested before both a stressed and an unstressed *u in ND, SD, Comorian, and Lower Pokomo. This interestingly parallels the behavior of *W in the same environment, and suggests that both have a similar provenance. However, we argued above that *W-loss was due to an assimilation process, i.e., feature identity between *W and *u (§4.2.1.2). We cannot argue for the same process here because of the lack of feature identity between *l and *u, or intermediate *y and *u. Nevertheless, it is likely that *W-loss set up the situation for *l-loss. Given the uniformity and regularity of *l-loss before *u, we assume an earlier innovation for this phase of loss than for others.

Loss of *l before *o is not as uniform as loss before *u and is more complex to state. For example, in Comorian most examples of *l before unstressed *o are realized without a reflex of *l, e.g., *lucelo : /utseo/ 'winnowing tray'. There are also a number of examples of *l before stressed *o which are also realized with Ø, e.g., /-hooa/ 'cough' (< *-kolola), /šondra/ 'wound' (< *kilonda), as well as examples where *l in a stressed syllable has not been deleted, e.g., /-lowa/ 'be wet' (< *-loWa), /-loza/ 'marry' (< *-loza), ɗomo/malomo 'mouth' (< *-lomo), ɗongo/malongo 'soil, clay' (< *Wulongo), etc. The reasons for these exceptional items are not clear. Strengthening (§6) may play a role in the latter two examples. The near-regular attestation of loss in unstressed environments indicates initiation of the innovation in this environment first, followed by generalization to the stressed syllable.

In ND, *l-loss in stressed and unstressed syllables before *o is regular, and is nearly so in SD, e.g., *kivyalo 'offspring' : /kizao/ (Am), /kivyao/ (Ung), *-poloza 'be withered'

[16]Albeit our claim here, B. Wald (pers. comm.) has pointed out on recordings that he made in the 1970s many instances of speakers of Standard Swahili dropping /l/ without any evidence of an intermediate stage.

: /-poza/ (Am), /-pooza/ (Ung). In SD there are a few cases of stem-initial *l/__*o in stressed syllables which do not delete, e.g., *-loWa 'be soaked' : /-lowa/. These may be borrowed from other dialects, or may be relics. There are far fewer exceptions of this sort in SD than in Comorian: Comorian has preserved twice as many stem-initial *l's in stressed syllables as SD—an indication, perhaps, that Comorian lagged behind SD in deleting stressed *l before *o.

The behavior of *l before *a is roughly similar. In Comorian, before unstressed *a, most cases of *l in the PSA lexis have been deleted, e.g., *-zyula 'pull up' : /-dzua/ (Ma), etc. In stressed syllables, most instances of *l have been preserved, e.g., *-laWa : /-lawa/ 'go away'. Whether stressed or unstressed, more examples of stem-initial *l have been preserved than lost, e.g., *-lalika : /-lalika/ 'invite', etc. Some examples of unstressed *l have been interestingly preserved where the stem-initial *l is untouched by change: /-lala/ (Ng) 'sleep', /-lela/ 'rear, raise' (Ng), but /-lea/ (Nz), /-lola/ (Ng, Nz, Ma) 'marry', /-lola/ (Ng) 'look at'.

In ND, and SD we find *l fairly consistently deleted in an unstressed syllable before *a. ND also has deleted *l if *a is stressed, whereas in SD *l is more often preserved than deleted in the same environment. There are a few counter-examples, though they are rare compared to the same situation in Comorian.

In ND *l-deletion has been extended to all vowel environments except before *i, e.g., *-lila 'cry' : /-lia/ (Am) , *-linda 'guard' : /-linda/. Furthermore, in some cases /y/ still appears as the reflex, usually before /e/ and sometimes /a/, e.g., *-lala 'sleep' : /-yaa/ (Am), /-yala/ (Ti), /-yaða/ 'put to sleep' (Pa, Si, Ti), *-leWa : /-yewa ~ -ewa/ (Am) 'be drunk', *-lemela : /-yemea/ (Am) 'be heavy'. Only Lower Pokomo, which preserves /y/ generally, deletes *l (via *y) before *i, e.g., *muWili 'body' : /mwii/. All other Sabaki dialects preserve a reflex of *l before *i.

The various facts outlined above suggest a more detailed two-part seriation for those dialects attesting loss. In each case we are assuming that loss is dependent on the intermediate shift of *l > y which first began in unstressed syllables and was extended in various degrees to stressed syllables, the details of which are not reflected in the following diagram:

*l > y /___ [-stress]	Sw (less Mw), Com, LP
*l > y > Ø/___u	Stage 1: Sw (less Mw), Com, (LP)
*l > y > Ø/___o	Stage 2: Sw (less Mw), Com
*l > y > Ø/___a	Stage 3: Sw (less Mw), Com
*l > y > Ø/___e	Stage 4: ND

Lower Pokomo participates in the above in attesting the intermediate shift of *l > /y/. It only shows deletion before *i (/y/ > Ø/___i) and before unstressed *u, e.g., *ɟjulu 'up' : /dzuu/, but not in Class 11 /yu-/ or before stressed *u, e.g., *nkulungu 'bushbuck' : /nkuyungu/. In ND, SD, and Comorian, the Stage 1 loss of *l before *u set the conditions for all subsequent loss. This is extended in specific vowel environments in varying degrees of generality, depending on vowel quality and variables of stress. Stage 2 probably

followed next, where deletion is generalized to all [+round] vowels, first before [-stressed] vowels and then before [+stressed] vowels. Stage 3 followed when the rule was extended to all [+back] vowels (*u, *o, *a), first before [-stressed] *a and then [+stressed] *a. Finally, Stage 4 included *l before *e in restricted cases.

As for subgrouping, only the hypothesized intermediate shift of *l > /y/ is helpful. The rest of the changes are specific post-PSA events that have little bearing on subgrouping, largely because they involve loss rather than shared innovation of replacement. The only question remaining is how Lower Pokomo fits into the overall picture. While it does share in the intermediate shift, and does show some evidence of *l-loss before unstressed *u, the Lower Pokomo change *l > y > Ø may well point to a separate development. It is noteworthy, however, that loss of *l before unstressed *u is one of the rare shifts that any Pokomo dialect shares with the Swahili dialect continuum, let alone Comorian. There is also the fact that only Lower Pokomo falls in with the others, suggestive of a late areal event.

Other changes affecting *l in Sabaki dialects are fairly straightforward: *l is realized as [r] (either phonemically /l/ or /r/, depending on the synchronic analysis of the dialects concerned) before the front vowels */i, e/ in Mwani, and before or after front vowels in Mijikenda, e.g., *-lima 'farm, cultivate' : /-rima/ (Mn, Gi), *-leWa 'be drunk' : /-rewa/ (Mn), /-rea/ (Gi), *-finika 'cover' : /-finikira/ (Gi), *-celela : /tserera/ 'come down' (Gi), *-gwila 'seize' : /-gwira/ (Du, Gi). [r] shows up less regularly in other Sabaki dialects; some of these may be borrowed, but others may just as well reflect older allophonic variation before the phonemicization of /l/ and /r/, e.g., in Swahili and some Digo dialects, after the massive infusion of Arabic loans.[17] Throughout the data in Appendices 3 and 4 there are also many cases of /r/ to be found where either /l/ or Ø would be expected. We have chosen not to mark these as skewed.

In Mwiini, *l is usually realized as /ɬ/, identified as a voiced alveolar lateralized fricative by Whiteley (1965:72) and as a flapped alveolar liquid by Abasheikh (1978:6-7). It is phonemically distinct from the flapped /r/ and the lateral /l/. /r/ is restricted to loan words, but *l gives both /ɬ/ and /l/, in most cases predictably. /l/ is restricted, occurring only: (a) adjacent to nonsyllabic consonants in inherited and borrowed lexis (e.g., /lbáwa/ 'feather < *luWaWa, múβli 'man'); (b) before glides, whether deleted or not (e.g. /-laaɬa/ 'be sick' < *-lwala); and (c) in certain morphophonemic alternations (e.g., /-pakiɬa/ 'load' plus the applicative suffix /-iɬa/ gives /-pakilila/ 'load for'). In all other cases, the reflex of *l is /ɬ/. Loan words also have a preponderance of /ɬ/ rather than /l/.

§4.2.2.2. *l-rules and chronology. In discussing the interaction of *l-loss with other Sabaki processes, we will for the most part, only consider the intermediate shift of *l > y > (Ø)/___u. Stratigraphy places this change fairly early, though not as early as *W-loss, which is more widely attested not only within Sabaki but in NEC as well:

---

[17]Thomason and Kaufman (1988:75) suggest that such phonemicization of earlier allophones often comes about under conditions of heavy contact.

*W-loss/___u	(1)	all Sabaki except Ti
*l-loss/___u	(2)	Sw (less Mw), Com, LP, part of UP

As for the various sorts of lenition that we have considered, it is difficult to say what stratigraphy might indicate, since the number of dialects affected by the early stages of *p- and *t-lenition are about the same as the number affected by *l > y > Ø, and there is no independent evidence, in the way of rule-ordering, to say more about the matter, since the operation of these various rules is independent of one another.

*l-loss does interact with *k-palatalization (§4.1.3.2), e.g., in SD. The latter does not apply in many cases where *l has been deleted, contrary to what we might expect, e.g., /kiungo/ (Ung) 'joint' < *kilungo. Were *l-loss an earlier rule than *k-palatalization, we would find such items matching items where palatalization has occurred (e.g., /čakula/ < *kyakulya), thus /kiungo/ would have become */čungo/. *k-palatalization is a rather late post-PSA change; *l-loss has to be relatively recent, innovating after the former ceased to be productive.

We have discussed *l-loss in terms of stress. The analysis by G. Philippson in §16 of this chapter suggests that penultimate stress has probably co-existed with tone at least since PSA, possibly much longer. Thus the invocation of "stress" in the analysis of the changes to *l and *g (see §4.2.3) is justified.[18]

The evidence from older written sources suggests that *l-loss started in Swahili (at least ND) not too long before A.D. 1700, and was ongoing throughout the 18th century (Miehe 1979; Knappert 1979).

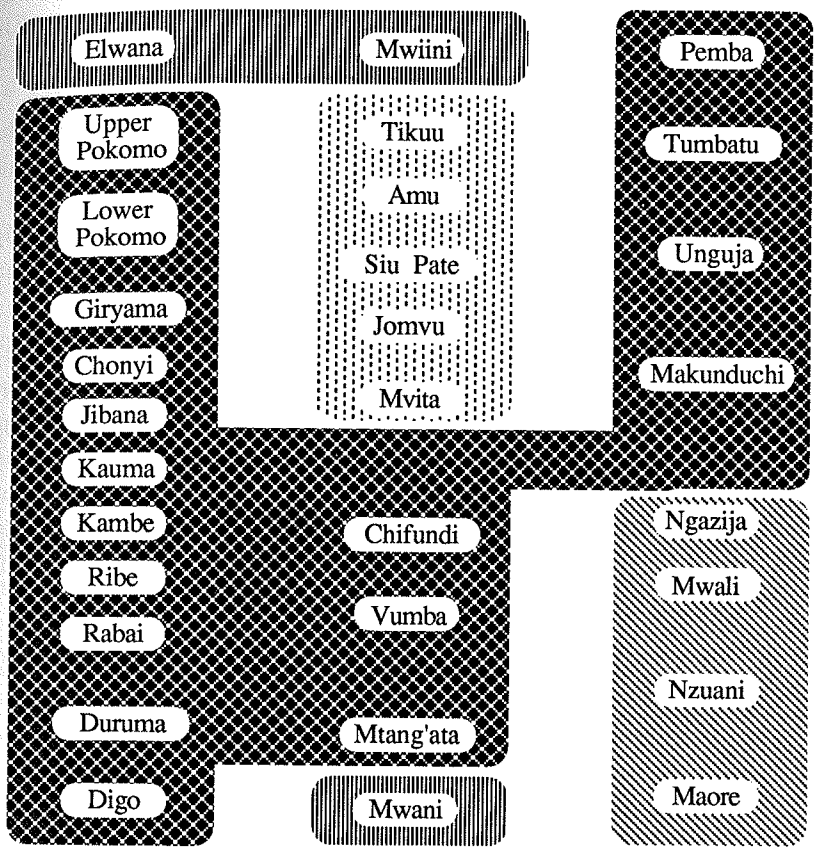
§4.2.3. Sabaki *g. The reconstruction of a voiced velar stop is well substantiated in the present-day Sabaki languages, but the situation is not uniform, and not all dialects have preserved *g. Elwana, Mwiini, and Mwani have uniformly lost *g in both stem-initial and intervocalic environments; exceptions are found in each, but most cases are attributable to borrowing. Two other dialect groups, Comorian and the rest of ND, also attest *g-loss, but less uniformly. It would appear that *g-loss here ceased to function before affecting all lexical items, or there are internal linguistic reasons, not all of which we have investigated, which explain *g-preservation where otherwise we should expect loss. Alternatively, these dialects may have come under extensive lexical influence from dialects which have preserved *g, namely, Pokomo, Mijikenda, and SD. These latter dialects provide the main internal evidence for *g (Chart 9).

*-gaw- 'divide'	-gawa (LP,Ung), -gaža (Gi), -gavya (Du,Ch)
*-gawany- 'share'	-aawanya (Mw), -avanya (Ti), -awanya (Si,Pa,Am,Mn), -gawanya (Mv,Ung), -(w)anya (Ng), -anyisa (Ng)
*-caagul- 'choose'	-tsagula (Gi), -ʈeua ~ -ʈaua (ND), -chagua (Ung), -ʈaɬa (Mw), -tsa(h)ua (Ng,Nz,Ma)

[18]Our analysis of the behavior of *l (and *g), and Philippson's analysis of prosodic features, were done independently.

Chart 9

Sabaki *g



Key

*g : g

*g : Ø

*g : Ø ~ g

*g : Ø ~ (g)

External support for a velar stop also comes from NEC. NEC languages attest a velar stop, with some exceptions, e.g., Shambala and Bondei which have /ɣ/, and a little further afield, Pare and Tuβeta. The others all attest /g/ (Nurse 1979a). Furthermore, virtually all of the languages of Guthrie's Zones D, E, F, G, and P attest a voiced velar stop. There are a few exceptions, such as Central Kenya Bantu (Kikuyu /ɣ/ (E50), Kamba Ø (E50) and Langi /ɣ/ (F30)). Also, most of Zones A, B, and C attest a voiceless stop, /k/ < *g. It is thus clear that PSA *g represents a fairly old situation, whereas languages with /ɣ/, or derivatives of /ɣ/, represent relatively more recent developments, such as Central Kenya Bantu or the large band of languages in the south of the Bantu range (see Guthrie 1967:75, topogram 10) in which CB *g has been deleted, presumably via /ɣ/ (*g > ɣ > Ø), as attested in Chaga where ɣ > h > Ø.

§4.2.3.1. Changes affecting *g. A central core of Sabaki, namely Pokomo, Mijikenda, and SD, has for the most part retained *g.[19] The other Sabaki dialects are *g-loss languages, or show various degrees of loss. The question about loss is whether or not the seriation we suggested for some of Bantu, namely, *g > ɣ > Ø or something similar, is valid for Sabaki. Elwana, Mwani, and Mwiini show the effects of complete loss of *g with no systematic evidence of having gone through some intermediate shift. Elwana apparently does attest a velar voiced fricative (Möhlig 1986c), but our only example of this (*ɟgego 'molar' : iɣɛɣɔ) may well be borrowed from Central Kenya Bantu. Without further evidence we do not want to make a claim for an intermediate stage in Elwana, Mwani, or Mwiini.

In ND, which has a mixed situation, with lexis attesting both loss and preservation of *g, there is no existing evidence that an intermediate change played a role. If we simply count the presence of *g in stem-initial or intervocalic environments, about 50 of the items attest a reflex of *g, the rest show deletion. Given that Mwiini is an almost unequivocal *g-loss language, we might want to conclude that Mwiini is the source of most of the lexis attesting *g-loss. But if we look at *g in stressed versus unstressed syllables, the results are different. In *stressed* syllables, about 65% of the tokens counted (approximately 32) show *g preserved, e.g., *-gula 'buy' : /-gura/, with about 35% showing loss, e.g., *-gumu 'hard' : /-umu/. In unstressed syllables (approximately 33 tokens) about 58% show *g-loss, e.g., *-guluk- 'fly' : /-uka/, and 42% show *g preserved, e.g., *-galuzy- 'change' : /-geuza/. This skewing, sensitive to [±stress], strongly suggests that a language-internal process of loss is responsible for the situation in ND, wherein *g first began to delete in unstressed syllables, and then at some point the change began to affect *g in stressed syllables. Given the mixed situation, it seems that the process of loss was not completed in ND, and that lexical items, with *g still preserved, are holdovers from an ear-

[19]While there are a few cases of lexis exhibiting loss in these dialects, they are the exception. In a few cases, e.g., in some SD dialects, alternations of /g ~ y/ have been noted, e.g., -mwaga ~ -mwaya 'pour' (Pe), mbegu ~ mbeu [mbeyu] 'seed' (Mt). This seems to be a rural Zanzibar/Pemba phenomenon, and until we have more data we will not make further comment.

lier time when the process of loss was active. It is, however, possible that complex loaning from multiple sources has contributed to the current status of *g in ND. Whatever the situation, there is no evidence in ND that *g was lost by going through an intermediate stage.

In Comorian there is a clear pattern of loss in intervocalic and stem-initial environments, though there are exceptions, with most occurring in Nzuani:

*-cagula 'choose'	-tsa(h)ua (Ng), -tsaua (Nz), -tsahua (Ma)
*-guza 'sell'	-udza (Nz,Ma), -huza (Ng)
*-laga 'take leave'	-laha ~ -laa (Ng), -lawa (Ma) but -laga (Nz)
*-Wulaga 'kill'	-uwa (Ng,Nz), -ula (Ma)
*mfigo 'kidney'	ntso (Ma,Nz,Ma)
*muzigo 'load'	mdzo (Ng,Nz), mudzo (Ma)
*mbegu 'seed'	mbeu (Nz,Ma)

But in Class 5 environments, *g is regularly preserved (see Strengthening, §6):

*i̯-govi 'skin'	gozi 'bark' (Ng,Nz)
*i̯-gogo 'log'	goho 'hollow tree trunk' (Ng), °gongo (°ng ¬ g) 'large piece of wood' (Ma)

The frequent attestation of /h/ (sometimes /w/) as a reflex of *g in Comorian suggests two possibilities as sources: (a) an intermediate stage, whereby *g > /h/ > Ø; or (b) /h/-insertion to break up vowel sequences. An intermediate stage for Ngazija is a possibility where *g first became /h/ before deleting. There is not much evidence at this point in our investigations to make a choice. However, the variable appearance of /w/ and /h/ across dialects, e.g., *-laga 'take leave' : /-laha/ (Ng) but /-lawa/ (Ma), and the absence of /h/ in Nzuani, suggest insertion to break up vowel clusters at a post *g-loss stage. Similar stray examples are seen in ND: *mutondo+igolo 'yesterday' : /mtondohoo/ (Ti), *kiloWo 'fishhook' : /čiɬoho/ (Mw), where in the latter case /h/-insertion occurred after *W-loss.

Finally, Mwani, a *g-loss language, attests some cases of /y/ < *g in intervocalic environments, e.g., *-mwaga 'pour' : /-mwaya/, *-saga 'grind' : /-saya/, *-βulaga 'kill' : /-ulaya/, suggesting perhaps an intermediate /y/, thus *g > y > Ø (also see -mwaga ~ -mwaya 'spill, pour' in rural Zanzibar). In summary, then, we distinguish three groups of languages (Chart 9):

1. Unequivocal *g-preservation	Pokomo, Mijikenda, SD
2. Unequivocal *g-loss	Elwana, Mwiini, Mwani
3. *g-loss (exceptions)	other ND, Comorian

It should be noted that even in *g-preserving languages, *g is lost in certain lexis across the whole Sabaki spectrum, e.g., *-(g)amba 'say', *-(g)enda 'go', *muWa 'sugarcane' (CB *-gùbá), plus others. Here we seem to be dealing with an early change, quite distinct from the *g-loss we see subsequently. One can imagine a situation in which *g-loss innovated quite early, while the Sabaki proto community was still homogeneous,

and had only affected a minimum of lexis before its breakup. After dialect differentiation, *g-loss was generalized in only part of the continuum, leaving the present-day distribution.

The only other shift affecting *g is found in Mijikenda and Maore. Before the front vowel *e but not before *i, *g palatalizes variably in Giryama: *-gez- 'try' : /-jeza/, *-genel- 'spread' : /-jenera/, but *-ogelel- 'swim' : /-ogerera/. In the rest of Mijikenda we find all the dialects palatalizing before *i and *e: /-jita/ 'cook' (cf. -gita Gi), /jiza/ 'darkness' (cf. giza Gi), *-ogelel- 'swim' : -ojerera (Du, Di, etc.). This is part of a general palatalization rule affecting *k in [+front] environments in Mijikenda (see §4.1.3.2).

In Maore we have *mugeni 'guest' : muǰeni. Rombi (1983) also indicates variation in Maore involving /g ~ j/ and /ng ~ nǰ/, e.g., /ngema ~ nǰema/ Class 9 'good', and in loans from Arabic, e.g., /gini ~ ǰini/ 'devil'. Whether this has anything to do with *k-palatalization is uncertain; more work in Comorian is necessary to understand what is happening.

There are sporadic examples in Swahili, usually involving Arabic loans, e.g., /-regea ~ -reɟea/ 'return', /rangi/ ~ /ranji/ 'color'. This is not significant for subgrouping.

§4.2.3.2. *g-loss and chronology. Because we are dealing with loss of a segment, and because there is no evidence that any of these dialects shared a common intermediate shift which might point to common origins, we will not consider this shift relative to other changes. Indeed, the distribution of *g-loss languages vis-à-vis others indicates separate development and thus convergence. Elwana is isolated from all other Sabaki *g-loss languages by Pokomo, which conserves the proto segment. Similarly, Mwani is separated from other *g-loss languages by SD; furthermore, Mwani *g-loss may relate to *g-loss in Makua, Makonde, and other adjacent Zone P languages. While it would be attractive to use *g-loss as evidence in support of closer ND and Comorian affiliations, loss is such a poor diagnostic for genetic connections that we don't want to make too much of it. Thus, all evidence points to late separate development. We will discuss *g again in §6 on *g-loss and Strengthening in Comorian.

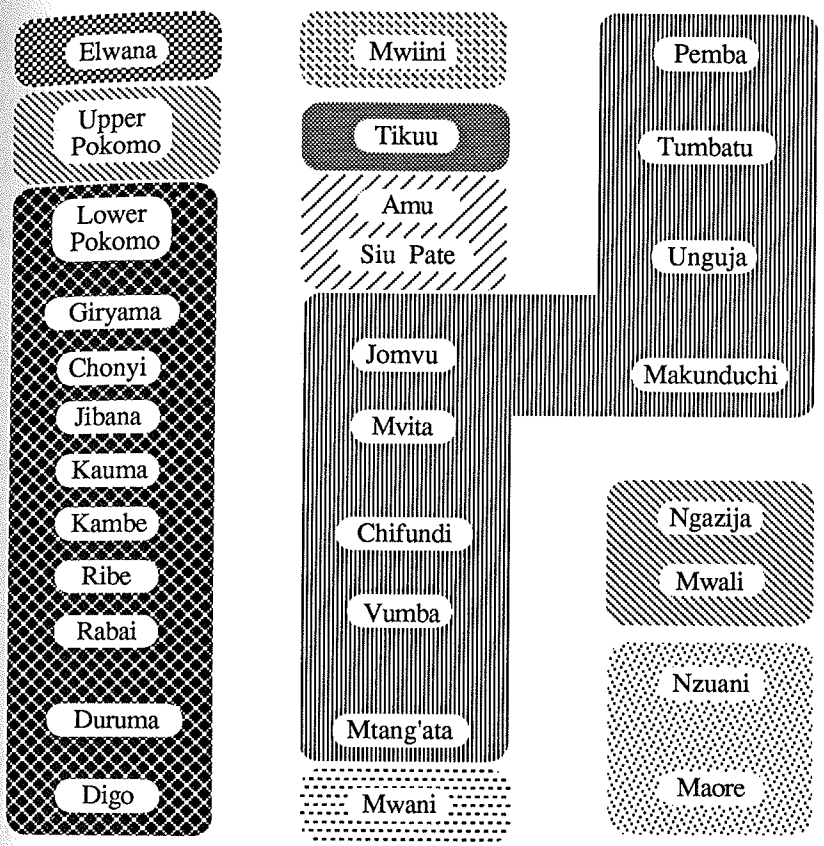
§4.2.4. PSA *j. The reflexes of PSA *j are quite varied from one dialect to another. They include palatal voiced stops, affricates, fricatives and glides: /ɟ/, /ǰ/, /ž/, and /y/, and two nonpalatals which we argue are derived from a palatal source: /dz/ and /z/. Correspondences and examples (Chart 10) are:

CB *j : PSA *j :

dz [d̪z]	MK, LP	Ø ~ y ~ ž	Ti
ǰ	UP	y ~ Ø	Am, Si, Pa
ɟ	Mv, Jo, all SD, Chi, Mn	Ø ~ y	Mw
ǰ	Ng, Mh	z	El
ž	Nz, Ma		

Chart 10

Sabaki *j



Key

*j : dz
*j : ɟ
*j : ǰ
*j : ž
*j : y ~ ž
*j : y ~ Ø
*j : Ø ~ y
*j : z
*j : ɟ

*maji 'water'	máàzɪ (El), maǰi (UP,Ng), madzi (LP,Gi,Du), maai (Mw), mai (Am,Si,Pa), mayi (Ti), maai (Mw), maɟi (Mv,Jo,Ung,Mak,Tu, Pe,Vu,Mt,Chi,Mn), maži (Nz,Ma)
*-vu̯j- 'leak'	-vudza (Gi), -vuya (Am), -vuuya (Mw), -vuɟa (Ung), -vuǰa (Ng), -vuža (Nz)
*i̯juWa 'sun'	dzua (LP,Gi,Ch,Du,Di), yua (Am,Si,Pa), yuva (Ti), iʋa (Mw), ɟua (Mv,Nga,Jo,Ung,Mak,Tu,Pe,Vu,Chi,Mt,Mn), ǰua (UP, Ng,Mh), žua (Nz,Ma)
*-i̯jal- 'be full'	-zala (El), -dzala (LP,Gi), °-odzala (°o) (Du,Ch), °-yela (°e ¬ a) (Mw), -yaa (Am), -ɟaa (Mv,Ung,Mak), -ǰaa (Ng), -iɟaa (Vu, Mt), -žaya (Nz), -iɟala (Mn)
*-i̯juW- 'know'	-iʋa (Mw), -yua (Si,Pa,Am), -yiva ~ -iva ~ -živa (Ti), -ɟua (Mv, Ung), -iɟua (Vu,Mt,Mak), -viɟiwa (Pe), -ǰuwa (Ng), -iɟula ~ -iɟiwa (Mn)

Our choice of a palatal noncontinuant as the source of the Sabaki reflexes is amply supported by the extant languages, and parallels Guthrie's reconstruction as well. Both an affricated palatal, [ǰ], and an imploded palatal stop, [ɟ], are attested in the Sabaki continuum, but it is not clear how we should specify the phonetic content of the PSA reconstructed segment, that is, as either an affricate or imploded stop. We will simply leave the proto segment unmarked for either implosion or affrication. Even though Guthrie reconstructs *j for CB, it is not clear whether PSA *j is an innovation or inherited segment. This problem is parallel to the one discussed for PSA *c (§4.1.4).

Most of the languages surrounding Sabaki, including its nearest relatives, attest /z/ as a reflex of CB *j. In fact the great majority of Bantu languages have /z/ or Ø (see Guthrie 1967:76-77, topograms 11 and 12). We suggested earlier that PSA *c may be an innovation, deriving ultimately from a reconstructed voiceless palatal continuant *š. Symmetry would then suggest that its voiced counterpart *j is also a PSA innovation, rather than a retention, and derivable from a voiced palatal continuant *ž. The evidence supporting this notion is more complex, largely because the distributional facts are different from those for CB *c, and, perhaps more importantly, because the reflexes of CB *c and *j do not always function symmetrically in the modern Bantu languages. Whatever shifts affected CB *c did not necessarily apply to CB *j. Thus, while languages in Zones H, K, and L demonstrate symmetry (CB *c : s, and *j : z), most others do not (e.g., Zone A languages have CB *c : s, and *j : Ø) (Guthrie 1967:76-77, topograms 11 and 12). The distributional facts of the modern reflexes are compelling evidence supporting reconstruction of a Bantu continuant. Furthermore, a continuant is more likely to delete than a stop or affricate.

Although most reflexes of CB *j are /z/ or Ø, that the original Bantu proto segment was a palatal is supported by scattered palatal stops and affricates in Zones B and K. But it is in the northeastern and eastern parts of the Bantu area (Zones D, E, F, G, P, and N) that

palatality is widely attested. Non-continuant palatals are almost exclusively found in the northeast. Given the distributional facts, two hypotheses are possible: (a) the noncontinuant palatals, e.g., /j/, are relics, or (b) they represent the products of an innovation in which a Bantu palatal continuant, possibly *ž, became [-continuant]. To decide this question we would have to go well outside of narrow Bantu, beyond the scope of this study. If we arbitrarily decide for the second possibility we would still have to decide at what point a hypothetical palatal continuant, *ž, become a noncontinuant. Sabaki is not the only language group that attests a noncontinuant palatal in the northeast; they are found in D20, D60, E10, E20, E30, E40, F30, etc. This points to an earlier innovation prior to PSA, and if so, would mean that PSA inherited its palatal noncontinuants not from CB, but from a later intermediate proto stage. This contrasts with the much stronger possibility that PSA *c is an innovation (§4.1.4). Here, however, we have to conclude that PSA *j is a retention, and that surrounding languages with /z/ are the innovators. In any case, we have subgrouping evidence, and we will leave the matter at that. For purposes of explicating Sabaki, we will assume a PSA palatal stop.

§4.2.4.1. **Sabaki *j and changes**. Several possible separate seriations can be proposed to cover the Sabaki dialect continuum. They involve (a) Lower Pokomo and Mijikenda, (b) the Swahili dialect continuum, (c) the Comorian group, and (d) Elwana:

a. *j > dz	Lower Pokomo/Mijikenda
b. *j > ʄ > ž > y > Ø	Swahili (ʄ: Mwani (Schadeberg pers. comm.)
c. *j > ǰ > ž	Comorian
d. *j > z	Elwana

It would appear that none of these are necessarily connected. The development of /dz/, shared by Lower Pokomo and Mijikenda, appears unrelated to the seriation attested in Swahili; there is no common shared intermediate stage, and no evidence that /dz/ in Lower Pokomo/Mijikenda developed from any source other than *j. Seriations (a) and (d) may be connected, namely *j > dz > z, thus linking Elwana with Pokomo, but this has to be ruled out on the basis of distribution. Elwana is separated from the Lower Pokomo/Mijikenda continuum by a number of Upper Pokomo dialects which conservatively attest a palatal stop, so it is more likely than not that Elwana /z/ represents an independent innovation, moreover unrelated and unconnected to the widespread Bantu correspondence in which CB *j : z. Elwana conserves *c, *nc, and *nj, only shifting *j, unlike other Bantu languages which attest the correspondence CB *j : z.

In (b) we assume that /ʄ/ is a Swahili feature that is retained in SD and Chifundi, and that Mwiini, Tikuu, Siu, and Pate (ND3) innovate a shift of Proto-Swahili *ʄ to /ž/, still retained by older Tikuu speakers; this was subsequently weakened to /y/ in Amu, Siu, Pate, and Mwiini. Among these, only Mwiini has deleted the reflex in most lexis, while the others show variable deletion of /y/.

The Comorian scenario assumes a different proto form, as suggested by the affricate /ǰ/ in Ngazija. However, the Comorian and Swahili seriations, (b) and (c), share a com-

mon development, namely /ž/. But the facts of distribution suggest a different scenario for Comorian. Nzuani/Maore have /ž/ while Ngazija has /ǰ/; thus we have no direct evidence that /ž/ in both groups has a common intermediate source. To claim that they do creates a much too complex early dialect grouping which cannot be easily supported. It would entail explaining why parts of both present-day groups remained conservative, namely, Ngazija in Comorian, and Unguja (and other SD) in Swahili, in retaining either /ǰ/ or /ʒ/, at a time when they were more closely geographically aligned. It is more likely and simpler to assume that Comorian /ž/ developed after Comorian separated from the mainland continuum, and was independent of anything the northern part of Swahili was doing. Of course, we could claim that Nzuani/Maore /ž/ was borrowed from ND, or vice versa; though this would establish connections, they would not be genetic.

Upper Pokomo, Mwani, SD, and Ngazija are conservative in retaining a palatal non-continuant, and because we cannot subgroup on the basis of retention we cannot assume that these share any closer genetic affiliation on the basis of this shared feature. Thus the overall picture Sabaki presents for *j is quite fractured and unhelpful vis-à-vis subgrouping and shared development.

Two other matters must be noted here:

a. Comorian has changes that affect *j after the Class 5 prefix *i̯; we will discuss this in §6.1 on Sabaki Strengthening.

b. Though the behavior of *nj somewhat parallels Sabaki's treatment of *j we will discuss this in §8.1.4 on Sabaki nasal-consonant clusters.

§4.2.4.2. *j, chronology, and rule relationships. Because both seriation and stratigraphy indicate that most changes affecting *j are late, there is not much to consider in this section vis-à-vis specific interactions with other rules so far considered. There is the matter of *g > j before front vowels in Giryama. This must follow the change of *j to /dz/ in Mijikenda (Gi), otherwise those *g's that palatalize would also shift to /dz/. This is a purely local Giryama rule.

Stratigraphy would give the same results. The shift of PSA *j to /dz/ is specific to Mijikenda and Lower Pokomo, and thus predates the more local palatalizing rule. Otherwise the rules affecting *j do not intersect with those that we have considered thus far for either the voiced or voiceless consonants.

In §4.1.4.1. we argued that Lower Pokomo, Mijikenda, Comorian, and ND may share a seriation of *c > ts [t̪s] > t̪. This development is not matched by what happens to *j. The only Sabaki languages that treat *j parallel to *c are Mijikenda and Lower Pokomo. If we are correct in hypothesizing that shifts affecting *j are late, as indicated by the diversity and unconnectablity of the changes, and that shifts to *c are early post-PSA changes, then the parallelism we see in Mijikenda and Lower Pokomo may represent an effort in those two to keep the reflexes of *c/j in balance. That is, *j went to /dz/ because *c > /ts/. Otherwise there is no apparent connection between the behavior of *c and *j, further supporting our contention that changes involving *c are earlier than changes involving *j.

§5.0. The Sabaki Fricatives *f, *v, *s, *š, *z

§5.1. Bantu Spirantization. All Sabaki languages have the fricatives /f, v, s, z, š/, or derivatives thereof (e.g., Tikuu, Siu, Pate, which have replaced *z by /ð/) in their phonemic inventories. These derive ultimately from the primary consonants *p, *t, *c, *k, *W, *l, *j, *g, by a rule we refer to as Bantu Spirantization. This is a widely attested *set of processes* in which the Bantu consonants become spirants, usually fricatives, but sometimes affricates. Although the term spirantization does not accurately characterize all the outputs of this process, it is convenient; a term used by others is assibilation.

Spirantization is conditioned by the so-called super-close Bantu vowels *i̧ and *u̧ following the affected consonants. The features of the resulting segment are a function of those of both the consonant and the conditioning vowel. Thus, most commonly, any combination of consonant and high super-close vowel ([+HV]) with the feature [+labial] present in either part of the structural description (the consonant or the vowel) yields a labial fricative, or affricate. If the feature [+labial] is not present, a nonlabial fricative, or affricate, results. Examples from Giryama illustrate the process for Sabaki (note also that the seven vowels of CB have reduced to five, a feature shared by most Bantu languages that have Bantu Spirantization, see §10 below):

Giryama Spirantization

CB		Giryama	CB		Giryama
*-pí̧dà	'pus'	ufira	*-pú̧ngàté	'seven'	fungahe
*-ti̧kù	'day'	siku	*-tù̧ngò	'genet cat'	fungo
*-ki̧na	'tree trunk'	sina	*-tákù̧n-	'gnaw'	-hafuna
*-ci̧d-	'grind'	-sira	*-cú̧í	'we'	siswi
*-bí̧n-	'dance'	-vina	*-bú̧	'ashes'	ivu
*-di̧bà	'pool'	zia	*-dù̧mbí	'contin. rain'	muvumbi
*-gi̧mà	'whole'	-zima	*-gù̧nd-	'stink'	-vunda

Giryama /f,v/ ([+labial]) derive either from *p/b ([+labial]) before [+HV], or from *t/d ([-labial]) and *k/g ([-labial]) before *u̧ ([+labial]). Giryama /s,z/ ([-labial]), on the other hand, derive from *t/d and *k/g before *i̧ ([-labial]). Although this process can be described by a simple, straightforward generalization, available evidence suggests that it is a series of connected changes which operated in somewhat discrete stages, affecting some consonants before others (Hinnebusch and Nurse 1981; Hinnebusch, Nurse, and Mould 1981:24ff).

The reflexes for Spirantization, as illustrated in Giryama, are widely attested in East Highlands Bantu (see Guthrie's Correspondence statements, 1969, vol. 2, and Nurse 1979a). One might therefore want to claim that PSA inherited its fricatives from a pre-PSA proto period, possibly PNEC. However, the matter is complex. Completely regular reflexes are found neither within Sabaki itself nor in NEC. The following table presents the full range of the reflexes of Spirantization for the Sabaki dialects. For comparison purposes, others are included: Seuta (Shambala, Bondei, Ngulu, and Zigula), and East

purposes, others are included: Seuta (Shambala, Bondei, Ngulu, and Zigula), and East Ruvu (Doe, Zalamo, Nhwele, Kutu, and Kami). Supporting data are to be found in Appendix 2; also see Nurse (1979a: 413-417).

If we compare closely the various elements of Sabaki Spirantization (see display on the following page) with the general situation in NEC, we find a mixed picture of regularity and variation. All of Sabaki, Seuta, and East Ruvu in general share the following correspondences (exceptions given in parentheses):

Sabaki, Seuta, East Ruvu:

CB *p/k :	f/ ___ *ụ	
CB *b/g :	v/ ___ *ụ	(NEC /v/ : Kami/Kutu /pf/)
CB *l :	v/ ___ *ụ	(NEC /v/ : Kami/Kutu /pf/ : Shambala /v ~ z/)
CB *t :	f/ ___ *ụ	(NEC /f/ : Shambala /s ~ š/)

So, before the back super-close vowel *ụ, the reflexes found in Sabaki are what we find generally in NEC, with Shambala attesting, we assume, irregular reflexes of Spirantization, and Kami/Kutu attesting regular but secondary reflexes, having innovated a devoicing and affrication process. This overall uniformity suggests that /f/ and /v/ were an early NEC, and therefore PSA, feature.

Before the front HV *į , the situation is more diverse, but even here there is some regularity between Sabaki and NEC; Sabaki shares with NEC the development of /s/ from *t before *į:

Sabaki and NEC:

CB *t : s/___*į (NEC /s/ : Shambala /s ~ š/)

Although Shambala, again as above, is the exception, regularity and distribution suggest that /s/ from *t is an inherited PSA feature. For the rest of the reflexes Sabaki is divergent, part of Sabaki patterning with NEC and part behaving differently. The dialects that behave idiosyncratically are Comorian and a portion of the Swahili dialect continuum. The question that now has to be addressed concerns the source of these apparent idiosyncrasies of Sabaki vis-à-vis NEC. Do they have their origin in Spirantization itself, and thus predate the formation of PSA? Or can they be attributed to later specific post-PSA developments, distinct and separate from Spirantization? We will discuss in turn each set of "irregularities" to attempt to form an appropriate answer:

a. /s, z/	:	CB *p, b /___ į	(Mw, Ti, Si, Pa, Am, Com) (§5.1.1)
b. /ts, dz/	:	CB *p, b /___ į	(Com) (§5.1.2)
c. /z ~ dz/	:	CB *d, g /___ į	(Com) (§5.1.3)
d. /š/	:	CB *k /___ į	(Sw: most of ND and Ung) (§5.1.4)
e. /ð/	:	CB *b, d, g /___ į	(Ti, Si, Pa) (§5.1.5)

Sabaki Spirantization

(see Charts 11 and 12)

CB	MK	Ng	Nz	Po	ND	Mv	SD	Ung	ST[1]	ER[2]
*p/__ i̯	f	s, ts	s, ts	f	s	f	f[3]	f	f	f
*p/__ u̯	f	f	f	f	f	f	f	f	f	f
*t/__ u̯	f	f	f	f	f	f	f	f	s, š	f
*k/__ u̯	f	f	f	f	f	f	f	f	f	f
*b/__ i̯	v	z, dz	z, dz	v	z, ð[4]	v	v	v	v	v (>pf)
*b/__ u̯	v	v	v	v	v	v	v	v	v	v (>pf)
*d/__ u̯	v	v	v	v	v	v	v	v	v, z	v (>pf)
*g/__ u̯	v	v	v	v	v	v	v	v	v	v (>pf)
*t/__ i̯	s	s	s	s	s	s	s	s	s, š	s
*k/__ i̯	s	s	s	s	š, s[5]	š	s	š	s	s
*d/__ i̯	z	z[6]	z	z	z, ð[4]	z	z	z	z	z (>ts)
*g/__ i̯	z	dz	dz[7]	z	z, ð[4]	z	z	z	z	?[8]

Notes:

[1]In Seuta, where /s/ and /v/ are the rule, Shambala is exceptional in attesting /š/ and /z/ respectively, in some cases; where no variant is given, Shambala has the general Seuta reflexes.

[2]In East Ruvu, Kutu and Kami attest the voiceless affricates, which we assume derive from the voiced fricatives via a devoicing and affrication process.

[3]Chifundi, according to Lambert (1958b), attests affricates instead of fricatives; M. Bakari (1985 and pers. comm.) records only fricatives, Walsh (pers. comm. 1987) cites affricatives for one dialect.

[4]Tikuu, Siu, Pate attest /ð/ < ND *z.

[5]Tikuu attests /s/; other ND /š/.

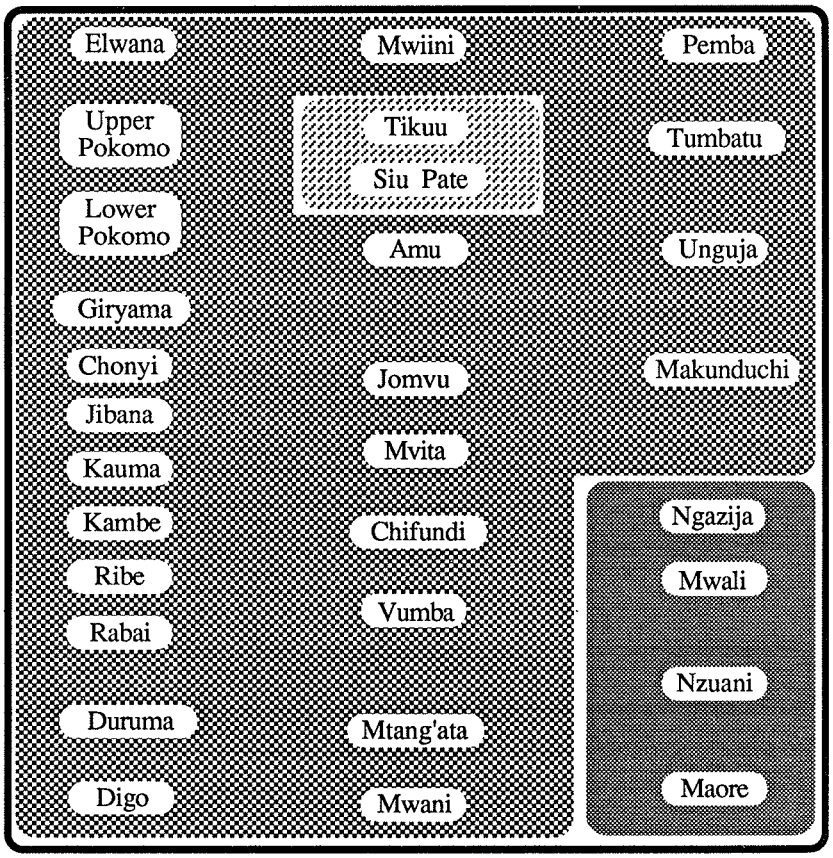
[6]*l regularly goes to /z/ before *i̯ in all Comorian dialects; however, there are cases in which /dz/ is the reflex.

[7]The few examples of reflexes of CB *gi̯ in Comorian are all /dz/; however, with a few exceptions, these follow nasals and thus derive from the Comorian rule in which *s and *z become /ts/ and /dz/ after *N: /mlandzi/ 'bamboo' < *mulanzi̯ (CB *-mudàngi̯), /-sindzia/ 'be sleepy' < *-sinzi̯la (CB *-ti̯ngi̯d-), ndzi 'fly' < *nzi̯ (CB *ngi̯). The only exception is an adjective: /-dzima/ 'whole, healthy' < *-zi̯ma (CB *-gi̯mà), but see /-dziro/ (Ng) 'heavy', /-zitro/ (Nz) (CB *-di̯tò), where the Ngazija item does not alternate but the Nzuani one does, e.g., /mzitro/ ~ /dzitro/); for an explanation of /dz/ in the latter two examples see §6.3.

[8]Presumably Kami and Kutu have a subsequent change in which *z becomes /ts/ (< CB *g/__ i̯), but the only examples in in our data base of *g/__i̯ occur following *N, which would protect the resulting /z/ from undergoing the devoicing/affrication process seen above.

Chart 11

Sabaki Spirantization 1



Key

*p, t, k ; b, d, g : f ; v / ___ *u̯

*t : s / ___ *i̯

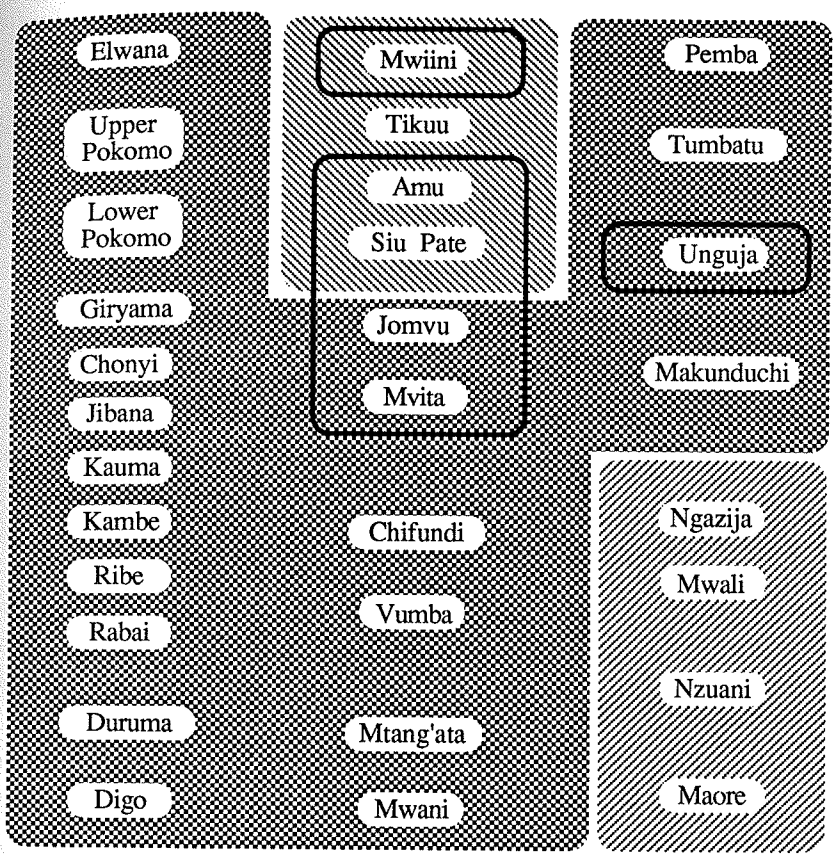
*d, g : z / ___ *i̯

*d, g : z ~ dz / ___ *i̯

*d, g : ð / ___ *i̯

Chart 12

Sabaki Spirantization 2



Key

*p ; b : f ; v / ____ *į

*p ; b : s ; z / ____ *į (z : ð Tikuu, Siu, Pate)

*p ; b : s ~ ts ; z ~ dz / ____ *į

*k : š / ____ *į

§5.1.1. Comorian/ND$_2$ /s, z/ : CB *p, b/___ i̧ (Coronalization). Instead of attesting the expected labial fricatives /f/ and /v/, Mwiini, Tikuu, Siu, Pate, Amu and Comorian attest /s/ and /z/ (/ð/ in Ti, Si, Pa):[2]

Comorian/ND$_2$:		**Sabaki/NEC:**	
/s/ :	CB *p /___ i̧	/f/ :	CB *p /___ i̧
/z/ :	CB *b /___ i̧	/v/ :	CB *b /___ i̧

-sikilia (Am) 'arrive', -siha 'come down' (Ng) (*-fi̧ka, CB *-pi̧k-)
-sinyanga (Am) 'mold pottery', -si̧nanga (Ti) (*-fi̧nya, CB *-pi̧ni̧-)
-sinya (Ng) 'pinch' (*-fi̧nya, CB *-pi̧ni̧-)
ikasia (Mw), k^{h}asi (Ti), nkasi (Ng) 'paddle' (*nkafi̧, CB *nkápí)
masiɬa (Mw), masia (Nz) 'pus' (*Wufi̧la, CB *-pídà)
lkusi (Mw) 'palm of hand' (*i̧koofi̧ 5/6 ~ 11/10, CB *-kóópi̧)

-zia (Am), -zila (Ma) 'be spoiled' (PSA *-vi̧la)
ngoði (Ti, Si, Pa), ngozi (Am, Ng, Nz), ngwezi (Ma) 'skin' (*ngovi̧, CB *-gòbi̧)
igozi (Mw) 'animal skin', gozi (Ng, Nz) 'bark' (*i̧govi̧, CB *-gòbi̧ 'skin, covering')
-zimba (Ng, Nz, Ma) 'swell' (*-vi̧mba, CB *-bí́mb-)
zi- (Am), zi- ~ s- (Mw), (i)zi- (Com) 'Class 8 nominal prefix' (*(i̧)vi̧-)

The examples above are typical, although in Comorian the situation is complicated by variation between /s, z/ and /ts, dz/ (cf. §5.1.2 below). We can hypothesize two possibilities for the source of /s/ and /z/ in these cases. In the first /s, z/ were either directly the result of Spirantization (Hypothesis A):

*p, b > s, z/___i̧

In the second, they derived in a two-stage process which first involved a regular form of Spirantization that yielded /f/ and /v/, and then a secondary shift that took /f/ and /v/ to /s/ and /z/ before a high front vowel, either [+HV], or [-HV], presumably at a much later period when PSA was differentiating into its constituent dialects (Hypothesis B):

Stage 1: pre-PSA *p, b > f, v /___ i̧
Stage 2: post-PSA *f, v > s, z /___ i (or i̧) Com/ND$_2$[20] (see Ch 5, §4.4)

A simple articulatory assimilation of /f, v/ to the features of the following high front vowel is the assumed basis for the shift at Stage 2. The fricatives /s/ and /z/ are more similar to /i/ (or i̧) in that all are more central in their articulatory characteristics than are /f/ and /v/. Though the vowel need not be [+HV] to explain the assimilation, we are allowing for the possibility that at the time Stage 2 innovated, the Sabaki dialects had not yet lost the super-close vowels.

---

[20]Note that we assume that ND$_3$ also shares this as an intermediate shift, and later shifts the resulting /z/ to /ð/ (see §5.1.5).

Both hypotheses are viable, and equally well motivated. At this moment it does not seem possible to choose between them. It could be the case that Comorian and ND_2 were already distinct when Spirantization, as an areal change, began to affect Sabaki, and *p and *b before *į were treated exactly parallel with *t/d and *g before *į. In such a case ND/Comorian share a rule in which *p/b, *t/d, and *g, (but not *k) before *į shifted to s/z across the board. However, for our purposes here, where we are assuming that PSA inherited its fricatives from a pre-PSA stage, or at least an early PSA stage, we prefer Hypothesis B. It should be noted that both hypotheses support a shared Comorian/ND_2 period. In the following pages we will refer to this shift as Comorian/ND_2 Coronalization. Also affected are sequences of *fy and *vy (e.g., *-fyonza 'suck' > -sonḍa (Am), *-vyala 'give birth' > -zaa (Am), -zaala (Mw), -za(y)a (Ng), -dza(a) (Nz), -dzaa (Ma). Though there is a similar shift in Giryama (see §5.3), in which *fy/vy > š/ž, it is only attested in this Mijikenda dialect, and is thus likely to be independently motivated.

§5.1.2. **Comorian /ts, dz/ : CB *p, b/___ į.** Comorian for the most part attests /s/ and /z/ corresponding to *p and *b/___ į and deriving from Comorian/ND_2 Coronalization, but there are also some items where we find /ts/ and /dz/ instead of expected /s/ and /z/. These need to be accounted for; examples are:

Comorian/ND_2:	**Sabaki/NEC:**
ts : CB *p /___ į	f : CB *p /___ į
dz : CB *b /___ į	v : CB *b /___ į

ntsimbo (Nz) 'stick' (*mfįmbo, CB *-pį́mbò), cf. simbo (Am)
ntso (Ng, Nz) 'kidney' (*mfįgo, CB *-pį́gò), cf. nso [ntso] (Am), iso (Ti)
kotsi 5/6 'slap' (Ng) (*įkoofį, CB *kóópį̀), cf. ukusi (Am), lkusi (Mwiini)
-tsitsa (Nz), -tsitsa (~ -dzitsa) (Ma) 'hide' (*-fįca, CB *-pį́c-), cf. -siṯa (Am)
-tsondza (Ng) 'suck', cf. -fyonza (Ung), -sonḍa (ND)

mwidzi (Ng, Nz, Ma) 'thief' (*mwįvį, CB *-yį́b-), cf. mwizi (Am)
-vudzia (Nz), -βudzia (Ma) 'blow on' (*-vụviḷa), cf. -vuzia (Am)

Although /ts/ and /dz/ could be direct reflexes of Spirantization, the evidence indicates that they are secondary developments, not necessarily associated directly with Spirantization. Examples with /ts/ and /dz/, relative to those with the expected /s/ and /z/, are few. The cases of /ts/ that we have extracted from our own data base or from Sacleux (Chamanga and Gueunier 1979) are listed above. In all but one instance /ts/ can be explained as deriving from a subsequent conditioned change in which /s/ became /ts/.

In Comorian, /ts/ derives from several sources: (a) from CB *c, e.g., /-itsi/ (Ng) 'raw' (CB *-bícì); (b) in causative constructions and presumably derived from consonant-glide clusters, e.g., /-rohotsa/ 'boil' < *-tokosya (CB *-tòkòt-) (cf. Causatives §5.3); and (c) from a general Comorian rule in which *s following a nasal becomes /ts/, e.g., /ntso/ (Ng, Nz) 'kidney' (CB *-pį́gò) (note synchronic variation in Ngazija: /siku ~ ntsiku/ 9/10). This final process explains the presence of /ts/ in other examples.

For the remaining, we have to appeal to somewhat ad hoc explanations. One of the exceptions, /-tsitsa/ 'hide', is obviously related to the shift of PSA *c (CB *c) to /ts/ in Comorian, e.g., CB *-pi̧ca > PSA *-fi̧ca > Com -sits- > /-tsitsa/, the final stage in the process being simply an assimilation of the stem-initial /s/ to the /ts/ in C_2 position. Or *-fi̧ča > fyiča > /-tsitsa/. The latter possibility is attested by /-tzondza/ (Ng) 'suck' < *-fyonza, parallel to the shift in which Sabaki *vy (< *bi̧a) becomes /dz/ (either directly or via an intermediate palatalized /zy/) in certain Nzuani and Maore forms, cf. /-dza(a)/ (Ma, Nz) < *-vyala 'give birth' (CB *-bi̧ád-), cf. Pokomo /-vyaa/.[21] It is also parallel to the shift of *sy (< CB *ti̧+a) to /ts/, e.g., /-rohotsa/ 'boil', < *-tokosya, (CB -tòkòt-) in causative constructions (see below §5.4).

The other exceptional item in our inventory is less easy to explain. There is no obvious surface-conditioning factor to explain the /ts/ in /kotsi/ 'slap' (PSA *-koofi̧, CB *-kóópi̧). If we were to set up an underlying nasal (which is a source of some instances of /ts/ and /dz/ elsewhere in the Comorian lexicon), e.g., *konsi, then the regular Comorian post-nasal affrication rule would apply, namely, *s > /ts/ after *N. There are examples in Ngazija, cited by Sibertin-blanc (1980:58-9), where non-etymological nasals in fact have been inserted: /-hundra/ 'find' (cf. /-kuta/, Ung), /ntsambu/ 'Cycas circinali fruit' (cf. /tapo/, Ung; /sapo/, Am), /-tumba/ 'jump from above' (cf. /-tupa/, Am), /-huntsanya/ 'collect' (cf. /-kusanya/, Ung), /-tsantsaua/ 'unravel' (cf. /-tatua/, Ung; /-tataua/, Pe). None of these have reconstructible proto nasals, as illustrated by the Swahili examples, and appear to be due to a process of nasal insertion which has operated idiosyncratically in Comorian. Unlike these examples, however, /kotsi/ does not have a surface nasal, so while the other examples are suggestive, we need to look elsewhere for an explanation.

The answer has to do with the long vowel Guthrie gives for the CB form (*-kóópi̧), even though he himself considers the reconstruction highly tentative (Guthrie 1970:300, CS 1156). Long vowels are attested in Mwiini, and therefore we have reconstructed them for PSA. If PSA long vowels were preserved until quite late, then we could have had a retained long vowel motivating the resulting fricative in Comorian /kotsi/. This is not as far-fetched an explanation as might first appear. Some instances of Comorian affricates develop before *VV, e.g., /-dziha/ < *-zi̧i̧ka (see §5.1.3), and before *y, which is a development of CB *i̧ + a, in other words, a VV sequence. However, in this case of /-kotsi/ the VV sequence precedes the conditioned segment.

§5.1.3. Comorian /z ~ dz/ : CB *d, g/___ i̧

Comorian:	**Sabaki/NEC:**
z ~ dz : CB *l /*g : /___ i̧	z : CB *l/*g /___ i̧

[21]Note that Comorian Class 8 forms do not fall under the domain of this rule: Nzuani /š-a/, /z-a/ 'finger' not /š-a/, /*dz-a/ (< *zy-a < zi̧-a), thus maintaining paradigmatic connection with the more common Class 8 /zi̧-/ which occurs before C-initial stems. However, application of the rule could well be blocked by an earlier syncopation or glide-deletion rule in which Class 8 /zi̧-/ before a V-stem root becomes /z-/.

The voiced affricate /dz/ has a number of sources in Comorian; most can be demonstrated to derive from /z/, while others derived from *j in postnasal positions, see §8.1.4. Nearly all the cases of /dz/ (: CB *g/___ į) found in the data base occur following a nasal, and so derive from the Comorian postnasal affrication rule: *s and *z > /ts and dz/ postnasally, for example:

mlandzi 'bamboo'	*mulanzį (CB *-dàngį́)
ndzi 'fly'	*nzį CB *ngį)
-sindzia 'be sleepy'	*-sįnzįla (CB *- tį́ngį̀d-)
-dzima 'whole, healthy'	*nzįma > ndzima 9/10; *įzįma > dzima 5 (CB *-gį̀mà)

The last item is an analogical reformation based on the Class 9/10 form, and/or the Class 5 strengthened form (for the derivation of the latter, see §6.3).

The other cases of /dz/ that concern us here are complex, but usually explainable in terms of *z in the following environments: (a) following and preceding *į (§6.3); (b) preceding *-y- in causatives (CB *į + V; see §5.3 Causatives); and (c) preceding *įį or *yV. In the first case, Comorian /z/ and /ǰ/ become /dz/:

dziwa 'milk'	*įzįWa, cf. hiziwa 'pool' < *kizįwa
dzitso 'eye'	*į-jį́co
mwidzi 'thief'	*mwįvį < CB *-yį́b- (*mwįvį > mwįzį > mwidzi)

In the second case, the causative morpheme *-y- triggers the strengthening, but only in Nzuani and Maore:

-žadza (Ma, Nz) 'fill' (caus.)	*-įjazya, but -ǰaza (Ng)
-rumidza (Ma,Nz) 'call'	*-tumizya, CB *-túm- 'send', but -rumiza (Ng)

In the final case, *VV or *yV conditions the shift:

-dziha (Ng, Nz, Ma) 'bury'	*-zį̀įka, CB *-dį̀į̀k-
-dza (Nz), 'give birth'	*-vyal-, CB *-bį́ád-, but -zaa (Ng)
-dzaya (Ma) 'give birth'	*-vyal-, CB *-bį́ád-, but -zaa (Ng)
-dzama (Nz) 'sink'	*-zyama, CB *-dį̀àm-, but -zama (Ng)
-dzu(w)a (Ma) 'raise'	*-zyula, CB *-dį̀ùd-
-dzuwa (Nz) 'create'	*-zyula, CB *-dį̀ùd-
shidza (Nz, Ma) 'darkness'	*kizya, CB*-yídį̀à-
hidza (Ng) 'darkness'	*kizya, CB*-yídį̀à-
m(u)dzo (Ng, Nz, Ma) 'load'	Com *muzyo (*g-loss) < *muzįgo, CB *mudį́gò

Here the sufficient condition for the shift is the presence of a following high front glide *yV (CB *į̀V), or following VV [+HV]. Both cases of affrication, given the similarity of the features of the conditioning environments, are most likely part of the same rule process, and will be considered as such here. A formulation of the rule, which is discussed in greater detail in §6.3, follows:

Comorian Affrication:

*z > dz / į ___ į
___ yV
___ įį

Here also the affricate is a post-PSA development, ultimately dependent upon Bantu Spirantization, and requiring [+HV] for its operation.

§5.1.4. Swahili /š/ : CB *k/___ į. ND (except Tikuu), Mwiini, Unguja (and some other SD in some lexis) depart from the normal pattern in attesting /š/ from *k/___ į, instead of /s/, the normal reflex of *k in this environment in the rest of Sabaki and NEC:

Swahili:	**Sabaki/NEC:**
š : CB *k /___ į (/š/ : Tikuu /s/)	s : CB *k /___ į

šina (Am, Mv,Ung,Mak,Vu), išina (Mw) 'base of tree trunk'; cf. isina (El), sina (LP,Gi) (*įšįna, CB *-kína)

mšipa (Ung,Mak,Vu) 'vein, tendon'; cf. msipa (Ti), musiha (Gi) (*mušįpa, CB *-kìpà)

-šinda (Ung,Vu,Am), -šiinda (Mw) 'press down'; cf. -sinda (Gi), -sindiha 'push' (Ng), -sindika (Ti) (*-šįnda, CB *-kínd-)

If PSA inherited all its fricatives from a pre-PSA era, we would expect all Sabaki to attest /s/ from CB *k/___ į, as NEC does generally. That this has not happened has several possible explanations; we list the following hypotheses:

a. All Sabaki, including PSW, underwent regular Spirantization and CB *k > s/___ į as generally occurred in NEC, but then PSW innovated a palatalization rule whereby PSA *s became /š/.

b. Bantu Spirantization shifted all of Sabaki *k/___ į to š, with all but parts of Swahili later innovating a process of depalatalization, namely, *š > /s/.

c. Bantu Spirantization treated (proto?) Swahili differently from the rest of Sabaki from the beginning and shifted *k/___ į directly to š, while in the rest of Sabaki and NEC *k went to /s/.

Options (a) and (b) imply inheritance of Bantu Spirantization. Option (c) implies inheritance of CB *k in PSA. Option (a) has some evidence in its favor. There is, in fact, a process specific to Swahili in which *s, which has derived from CB *t and CB *c, does become /š/ after *į :

*įsįmo 'hole' >	šimo (Ung,Am); but cf. kisima (Ung,Am) 'well' (CB *-tímá 'water well')
*įsįki 'tree stump' >	šuku (Am); but cf. kisiki (Ung,Am) 'stump' (CB *-tįkí)
*įsoWe 'your father' >	išoe (Ti); (CB *-yį́có)

In these examples, palatalization is triggered by the preceding Class 5 prefix *i̧-: *s > š /i̧—. We can refer to this as Class 5 Palatalization of *s.

We find other instances of /š/ which are reflexes of CB *k /___ i̧ , and presumably we could extend this rule to these cases as well: CB *-kɪ̨na 'base of tree trunk' : šina (Ung, Mak, Am). However, there are lexical items that have /š/ from *k /___ i̧ which cannot be derived by this rule:

CB *nkɪ̨ngò 'neck' :	šingo (Ti,Am,Pa,Jo,Mv,Ung), šiingo (Mw), but singo (UP,Gi,Ch,Di,Chi,Vu,Mt,Pe,Tu,Mak,Mn), ntsingo (Ng,Mh,Nz), tsingo (Ma)
CB *-kéékɪ̨- 'pass the night' :	-keša (Mw,Am,Mv,Ung,Mak), but -kesa (LP,Ti, Vu), -česa (Gi,Di,Chi)
CB *-yókɪ̨ 'smoke' :	mooši (Mw); moši (Am,Pa,Si,Mv,Ung), but -mosi (LP,Gi,Du,Ti,Mak,Pe,Mt,Nz), móòsu (El), ɟosi (Ng), mwesi (Ma), rosi (Mn)

There is no obvious conditioning factor to explain /š/ in these items. We could set up a generalized form of the palatalization rule that shifts *s to /š/ in Swahili. This will not work without distinguishing *s's that derive from CB *t and CB *c, which do not palatalize, from those that derive from CB *k, which do palatalize, e.g., /-sinzia/ (Ung), /-sinɖia/ (Am) (< *-sɪnzɪ̨la 'be sleepy', CB *-tɪ́ngì-); /siku/ (Ung, Am) (< *nsɪ̨ku 'day', CB *tɪ́kù); /-singa/ (Ung) (< *-sɪnga 'rub', CB *-cɪ̨ng-). These examples rule out a generalized palatalization rule, i.e., Option (a). To make Option (a) work we would have to assume that *s from CB *k was somehow different phonetically from *s from other sources, possibly [sy] or [š] (cf. the dialectal variation in British English: [ɪsyu] ~ [ɪšu] 'issue'). This leads us directly to Options (b) and (c), which assume that the reflexes of CB *k were distinct from those of CB *t and CB *c from the beginning, namely, the former became *š, and the latter became *s. The problem is in deciding between the two, and this is essentially a question of deciding at what point in the history of these languages CB *k began to be treated distinctly: either late in the history of Sabaki, after the various dialects differentiated and then only affecting Swahili (Option c), or earlier, affecting all of the PSA continuum (Option b). There is no easy way to decide between these two; however, we prefer Option (b) for the following reasons:

Both Options (b) and (c) assume that Bantu Spirantization was not a uniform East Highlands Bantu sound shift, but rather one that spread areally and reached NEC and its subgroups somewhat late in the history of eastern Bantu. We argue against Option (c), because it requires us to believe that only CB *k before *i̧ remained in Sabaki until the Swahili period without spirantizing, while all the other consonants in Swahili and elsewhere in Sabaki did. If we assume that CB *k did spirantize early, as seems likely given all the evidence and understanding we have of Bantu Spirantization, we still have to assume its reflex was different from those for CB *t and CB *c, and this takes us to Option (b). If

CB *k spirantized at more or less the same time as the other obstruents, it had to have become a fricative other than *s, and that fricative had to have been *š. Thus we are assuming that all Sabaki attested this innovation, with non-Swahili dialects, plus Tikuu, innovating a depalatalization shift at a later date: *š > s.

Under this option, then, Swahili /š/ : CB *k/___ į is a retention from an earlier proto period. Swahili is generally conservative vis-à-vis the rest of Sabaki. It tends to preserve PSA forms where other Sabaki dialects innovate, e.g., preservation of stopped reflexes of *c and *j and nonlenition of *p, *t, and *k.

We assume the following scenario for CB *k/___ į :

Pre-PSA/NEC:	Spirantization spreads as an early Bantu proto/areal innovation
Early PSA:	As Spirantization spreads to Sabaki languages *k shifts to *š/___ į
Post-PSA:	Depalatalization: *š > s (non-Swahili Sabaki, most SD, Tikuu)
ND, Mw, Ung:	Retention of PSA *š as /š/

A final note on this analysis is in order here. It works well enough for ND (less Tikuu), Mwiini, and Unguja, but it doesn't explain what is happening in non-Unguja SD dialects. In some cases, some SD dialects pattern with ND and Unguja (cf. *įšįna 'base of tree trunk' above), and in other cases do not (cf. *mošį 'smoke' above).[22] An added complication is the /č ~ š/ variation in these dialects (see §4.1.4.2). However, the fact that the etymologies for these SD dialects do match the NEC pattern (s : CB *k /___ į) in some cases (e.g., *mošį) suggests that it is Unguja which may be skewed, falling in with ND and Mwiini as it does in other areas of the phonology and morphology. Where the rest of SD agrees with Unguja we may be seeing the effects of a later situation in which SD speakers are modifying their speech pattern in favor of Unguja (= Standard Swahili) as they have been doing presently.

§5.1.5. Tikuu, Siu, Pate /ð/ : CB *b, d, g/___ į. This correspondence is due to a late dialect-specific shift in which *z becomes /ð/ (: *b, d, g /___ į). This affected all /z/ deriving either from the application of Spirantization or from the local ND/Comorian shift of *f/v > s/z:

Tikuu, Siu, Pate:	**Sabaki/NEC:**
/ð/ : CB *b, *d, *g /___ į	/z/ : *b, *d, *g /___ į

mfuði 'smith' (*mufuzį, CB *-tú̯d-)
məði 'parent' (*muvyazį, CB *-bį́ád-)
-vuðia 'blow on' (*-vu̯vį̱la), cf. -vuzia (Am), -vuvia (SD)

For the connection of this change with Northern Sabaki dentalization, see Nurse (1985b). The [+HV] vowel feature is not necessary to condition this change.

---

[22]Further complicating the picture are some data from Pemba (Khamis 1984): where Unguja has /š/, Pemba has /s/ (/šingo/ vs. /singo/ 'neck', /-šangaa/ vs. /-sangaa/ 'be surprised', /matamši/ vs. /matamsi/ 'pronunciation'); also where Unguja has /s/, Pemba has /š/ (/matusi/ vs. /matuši/ 'insults', /sindano/ vs. /šindano/ 'needle', /sikio/ vs. /šikio/ 'ear'). We are not sure how totally regular this situation is.

§5.2. Bantu Spirantization and chronology. The various subparts of Bantu Spirantization can be ordered relative to one another using stratigraphy. Thus most of what has happened predates PSA:

1a.	*C [αvc] >	[+fric, αvc]	/___ ụ	Pre-PSA
1b.	*p, *b >	f, v	/___ į	Pre-PSA
1c.	*t, *d >	s, z	/___ į	Pre-PSA
1d.	*g >	z	/___ į	Pre-PSA
2.	*k >	š	/___ į	Early PSA
3.	*f, *v >	s, z	/___ i	Post-PSA: Comorian/ND_2 Coronalization
4.	*š >	s		Post-PSA: all Sabaki (including Ti); /š/ retained in Ung and Am, variably in SD)
5.	*z >	ð		Post-PSA: Tikuu, Siu, Pate only
6.	*z >	dz	/ *į ___ *į	Post-PSA: Comorian[23] (see also §6.3)
			___ *y	(see §5.3)
			___ *įV	(see §6.3)

The processes represented by Rule 1 are widely attested outside of Sabaki and are uniformly attested in Sabaki, and thus are probably inherited from pre-PSA periods, if not NEC then earlier. They are therefore of no value in subgrouping Sabaki, since the resultant fricatives are all retentions. Some further chronology for this set of rules can be established by two Southern Cushitic loans, namely, *ʔiliba 'milk' : PSA *įzįWa; *ʔarigʷ 'banana' : PSA *įzįgu. These help substantiate our claim that Bantu Spirantization is a NEC event. Their distribution indicates that their borrowing could only have occurred in the course of early PNEC contact with early Southern Cushitic speakers in East Africa. All NEC languages show them in forms that have undergone Bantu Spirantization, supporting the reconstruction of *įzįWa and *įzįgu. If these loans had entered PNEC after Spirantization had applied, or if Spirantization were a wholly inherited feature, predating PNEC, we would not find spirantized reflexes of these loans today. Therefore, Bantu Spirantization can only have begun to affect PNEC after the loans were integrated in the PNEC lexicon, and at a time when it was distinct from other East Highlands Bantu languages. Thus, Bantu Spirantization in PNEC, whatever its ultimate source in East Highlands Bantu, and however it might have begun to do its work in PNEC, cannot have been wholly inherited, and must still have been active after PNEC communities were in situ.[24]

---

[23]This rule also affects *j (see §6.3 and 11.4).

[24]Bantu Spirantization, we argue, is a process that occurred in East Africa after contact with early Southern Cushitic peoples. This would also give us an approximate dating procedure. Archaeological evidence suggests that Bantu speakers first entered East Africa in the closing centuries B.C. and encountered Southern Cushitic communities already in place. Although we have no archaeological evidence that specifically identifies the PNEC community as separate from other Bantu-speaking groups at that time, this general time framework allows us to say that Bantu Spirantization, at least in PNEC, occurred at some point after the late centuries B.C., and probably not long after that.

Rule 2 (/š/ < *k) is, as argued by our analysis, a PSA innovation. Its local distribution points to PSA operation of Bantu Spirantization. Any other analysis involves unacceptable complication to explain away the exceptional behavior of CB *k in PSA before *į.

Rules 3 through 6 are all post-PSA events of local distribution and with limited subgrouping value. Rule 3 is shared by ND_2 and Comorian; it contributes to the body of evidence establishing Comorian's identity as Sabaki. Rule 3, furthermore, may attest the presence of *į, though the shift of labial to coronal consonants here can just as easily be explained by a [-HV] high vowel.

Rule 4 which is dependent on the analysis required by Rule 2 does subgroup all but Mwiini, Pate, Siu, Amu, and Unguja.

Rules 5 and 6 provide specifically local isoglosses, respectively distinguishing some ND (Ti, Si, Pa) from the rest of the Swahili dialect continuum, and Comorian from the rest of Sabaki. Rule 6, moreover, requires for its operation the [+HV] vowel *į, and points to the late retention in PSA of seven vowels (see §5.4, and §10.1 for further discussion).

Several of these rules are ordered relative to several other shifts previously discussed. The Comorian/ND_2 Coronalization shift of *f/v > s/z must be ordered before Comorian Affrication. This ordering fits that predicted by stratigraphy: the first is shared by two major groups; the second is specific to Comorian. An example illustrating the derivational relationship between the two rules is:

*mubįádį 'parent'	CB
*muvyazį	PSA
muzyazį	(1) Comorian/ND_2 Coronalizaton
mudzazį	(2) Comorian Affrication (z > dz /___ y)
mudzadzį	(3) Ad hoc assimilation (z > dz /dzV___)
m(u)dzade	(4) Ad hoc deaffrication (Ma) (dz > d); final output of all rules

Comorian Affrication also postdates Comorian *g-loss, a shift discussed in §4.2.3 :

*mudįgò 'load'	*?arig^{w} 'banana'	
*muzįgo	*įzigu	PSA
muzįØo	įzįØu	(1) *g-loss
mzyo	įzyu	(2) vowel-gliding (i > y/___ o)
mdzo	dzu	(3) Comorian Affrication; final output

Stratigraphy also fits this etymology. Both etymologies support the analysis above that the Comorian affricates are late changes independent of Bantu Spirantization. The Tikuu-Siu-Pate shift in which ND *z shifts to /ð/ follows the more widely attested Comorian/ND_2 Coronalization shift in which *v/___ i became /z/; again stratigraphy and rule etymology match:

*-vụvịla 'blow on'	Sabaki
-vuzila	(1) Comorian/ND_2 Coronalization (+ other rules)
-vuðila	(2) Tikuu-Siu-Pate *z > ð

§5.2.1. *p-lenition and Spirantization. Lenition weakens *p in intervocalic environments, but not after a nasal nor after *į (see §4.1.1). The other phonological process affecting CB *p is Spirantization, which created PSA *f from CB *p/___ į. As seen in the following derivation, and as assumed by our analysis, Spirantization, or at least the part of the process which operated on CB *p, is historically prior to *p-lenition:

CB *-píc- 'hide'	*-píti- 'let pass'		
-fic-	-pisi-	(1)	Spirantization
---	-hisi-	(2)	*p-lenition
-fitsa	-hisa	Output:	Giryama

The reverse order gives a different and incorrect result:

CB *-píc- 'hide'	*-píti- 'let pass'		
-hic-	-hiti-	(2)	*p-lenition
---	-hisi-	(1)	Spirantization
*-hitsa	-hisa	Output:	Giryama

These examples are typical of Sabaki. Here the correct rule-ordering is consistent with the prediction stratigraphy would make about these two processes. Spirantization in East Highlands Bantu is a much more widely attested shift and thus presumably earlier than *p-lenition. Both stratigraphy and rule order support our earlier contention (§4.1.1) that *p-lenition is an areal change which spread late into the PSA area as the various groups of the dialect community were beginning to diverge and go their own ways. We also concluded that Bantu Spirantization, as far as Sabaki is concerned, is a mix of genetic and areal factors, as indicated by Sabaki's treatment of CB *k before *į (§5.1.4) which affected Sabaki somewhat later than it did languages to the south of the group.

Although the etymology we have outlined just above indicates that Spirantization must be ordered prior to *p-lenition, there are some data that show that possibly these two processes may have operated at more or less the same time, or at least within such a time frame to allow the reverse of the usual rule-ordering. Within Sabaki, for a few lexical items (there may be others, but not in our reconstructed PSA lexis), we get a curious mix of reflexes, some of which are consistent with an earlier operation of *p-lenition vis-à-vis Spirantization, and some of which are consistent with the regular rule order, as noted in the above etymology. Note the Sabaki reflexes of CB *-pígid- 'sweep'; the forms are listed in columns according to whether they undergo Spirantization or not; if available both the verb and the derived noun are given:

[+spirantization; -lenition]	[-spirantization; (+lenition)]
-fyagia (Mak)	-ϕea; yuϕeo (LP)
-fagia; ufagio (Ung)	-βe(y)a; peleo (Ng)
-fyela (El)	-pea; upeo (Am)

The items in the left-hand column show normal Spirantization, where CB *p > f/__ į; those on the right have not undergone Spirantization, and consequently, where *p-lenition

has applied in Sabaki, they show weakened reflexes, e.g., Pokomo and Comorian. Rule-order 1-2 applies to the forms in the first column, and order 2-1 in the second. Such a situation can arise where two changes spread at different rates. This is intriguing evidence supporting the possibility that even some pre-PSA stages of Spirantization (e.g., Stage 1b, see §5.2) were operating in Sabaki after the PSA community was differentiating and at approximately the same time that *p-lenition was operative.

At least two other examples of this can be cited: *-fyokoca 'twist'; *-fyuka 'go off on a tangent, take a short cut, deviate':

[+spirantization, -lenition]	[-spirantization, +lenition]
-sokọ̱a (Am)	-ɸyokotsa (LP)
-fyokota (Ung)	-vuha (Nz)
-fyuka (Ung)	

Here Lower Pokomo /-ɸyokotsa/ and Comorian /-vuha/ are consistent with *p-lenition; the others in the set derive from Spirantization.

Another interesting case is the palatal /š/ in Giryama. In most cases, /š/ transparently derives from *fy, as the following illustrates:

*kifyu 'knife'	kišu (Gi), kifyu (Vu), kisu (Am), kyifyu (Du)
*-fyuka 'deviate'	-šuka (Gi), -fyuka (Mv, Ung)
*-fyoma 'read'	-šoma (Gi), -fyoma (Du, Vu, Pe, Mt)

However, there is a lexical item whose derivation is not straightforward:

*-pya 'new'	-ša (Gi), -ɸya (LP), -via (Nz), -βya (Cl. 3) ~ -pya (Cl. 9/10) (Di, Du)

Given both sets, it would seem that Mijikenda /š/ has two sources: *fy, and *py. The shift of *fy to /š/ parallels the Mijikenda shift of *vy to /ž/, e.g., *-vyala > /-žala/ (Gi), whose motivation is similar to that seen for Comorian//ND_2 Coronalization (§5.1.1). The second source, *py, is a complicated matter. What is happening here? Has *py directly shifted to /š/ in Giryama, or are intermediate stages involved? Data in Digo and Duruma suggest that lenition (§4.1.1.1) played a role in its derivation: Digo/Duruma /-βya/ is a product of lenition: PSA *-pya > /-ɸya/ (via lenition) > /-βya/; also -βiera 'sweep', ruβiero 'broom' (CB *-pɟ́ágid-). This leads us to think that /-ša/ 'new' in Giryama derives from an intermediate fricative + glide sequence, namely, /ɸ/ via lenition, or possibly an intermediate /f/ via spirantization. If the latter, then spirantization is the source: *py > fy > /š/, (see Giryama /-hopa/ 'be soft, impressionable' and /-hoša/ 'make an impression', where /-hoša/ < -hofya < *-hopya < ?CB *-tòpè 'mud'). We are unable to decide between the alternatives here. Nevertheless, the opposition in the data shows the same kind of rule interference, and is further indicative of late operation of at least parts of Spirantization within Sabaki. We also find similar oppositions in eastern Bantu, as evidenced in Guthrie (1969-71). This disjunction is evidence that *p-lenition and Spirantization of CB *p were not as uniform in their application as stratigraphy might indicate (see beginning of §5.2).

§5.3. The PSA Causative *-y-.[25] We have reconstructed the PSA causative as *-y-, an inherited feature from a pre-PSA stage corresponding to CB *į. Here we are only concerned with *-y- as it interacted with stem-final consonants, and not with the productive "modern" causative, e.g., Swahili /-iš-/ (-pik-iš-a 'cause to cook') which derived from stem-final stative/neuter (*-ik + *-y-). The behavior of this morpheme closely parallels the spirantizing effects of the HV *į. There are two notable differences: (a) in some cases affricates are produced, e.g., /ts/ and /dz/ in Comorian; and (b) *-y- shows up on the surface as /y/ in some dialects. Otherwise, the effects of Bantu Spirantization and *-y-causative are parallel, most notably in the treatment of CB *k. The following table lists the full set of reflexes where known, and gives the PSA and CB reconstructions:

Sabaki Spirantization in Causatives

CB	PSA	MK	Ng	Nz/Ma	Po	ND	SD	Ung
*pį /__[+caus]	*fy	š	s	s?	fy?[26]	s	fy	fy
*tį /__[+caus]	*sy	s	ts	ts	s	s	s	s
*kį /__[+caus]	*šy	š ~ s[27]	s	s	s	š	š	š
*bį /__[+caus]	*vy	ž	z	dz	vy	z, ð	vy	vy
*dį /__[+caus]	*zy	z	z	dz	z	z, ð	z	z
*gį /__[+caus]	-	-	data not available		-	-	-	-

The motivation for reconstructing *-y- as the PSA causative is twofold: 1. /y/, its reflex, is directly attested mostly following labials in SD and Mijikenda, and in other contexts in Giryama (see Deed 1964 for other examples):

SD/Mijikenda

-pofya (Vu) 'woo'	*-pofya
-ogofya (Ung) 'frighten'	*-ogofya, CB *-yògòp-
-levya (Ung) 'make drunk'	*-levya
-govya (Ch) 'hang up'	*-govya, CB *-gòb-
-enendzya (Gi) 'push along'	*-enend-, CB *-gènd-
-hambya (Gi) 'cause to travel'	*-tambya, CB *-támb-[28]

---

[25]A caveat is in order here; note that the *-y- (+causative) derived from CB *į must be considered distinct from *y (-causative, CB *i), e.g., *kyakulya 'food', which does not trigger spirantization. We do not mark the difference in the text nor in Appendix 3.

[26]For Nzuani and Pokomo (LP), we do not have direct evidence for the causative reflexes of PSA *fy. We surmise their shapes on the basis of behavior in other environments; in Nzuani, /s/ < *fy generally, e.g., /-sonza/ (Nz) < *-fyonza 'suck'. For Pokomo /fy/ is a reasonable guess, given that we have /vy/ < CB *bį (causative).

[27]Giryama (Deed 1964) has a nearly equal number of causatives with /s/ and /š/ (: CB *k); we assume those with /š/ are recent formations due to borrowing, e.g., /°-tangamuša/ 'cheer up' (< ND; cf. Ung /-čangamka/, /-čangamša/), or are the result of paradigmatic leveling with inherited forms with /š/ that derive from *fy (: CB *p), e.g., /-aruša/ 'whiten' (*-elufya); cf. Swahili /-eupe/ 'white', also see Giryama /aša/ < afya (Arabic) 'health'.

[28]Spirantization is blocked by the nasal in this example.

Other reflexes in Giryama also reflect derivation from C+y sources, wherein /š/ (< *fy), and /ž/ (< *vy) are parallel to changes we see in noncausative environments, e.g., /-šolera/ < *-fyola 'scold', /-žala/ < *-vyala 'give birth' (cf. -vyala (Ch, Vu); ž- Class 8 < *vị /___V-stem (see §11.7):

Mijikenda (Giryama)

-ogoša 'frighten'	*-ogofya, CB *-yògòp-, cf. -ogofya (Ung)
-gaža 'divide, share'	*-gavya, CB *-gàb-, cf. -gavya (Ch)
-goža 'hang up'	*-govya, CB *-gòbé 'hook'
-loža 'make wet'	*-lovya, CB *-dòb-

2. *-y- best explains the behavior of the reflex languages, specifically, the derivation of Comorian affricates in causative verb forms. In some dialects of Comorian, affricates turn up in just those environments where CB has *ị followed by another vowel (*ị + V and *y +V):

Comorian

-dza (Nz) 'give birth'	*-vyala, CB *-bíád-
-dzama (Nz) 'sink'	*-zyama, CB *-dìàm-
-zama (Ng) 'sink'	*-zyama, CB *-dìàm-
-dziha (Ng, Nz) 'bury'	*-zịịka, CB *-dìịk-
-tsondza (Ng) 'suck'	*-fyonzya, cf. -fyonza (Ung)
shidza (Nz) 'darkness'	*kizya, CB *-yídìà
hidza (Ng) 'darkness'	*kizya, CB *-yídìà

So the following causatives in Comorian are best analyzed as derived from a proto consonant + glide combination: /ts/ (Ng, Nz, Ma) derives from *sy (CB *t + ị caus.), and /dz/ (Nz, Ma) from *zy (CB *d + ị caus.) and *vy (CB *b + ị caus.):

-rotsa (Ng, Nz, Ma) 'submerge'	*-tosya, cf. -tota (Ung) 'be soaked'
-rahatsa (Nz) 'clean'	*-takasya, cf. -takatifu (Ung) 'holy', < Arabic
-βaβatsa (Ma) 'feel, caress'	*-papasya, CB *-pápat-
-rohotsa (Ng, Nz, Ma) 'boil'	*-tokosya, CB *-tòkòt-, cf. -hokosa (Gi, LP)
-vungudza (Nz) 'diminish'	*-punguzya, cf. -βunguza (Ng)
-žadza (Nz) 'fill'	*-jazya, CB *-yíjad-, cf. -jaza (Ng)
-lodza (Nz, Ma) 'marry (caus.)	*-lozya, cf. -loza (Ng)
-lovya (Nz, Ma) 'soak'	*-lovya, cf. -loza (Ng)
-haradza (Nz) 'forbid'	*-katazya, cf. -kahaza (MK)

Otherwise, the reflexes of the *-y- causative parallel those produced by Bantu Spirantization:

(a) /š/ (ND, Ung; not Ti) and /s/ (other Sabaki) < *šy : CB *k /___ ị + a;
(b) /s/ (Sabaki except Com) < *sy : CB *t /___ ị + a;
(c) /z/ (Sabaki except Nz, Ma) < *zy : CB *d /___ ị + a

(a) **/s/ and /š/ : PSA *šy : CB *k**

*-Wuušy- 'rise up'	-usa (Gi,Ti), -uša (Ung,Am), -wiša (Vu). (CB *-búúk-)
*-keešy- 'pass the night'	-kesa (LP,Ti), -česa (Gi), -keša (Am,Ung), -šesa (Ng). (CB *-kéékį-)
*-lamušy- 'wake up'	-lamusa (Gi,Ng), -amša (Ung). (CB *-dàmuk-)
*-cešy- 'laugh at'	-tsesha (Gi), -t̪esa (Ti), -t̪eša (Am), -češa (Ung), -tsesa (Ng,Nz), -tsesedza (Ma); cf. *-ceka 'laugh'

(b) **/s/ : PSA *sy : CB *t**

*-pukusy- 'strip off'	-βukusa (Di), -pukusa (Am,Ung). (CB *-púkùt-)
*-tosy- 'drown, sink'	-hosa (Gi), -tosa (Am,Ung), but -rotsa (Ng,Nz,Ma). (cf. -tota Ung 'be soaked')
*-rohosya 'boil'	-hokosa (LP,Gi), -tokosa (Am,Ung), but -rohotsa (Ng,Nz,Ma). (CB *-tòkòt-)

(c) **/z/ : *PSA *zy : CB *d**

*-ongezy- 'increase'	-ongeza (LP,Gi,Am,Ung), -oogeza (El), -enǰeza (Ng). (CB *-yòngìd-)
*-pozy- 'cool'	-hoza (Gi), -poza (Am,Ung). (CB *-pòd-)
*-lozy- 'marry (caus.)'	-yoza (LP), -loza (Gi), -oza (Am,Ung), -loza (Ng), but -lodza (Nz,Ma)

The analysis assumes that the CB causative morpheme was *-y- (see Meeussen 1967) and that PSA inherited from a pre-PSA stage a series of fricatives followed by *y: *fy, *vy, *sy, *šy, and *zy. The support for *fy and *vy is ample and obvious, but the main motivation for *sy and *zy is the Comorian affricates, which at this point seem to be best explained by a glide. Were it not for our assumption that Comorian is part of Sabaki, we would not necessarily need to posit glides. Finally, *šy, i.e., *š plus *y, is reconstructed simply to preserve patterning with the rest; it is not necessary to posit *y here, since affricates in Comorian do not arise from /š/. It is, however, both sufficient and necessary to specify this segment as a palatal in order to keep the reflexes of CB *t and CB *k separate, as described earlier for Bantu Spirantization in PSA. Reflexes of *t here and in Bantu Spirantization are nonpalatals, whereas those derived from CB *k are palatals in some dialects but not in others. The use of /y/ here is intended to capture the probability that the two original segments were phonetically different. The rules necessary to derive the modern reflexes follow:

1.	CB *p/b /___ i̧ + V	:	*fy/vy	Pre PSA
	CB *t/d /___ i̧ + V	:	*sy/zy	Pre PSA
2.	CB *k /___ i̧ + V	:	*šy	Pre-PSA or early PSA
3.	PSA *fy/*vy	>	sy/zy	Com/ND_2 (Coronalization; see §5.1.1)
4.	Com *sy	>	ts	Com (Com Affrication §6.3)
	Com *zy	>	dz	Nz/Ma causatives only
5.	PSA *sy, zy	>	s, z	PSA
6.	PSA *fy	>	š	MK (Gi, Ch, Ka; not Di, Du)
	PSA *vy	>	ž	Gi (not Ch, Di, Du, Ka, Ra)
7.	PSA *šy	>	s	Sabaki (not Si, Pa, Am, Mw, Mv, Ung)
	PSA *šy	>	š	Mw, Si, Pa, Am, Mv, Ung

There is some lexical evidence that Rules 4 and 5 overlapped in Comorian, or that in some dialects at least there was some conflict in application. Sacleux (Chamanga and Gueunier 1979), for example, lists /-rosa/ and /-rotsa/ 'drown' (*-tosya; cf. Ung /-tota/ 'be soaked'). Our data, however, did not reflect this variation within dialects. Such variation can, however, arise from paradigmatic leveling on analogy with causatives derived from CB *k which have /s/ in Comorian, e.g., /-tsesa/ (Nz) 'make laugh' (*-cek-). Analogy operating in the causative paradigm is a likely source of irregularity vis-à-vis these rules in other Sabaki dialects as well, e.g., /-piša/ (Ung) 'let pass' (*-pisya, CB *-pít-) should be /*-pisa/ according to the rules, but is not. Causatives with /š/ in Swahili predominate statistically so reformation along such lines is easily the explanation.

§5.4. PSA seven vowels, and vowel neutralization. Usually associated with Spirantization is Bantu Vowel Neutralization, in which the two super-close Bantu vowels *i̧ and *u̧ merge with *i and *u, respectively, thus reducing the original Bantu seven-vowel system to five-vowel systems in the reflex languages. Vowel reduction does not appear to have occurred anywhere in East Highlands Bantu unless Spirantization has operated, and generally, when Spirantization operates, Vowel Neutralization follows. Bantu Spirantization has the effect of increasing the consonant inventory of the reflex languages, Vowel Neutralization decreases the vowel inventory. The increase in the consonant inventory makes redundant the extra vowels whose lexical function is assumed by the new fricatives and spirants. Thus, when Bantu Spirantization operates, we generally find five-vowel systems, as we have in the present-day Sabaki languages. There are, however, exceptions: Sukuma and Nyamwezi show some effects of Spirantization, but still have seven vowels. Several Lake Corridor languages, e.g., Safwa (M25) and Nyakyusa (M31), have Bantu Spirantization and also retain seven vowels. Meinhof (1932:145) reported that some speakers of the Konde dialect of Nyakyusa were merging the close and open high Bantu vowels, while Ndali, another Nyakyusa dialect, has spirants and five vowels (Nurse 1979a:447). The Southern Highlands all attest Bantu Spirantization, but split down the

middle for the vowels, some having five vowels, some seven (Nurse 1979a:440f., Ngonyani 1989).

The reflex Sabaki languages are all five-vowel languages, with the exception of Elwana. Thus, the fact that virtually all extant Sabaki and NEC languages have spirantized consonants and five vowels might suggest that PSA inherited a five-vowel system, as well as Bantu Spirantization. We argue against this possibility, and we have outlined the evidence for this in §10. Our major argument hangs on the fact that Bantu Spirantization was both a partially inherited feature from NEC and partially a PSA event, and therefore PSA had to have seven vowels. An other major argument hinges on the need to posit *į and *ų to explain Strengthening in Comorian and SD (§6 below), and *į to explain Comorian *z-affrication (§6.3).

Because seven vowels were needed until quite late in the history of Sabaki, the rule that ultimately reduced the seven to five was equally late. We have no evidence nor thoughts on how this might have happened, that is, in piecemeal fashion or in one fell swoop. The nine phonetic vowels of Elwana, in spite of our uncertainly about their precise phonemic status, suggest that it was not a totally uniform process. Also, a seven-vowel system may have lasted longer in some dialects than in others, e.g., Comorian versus nearly all other Sabaki dialects.

§6.0. Strengthening (Comorian and SD)

SD and Comorian attest a phenomenon in which certain consonants, which otherwise would show the effects of other changes, such as segment-lenition or even deletion, do not undergo the expected shifts. Typically, where this happens the segment is preceded by a super-close vowel ([+HV]), usually CB *į but sometimes CB *ų. Most examples in Comorian are taken from Class 5, whose prefix we have reconstructed as PSA *į-. In Comorian, alternations are common in Class 5 and in the adjectival paradigm (Chart 13):

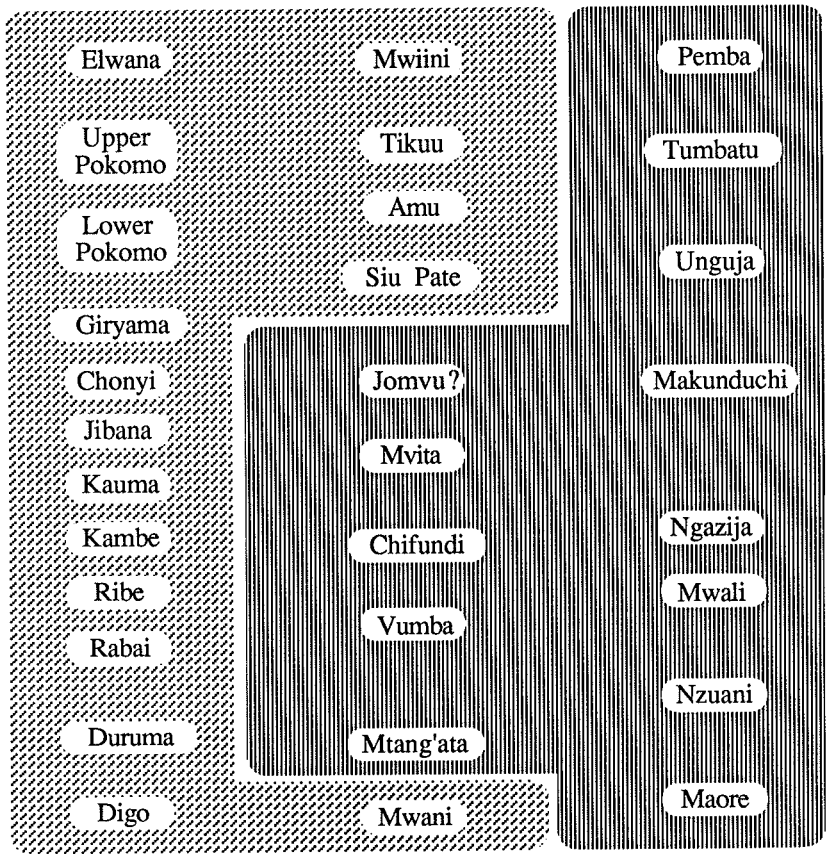
Comorian Alternating Forms

paa/maβaa (Ng) 'roof'	*įpala/ma-
pukuli/mavukuli (Nz) 'rat sp.'	*įpuku/ma-, CB *-púkù
trumbo/marumbo (Ng) 'belly'	*įtumbo/ma-, CB *-tùmbò; cf. humbo (Gi)
kapwa/mahapwa (Nz) 'armpit'	*įkwapa/ma-, CB *ku-yápa 15
bi/mai (Ng) 'bad (adj.)'	*-Wi(Wi), CB *-bíbì; cf. -wi (Am)
bwe/mawe (Ng, Nz) 'stone'	*įWe, CB *-bùè; cf. iwe (Gi)
dongo/malongo (Ng) 'soil, clay'	*įlongo, CB *-dòngó; cf. ulongo (Gi)
gamba (Ng) 'turtle shell'	*įgamba, but nyamba ~ ŋamba 'turtle'
gandzi/mahandzi (Ma) 'throat'	(Blanchy 1987)

The PSA voiceless stops regularly undergo lenition in Comorian: *p > /β/ (Ng, Ma), /v/ (Nz), *t > /r/, *k > /h/, but not those following *į. Similarly, PSA *W, PSA *l, and

Chart 13

Sabaki Strengthening



Key

PSA *W : [-stop] (e.g. /w ~ Ø/, etc. / *į ___

PSA *W : ɓ / *į ___ e.g. *-įWa 'steal', *įWe 'stone'

PSA *g become /w/ (or Ø), /l/ (or Ø), and Ø respectively, but following *į these are realized as imploded /ɓ/, /ɗ/, and /g/ respectively.[29] Thus in Class 5 we find synchronic alternations, as exemplified immediately above, involving these pairs: /p ~ β or v/, /t ([t] ~ [tɾ]) ~ r/, /k ~ h/, /ɓ ~ w or Ø/, /ɗ ~ l/, and /g ~ Ø or h/. There are few examples of alternating /g ~ Ø/ because of paradigmatic leveling. Some examples of nonalternating, non-Class 5 items, and reflexes of *g are:

Comorian Nonalternating Forms

-piha 'cook'	*-įpik-, CB *-yìpik-
-hiba (Nz) 'steal'	*-įW-, CB *-yíb-; cf. -iwa (Am, Mn)
mvuba 'bellows'	*muvụwa, CB *-gùbà; cf. muvuwo (Gi)
ifuba 7 (Ng), šifuba 7 (Nz) 'chest'	*-fụWa, CB *-kúbà; cf. kifua (Ung)
uwade/mbwade 11/10 'sickness'	*-lwal-, CB *-dúáídé ; cf. -lwala (Mn)
-du 'black'	*-įlu, CB *-ídu; cf. -iru (Gi)
goho 5 (Ng) 'hollow trunk'	*įgogo, CB *-gò(ò)gó; cf. -haha (Ng) 'go bad' (CB *-gàg-)
gana 5 (Ma) 'one-hundred'	*įgana, CB *-gànà
gamba 5 (Ng) 'turtle shell'	*įgamba; cf. amba (Ti), ambe 'turtle' (Mn)

Though there are sporadic traces of Strengthening in other Sabaki dialects, it is SD which most clearly shows evidence of the phenomenon. Some clear-cut examples follow:

Southern Dialects

biwi/ma- 'refuse heap' < *įWiWi, CB *-bí 'bad'; cf. wiwi ~ ʋiʋi (Am, LP)

buyu/ma- 'baobab fruit' < *įWuyu, CB *-bùyú; cf. uyu 14, muyu 3 (Gi)

bega/ma- 'shoulder' < *įWega, CB *-bègà; cf. mwega (Pe) 'support'

domo/ma- 'large lip' (Ung, Mak), muomo 'mouth' (Mak) < *-lomo, CB *-lòmò; cf. omo/ma- 'prow' (Am),

dume/ma- 'male (animal)' < *-lume, CB *-dúmè; cf. -lume (Gi)

dumbu (Mak, Pe, Vu, Mt) 'co-sibling' < *įlumbu, CB *-dùmbù; but umbu (ND, Ung)

-iba 'steal' (Ung, Mak, Vu, etc.) < *-įWa, CB *-yíb-; cf. -iwa (Mn)

-ziba 'stop up' (Ung, Mak) < *-zįWa, CB *-díb-; cf. -ziʋa (ND), -zia (Gi); also kizibo 'stopper'; cf. kiðivo (Ti)

ɟibwe ~ bwe (Vu, Mak, Chi, etc.) < *įjįWe, CB *bùè

SD should pattern with the rest of the nonstrengthening Sabaki dialects and attest /l/, /w/, or Ø in these examples, as the comparison items show. That they do not, points to the effect of the super-close vowel *į. Other examples are less straight-forward, largely be-

[29]We have not marked implosion in examples unless they are enclosed within phonemic brackets since it is predictable in both Comorian and SD. In Ngazija there is an exception: /b/ and /ɓ/ do contrast postnasally (see §8.2).

cause paradigmatic leveling has altered the picture, but there are still many examples, some involving alternation, which also show the effects of the *ị:[30]

Unguja

-bovu 'rotten (adj.)', mwovu 1 'evil person' < *-Wovu, CB *-bòd-; cf. -ovu (Am)
-bivu 'ripe, cooked' < *-Wivu, CB *-bídù; cf. -wivu ~ -ʋivu (Am, LP)
bongo/ma- 5/6 ~ ubongo 14 'brain' < *Wuwongo, CB *-bòngó; cf. ongo/ma- (MK)

udongo 14 'soil, clay', dongoa/ma- 5/6 'lump, mass', donge/ma- 5/6 'small rounded lump' < *Wulongo, CB *-dòngó; cf. ulongo (MK), uʋongo (Am)
mdomo 3/4 'mouth, lip', domo/ma- 5/6 'big lip' < *-lomo, CB *-dòmò; cf. mulomo (MK), muyomo (LP)

-jigamba 'boast' < *-amba, CB -gàmb-; cf. -amb(i)a (Ung)
-jidamka 'wake up early' < *-lamuk-, CB *-dàmuk-; cf. -amka 'wake up' (Ung)

Many of these do not have alternating stems within Swahili. Thus, for example, /-ɓovu/ is the form of the adjective throughout its paradigm (/miɓovu/ 4, /ɓovu, maɓovu/ 5/6, /kiɓovu/ 7, etc., but cf. /mw-ovu/ 'evil person'); this is also true for /-ɓivu/, but cf. /-iv-/ 'be ripe'. The fact that they do have /ɓ/ here can be explained through analogical leveling on the model of the strengthened Class 5 forms: /-ɓovu / and /-ɓivu/ (or on the basis of Class 9/10 forms: /mbovu, mbivu/). For Swahili, then, we would have to set up alternating forms, e.g., *-ovu 1/2, 3/4, etc. ~ *-ɓovu 5 ~ *mbovu 9/10. Subsequent reformation gives us today's /-ɓovu/. Even now on the basis of /mwovu/ we could argue for an underlying synchronic /-ovu ~ -ɓovu/.

For the Class 14 items we can appeal to a similar mechanism. The only explanation for the stops in these forms is that they are modeled on the Class 5 forms with the strengthened consonant; note the Mijikenda item /ongo, maongo/ 5/6, and the co-existent /ɓongo/ 5 (Ung). Presumably, then, an item like /uɓongo/ is a derivative form based on the Class 5 item /ɓongo/. For /uɗongo/, the Class 5 strengthened form /ɗonge/ served as a model.[31]

---

[30]The effects of strengthening in Sabaki are not uniform. There are many counter-examples, those where plosives should be attested and are not, and vice versa. For example, /wimbi, mawimbi/ (Ung) 'wave' < *ịWimbi (CB *-bìmbì) should be *ɓimbi/mawimbi. Irregularity in Unguja could also be traced to borrowing from ND where strengthening is not attested, beyond a few isolated examples, which may be borrowings themselves from SD. Also, while Comorian shows alternating Class 5 and Class 6 forms, SD usually does not. Thus, irregularity and non-alternation can be explained by paradigmatic leveling. If we assume regular operation of strengthening in SD, then we probably once did have a singular *ɓimbi alternating with /mawimbi/. If the plural were the more frequently used form, then a reformation of the singular, from *ɓimbi > /wimbi/, would make sense. For Comorian, Tucker and Bryan (1970) give examples of the sort /ɓega, maɓega or maega/, which illustrate leveling. A different case, which cannot directly be traced to *ị or to analogy, is exemplified in the rural dialects of Zanzibar and Pemba: /-ɗalia/ 'go to bed' (cf. Ung -lala), /-ɗawa/ 'go out' (cf. Com -lawa), /-ɗilia/ 'eat' (cf. Ung -la), /-ɗipa/ 'pay' (cf. Ung -lipa), /-ɗomba/ 'ask' (cf. Ung -omba), /ɗumbi/ 'dust' (cf. Ung vumbi), /mwiɗa/ 'song' (cf. MK /mwira/).

[31]The items /ɗonge/ 'small rounded lump', /ɗongoa/ 'lump, mass', and /uɗongo/ 'clay, soil' are semantically related: *ɗonge: *something done with clay > piece of clay > lump, mass, small rounded lump; and

Furthermore, a connection with Class 5 is reinforced paradigmatically since the plural of Class 14, where it exists, is Class 6, the plural of Class 5. We should be careful to note, however, that we are not presuming that all imploded stops in Swahili derive from strengthening, or from paradigmatic sources. Some are due to borrowing, and the source of others is unclear.

In the rest of Sabaki, with the exception of Mvita (and presumably Jomvu), the evidence for Strengthening is not at all regular. The irregular distribution suggests borrowing, presumably from SD. A few examples of such items follow:

kibago (LP) 'stool' < *kiWago CB *-bàgò; cf. *-ị̀Wa, -iwa 'steal' (LP)
kibao (Mn) 'board' < *kiWago CB *-bàgò; cf. *-ị̀Wa, -iwa 'steal' (Mn)
bua/ma- (Am, Ch) 'stem' < *ị̀Wua
bawa/ma- (LP), ibava (Ti) 'wing' < *ị̀WaWa CB *-bàbá
povu 5 (LP) 'foam' < *ị̀povu CB *-pódṳ̀
buyu (LP) 'baobab fruit' < *ị̀Wuyu CB -bùyú
fudu((fu)du) (LP, Pa, Si, Ti) 'tortoise' < *ị̀fṵlu ~ *ị̀fṵvṵ CB *-kṵ́dù ~ *-kṵ́dṳ̀; cf. kifuvu-mayoṇḍi (Am)

These are idiosyncratic attestations of strengthening. Pokomo shows the most examples, but in general it attests a mass of Swahili loan words, which could be the source. Guthrie (1967-71, vol. 2, p. 46) reports that CB *p and CB *d are retained after *ị̀, but lists only one example: /puya, maɸuya/ < *mpuda 9/10, CB *-pùdà. Our data do not support his analysis. Other than the above exceptions, Pokomo regularly undergoes lenition of *p > ɸ and *l > y and other changes in the environment where we expect strengthening. Strengthening is only regularly and for the most part, consistently attested in SD and Comorian. There are, however, some counter-examples where the change has not applied, e.g., /ɟiwe/ 'stone' (Ung) < *ị̀We, (CB *-bùè), but other SD and Comorian have /bwe/ (Mak, Tu, Com). Such counter-examples are either due to analogical back-formation on the plural form (e.g., /mawe/) or are loans, in the case of Unguja, from ND dialects which do not attest Strengthening, beyond a few scattered examples.

Mvita, which otherwise usually patterns with ND, shows evidence for Strengthening: /-iɓa/ 'steal', /-ziɓa/ 'stop up', /mwiɓa/ 'thorn'. These could be borrowed from SD, but may as well have the same source as that for SD. We reserve judgment.

Outside Sabaki, in NEC, we again have sketchy evidence of Strengthening.[32] Our data searches have come up with a few sporadic examples:

---

*ɗongoa: *something removed from soil > lump of soil > lump, mass; Sacleux (1939) also lists /ɗongo/ 'lump of dirt'.

[32]Nurse (1979a) observed for Seuta that "all these languages are like the Swahili dialects in that they can produce /ɓ, ɗ/ from PB sequences of *ị̀ + *b/*d + V" (p. 414); for East Ruvu: "some /ɓ, ɗ/ in the same environment as described for Sabaki and Seuta" (p. 417); and the same sort of generalization for Lugulu and West Ruvu (p. 419). Guthrie (1971, Vol. 2, p. 49) claims CB *b > ɓ and *d > ɗ in Class 5 in Nhwele (Kwere, G32) versus Ø and /l/ respectively in other environments, but supplies no examples.

dibwe 'stone' (Kwere, Kami, Kutu, Sagara), but -we (Seuta)
-iba 'thorn' (Zaramo, Kutu), but -iwa (Seuta)
-bongo 14 'brain' (some Seuta, some Ruvu), but -ongo 14 (Shambala, Doe, Kwere, Luguru, Kaguru)
-devu 7 'beard' (Shambala, some Ruvu, Luguru, Gogo), but -(l)evu 7 (Bondei, Sagara)
-bovu 'rotten, bad' (Shambala, Bondei)
-donda 'wound' (Kami, Gogo), but -(l)onda (Seuta, Ruvu)
-bua 'stalk, stem' (Seuta, Ruvu)
-ziba 'stop up' (Bondei, Zaramo, Kami)

In Pokomo, and frequently throughout Seuta and Ruvu, the Class 5 verbal prefix is /ɗi-/ (< *i̧li-); this is also found in other environments, e.g., in Class 5 monosyllabic stems in Nhwele (Kwere).

Further afield, beyond NEC, we also find strengthening in Class 5 nouns in Nyanja :

tsiku/masiku 'day' (*-tḭ́kù)
phili/mapili 'hill' (*-pìlì)
khasu/makasu 'hoe' (*-kácù)
dzila/mazila 'egg' (*-li̧là)
thako/matako 'buttock' (*-tákò)

Ganda gemination is similarly motivated (Meeussen 1955, Mould 1974, Clements 1986:64ff.):

eɲɲembe 5 'horn' (CB *-pèmbè)
essiga/amasiga 5 'cooking-stone' (CB *-pi̧gà)
ettaka 5 'soil' (CB *-tàkà)
effumu 5 'spear' (CB *-tu̧mò)
-bba 'steal' (CB *-yi̧b-)
-tta 'kill' (CB *-yi̧t-)
amazzi 'water' (CB *-yígi̧)
oluggi 'door' (CB *-yi̧gì)

Here *i̧ conditions the shift of /s/ and /z/ to affricates, voiceless stops to aspirated voiceless stops in Nyanja, and nongeminated segments to geminated ones in Ganda.[33] Strengthening in Comorian and SD is parallel.

Because of such data, there are two issues that require discussion. One is whether Strengthening was a pre-PSA phenomenon; the other concerns the nature of the underlying conditioning mechanisms of Strengthening itself (§6.1 below). As to the first, we assume that Strengthening is post-PSA, having independently innovated and affecting only part of the Sabaki dialect continuum, namely, SD and Comorian, presumably at a time when their respective speech communities were still in contact. Because of the lack of pattern and regularity in the attestations in e.g., Mijikenda, Pokomo, and ND, in contrast to the consistency of SD and Comorian, we have to conclude that Strengthening did not affect those dialects (although the evidence in Pokomo is somewhat more consistent). To assume otherwise, we would have to explain how the effects of Strengthening were undone in these dialects. Analogy can only go so far in explaining how dozens, perhaps hundreds, of

[33]For a parallel shift in Comorian where *z > /dz/, see §6.3.

forms no longer show the effects of the change. We therefore conclude that they were not there in the first place, and that the few examples that we do have are more consistent with a borrowing hypothesis. There is, however, the evidence provided by non-Sabaki NEC languages. Here the data is more difficult to judge. Seuta and Ruvu, with some exceptions, attest a strengthened Class 5 verbal prefix (along with Pokomo). But otherwise, only about a half-dozen items out of a data base of over a thousand show evidence of Strengthening, and even in these few items the various Seuta and Ruvu dialects do not behave consistently. Borrowing from SD is a likely explanation: these languages are adjacent to SD. They have, moreover, small communities of speakers and have come under much influence from Swahili. Based on this kind of evidence, we are unwilling to claim that a productive form of Strengthening pre-dated PSA, although it is possible that some sort of drift-like tendency for Strengthening did pre-date PSA. We are willing to concede that the sporadic nature of the data outside of SD and Comorian is evidence of this. However, only Comorian capitalized on this in a regular fashion in NEC proper, and only SD in a less regular way.

Finally, because of the relatively remote connection between Sabaki and the Nyanja dialect cluster, and the somewhat different manifestation of the phenomenon in Nyanja, we conclude that independent convergent innovations are responsible, though we leave open the possibility for genetic causation.

§6.1. **Stage 1 Strengthening (Comorian and SD).** In this section we consider the motivating mechanisms of Strengthening, and its multi-staged development, first as a phonologically conditioned set of changes and then as a morphologically conditioned one.

We considered two hypotheses about the nature of Strengthening and its conditioning mechanisms in Sabaki. These have to be seen in relation to other changes which have operated in Sabaki. These would include Spirantization and the various changes affecting *W, *l, *g, *p, *t, and *k. The two hypotheses are that either a super-close *į was the motivating operator, or that loss of a Class 5 vowel prefix (either *į or *i) was responsible.

There is evidence for both in the data. The second hypothesis is attractive because most examples of strengthened consonants occur in initial position in Class 5 nouns, after a reconstructed *į which has been deleted, thus hypothetically *įjįWe 'stone' > įwe > Øwe > /ɓwe/, where /w/ > /ɓw/ as a consequence of V-loss. There are also a number of older *į-initial verb stems, which in Comorian have both lost the initial vowel and have strong rather than weak reflexes of the post-vowel consonant, e.g., /-piha/ (not *-βiha) < *-įpik-. Other examples can be cited: /mwaɗe/ (Ng) 'sick person' < *-(lw)aįl- (CB *-dúáįd-), /-ɗutsi/ (Ng) 'very black' < *-įlu + įfį (CB *-yį́d- + -pį̱į̱pį́). However, this vowel loss cannot account for those SD or Comorian forms where strengthened stops are attested without vowel loss: /-(h)iɓa/ (Ung, Com) 'steal', /muvuɓa/ (Com) 'bellows', /šifuɓa/ (Ng) 'chest', /-ziɓa/ (Ung) 'stop up', /mwiɓa/ (Ung) 'thorn'. All these examples are reconstructible with *į or *ų. We conclude that the minimum constraint on the form of the rule is the [+HV] quality of the vowel. Loss only makes sense if it occurs after strengthening.

Although the articulatory basis of strengthening is not totally clear, two aspects of the process stand out. One is that it takes place in the context of close vowels, that is, vowels with maximal tongue height, and therefore minimal oral aperture. The other is that this environment retains or produces stops, that is, consonants with total closure of the oral tract. Thus the height of the vowels is central to the process. Note also that the associated segments in all the languages here are "long": Luganda geminates : Nyanja affricates : SD/Comorian implosives.

Stratigraphy indicates that Strengthening within Sabaki is multi-staged. SD and Comorian both attest a stage which led to the development of the following alternations: /w or Ø ~ ɓ/, /l ~ ɗ/, and /Ø(h) ~ g/. These are still seen most clearly in Comorian, and only somewhat opaquely in SD, where paradigmatic leveling has obscured matters. This first stage we argue was motivated by [+HV]. The next development is only attested in Comorian and resulted in the /p ~ β (v)/, /t (tr) ~ r/, and /k ~ h/ alternations. The following outlines the main divisions of Strengthening Stage 1, and its interactions with other rules:

Comorian: Stage 1 Strengthening after *į and *ų

*ki-fųWa 'chest'	* -įWa 'steal'	*-lwaįl- 'sickness'	
ki-fųba	-įba	-lwaįd-	Strengthening
ki-fuba	-iba	-lwaid-	*7 > 5 Vowels
ši-fuba	-(h)iba	-wad-	Other rules
ši-fuba (Nz)	-(h)iba (Com)	-wad- (Com)	Output

Comorian and SD: Stage 1 Strengthening after *į[34]

*-įWa 'steal'	*į-Wici 'raw'	*į-lomo 'mouth'	
-įba	įbici	įdomo	Strengthening
-iba	ibici	idomo	*7 > 5 Vowels
---	-bici	-domo	V-loss
-iba (SD)	-biči (SD)	-domo (SD)	Output
-(h)iba (Com, SD)	-bitsi 5/-itsi 6 (Com)	domo/malomo (Com)	

At an early point in Stage 1, only Comorian seems to have strengthened *l in non-Class 5 environments, and only Comorian has examples of strengthened consonants after *ų. SD only agrees with Comorian in shifting *w to /ɓ/ after *į and after the Class 5 prefix, and in shifting *l to /ɗ/ following the Class 5 prefix. Another difference is that Comorian kept alternations, e.g., /ɗongo ~ malongo/ 'soil, clay', but SD did so only residually, e.g., /-ɓovu/ 'rotten' ~ /-ovu/ 'evil', having leveled out the paradigm in favor of the stop. Comorian also differs from SD in affricating *z in certain specific environments, but there is evidence to indicate that this is a relatively late change in Comorian (see §6.3). The rest of the changes in Comorian are part of a second stage of strengthening.

[34]Although not exemplified here, we assume the Strengthening influence on *g is also Stage 1.

§6.2. Stage 2 Strengthening (Comorian). The motivation for this multi-stage treatment is twofold: the differential behavior of Strengthening in SD vis-à-vis Comorian, and the nature of the phenomenon itself. Stage 1 affects *W and *l in both dialects, shifting them to /ɓ/ and /ɗ/. At Stage 2 the change is generalized to include *p, *t, and *k. These undergo lenition in intervocalic contexts and are realized as /β ~ v (Nz)/, /r/, and /h/, but in Class 5 environments after *į they are preserved as stops, as illustrated above (§6.0).

Lenition in Comorian is also multi-staged. *p-lenition is shared with Mijikenda and Pokomo, and many NEC languages. *t-lenition is shared with Mijikenda and Pokomo alone, and *k-lenition is specific to Comorian. While lenition of *p and *t is found in Comorian because of its early Sabaki connection, *k-lenition is strictly a Comorian innovation, albeit a generalization of *p/*t-lenition. In Comorian, however, the behavior of *k would dictate the presence of *į to explain its preservation in Class 5 nouns and in many verbs, thus calling for the presence of seven vowels, at least *į, much later than its presence might be otherwise indicated. But because stratigraphy tells us that vowel-neutralization, and *į-loss in Class 5 and *į-initial verbs, are both early post-PSA shifts, we must consider the possibility that *k, and possibly *t, do not undergo lenition in Class 5 nominals for nonphonological reasons.

On the working assumption that we are correct in assuming that the high vowels had neutralized and the Class 5 nouns had lost their prefix by at least the late stages of lenition, we can appeal to morphological, rather than phonological, conditioning to explain what is happening. Stage 1 earlier created a considerable amount of morphophonemically conditioned alternation with /ɓ ~ w(Ø)/, /ɗ ~ l/, and /g ~ Ø(h)/ alternating in singular and plural Class 5/6 nouns. When lenition began to operate at a later period, this paradigm was in force, and a morphophonological condition was added requiring that stops (or affricates) occur in Class 5 initial position and weakened consonants in intervocalic positions. The language code required this: speakers knew that this was the case, and consequently Class 5 initial *p, *t, and *k did not undergo lenition. Thus, at Stage 2 Strengthening, instead of a sound change having been added to the grammar, a condition on the operation of the various lenition rules, as they were innovated, is the explanation for the alternations of /p ~ β(v)/, /tr ~ r/, and /k ~ h/:

Comorian Consonant Lenition (multi-staged)

*p, *t, *k > β(v), r, h / V ___ V {not applicable to Cl. 5 stem-initial *p, t, k}

*į-paa/ma-paa	*į-tako/ma-tako	*į-kala/ma-kala	
ipaa/mapaa	itako/matako	ikala/makala	*7 > 5 Vowels
paa/mapaa	tako/matako	kala/makala	Class 5 V-loss
paa/maβaa	taho/maraho	kala/mahala	*p, *t, *k-lenition
---	traho/maraho	kaya/mahaya	*t > tr; *l > y
'roof/s'	'buttock/s'	'charcoal'	

However, there is some evidence, albeit just a few examples, that the preservation of the voiceless stops is directly attributable to *į and *ų, apart from morphological motivation: /-piha/ 'cook' < *-įpika (CB *-yįpík-), /uku, mauku/ (Nz), /-siku/ 'night' (Ng) < *Wusįku (CB*-tįkù), /-para/ 'get' < *-įpata, etc. These indicate that, even as late as the innovation of *p, *t, *k-lenition in Comorian, the preservation of the voiceless stops was also phonological. In these items there can be no paradigmatic motivation operating to explain the preservation of the voiceless stop. The [+HV] quality of the vowel itself can be the only explanation for the preservation of the voiceless stops in these examples. These further suggest that in Class 5 the prefix *į itself, rather than morphological motivation, explains the preservation of nonlenis reflexes of *p, *t, and *k, and, furthermore, that seven-vowel neutralization and Class 5 V-loss followed lenition in Comorian. In this case we also need a condition on the rule:

Comorian Consonant Lenition Alternative (multi-staged)

*p, *t, *k > β(v), r, h / V ___ V {the rule does not apply if V_1 is *į }

*į-paa/ma-paa	*į-tako/ma-tako	*į-kala/ma-kala	
įpaa/maβaa	įtaho/maraho	įkala/mahala	*p, *t, *k-lenition
ipaa/maβaa	itaho/maraho	ikala/mahala	V-neutralization
paa/maβaa	taho/maraho	kala/mahala	Class 5 V-loss
---	traho/maraho	kaya/mahaya	*t > tr; *l > y
'roof/s'	'buttock/s'	'charcoal'	

We reject the possibility under either scenario that Comorian had a condition-free lenition rule which shifted all voiceless stops to /β(v), r, h/, and that Strengthening later on reversed the effects of lenition, changing them back to /p, t(r), k/. Strengthening in this case would require an especially unnatural, complex rule, involving the change of multiple features.

§6.3. Comorian *z-affrication. In Comorian *z and *j are both shifted to /dz/. This is a type of strengthening, thus its treatment here; but unlike Strengthening a further condition is required on the operation of the rule. It is also part of a larger process affecting *z before *įį and *yV (see §5.1.3); it also affects PSA *v that shifts to /z/ via Comorian/ND_2 Coronalization. Note the following data:

dzitso 'eye'	*įįįco	dzinyo (Ma) 'tooth'	*įįįno
dzina 'name'	*įįįna	dziho 'fireplace'	*įįįko
dzimu (Nz) 'ogre'	*įzįmu	dzu 'banana'	*įzįgu
dziwa 'milk'	*įzįWa	mwidzi 'thief'	*mwįvį

The change is only triggered if the segment is preceded and followed by *į. Thus the following are not affected: /-ǰaa/ 'be full' < *-įjala), /ǰembe/ 'hoe' < *įjembe, /ǰungu/ 'big cooking pot' < *įjungu, /ǰua/ 'sun' < *įjua, /ǰioni/ 'evening' < *įjilo, /ǰintu/ (Nz) 'giant' < *įjintu (where -ji- is [+augmentative]), /hiziwa/ 'pool' < *kizįwa (but dziwa 'milk' < *įzįwa).

There are a few exceptional items; for example, /madzi/ < *ma-zį 'excrement' attests the shift but does not have the requisite initial vowel, at either a pre-PSA or PSA level. Our guess for an explanation here is that because the item is a Class 6 plural it is relatable to a potential Class 5/6 pair, namely, *į-zį/*ma-zį 'piece of excrement/excrement'. If the pair existed, it was liable to analogical reformation on the basis of the putative singular: *į-zį/*ma-zį > *dzi/*ma-zi > *dzi/ma-dzi (by analogy) > Ø/ma-dzi (loss of the singular). On the other hand, /madzi/ could derive from an intermediate PSA *maįzį or *maazi. If the first case, then *z-affrication would apply; in the second case, the previous *VV would derive /dz/ via the same putative process that would have given rise to Comorian /kotsi/ 'slap' (PSA *įkoofį) (see §5.1.1). The interaction of *z-affrication, with other rules needed to derive forms in this paradigm, is illustrated below:

*įzįgu	*įjįco	*mwįvį	*muzįgo	
---	---	mwįzį	---	Comorian/ND$_2$ Coronalization
įzįu	---	---	muzįo	*g-loss
įzyu	---	---	muzyo	Gliding
įdzu	įdzįco	mwįdzį	mudzo	*z-affrication
idzu	idzico	mwidzi	---	*7 > 5 Vowels
dzu	dzico	---	---	Class 5 V-loss
---	dzitso	---	---	*c > ts; other rules
dzu	dzitso	mwidzi	m(u)dzo	Output
'banana'	'eye'	'thief'	'load'	

The derivation is important in supporting our claim that seven vowels lingered on in Sabaki after the PSA period. *z-affrication applies after Comorian/ND$_2$ Coronalization (§5.1.1), and after *g-deletion. Both processes postdate PSA. Thus *z-affrication is an even later process, and since it is triggered by a [+HV] vowel, it follows that Comorian had seven vowels for some time after PSA.

Finally, *z-affrication is also unlike Strengthening in that only *z and *j are affected, whereas *s and *c are not. Strengthening in Comorian is symmetrical. But z-affrication is like Strengthening insofar as [+HV] is the motivation for the shift. We should also note the possibility that *j in Comorian in these cases may have first shifted to /z/ before becoming /dz/. Also, *z-affrication has the same result as the Comorian rule, most readily seen in causative formations, which operates on fricative glide clusters, e.g., (*fy, *vy, *sy, *zy, etc.), creating affricates (see §5.3). Both cases of affrication in Comorian depend on similar conditioning elements, *į ___ *į , ___ *yV (< *įV), and ___*įį , which suggests that, although the causative rule is more general, they are related processes, having likely operated at approximately the same period in the history of Comorian.

§6.4. **Strengthening and analogy**. A synchronic morphophonology of Comorian would have to include a productive rule expressing the relationship between the various alternating forms in the Class 5 paradigm. However, there is evidence that analogy is operating as well to undo the effects of the various lenition rules. The reflex of *-zito 'heavy' in

Ngazija is /-dziro/, a nonalternating form in all classes; but in Nzuani, /-zitro/ alternates: /dz/ in Class 5 and Class 9/10, /z/ elsewhere. The nonalternating Ngazija item is the result of paradigmatic leveling in favor of the Class 5 (and Class 9/10) forms. Tucker and Bryan's (1970) data compared with our own field-collected material show substantial variability, indicating the conflicting operation of rules and analogy. Note: /dzinyo, madzinyo ~ manyo/ (TJH) 'tooth' < *ɟiɟino, /bega, maega/ (T and B) versus /bega, mabega/ (TJH) 'shoulder' < *ɟiWega, /dziwa, maziwa/ (T and B) versus /idzia, zidzia/ 7/8 (Ng) and /dzia, madzia/ (Nz) 'lake, pool' < *ɟiziWa (CB *-dị̀bà).

§6.5. Strengthening as a process. The analysis we have just proposed for Strengthening argues that (a) *ɟ is the conditioning environment; (b) Stage 1 involves a change in the features of inherited PSA segments, namely, *W (Comorian /w/) and *l, wherein nonstops become stops, with *g preserved and unaffected by *g-loss; and (c) Stage 2 preserves stops which otherwise would have undergone lenition. This analysis is partially a consequence of some earlier decisions concerning the nature of the PSA consonant inventory, namely that PSA inherited *W, *l, and *g, and *p, *t, and *k from a pre-PSA period. If East Highlands Bantu had reconstructed stops, as suggested by Guthrie's CB *b, *d, *g, Strengthening would look a lot neater: the rule essentially would say *ɟ preserves stops, voiced and voiceless (as in fact it does with *g), which otherwise undergo lenition in other environments. This is not our analysis largely because of the paucity of evidence for East Highlands Bantu having reconstructible voiced stops. This analysis is also congruent with what has apparently gone on in Nyanja, where weaker consonants become stronger (e.g., /p/ > [p^h], /s, z/ > [ts, dz]) and with *z-affrication (§6.3).

As for the underlying motivation for both Strengthening and Bantu Spirantization, apart from strictly phonetic/phonological motivation, we considered the possibility that Strengthening was a consequence of Class 5 Vowel-loss, presumably a late post-PSA shift, but earlier than Strengthening, as indicated by stratigraphy. The strengthening of the initial Class 5 segment would then have been compensatory and have served as a paradigmatic marking function when the overt class marker was lost through phonological change. We rejected this analysis because of the demonstrated need for [+HV] in cases where *z-affrication applies, and because of forms where strengthened consonants exist side by side with reconstructible *ɟ or *ų, e.g., /-iβa/ (Ung, Com) 'steal' < *-ɟiWa), /muvuβa/ (Com) 'bellows' (*muvųWa). The preponderance of phonological evidence indicates the need for [+HV], whereas an analysis based on stratigraphical evidence would point in the other direction.

§6.6. Summary: Strengthening and chronology. Forms of strengthening comparable to those in Sabaki occur only spottily, albeit widely, in eastern Africa (Ganda, Chewa/Nyanja). We have no evidence that there is any genetic connection between these scattered local attestations. In NEC, it is found outside Sabaki only in a very few lexical items in Seuta and Ruvu, and the origin of these is not clear. We are not persuaded we can reconstruct Strengthening for PSA, on the basis of what we find in Seuta/Ruvu or

Pokomo. This distributional information suggests that Strengthening in Sabaki is a relatively late innovation, restricted to Comorian and SD. Evidence of rule-ordering confirms this. Strengthening only interacts with local, post-PSA processes, such as *g-loss or Comorian/ND_2 Coronalization. Thus, both distributional and internal evidence support our contention that Strengthening is a local, post-PSA innovation shared by Comorian and SD.

§7.0. The PSA Nasals

In this section we combine discussion of nasals and nasal-consonant clusters (prenasalized stops, etc.). PSA inherited from pre-PSA Bantu a full set of nasals, with the possible exception of the velar nasal *ŋ, and a full set of voiced and voiceless prenasalized stops, e.g., *mp, *mb, etc., plus nasal + fricative sequences, e.g., *ns, *nz, etc. That PSA inherited nasals and nasal-consonant clusters is noncontroversial. Even a cursory comparison of the PSA lexis (App. 2) with CB forms will support this.

The PSA nasal stops *m, and *n, directly corresponding to CB *m and *n, are involved in little change. Where change is seen in these cases, it is for the most part either reflective of quite natural low-level phonetic change or is strictly local, and thus post-PSA change. Nasal-deletion commonly affects *m and *n in nasal environments, e.g., *n-nane 9/10 'eight' > /nane/. While the status of *m and *n is straightforward, the status of the nonanterior nasals *ny [ɲ] and *ŋ is more complicated and requires some explication.

Nasal + voiced consonant clusters are generally quite stable throughout Bantu. Postnasal consonants, especially voiced ones, do not exhibit the same range of change as consonants in other environments. They are affected only by relatively minor processes such as Meinhof's Law or the Kwanyama Law (Meinhof 1932:183-184). In general, PSA *mb, *nd, *nj, *ng, etc. (hereafter collectively *NÇ) are well attested and inherited.

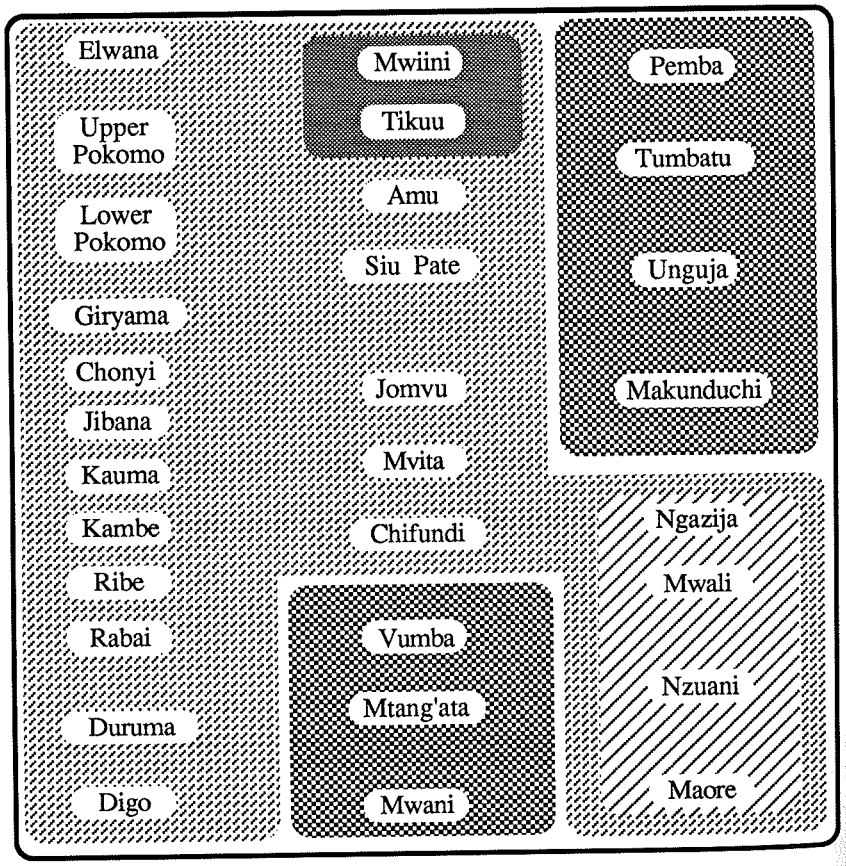
All Sabaki languages have segments, some with nasal onset, some not, corresponding to CB nasal + voiceless consonant sequences (hereafter *NÇ). Here, as in Bantu generally, much change is apparent. Thus we need to justify the reconstruction of PSA *NÇ (*mp, *nt, *nc (=[ɲc]), *nk (=[ŋk]), etc.) and to outline the changes the various Sabaki dialects attest.

§7.1. The PSA nasals *m, *n, *ny [ɲ], *ŋ. The following distribution of reflexes obtains (Chart 14):

*m	:	m	all Sabaki
*n	:	n	all Sabaki
*ny	:	n/___ w	El, Po, MK, ND, Com
	:	ny ~ n̪ ~ n	Mw
	:	n̪	Ti
	:	ny	all Sabaki (except as above)
*ŋ	:	ŋ	Po, MK, SD, ND
	:	ŋ ~ ny	Com
	:	ng (?) ~ ŋ	El

Chart 14

Sabaki Nasals



Key

*m : m ; *n : n ; *ny : ny (except as noted below); *ŋ : ŋ (except as noted below)

*ny : n / ___ w

*ny : ny / ___ w

*ny : ny ~ n ~ n̪

*ŋ : ŋ ~ ny

*-mal- 'finish'	-mala (El,Ch), -maɬa (Mw), -marigiza (Gi), -maliza (Am,Mv, Ung,Ng). (CB *-màd-)
*-lum- 'bite'	-lúma (El), -luma (UP,Gi,Ch,Ra,Du,Mn), -luma (ND obs), -uma (LP,Am,Pa,Si,Ti,Jo,Mv,Ung,Vu,Chi,Mt,Ng), -ɬuma (Mw). (CB *-dúm-)
*-neen- 'speak'	-nena (LP,Gi,Am,Si,Pa,Ti,Mv,Ung,Vu,Ng), -neena (Mw). (CB *-néén-)
*-nene 'fat'	-nene (El,LP,Gi,Mw,Am,Mv,Ung,Ng,Mn). (CB *-nénè)
*-nyw- 'drink'	-nwa (El), -nwa (UP,LP,Gi,Ch,Ra,Du,Di,Am,Si,Pa,Jo,Mv, Chi,Ng,Mh,Nz,Ma), -na (Mw), -n̪wa (Ti), -nwa (Jo), -nywa (Ung,Vu,Mt,Pe,Tu,Mak,Mn). (CB *-nyú-)
*-ŋal- 'shine'	-ngangala (El), -ŋala (Gi,Du), -ŋaa (Am,Jo,Nga,Ung), -ŋaa ~ -nyaa (Ng), -ŋanira (Mn)
*-ŋol- 'uproot'	-ŋola (Gi), -ŋoa (Am,Jo,Ung)

The few changes we see here are largely local and for the most part not shared across language boundaries. The nonvelar nasals are stable, but in some environments in some languages *ny [ɲ] has depalatalized before *w, e.g., *-nyw- 'drink' > -nwa (El, Po, MK, ND, Com). This is a near pan-Sabaki change, but is not attested in SD.

Another change, strictly local, is the shift in which *ny becomes /n̪/ regularly in Tikuu, and haphazardly in Mwiini, and in the latter we find a mix of /ny ~ n̪ ~ n/.[35] We have examples for both dialects:

*-nyw- 'drink'	-n̪wa (Ti), -na (Mw)
*-nyol- 'shave'	-n̪oa (Ti), -nooɬa (Mw)
*nyama 'meat'	n̪ama (Ti, Mw)
*nyumba 'house'	n̪umba (Ti), n̪uumba ~ nuumba (Mw)
*nyoka 'snake'	n̪oka (Ti, Mw)
*nyota 'star'	noota (Mw), n̪oča (Ti)
*nyuni 'bird'	n̪yunyi (Mw), n̪oni (Ti)

Comorian shows evidence of a rule which has shifted *ŋ to /ny/, e.g., /-nyaa/ < *-ŋal- 'shine', /-nyoa/ < *-ŋola 'uproot', /nyombe/ (Nz, Ma) < *ŋgombe 'cow'. Sacleux (Chamanga and Gueunier 1979) records some items with /ŋ/, e.g., /-ŋaa/ 'shine', /-ŋoa/ 'uproot', and others with variation, e.g., /ŋamba ~ nyamba/ 'turtle'. The data here may reflect the idiosyncrasies of Sacleux's Zanzibari Comorian consultants. In any case, the shift of *ŋ > /ny/ postdates the split of Comorian from the mainland languages.

[35]The mixed situation in Mwiini may reflect imperfect data; the incidence of dental /n̪/ in some items is identical to that in Tikuu, so it must derive from either an earlier shared phonological innovation or borrowing of individual items from Tikuu.

§7.1.1. The velar nasal *ŋ and Meinhof's Law. The velar *ŋ is problematic, but the data is sufficient to indicate that some velars were inherited (e.g., *-ŋol- 'uproot', *-ŋal- 'shine'). However, some evidence within Sabaki points to *ng as a source for other instances of /ŋ/; note the alternations of /ŋg ~ g/ and /ŋ ~ g/ both within languages and across languages:

*ŋgamba 'sea turtle' ŋamba (Am,Ung), amba ~ nyamba (Ng), ngaamba 'shrimp' (Mw)

*ŋgombe 'cow, cattle' ŋgobe (El), ŋgombe (LP), ŋombe (UP,Gi,Du,Am,Mv,Ung, Vu,Mn), ŋoombe (Mw), mbe (Ng), nyombe (Nz,Ma), kigombe ~ kagombe (Gi,Ti,Pa,Si)

*igombe 5 large cow: igombe (Ti), gombe 'old, fat cow' (Ng). (CB *-gòmbè, *-ŋòmbè < Central Sudanic)

*ŋgonzį 'sheep' ŋgozi (El,LP), ŋozi (UP), ŋondzi ~ gondzi 5/6 (Gi), ŋondzi/magondzi (Du); ŋonḍi 'sheep sp' (Am), ŋonzi (Ung,Vu), gondzi 5/6 (Ng,Nz,Ma). (PSC *-gwand-)

The alternating forms, e.g., /ngaamba/ (Mw) vs. /ŋamba/ (Am, Ung), /ŋombe/ (Pa) vs. /kigombe, kagombe/ (Pa), etc., plus the the Lower Pokomo and Elwana attestations of /ŋg/, call for PSA *ng.[36] They also suggest the operation of Meinhof's Law, i.e., Ç > Ø/N ___ VNÇ), thus *ŋgombe > /ŋombe/. However, the evidence is mixed, e.g., /ŋgambi/ 'agreement' (Ung) (CB *-gàmb- 'speak') is left untouched. While there is no synchronically active form of Meinhof's Rule, the evidence indicates that it may once have been active in Sabaki. Elsewhere in our reconstructed lexis (App. 2), we find for the most part forms which have been inherited from a stage when it was active, and left only reduced, nonalternating forms that are reconstructed without a postnasal consonant:

*mamba 'scale' mamba 'crocodile' (Du,Du,Ung), mamba 'mudfish' (LP), mamba 'crocodile scales' (Gi), maamba 'a large shark' (Mw), mamba 'scale' (Am,Mn), baamba/maamba 5/6 'scale' (Ng), mamba 'fish scales'

*mimba 'fetus' miiba (El), mimba (LP,Gi,Am,Mv,Ung,Nz,Mn), mimba ~ mba (Ng). (< ? CB *-bį́mb- 'swell'; cf. PSA *-vį́mb-)

*nundu 'hump' nundu (LP,Gi,Am,Mv,Ung). (CB *-dúndù)

*nungu 'porcupine' nungu (LP,Gi,Am,Mv,Ung), noogo (El), nungu 'spiny blowfish' (Nz). (CB *-dùngú)

---

[36]We are reluctant to accept the Elwana data as evidence for *ng since the items cited here are not regular. Elwana normally denasalizes voiced prenasalized consonants, as it has done here in C_2 position, and as it usually does in C_1 position, e.g., /gooba/ < ngomba < *nkomba 'bushbaby'. These items are a puzzle, in that denasalization has not affected C_1 position. While they may be borrowings, why it is that denasalization has applied to C_2 position but not C_1 position?

This set and other alternations suggest a pre-PSA period, when Meinhof's Law was active (see e.g., PSA *ndongo 'dirt' ~ *Wulongo 'clay' > /ulongo ~ nongo/ (Gi), /udongo ~ nongo/ (Ung); *mbango 'wart hog' : mbango (LP, Am), but mbango ~ mango (Ung) and mango (Mn) from CB *-bàngà 'tusk'). Whether PSA simply inherited the mixed situation we see today or had a more active form of the process is probably not determinable on the basis of current evidence and without a fuller analysis of data in NEC and adjacent groups. The facts outlined here are not helpful for subgrouping.

§8.0. Proto-Sabaki Nasal + Consonant Sequences

§8.1. The PSA voiced N + Ç sequences: *mb, *nd, *nj [ɲj], *ŋg. Examples of correspondences follow (Chart 15); *nj is treated in §8.1.4:

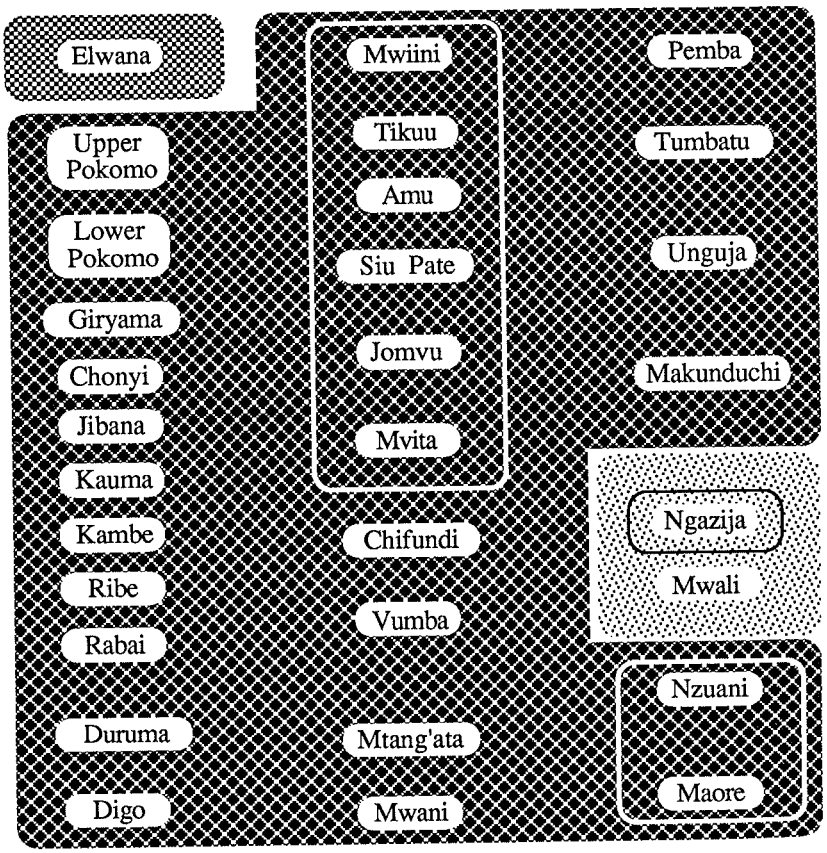
*mb :	b	El
	mb	Po, MK, ND, SD, Nz, Ma, Mn
	mb [mɓ]	Ng, Mh
	b/NÇV___	Ng (Kwanyama Law)
*nd :	d	El[37]
	nd	Po, MK, SD, Mn
	nd [nɗ]	Ng, Mh
	ndr	ND_2, Nz, Ma
	d/NÇV ___	Ng (Kwanyama Law)
*ng :	g	El
	ng	Po, MK, ND, SD, Com, Mn
	g/NÇV ___	Ng?

*mbegu 'seed'	beu (El), mbeǰu (LP), mbeu (UP,Gi,Ra,Du,Di,Ti,Pa,Si, Am,Mv,Chi,Vu,Tu,Nz,Ma,Mn), mɓeu (Ng); mɓia (Mh), mbegu (Ung,Mak), mbegu ~ mbeu (Pe), mbeyu (Du). (CB *-bégú)
*mbwa 'dog'	bóa (El), mbwa (LP,Am,Si,Pa,Mv,Chi,Ung,Mak,Ng,Mh, Nz,Ma), mbaa ~ mbwa (Mw), imbwa ~ mbwa (Ti), imbwa ~ umbwa (Mn). (CB *-búa)
*ndema 'bat sp.'	ndema (Gi,Ung), ndema (Ng). (CB *-déma)
*ndani 'inside'	ndani (LP,Gi,Du,Mv,Ung), ndani (Ng), ndrani (Ti,Am,Nz, Ma). (CB *-dà)
*-tundu 'hole, etc.'	t^hundu (LP,Gi,Am,Mv,Ung), t^hundru (Am), trundru/ marundru (Ma), kitundu 'pen' (Mn). (CB *-túndú)

[37]We follow Maddieson and Sands (field work 1991) in not marking dentalization here.

Chart 15

Sabaki *mb, *nd, *ng



Key

*mb, *nd, *ng : b, d, g

*mb, *nd, *ng : mb, nd, ng

*mb, *nd, *ng : mɓ, nɗ, ng

Kwanyama Law

*nd : ndr

*nguWo 'clothing'	guo (El), nguuo (Mw), nguvo (Ti), nguo (LP,Gi,Du,Am, Si,Mv,Ung,Ng,Nz,Ma). (CB *ngúbò)
*Wucungu 'gall'	učoogu (El), utsungu (Gi), uṯungu (Am), učungu (Ung), tsungu (Ng). (CB *-cúngú)
*-fung- 'tie'	-fuga (El), -fuunga (Mw); -funga (LP,Gi,Du,Am,Mv,Ung, Vu,Ng,Nz,Ma,Mn). (CB *-tíų̣ng-)

There are several changes apparent in this data: (a) Elwana has denasalized its prenasalized stops; (b) some Comorian dialects and ND have a "rhotacized" articulation of *nd with an r-like offglide; and (c) at least one dialect of Comorian attests the Kwanyama Law (not exemplified above, see §8.1.3). Changes affecting *nj are treated in §8.1.4.

§8.1.1. **Elwana denasalization**. Elwana denasalizes the voiced prenasalized stops. In word-initial position this involves complete deletion, but in medial position Möhlig (1986a) records nasalization as consonant nasalization, e.g., /mooo̰u/ 'person', /siɓa/ 'lion', etc. In data recorded by Maddieson and Sands (fieldwork 1991) there is no nasalization; Maddieson (pers. comm.) reports that the nasal has been fully deleted. We do not follow Möhlig in marking vestigial nasalization. Our data also indicate that where deletion has occurred the vowel is frequently long:

beu 'seed'	*mbegu, CB *-bégú	-aaba 'say'	*-gamba, CB *-gàmb-
-ɛ́ɛ̀da 'go'	*-enda, CB *-gènd-	jovu 'elephant'	*njovu, CB *-jògṳ̀
-vuuja 'break'	*-vų̣nja, CB *-bṵ́ṇj-	gila ~ gela 'path'	*ngila, CB *-jìdá
ge 'scorpion'	*nge	gwena 'crocodile'	*ngwena, CB *-gùènà

Elwana also has voiced its N + Ç sequences, and these are also affected by the denasalization rule:

bula 'nose'	*mpula, CB *mpùdà	jopa 'bottle'	*ncupa, CB *-cúpà
moodu 'person'	*muntu, CB *muntù	gooba 'bushbaby'	*nkomba
-nyiga 'give'	*-ninka, CB *-nḭ́nk-		

Thus for Elwana we need the following rules:

*NÇ > NÇ (see §8.2.1 for a similar process in Comorian)
*NÇ > ØÇ

Elwana voicing must have preceded denasalization. Some nearby Central Kenyan languages have voiced the voiceless prenasalized stops and are undergoing a similar denasalization process. Although these communities are not in direct contact, it is possible that these changes in Elwana are part of a general shared areal phenomenon. Another possible explanation is suggested in Nurse (1983a) where this is attributed to a historical Cushitic (Dahalo) substratum.

§8.1.2. **Comorian/ND$_2$ retroflexion/rhotacization**. In some of the Comorian dialects (Nzuani, Maore) and in ND (Mwiini, Tikuu, Siu, Pate, Amu, Mvita), *nd is realized with an r-like offglide [nd^{r}] which we write "ndr":

*ndala 'frond, fiber'	ndrala (Nz, Ma), ndraa (Mv), ndrao (Am)
*ndi 'be (emphatic)'	ndri- (Am, Ti), ndri (Nz)
*luWanda 'open space'	bandra (Nz, Ma)
*Wuvundo 'stink'	uvundo (Ng), uvundru (Nz), uvundro (Am,Ti,Ma)
*-en(en)da 'go'	-eendra, °-ineendra (°i) 'walk' (Mw); -endra ~ -enendra (Am, Ti, Si, Pa); -endra (Nz,Ma)

In the northern-most ND this is a regular process in all dialects, including Mwiini, and has been attested in Mvita (TJH field notes 1988; Wald, pers. comm.). In Comorian it is restricted to etymological *nd in Nzuani and Maore (Sibertin-blanc 1980; Blanchy 1987). In Ngazija [ndr] is attested only as a reflex of *nt, e.g., /ndrovi/ (Ng) versus /ntrovi/ (Nz) 'banana', /mndru/ [mdru] (Ng) versus /mntru/ 'person'. In Comorian this retroflexed pronunciation of *nd contrasts with /(n)ḍ/ found in borrowed lexis, e.g., nḍovu (Nz, Ng) 'elephant', and in ND with /nḍ/ derived from *nj/*nz.

Rhotacization in these dialects is perceptually motivated. When /nḍ/ developed out of */nj, nz/ in ND, or entered the phonology through loaning (Comorian), rhotacization affected the inherited */nd/, giving /nd/ [ndr]. The effect is to keep the reflexes maximally distinct. The rhotacization of /nd/ [ndr] from *nt in Ngazija is similarly motivated. When *nt became voiced, rhotacization kept distinct its reflexes and those of older *nd, which are typically imploded (e.g., /ndevu/ [ⁿɗevu]. Similarly, in Ngazija *t became /tr/ (or /r/) in contrast to /ṯ/, which typically occurs in loans.

At first sight these events in ND and Comorian may not appear to be connected. However, between 25 and 30 of all East African Bantu languages have merged */nt/ and */nd/, and only some dialects of one, namely, Chaga, (see Nurse 1979a:393-394) have taken similar steps to keep the two in any way distinct. Against this statistical background, the development of rhotacization is exceptional, and its appearance hardly seems a coincidence in two sets of dialects which also share other innovations. This shared development may be sociolinguistic in nature; that is, once this strategy had been innovated by one of the dialects involved, it was then used by the others, though in different ways.

§8.1.3. The Kwanyama Law in Comorian. Part of Comorian attests a phenomenon whereby voiced prenasalized stops in C_2 are denasalized when C_1 is a voiceless prenasalized stop (*NÇVNÇV > NÇVØÇV). This is often referred to as the Kwanyama Law (Herbert 1977, M. R. Johnson 1979, Meinhof 1932:184, Tucker and Bryan 1970). Our data show this for Ngazija only. A few examples follow:

*nkondo 'war'	nkodo [nkoɗo] (Ng), but nkondro (Nz), kondro (Ma)
*lukunde 'cowpea'	nkude [nkuɗe], but nkundre (Nz) (< plural form *nkunde)
*mpembe 'corner'	mbebe [mbeɓe] (Ng)
*ịtunda 'fruit'	ntruda [ntruɗa] (Ng) 'seed' (?) (note lack of voicing in C_1)

This dissimilation process is similar to Meinhof's Law, except for the syllables affected and differences in the conditioning environments; in Meinhof's Law the C_1 voiced prenasalized stop totally assimilates the surrounding nasality (*NCṼN(C)V > NC̃Ṽ N(C)V > NNṼN(C)V > NØVN(C)V, e.g., *ŋgombe > ŋombe). The motivation for both processes is quite natural: simplification of the articulatory motions involving velic closure. However, Meinhof's Law is more widely distributed and more commonly attested in Bantu than Kwanyama. In eastern Bantu we find this attested in only one Comorian dialect, and elsewhere only in southwestern Bantu. Because of the idiosyncratic nature of this process, it is tempting to perhaps argue for some relationship between southwestern Bantu and Comorian. However, this possibility is easily discounted. It is, first of all, ruled out by geography; there is too much distance and too many languages lying in between to argue for an areal influence. Secondly, attestation in only Ngazija (and Mwali?) suggests recent independent innovation following the differentiation of Comorian dialects, thus ruling out any genetic relationship.

§8.1.4. Sabaki *nj. The correspondences for *nj are (Chart 16):

*nj :	ǰ	El
	nj [nǰ]	UP, SD, Mn
	nd̪	ND_3, Mw
	nj ~ nd̪	Chi
	ndz	LP, MK, Com

*njovu 'elephant' ǰovu (El), ndzovu (LP,Gi,Du), njovu (Vu), nd̪ovu (Am,Mv)

*njala 'hunger' ǰala (El), ndzaa (LP), ndzala (Gi,Du), nd̪aɬa (Mw), nd̪aa (Am,Mv), njaa (Ung), njaya (Tu,Pe), ndza(y)a (Ng), ndza (Nz,Ma), njala (Mn)

*njiWa 'dove sp.' njiβa (UP), nd̪iva (Am,Mv), njiwa (Ung,Mn), nd̪iwa ~ njiwa (Chi), °ngiwa (Vu), ndziwa (Ng), ndziaa (Nz)

*-vn̪j- 'break' -vuuǰa (El), -vunja (UP,Ung,Mn), -vundza (Gi,Du,Ng,Nz, Ma), -vund̪a (Am,Mv), -vuund̪a (Mw)

*(i)nje 'outside' ndze (LP,Gi), nd̪e (Mw,Am,Si,Mv), ind̪e (Ti), nje (Ung,Vu), panja (Mn), βondze (Ng), mwendze (Nz)

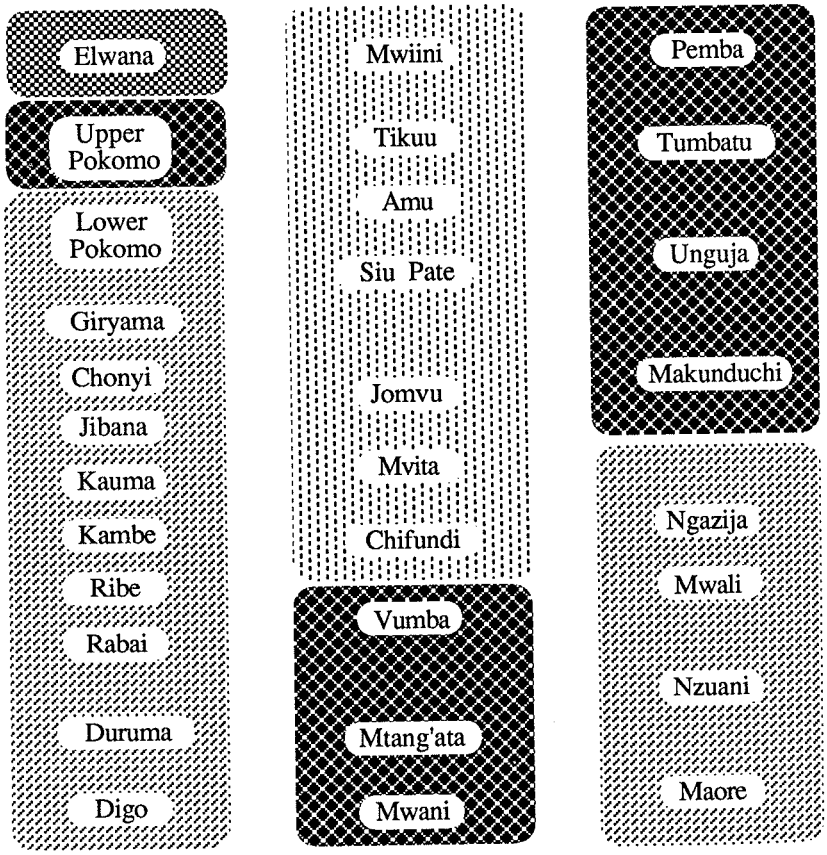
Elwana, Upper Pokomo, Mwani, and SD are conservative in retaining a palatal reflex. Mijikenda, Lower Pokomo, and Comorian attest /ndz/, and all ND, including Mwiini, have /nd̪/. We have two hypotheses, one which holds that /nd̪/ derives from an intermediate *ndz:

Hypothesis 1

Stage 1:	*nj	>	ndz	MK, LP, Com, ND_3, Mw, Mv, (Chi)
Stage 2:	*ndz	>	nd̪	ND_3, Mw, Mv, (Chi)

Chart 16

Sabaki *nj



Key

*nj : ǰ

*nj : nǰ

*nj : ndz

*nj : nd̠

That an intermediate *ndz is the source of /nd̪/ is reasonable, given that in Mijikenda and Lower Pokomo /ndz/ has a dental articulation [nd̪z]; a simple shift in the manner of articulation would yield /nd̪/. In the alternative hypothesis, the ND phenomenon is independent of what happens in Mijikenda, Lower Pokomo, and Comorian, and involves *nj shifting directly to /nd̪/:

Hypothesis 2

Stage 1: *nj > ndz MK, LP, Com
Stage 2: *nj > nd̪ ND_2, Mw, Mv, (Chi)

Nurse (1985b) suggests that this parallels what happens to *c (§4.1.4.1) and *nc (§8.2.3).

While Comorian, Mijikenda, and Lower Pokomo apparently share the same change (*nj > ndz) (Hypothesis 1), there are indications that perhaps the source of /ndz/ in Comorian is not immediately *nj, but rather an intermediate *nz. In other cases, the fricative *z is the source of [dz]: (a) after a nasal *z > dz (/ndzi/ (Ng, Ma) 'fly' < *nzi̱ (CB *ngi̱); (b) between super-close front vowels *z > dz (see §6.3) (/dzima/ 5 (Ng) 'whole' < *i̱-zi̱ma (CB *-gi̱mà); and (c) the causative morpheme is realized as /dz/ in Maore and Nzuani (see §5.3) (/-žadza/ (Ma, Nz) ~ -ǰaza (Ng) 'fill' (caus.). However, we have no direct evidence that in Comorian *nj first went to *nz before becoming /ndz/.

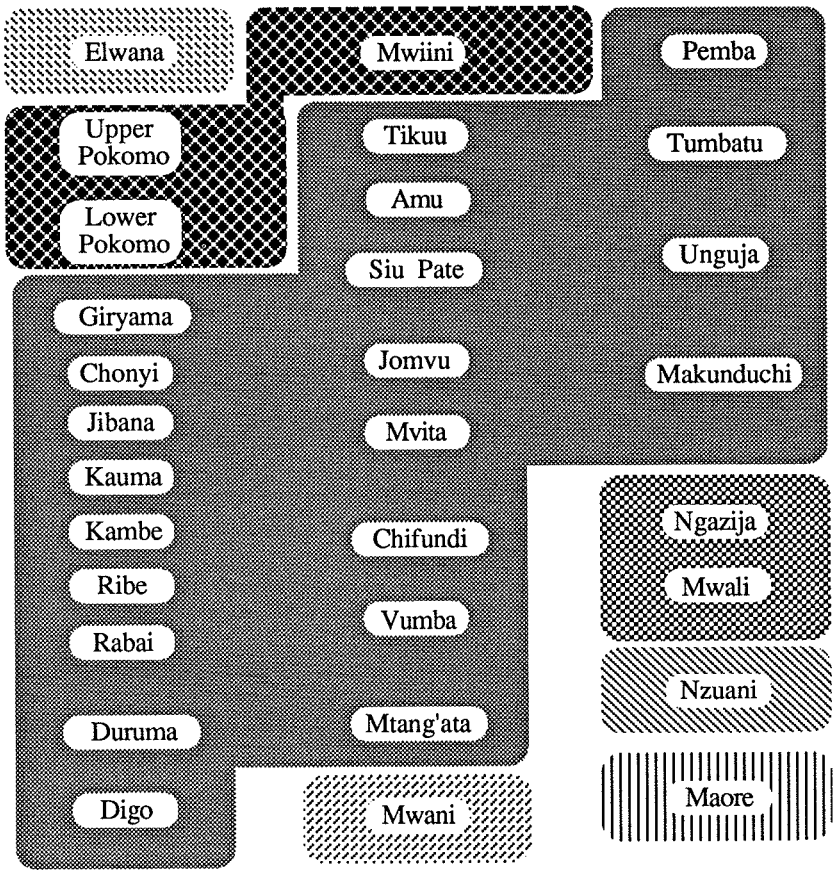
Although the particular steps of this seriation in Hypothesis 1 are in doubt, all of these languages behave similarly in that the shift or shifts involve depalatalization. For subgrouping purposes, Mijikenda, Lower Pokomo, and Comorian clearly fall together. Whether the ND phenomenon is interpreted as a separate act or not, there is likely to be a historical connection between what happened in ND and the others, given their distribution and the similarity of the depalatalization. The shift of *nj to /ndz/ and /nd̪/ is essentially one entailing the depalatization of *nj, and subgrouping implications are the same whether or not ND actually shared with Lower Pokomo, Mijikenda, and Comorian the shift to /ndz/. For subgrouping purposes, what is important is the shift *nj > [-palatal], and not the specific end-results in each group. This fact, and the distribution of the languages involved, strongly supports a historical connection between these events in ND, on the one hand, and Lower Pokomo, Mijikenda, and Comorian on the other. Before going on to the next section, one final note is in order. In the synchronic phonological systems of the Sabaki languages, *N + Ç sequences are phonetically prenasalized stops [^{N}Ç] unless the nasal is stressed, e.g., *mbwa > ḿbwa versus *mbuzi̱ > [^{m}búzi] (Ung).

§8.2. The PSA sequences of N + Ç: *mp, *nt, *nk. Correspondences for these follow (Chart 17):

*mp :	b	El
	m	Mn
	p^h	MK, SD, ND
	mp [m̥p^h]	UP, LP, Mw

Chart 17

Sabaki *mp, *nt, *nk



Key

*mp, *nt, *nk : $m\underset{\circ}{p}^h$, $n\underset{\circ}{t}^h$, $\eta\underset{\circ}{k}^h$

*mp, *nt, *nk : p^h, t^h, k^h

*mp, *nt, *nk : mb, ndr, nk

*mp, *nt, *nk : mp, ntr, nk

*mp, *nt, *nk : p, tr, k

*mp, *nt, *nk : b, d, g

*mp, *nt, *nk : m, n, Ø ~ k

	mp	Nz
	p	Ma
	mb	Ng, Mh (in contrast to /mɓ/ < *mb)
*nt :	d	El
	n	Mn
	t^h	MK, SD, ND
	nt [$\underset{\circ}{n}t^h$]	LP, UP, Mw
	ntr	Nz
	tr	Ma
	ndr	Ng, Mh
*nk :	g	El
	Ø	Mn
	ŋk [$\underset{\circ}{ŋ}k^h$]	LP, UP, Mw
	k^h	MK, SD, ND
	nk	Ng, Nz, Mh
	k	Ma

*mpula 'nose' bula ~ bóla (El), mpula (UP), mpuya (LP), p^hula (Gi,Ch,Ra, Du); p^hura (Di), mpuɬa (Mw), p^hua (Ti,Si,Pa,Am,Jo,Mv,Jo, Vu,Chi,Mt,Pe,Ung,Mak), p^huya (Tu), mbua (Ng,Mh), mpua (Nz), pua (Ma), mula (Mn)

*muntu 'person' mudu ~ moodu (El), muntu (LP,UP), mut^{h}u (Gi,Ch,Du,Di), mt^{h}u (Am,Jo,Mv,Ung,Vu,Mak,Tu,Mt,Vu), muuntu (Mw), nc^{h}u (Pa,Si,Ti), nt^{h}u (Jo,Pe), mndru [mdru] (Ng), muntru (Nz), m(u)tru (Ma), munu (Mn)

*nkala 'crab' k^{h}ala (Du), nkaɬa (Mw), k^{h}aa (Am,Mv,Ung), kala (Ma), nkaa (Nz). (CB *-kádá)

*nkuku 'chicken' nkuku (LP,Mw), k^{h}uku (Gi,Du,Am,Mv,Jo,Ung,Vu), nkuhu (Ng,Nz), kuhu (Ma), uku (Mn)

It is the attestation of a nasal in Pokomo, Mwiini, and Comorian that allows the assumption of PSA *N + $\underset{\circ}{C}$ sequences, plus the fact that Sabaki's closest genetic affines, Seuta and Ruvu, also attest N + $\underset{\circ}{C}$. There is also textual evidence; older ND_3 texts write the nasal: *wantu* 'people', *yonte* 'all (Cl. 4)', *ntano* 'five' (see Knappert 1968, 1979).

Throughout Bantu there is a range of reflexes attested for *N$\underset{\circ}{C}$ clusters (see Kerremans 1980, Nurse 1987 for surveys). As in Bantu, there is considerable variety in Sabaki. We can distinguish several major processes responsible for this variety: voicing of voiceless prenasalized stops, assimilatory devoicing and aspiration, loss of the voiceless consonant, and loss of the nasal consonant.

§8.2.1. Postnasal voicing in Elwana and Comorian. Voicing is found in just two Sabaki dialects, Elwana and Comorian (Ngazija/Mwali). Elwana voicing is general, and the domain of the rule is *mp, *nt, *nc, and *nk, but in Ngazija/Mwali only *mp and *nt are affected:

*muntu 'person'	mudu ~ moodu (El), mndru [mdru] (Ng,Mh)
*mpula 'nose'	bola ~ bula (El); mbua (Ng,Mh)
*ncupa 'bottle'	jopa (El); ntsupa (Ng)
*nkomba 'bushbaby'	gooba (El)
*nkuku 'chicken'	nkuhu (Ng)
*-nunka 'smell'	-nuuga (El), -nunka (Ng,Nz)

Ngazija /mb/ and /nd/ are voiced prenasalized plosives and contrast with /mɓ/ <*mb, and /nɗ/ < *nd. We generally have not marked implosives in this study.

The Elwana shift may be related to that found in the Central Kenya languages, in which *NC̥ is voiced. However, in Comorian this must be a recent innovation, given its attestation in just two dialects and the fact that it does not apply across the board. There is little reason to believe that Elwana and Ngazija share this shift for genetic reasons at any level (see Möhlig's hypothesis that postnasal voicing is an early "Stratum I" change (1977 - 1986a). Given the overall distribution and variety of change affecting *NC̥ and *NÇ, and the fact that such changes frequently cut across genetic boundaries (e.g., postnasal voicing shared among Chaga, Taita, Gweno, and the northern Pare dialects), it makes little sense to claim that noncontiguous examples of the same shift are due to common genetic origin, especially in the absence of other evidence that points to genetic relatedness, and further, in view of the natural assimilatory mechanisms underlying the processes affecting *NC̥ (Hinnebusch and Huffman 1987). We conclude that postnasal voicing assimilation in Elwana and Comorian are independent, late innovations. Likewise, for Comorian there is the remote possibility that voicing could be due to influences from a Malagasy substratum, or from Zone P languages where postnasal voicing is the norm, since historically the Comoros absorbed fair numbers of such speakers (Simon 1988). The rules that apply here are:

*NC̥ (C̥ = [-velar]) > NÇ	Ng, Mh
*NC̥ > NÇ > Ç	El

§8.2.2. Nasal-devoicing and aspiration. The other Sabaki reflexes not covered by postnasal voicing result from a process of nasal-devoicing and aspiration: *NC̥ > NC̥^{h}. This formula implies that aspiration is concomitant to devoicing, either as an automatic consequence or as a metathesis of pre oral-stop voicelessness to post oral-stop voicelessness. This somewhat oversimplifies the phonetic processes involved; for a detailed discussion see Hinnebusch (1975), or Hinnebusch and Huffman (1987).

Nasal-devoicing is attested in Mwiini (Kisseberth, pers. comm. 1975; TJH, field notes 1988) and in Pokomo (TJH, field notes for LP and UP; also Nurse field notes, Ehret field notes; Ipu 1982; and Meinhof 1932:33-34). In Pokomo, in word-initial position *N

is devoiced after pause; in postvocalic position it is only variably devoiced. Nasal-devoicing has been reported for a number of NEC groups and languages: Seuta (Nurse 1979a), Shambala (Tucker and Bryan 1957), Bondei (Ladefoged and Maddieson 1986:47-49, TJH field notes 1985), and Zaramo (Meinhof 1932:33). In Ruvu, nasal-devoicing is reported for Doe and variably in Gogo (Nurse 1979a), but for the most part Ruvu dialects attest /N^{h}/, where the voiceless consonant has been deleted.

We also also argue that aspiration in Swahili and in Mijikenda historically derives from an intermediate stage involving nasal-devoicing, and nasal-loss:

*NC̥	>	NC̥^{h}	Po, MK, ND, SD, Seuta
*NC̥^{h}	>	ØC̥^{h}	MK, SD, ND (not Mw)

This is supported both by seriation (loss of a nasal is more likely if it is voiceless) and by stratigraphy: languages to the north and south of both dialect continua attest nasal-devoicing, while at least one Swahili dialect, Mwiini, still attests the voiceless nasal. However, the distribution of nasal-devoicing, since it cuts across group boundaries, also suggests that it is a late areal shift. This is consistent with what we see with most change affecting *NC sequences. There is little support for Möhlig's hypothesis (1977) that this is an early Bantu shift (see Hinnebusch and Huffman 1987).

In Comorian, *NC̥ clusters behave differently in a number of ways from mainland Sabaki languages. None of the dialects shows convincing evidence of having undergone either nasal-devoicing or aspiration. Where Comorian preserves nasals on the surface they show up regularly in postvocalic environments, thus, for example, in Ngazija, /nkodo/ [kodo] 'war' ~ /(y)enkodo/ [(y)eŋkodo] 'the war' < *nkondo. Even in cases where the nasal is never reported, as in Maore where *NC̥ becomes /C̥/, aspiration has not been recorded. These facts, plus postnasal voicing in Ngazija/Mwali (*mp > /mb/, *nt > /ndr/) and the conservative Nzuani case which preserves N + C̥ sequences—phonologically /NC̥/ in all contexts, with no evidence of nasal-devoicing—support the contention that Comorian has not shared with the rest of Sabaki an intermediate shift of *NC > /NC̥^{h}/ and that, moreover, changes involving N + C̥ sequences are late post-PSA changes.

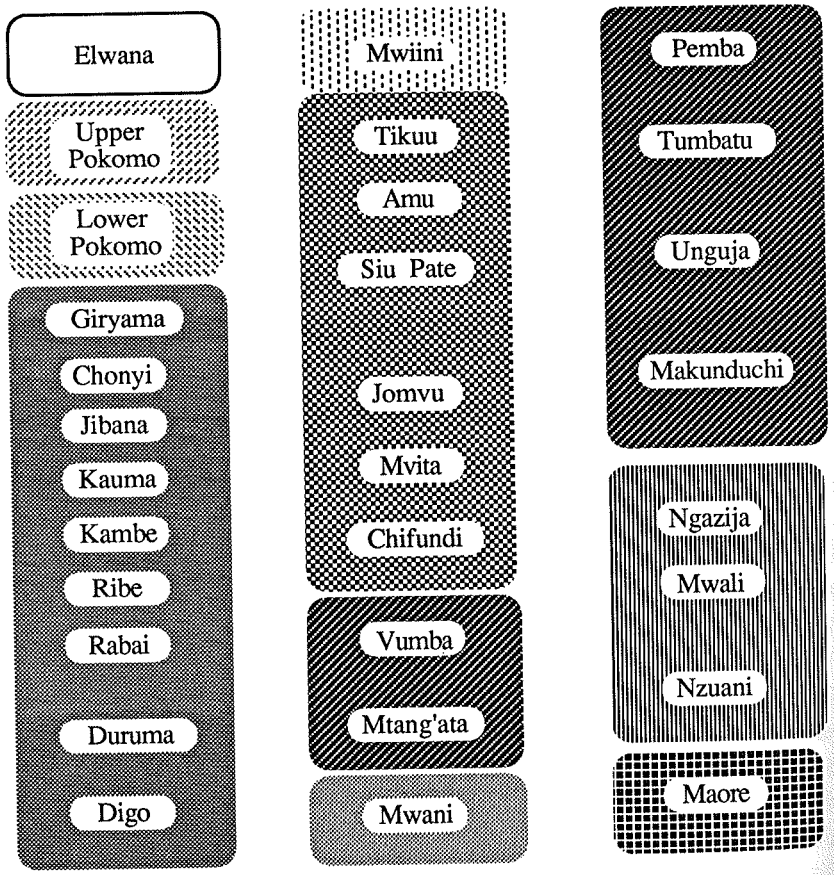
§8.2.3. PSA *nc. The behavior of Sabaki *nc parallels that of *nj in terms of affrication and dentalization (Chart 18); reflexes for *nc with polysyllabic stems are:

*nc :	ǰ	El	nt̪ [nt̪^{h}]	Mw
	nč [nč^{h}]	UP	č^{h}	SD
	nts [nt̪s^{h}]	LP	nts	Ng, Nz
	ts [t̪s^{h}]	MK	ts	Ma
	t̪^{h}	ND, Chi	s	Mn

*(i)nci 'earth, etc.' °ńčɪ (°č ¬ ǰ) (El), ntsi (LP,Ng,Nz,Ma), ts^{h}i (Gi,Du), it̪^{h}i (Ti), nt̪^{h}i (Mw,Am,Si,Mv), nč^{h}i (Ung), č^{h}i (Mak,Tu,Mt), t̪^{h}i ~ nt̪^{h}i (Chi). (CB *-cí)

Chart 18

Sabaki *nc



Key

- *nc : j
- *nc : nt [nt^{h}]
- *nc : nč [nčʰ]
- *nc : čʰ
- *nc : nts [ntsʰ]
- *nc : nts
- *nc : ts [tsʰ]
- *nc : ts
- *nc : tʰ
- *nc : s

*ncupa 'bottle' ǰopa (El); °nt̪upa (°t̪ ¬ č; °p) (LP); °t^{h}upa (Du); °t̪^{h}upa (Ø ¬ n) (Mw); t̪^{h}upa (Am,Mv); č^{h}upa (Ung); °t̪^{h}upa ~ č^{h}upa (Vu); ntsuβa (Ng); ntsuva 'calabash' (Nz). (CB *-cúpà)

*ncungu 'ant sp.' ntsungu (LP), ts^{h}ungu (Gi), t̪^{h}ungu (Am), č^{h}ungu (Ung), sungu (Mn)

We will assume the same seriations and analyses here that we proposed for *c (see §4.1.4.1) and *nj (see §8.1.4):

Hypothesis 1

Stage 1:	*nc	>	nts [nt̪s]	LP, MK, Com, ND$_3$, Mw, Mv, Chi
Stage 2:	*nts	>	nt̪	ND$_3$, Mw, Mv, Chi

Hypothesis 2

Stage 1:	*nc	>	nts [nt̪s]	LP, MK, Com
Stage 2:	*nc	>	nt̪	ND$_3$, Mw, Mv, Chi

The same processes that applied to *c and *nj are also responsible for the changes here, and the changes are parallel. Moreover, the same languages undergo the same changes. Also, the same languages are conservative (El, UP, SD). This can be no accident and must be related historically. After *nc becomes /ts/, Nasal-devoicing/aspiration and nasal-loss apply accordingly, e.g., *ncungu > ntsungu > n̥ts^{h}ungu > /ts^{h}ungu/.

§8.2.4. **Aspiration-migration.** For some Swahili dialects, it has been noted that aspiration can be leftward shifted (Nurse 1982b:80, Lambert 1957:22, 1958a:17, Lodhi and Engstrand 1984); thus:

*kintu 'thing'	>	kit^{h}u	>	k^{h}itu (Am,Mv,Vu), č^{h}itu (Chi)
*-ikalanči 'sit'	>	-ket̪^{h}i	>	-k^{h}et̪i (Am)
*-kanta 'cut'	>	-kat^{h}a	>	-k^{h}ata (Am,Mv,Vu,Chi)

Lambert notes that this shift to the left is to the stressed syllable, and this would appear to be the case, since we do not find examples such as *k^{h}uket̪i 'to sit on the ground'. We are not able to report whether this shift takes place in Mijikenda languages.[38]

§8.2.5. **Mwani *NÇ > N ~ Ø.** The reflexes of *NÇ in Mwani reflect its isolation from the rest of Sabaki, and the contact it has with non-Sabaki languages. The [-velar] voiceless stops are deleted, thus *mpula 'nose' > /mula/, *mpwani 'coast' > /mwani/, *muntu 'per-

[38]Aspiration in Sabaki languages is undergoing readjustment, and is not always etymologically related to *NÇ; Polomé (1967:41) reports for Swahili that "there appears to be a general tendency to connect aspiration with stress and initial position, e.g., *mwítu* 'forest', with [t]; *mwitúni* 'in the forest', with [t^h] in the stressed syllable; *una takataka* 'you are dirty', with [t^h] in word-initial position and in the stressed penultimate; *kahawa*, with [k^h] in initial position, etc." Whiteley (1958) in a description of Pemba, reports /k^{h}eča/ 5/6 'small bitter orange', /t^{h}uka 5/6 'main crossbeam of a roof', /t^{h}umbo/ 5/6 'a youngster', but /tumbi/ 5/6 'bag made of coconut palm fronds'.

son' > /munu/. The velar sequence *nk is completely deleted, e.g., *nkuku 'chicken' > /uku/, *-enkundu 'red' > /-eundu/, *nkata 'headpad' > /ata/.

These changes are unrelated to the usual Sabaki changes, but are akin to what occurs in Ruvu, e.g., *NC̥ > /N^{h}/, e.g., Kagulu: *muntu > /mun^{h}u/, *kuku > /ŋ^{h}uku/; no doubt Mwani was influenced by Ruvu languages at a time when it still had /NC̥/, but not /C̥^{h}/.

In Mwani, *nc becomes /s/ in word-initial position, and /ns/ in medial position:

*pa+nci 'ground'	pansi	*nci+ni 'down'	sini
*-once 'all'	-onse	*nsiku 'day'	siku ~ suku

Again, this shift to /s/ is unrelated to events in Sabaki and would appear to be an areal adjustment to nearby languages, either Ruvu or Zone P languages.

§8.3. The PSA nasals *mf/ns, *mv/nz. Before listing correspondences and examples, note should be taken that orthographic practices of various sources differ in how the reflexes of these have been recorded. Deed (1964), for Giryama, consistently records "ndz" for *nz and *nj, while all of Hinnebusch's Mijikenda consultants used "nz" in writing the reflexes of both *nj and *nz (TJH field notes 1972-73). Here perhaps Swahili orthography, or possibly of the Giryama bible, is a variable influencing Mijikenda spelling. Ottenheimer (1986) for Nzuani records "nz" for both *nz and *nj, and "nts" for *ns, while others (Sibertin-blanc 1980, Rombi 1983, Lafon 1987, etc.) use "ndz" and "nts" for all Comorian dialects, including Nzuani. The reasons for these conflicting orthographic practices are no doubt complex, but several factors are at play. One is that in Sabaki the fricatives /s, z/ when preceded by a nasal are usually realized phonetically as [nts] and [ndz], where the fricative is affricated to some extent, e.g., /nso/ [ntso] 'kidney' (Am). Another is that many of these dialects have plain voiceless and voiced affricates, /ts/ and /dz/. This establishes the precedent for phonemicizing any occurrence of [ts] and [dz], preceded by a nasal or not, as /ts/ and /dz/, thus /ns/ > /nts/ and /nz/ > /ndz/.

Our practice in this work is to phonemicize the reflexes of *nz and *ns, where they have not been affected by Nasal-deletion, as /ndz/ and /nts/ for all of Sabaki except the Swahili dialects and Mwani. This assumes that /ns/ [nts] and /nz/ [ndz] in the former have been reanalyzed as /nts/ and /ndz/.[39] Some examples are:

(a) *mf :	mf ~ nts	Ng, Nz
	f (~ s)	El, UP, MK, ND, SD, Nz, Ma, Mn
	mf [m̥f]	LP, Mw[40]

[39]The phonemic status of *ns and *nz requires further study for each Sabaki dialect. We also note that this decision is contradicted by Ipu's analysis (1982) of LP. He phonemicizes postnasal fricatives as /ns/ and /nz/ where they alternate with /s/ and /z/, thus /nsoho/ [n̥ts^{h}oho] 'breath' ~ /-soha/ 'breathe', /-ziho/ 'heavy' ~ /nziho/ [ndziho] 9/10); it is not certain how he analyzes non-alternating reflexes of *ns and *nz. We have not followed Ipu in this case in order to simplify the diachronic treatment; however, this has minimal implications for diachrony.

[40]Our major source for Mwiini data, Kisseberth, writes reflexes of *mf and *mv with "nf" and "mv ~ nv"; we assume in each case that the nasal is labiodental and we write them throughout as "mf" and "mv".

*mfimbo 'stick' fimbo (Ch,Ung,Mv), simbo (Am), usimbo/simbo (Ti), °simbɔ ~ °sembɔ (El), °ntsimbo ~ nsimbo (LP), ntsimbo (Nz)

*mfigo 'kidney' figo (UP,Gi,Di,Ch,Du,Ung), nso (Mw,Am,Si,Pa), iso (Ti), imfyo (Mn), ntso (Ng,Nz), tso ~ ntso (Ma)

*mfwisi 'hyena' °mfwisi ~ fwisi (UP), mfwisi (LP), fisi (Gi,Du,Mv,Jo, Ung,Mn)

*mfungate 'seven' °mfungaṛe ~ fungahe (UP), mfungahe (LP), fungahe (Gi,Du), fungate 'honey moon' (Am,Mv,Ung), mfukare (Ng), fukari (Nz), fukatre ~ fukare (Ma)

*mfuno 'duiker' funo (El,UP,Gi,Am,Ung)

(b) *ns : nts [n̥t̪s^{h}] ~ s LP
s El, UP, MK, ND including Mw, SD
ts^{h} Du, Ra
nts ~ s Ng
s ~ ts Ma
s ~ nts Nz
s ~ ns Mn

*nsi̯ku 'day' siku (El,UP,Gi,Ch,Di,Am,Mv,Jo,Ung,Vu), ntsiku [n̥ts^{h}iku] (LP), ts^{h}iku (Du,Ra), sku (Mw), usiku/ntsiku ~ siku (Ng), °suku (°u ¬ i) (Nz), siku ~ °suku (°u ¬ i) (Mn)

*nsi̯mba 'lion' siiba (El), siimba (Mw), simba (UP,Gi,Ung,Am,Mv,Mn), ntsimba (LP), ts^{h}imba (Ra,Du), °simba (°s ¬ ts) (Ng,Nz,Ma)

*nši̯ngo 'neck' siigo (El), singo (UP,Gi,Ch,Di,Chi,Vu,Mt,Pe,Tu,Mak,Mn), ts^{h}ingo (Du,Ra), tsingo (Ma), šingo (Ti,Am,Pa,Si,Jo,Ung), šiingo (Mw), šingo 5/6 (Mv), ntsingo (Ng,Mh,Nz)

(c) *mv : v El, UP, Ti, Mak, Tu, Ma
mv ~ v Pe
mv LP, MK, ND, Ung, Ng, Nz, Mn

*mvula 'rain' vúla (El), vula (UP), mvuya (LP), mvula ~ vula (Gi), mvula (Ch,Du,Mn), mvula ~ mvura (Di), mvuɬa (Mw), vua (Ti,Mak), mvua ~ vua (Si), mvua (Am,Pa,Mv,Jo,Chi,Ung,Mt,Ng,Mh), vuya ~ mvua (Pe), vuya (Tu), mvua (Vu), mvua ~ vua (Nz), vua ~ vwa (Ma). NB: /mv/ is phonetically [ɱv]; some sources write "nv"

(d) *nz	:	z	El, UP, Mak, Tu, Pe
		nd̪	Am, Si, Pa, Ti
		nz	Mw, Ung, Mv, Mn
		ndz	LP, MK, Ng, Nz, Ma

*nzige 'locust'	ndzige (LP), ndzije (Gi), nzie (Vu), nd̪ie (Am), nzige (Ung,Mv)
*-anzy- 'begin'	-andza (Gi,Du), -and̪a (Am), -anza (Ung,Mv,Mn), -aza (Mak), -aanza (Mw)
*-fyonz- 'suck'	-sond̪a (Am), -fyonza (Ung), -fyonza ~ -fyoza (Mak), -fwoza (Pe), -tsondza (Ng), -sondza ~ -tsondza (Nz), -sondza (Ma)
*mpanzi 'grasshopper'	mpazi (UP), mpandzi (LP), pʰand̪i (Amu), pʰanzi (Ung), pandzi/ maβandzi ~ mapandzi 5/6 (Ng)
*(i)nzį fly	nindzi (LP), indzi (Gi,Du), ind̪i (Ti), nd̪i (Am), nzi (Mw, Mv,Ung,Pe,Mt), nzi ~ zi (Mak), zi (Tu), ndzi (Ng,Nz)

In most Sabaki dialects the nasal has been lost before voiceless fricatives, thus *NF > ØF (El, UP, MK, ND, SD, Mn; F = voiceless fricative).

The exceptions, where the nasal has been retained, are Lower Pokomo and at least one dialect of Comorian, namely, Ngazija, and variably Nzuani. In Lower Pokomo the nasal is devoiced, e.g , *fisi > /mfwisi/ [m̥fʰwisi] 'hyena'; in the case of reflexes of *ns, the nasal is devoiced and the fricative is affricated, e.g., *nsiku > /ntsiku/ [n̥tsʰiku] 'day'. In some cases, however, the nasal has been deleted, e.g., *nšį̨ngo > /singo/ 'neck'. Möhlig (1984/5:285) also notes variation in terms of nasal-loss in Lower Pokomo. In Ngazija and Nzuani, *ns > /nts/, e.g., *nšyį̨ngo > /ntsingo/, *mfigo > nsigo > /ntso/ 'kidney', a change shared with Maore, which later deletes the nasal, e.g., *nšyingo > ntsingo > /tsingo/. In one or two cases in Ngazija, /ts/ alternates with /s/, e.g., /usiku, ntsiku ~ siku/ 'day', and in other cases there is deletion but no affrication, e.g., /suku/ 'day' (Nz, Ma). Only Ngazija shows retention of the nasal before *f, e.g., /mfukare/ 'seven', /mfia/ 'breath'; the others, Nzuani and Maore, delete the nasal, e.g., /fukari/ (Nz) and /fukare/ (Ma), /fuwu/ 9/10 'bean'.

The Lower Pokomo situation in which *NF > /N̥F ~ F/ suggests this as a path to loss of nasal before fricative. This is the likely explanation for loss in this context in Upper Pokomo, possibly influenced by Elwana, which attests nasal-loss in all environments. That Elwana played a role here is evident in the fact that Upper Pokomo has lost the nasal before both voiced and voiceless fricatives, while retaining a nasal before both voiceless and voiced oral stops, a pattern not found elsewhere in Sabaki.

While this is a reasonable hypothesis for Pokomo, there is no evidence that Nasal-devoicing underlies loss of *N elsewhere in Sabaki. The incompatibility of nasal plus

voiceless fricative is itself enough to motivate the loss. The feature identity of both segments can easily lead to cluster-simplification.

Evidence in Mijikenda indicates that we should reconstruct a nasal for Proto-Mijikenda, at least before Proto-Mijikenda *s. Duruma and Rabai both attest a dental affricate where other dialects have an alveolar fricative, e.g., *nsiku > /ts^{h}iku/ [t̪s^{h}iku] (Du, Ra), but /siku/ elsewhere. The most straightforward explanation for the affricate is postnasal affrication, which requires the presence of either a synchronic or a diachronic nasal. There is no evidence for an underlying synchronic nasal, thus we posit Proto-Mijikenda *nsiku, etc.

Before voiced fricatives the nasal is widely attested, having only consistently deleted in Elwana and Upper Pokomo, with some apparently scattered attestation of loss in other dialects, e.g., /vua/ (Ti, Mak), /vuya/ (Tu, Pe), /vua ~ vwa/ (Ma) < *mvula 'rain', /-aza/ (Mak) < *-anza 'begin', /-fwoza/ (Pe) < *-fyonza 'suck'. Pemba attests both /vuya/ and /mvua/, where the latter may be a borrowing. In Maore, nasal-loss before /v/ would appear to be general, e.g., /vi/ 'white hair' /vau/ 'garment'. Otherwise Maore retains a nasal before voiced consonants (*nz : /ndz/, etc.).

Finally, in ND$_3$, but not Mwiini nor Mvita, *nz > /n̪d̪/, while in Lower Pokomo, Mijikenda, and Comorian (Ng, Nz, Ma) *nz > /ndz/ [n̪d̪z]. As argued earlier for ND /d̪/ (see §8.1.4), /n̪d̪/ may derive from an intermediate /ndz/, thus *nz > ndz [n̪d̪z] > n̪d̪. Given that the languages involved, except present-day Comorian, are in roughly the same geographic area, and given the phonetic reasonableness of the seriation (affrication and stopping of a fricative), a case for shared history is supported. However, since Mwiini and Mvita did not undergo this process, it must be later in ND than the change affecting *nj, although probably an extension of it. Again, as in §8.1.4., the similarity of what happens in Lower Pokomo-Comorian and (parts of) ND supports the case for shared history.

§8.4. Monosyllabic nominal stems and nasal rules. The nasal rules discussed above were described for polysyllabic stems. For monosyllabic stems the situation is different. For example, nasal-loss which applies regularly in Swahili before voiceless consonants does not affect nasals with monosyllabic stems, thus /n̩t^{h}a/ 'wax', /m̩p^{h}ya/ 'new', or Digo /n̩ts^{h}a/ 'top twigs in a tree' versus /ts^{h}ano/ 'five'. The rule does not delete syllabic nasals which are stress-bearers.

The Sabaki dialects differ in their strategies for dealing with nasals with monosyllabic noun stems, and differ in the morpheme structure of their lexicons. In some cases nasal deletion in the usual environments does apply but extra elements are added to the noun. In Upper Pokomo, /ni/ has been added to monosyllabic roots, either as a stress- or prosody-bearer, or to preserve bisyllabic lexical structure, or both, e.g., /nisu/ < *nswi 'fish , /nizi/ < *nzi 'fly', /ninge/ < *nge 'scorpion', etc., in contrast to Lower Pokomo: /ntswi/ < *nswi 'fish', /ntsa/ < *nca 'point', /nzi/ < *nzi 'fly'.

A similar phenomenon is found in Mijikenda where certain monosyllabic adjectival stems have an epenthetic /i/ in Class 9/10 forms, and are consequently treated as V-initial stems, e.g., Giryama and Chonyi /-re/ < *-le 'long', /mure/ 1/3, /are/ 2, /kire, vire/ 7/8, /rire, mare/ 5/6, /lure/ 11, /kare/ 12, /hare/ 16, but /nyire/ 9/10. The adjective /-ša/ < *-pya

'new' behaves similarly, except that the 9/10 form is /mbiša/ (Gi), /m̥bya ~ nyiβya/ (Di). Two other monosyllabic adjectives, /-dzo/ 'good' and /-fu/ 'hard', have /mbidzo/ and /mbifu/ (but Digo /nyifu/) in their 9/10 shapes. Loss of nasal, epenthesis, and analogy operate here to explain these forms. Giryama and Chonyi show evidence of nasal-loss and i-epenthesis in the case of monosyllabics: Giryama /n̥tsʰi ~ tsʰi/ 'land', Chonyi /itsi/. It would appear that Class 9/10 adjectives were similarly treated, thus Giryama *nša [+9/10] > /ša/ > iša > /mbiša/. The final stage is due to analogy. Once the adjective looked like a vowel-stem, it began to be treated like other synchronic vowel-stem adjectives: /-iri/ 'two', /-i/ 'bad', /-itsi/ 'unripe' (all *W-initial stems), and /-iru/ 'black' (a *V-initial stem) are in their 9/10 shapes, /mbiri, mbi, mbitsi, and nyiru/ respectively. Two patterns are thus available. Giryama and Chonyi treat intermediate *-idzo, *-iša, and *-ifu following the first, and *-ire the second. Digo, on the other hand, favors the second pattern: /m̥bya ~ nyiβya/ 'new', /nyifu/ 'hard' (for further detail and discussion see Hinnebusch 1973).

Monosyllabic Class 9/10 nouns in Tikuu have an initial vowel /i/: /iţ̪ʰi/ 'country' (cf. Amu /ntʰi/), /isi/ 'fish' (Amu /nswi/), /inḏi/ 'fly' (cf. Amu /nzi/), /imbwa/ 'dog' (cf. Amu mbwa) (Nurse 1982b has a fuller list). The simplest explanation for this initial /i/ is that it is a relic of the old Class 9/10 pre-prefixal vowel, retained from PSW, PSA, and PNEC (see Chap. 4, §2.1.1 and 2.1.2) for the same kinds of reasons that Upper Pokomo has added a new initial /ni-/. The pre-prefix is also preserved in Sabaki in all classes in Pokomo and in Comorian. In NEC, most of Ruvu still has the pre-prefix in all classes. A similar analysis is possible for Class 5 monosyllabic nouns in some Sabaki languages.

An alternative analysis of this Class 9/10 vowel in Tikuu would see it as epenthetic. The addition of /i/, a [+coronal] vowel, might be expected, since the most frequently occurring consonants in these Class 9/10 nouns are coronal, namely, [t̪, t, ḏ, d]. Similar epenthesis is seen in some nonstandard varieties of upcountry Swahili, e.g., /umbwa/ 'dog', /inči/ 'country', /inje/ 'outside', where the features of the nasal predict the features of the epenthetic vowel. In Mwani, where we find both /imbwa/ and /umbwa/ 'dog', we can assume something of the same sort is happening. A further Mwani item attests to the contemporary nature of the process. Mwani is a *g-loss language. After this late change created a monosyllabic stem from a bisyllabic stem, an epenthetic vowel was added to the nasal prefix: *mfigo 'kidney' > mfyo > /imfyo/.

Polomé (1967) suggests that the addition of an i-colored vocalic peak to monosyllabic nouns in some Swahili dialects is motivated to mark the syllabicity of the nasal; this then becomes identified with the phoneme /i/. Another way of looking at it is to say that vowel-addition in such cases assumes the absence of syllabicity-marking in the nasal. The fact that /i/ is then found before [+labial] segments, as in Tikuu /imbwa/, or variably in Mwani, is a matter of regular morphological marking.

In Chifundi we find vowel-lengthening and nasal-desyllabification, or nasal-loss (Bakari 1985:76-77), thus *(i)mbwa 'dog' > /mbwaa/, *(i)nzi̧ 'fly' > /nzii/, *(i)nswi̧ 'fish' > /swii/, *(i)nci 'country' > /t̪ʰii/. Vumba behaves similarly, e.g., *(i)mbwa > /mbwaa/, but *(i)nswi̧ becomes /swi/. In other dialects there is no restriction on having monosyllabic

nouns in the lexicon, e.g., *(i)nchi 'country' > /čʰi/ (Mt, Mak, Tu), /tsi/ (Ma); *(i)nzį > /zi ~ nzi/ (Mak), /zi/ (Tu). H. Ipu (pers. comm.) reports a similar process in Lower Pokomo.

In sum, what we see here is a range of strategies to prevent the reduction of nouns to a monosyllabic shape, all apparently utilized for reasons of accent/stress assignment and morpheme structure requirements. These strategies are: retention of the initial vowel from the older pre-prefix (Ti); insertion of syllabic material (UP); insertion of syllabic material and analogy (MK; see Hinnebusch 1973); nasal-syllabicity (Ung); prefixless stressed monosyllabic nouns (Ma); and final-vowel lengthening (Chi, Vu, LP). We assume all of these postdate the breakup of the PSA dialect continuum. Thus our analyses here have few implications for subgrouping. In Appendix 2 these nouns are listed with an initial vowel in parentheses, e.g., *(i)mbwa.

§8.5. **Nasal rules and chronology**. We have discussed a number of rule processes affecting *NC in Sabaki. Their forms (somewhat simplified) and distributions, along with a statement of their value for subgrouping, are:

Kwanyama Rule	Ng, Mh (independent innovation)
Meinhof's Law	PSA (< pre-NEC period; appears to affect *ng only)
*NC̥ (C̥ = [-velar]) > NC̥	Ng, Mh (independent innovation)
*NC̥ > NC̥ > ØC̥	El (independent or areal innovation, or substratum)
*NC (C = [±stop]) > ØC	El (all environments)
*NC̥ > N̥C̥ʰ	Po, MK, ND, Mw, SD (NEC areal)
*N̥C̥ʰ > C̥ʰ	MK, ND (not Mw), SD (post-PSA innovation)
*NC̥ > ØC̥	Ma (independent innovation)
*nj > ndz	MK, LP, Com, possibly ND_3, Mw, Mv, Chi (late PSA innovation)
*nj > (ndz) > nd̯	ND_3 Mw, Mv, Chi (late PSA innovation)
*nz > ndz	MK, LP, Com, possibly ND_3 (not Mw and Mv) (late PSA innovation)
*nz > (ndz) > nd̯	ND_3 (not Mw and Mv)
*nd > ndr	ND (including Mw), Nz, Ma (post-PSA innovation)
*NF > N̥F	Po (F = voiceless fricative)
*mf > Øf	Sabaki (except LP, Mw, Ng, Mn) (post-PSA innovation)
*ns(y) > Øs	Sabaki (except LP, Ng, variably Nz, Ma) (post-PSA innovation)
*mf > ns	Com, ND_2 (parallel; to *fi/vi > si/zi, see §5.1.1)
*mf > ns > nts	Com (independent innovation)
*nts > Øts	Du, Ra, Ma (independent innovation in each)
*mv > Øv	El, Ti, Ma, variably in rural SD (post-PSA innovation)
*nz > Øz	El, UP, variably in rural SD (post-PSA innovation)

Most of these are post-PSA rules, and in many cases innovated long after the development of individual dialect/language groups, as indicated by their limited and restricted distribution. This is true even in the case of apparently widespread Bantu rules. Consequently, their operation impinges very little on the operation of other rules. Further, in many cases their domain and structural description are independent of other rules, and thus need not be ordered in respect to other rules. There are, however, examples where rule processes are related, but they generally involve other nasal rules. For example, Elwana postnasal C-voicing feeds nasal-deletion:

*nkomba 'bushbaby'	
ŋgomba	Postnasal voicing
ØgooØba	Nasal-loss (and length adjustment)
gooba	Output

Nasal-loss in Swahili and Mijikenda dialects was made possible by nasal-devoicing and aspiration. Otherwise it is unlikely that the nasal would have deleted:

*mpanya 'rodent species'	
m̥pʰanya	Nasal-devoicing/aspiration (Sw/MK/Po)
Øpʰanya	Nasal-loss

But as far as we can tell, there is no evidence to indicate that nasal-loss in Maore involved either nasal-devoicing or aspiration.

There are also a few examples which illustrate the potential interaction with other rules. Thus the following Mijikenda derivation involves lenition and nasal-loss:

*mpepo 'wind'	
m̥pʰepo	Nasal-devoicing/aspiration (Sw/MK/Po)
m̥pʰeɸo	*p-lenition (MK/Po)
Øpʰeɸo	Nasal-loss

In Bantu languages, the least marked feature specification of a consonant in postnasal position is an oral stop. The nasal stop, therefore, prevents lenition from applying. This is clear from Comorian, where lenition does not affect postnasal stops. This is good evidence that nasal-loss is later than *p-lenition. The stopness of the nasal is also a factor in understanding the widespread application of nasal-loss before fricatives, and where loss does not occur, we usually find fricatives becoming affricates: *ns > [nts], *nz > [ndz], *mf > [mpf], etc.

If nasal-deletion occurred chronologically earlier than *p-lenition all *p's would have weakened, an incorrect derivation:

*mpepo 'wind'	
m̥pʰepo	Nasal-devoicing/aspiration (Sw/MK/Po)
Øpʰepo	Nasal-loss
*ØɸeɸoØ	Lenition

There are conditions, however, where even this ordering will give the correct output. One could reasonably argue that lenition only affects consonants in intervocalic position, thus exempting word-initial consonants. Either a preconsonantal nasal or a word-initial consonant constitute a natural, phonetically plausible barrier to lenition. Thus, the question of ordering is moot with such a constraint on the operation of Sabaki lenition rules. Therefore, Sabaki lenition could apply before or after nasal-loss. So in this case we do not have strong evidence for a particular rule-order, and no predictability for chronology. Stratigraphy tends to favor the earlier innovation of lenition, assuming, as we always do in applying notions of stratigraphy, that a particular process is a genetic-areal one. It is found in Pokomo, Mijikenda, Vumba, Chifundi, and Comorian, while nasal-loss (only that part of it which is contingent upon nasal-devoicing/aspiration) is specific to Mijikenda and Swahili (not Mwiini). So while we have problems with our methodology here, in that it does not provide clearcut chronology, one possible rule-ordering (lenition + nasal-loss), plus stratigraphy, points to the late presence of a nasal in Mijikenda. We should note that the implications of this discussion for subgrouping are minimal, since Bantu nasal-loss is due to natural phonetic processes, and innovations leading to loss are weak subgrouping criteria.

Another example of rule interaction is found in Mijikenda. In Duruma/Rabai there are affricates which can only be explained by the diachronic presence of a nasal:

*nsiku 'day'	
ntsiku	*s-affrication (Du/Ra)
nts^{h}iku	Nasal-devoicing/aspiration
Øts^{h}iku	Nasal-loss (cf. other MK /siku/)

This would indicate the presence of a nasal in Mijikenda long after the breakup of the group into its separate dialects, and late application of nasal-loss before *s. This also attests to the late operation of nasal-loss and aspiration in Sabaki across the board.

A final example of ordering comes from Comorian. Comorian (and ND_2) Coronalization, in which *f and *v, derived earlier by Bantu Spirantization, shift to /s/ and /z/ before the high front vowel /i/, feeds several nasal rules in Ngazija and Nzuani:

CB:	*-pi̧gò	*-pímbò	*-pu̧ngàté	*-pu̧ngàté	
	'kidney'	'stick'	'seven'	'seven'	
PSA:	*mfi̧go	*mfi̧mbo	*mfu̧ngate	*mfu̧ngate	
Com:	*mfi̧go	*mfi̧mbo	*mfu̧kare	*mfu̧kare	
	nsigo	nsimbo	---	---	(1) *mf > ns/ __ i
	--	--	---	fukare	(2) Nasal-loss/ __*f
	ntsigo	ntsimbo	---	---	(3) s-affrication
	ntso	ntsimbo	mfukare	fukari	(4) Other rules
	ntso	ntsimbo	mfukare	fukari	Output
	(Ng, Nz)	(Nz)	(Ng)	(Nz)	

The distribution of nasal-loss in the Comorian dialects indicates that loss before *f is regular in Nzuani and Maore, but not in Ngazija, and that loss before *s only occurred in

Maore after /s/ (from *f or *s) was affricated (*nšyingo 'neck' > /ntsingo/ (Ng, Nz, Ma) > Øtsingo). This evidence shows that *m-loss before a fricative in Comorian preceded *n-loss, and that, furthermore, nasal-loss before fricatives was late in Comorian and probably independent of nasal-loss in other Sabaki languages, apart from what might be explained by drift. This is supported in the derivational history above. Were nasal-loss before *f an early Bantu change inherited by Comorian, as its uniform distribution in Bantu might suggest, it would have applied to forms such as *mfigo and *mfimbo before the strictly local (Com, ND) change in which *f > s/___ i. Thus, *m-deletion must be later and specific to Comorian.

In the above list of rules, we have identified languages and/or language groups which attest a nasal-loss rule. Before the fricatives *f and *s the rule affects nearly all Sabaki dialects; the exceptions are Lower Pokomo, Comorian, and Mwiini, where loss is only variable. Moreover, loss before *f is more general than before *s. In contrast, loss before voiceless stops occurs less widely: Swahili (not Mwiini), Mijikenda, Pokomo, and Maore. This would indicate that nasal-loss is not a single process within Sabaki.

This is supported by the observation that nasal-loss before voiceless fricatives is common outside Sabaki, while loss before voiceless stops is less frequent in comparison. Also, there are no Bantu languages that we know of which attest nasal-loss before voiceless stops that do not also attest loss before voiceless fricatives. This suggests that nasal-loss before voiceless fricatives is a Bantu process, that is, liable to occur whenever nasal + voiceless fricative combinations are created, for example, as they are by Spirantization. We previously concluded that Spirantization is an early PSA process, portions of which are undoubtedly pre-PSA, even pre-NEC. It follows then that, as the process created nasal + voiceless fricative combinations, these became candidates for change as the combinations themselves appeared, thus CB *nkɨ́ngò 'neck' > *nšyɨngo (PSA) > /singo/ (LP); CB *ntɨ́kù 'day' > *nsɨku (PSA) > /siku/ (El, UP, Gi, ND, SD) and > /ntsiku/ (LP). Loss of a nasal in these cases is a natural consequence of a phonologically marked combination of voiced nasal stop + voiceless oral fricative; loss leads to a simplified phonological inventory. If, on the other hand, loss does not occur, for whatever internal phonological reasons, then affrication of the fricative occurs, e.g., *nsɨku (PSA) > /ntsiku/ [n̥tsʰiku] (LP), /tsiku/ (Du/Ra). While nasal-loss before voiceless stops is due to equally natural and plausible mechanisms, processes other than loss can come into play, creating greater variety in reflex type. For Bantu overall, the following chronology for change involving loss of the nasal element in *NC clusters is clear: (a) loss of the nasal first before *f and then before *s, (b) loss before a voiceless consonant, (c) loss before voiced fricatives, and (d) finally before voiced stops (a counter-example is Upper Pokomo, which has deleted *N only before fricatives).

The distribution of nasal-loss within Sabaki is probably not due to innovation during a common proto period. Rather it more likely started as the Sabaki subgroups began to go their own ways. This hypothesis is supported by the differential behavior within otherwise well-established groups of languages: Mwiini with nasal-loss before nonlabial voiceless

fricatives but not before voiceless stops in contrast to the other Swahili dialects with both; Mijikenda with evidence of a nasal before fricatives in Duruma/Rabai, and thus at an earlier Proto-Mijikenda level; Comorian with nasal-loss in Maore, but only variably before fricatives in Ngazija and Nzuani; Lower Pokomo with well-attested nasals before fricatives which only now appear to be deleting, while Upper Pokomo has already lost the nasal in that environment; Mwani with deletion apparently before *s but not before *f. Thus the evidence is strong that nasal-deletion before fricatives in Sabaki, although identical to what happens elsewhere in NEC and elsewhere in Bantu, must be considered an independent process.

§9.0. Consonant-Glide Sequences

Sabaki consonant-glide (CG) sequences undergo little change, and the change that is apparent is either restricted to specific subgroups, e.g., Mijikenda Labialvelarization, Mwini W-absorption, or is sporadic.

§9.1. Mijikenda Labialvelarization. In the Mijikenda dialects, certain sequences of C + w are affected by a synchronically active process which turns them into labiovelar coarticulated stops (Hinnebusch 1973; Kelly, forthcoming):

/kw/ → k͡p
/gw/ → g͡b
/mw/ → ŋ͡m
/bw/ → g͡b

The following examples are from Giryama. Nasals, written /N/, are realized as homorganic with the following segment:

/kwapa/ → [k͡paha] 'armpit'
/-gwa/ → [-g͡ba] 'fall'
/mwaka/ → [ŋ͡maka] 'year'
/-imbwa/ → [-iNg͡ba] 'be sung'

While the process is undoubtedly active, it is not fully productive, in that not all forms which could undergo the change do so. In all dialects the process usually occurs only before the [-round] vowels /i, e, a/ and not before /o, u/, with the exception of Digo, which has optionally generalized the process to cases of /labial nasal + velar/, namely, [m̩k͡pono] 'arm', [m̩k͡podzo] 'urine', [m̩k͡poβa] 'basket', [m̩g͡buβa] 'barren woman', [m̩g͡baŋga] 'doctor'. Kelly (forthcoming) recorded these as well, but with a homorganic nasal, e.g., ŋ͡mk͡pahe 'bread'; these were also heard in their plain form, e.g., [m̩kono], [m̩kodzo], etc. Variation of various sorts within idiolects, from idiolect to idiolect, and from dialect to dialect, was noted. In Giryama, labiovelars are produced with considerable lip-smacking and suction; this was not noted in the other dialects. Differences were also noted in the degree of lip-rounding associated with labiovelars. It is most prominent before front vowels, e.g., k͡pʷenda 'to go' (Gi), [k͡pʷiva] ~ [kwiva] 'to ripen' (Ch), [kuwima] ~ [k͡pʷima] 'to stand' (Di). In Duruma, lack of lip-rounding was reported (L. C. Jacobsen

pers. comm.). Other interesting variation was recorded. Some examples follow. In Giryama, compared to Chonyi, we also find optional vowel deletion:[41]

	Giryama	Chonyi
/kwambira/ 'to tell'	[k͡pwambira ~ kumbira]	[k͡pambira]
/kwanyesa/ 'to show'	[k͡panyesa ~ kunyesa]	[ku^{ʷ}onyeša]
/kwalaga/ 'to kill'	[k͡palaga ~ kulaga]	[ko:laga]
/kwangiza/ 'to put into'	[k͡pangiza ~ kungiza]	[k͡p^{ʷ}enjiza]
/kwakuha/ 'to be satisfied'	[k͡pakuha ~ kukuha]	[k͡p^{ʷ}ekuha]
/kwanula/ 'to lift'	[k͡panula ~ kunula]	[k͡p^{ʷ}enula]

Vowel deletion here has the effect of limiting the application of labialvelarization, and results in more uniform marking of the infinitival affix /ku-/ in the whole paradigm. Compared with other Mijikenda dialects, Giryama sequences of infinitival /ku-/ plus vowel stem with /i/ and /e/ tend not to undergo the change, and are realized as [ku^{ʷ}i] and [ku^{ʷ}e]. In Chonyi, /kw/ can be realized as [k͡p^{ʷ}], [k͡p], or [kw]; /mw/ as [mw̃], where the [w] is nasalized, or [ŋ͡m]; however, /gw/ is always realized as [g͡b]. In Digo, /kw/ was usually heard as [ku] in citation forms, but in sentential context it was almost always realized as [k͡p^{ʷ}] or [k͡p], depending on the quality of the following vowel. The Digo sequence /mw/ was sometimes heard with some nasalization of the approximant, but never as [ŋ͡m], in the speech of several consultants.

Labialvelarization is a process in which sequences of a [+stop] consonant plus an approximant are reduced to single segmental units. Underlying the shift is an absorption process in which /w/ assimilates the manner of articulation features (e.g., [-stop] → [+stop]), voicing, and nasality features of the preceding stop, while providing either the feature [+labial] or [+velar] to the resulting segment. Thus:

/CVV/ → CGV → C(^{ʷ})V
/guV/ → gwV → g͡b(^{ʷ})V
/kuV/ → kwV → k͡p(^{ʷ})V

In the previous examples, the labiovelar /w/ provides the feature labiality to the resulting segments, but it provides velarity in the following:

/muV/ → mwV → ŋ͡m(^{ʷ})V
/buV/ → bwV → g͡bV

In Digo cases where /ṃkV/ and /ṃgV/ are realized [ṃk͡p] and [ṃg͡b], e.g., [ṃk͡pono] 'arm', [ṃk͡podzo] 'urine', [ṃk͡poɓa] 'basket', [ṃg͡buɓa] 'barren woman', [ṃg͡baŋga] 'doctor', the velar absorbs labiality from the preceding syllabic nasal:

---

[41]The following data were elicited in 1972/73 by TJH. Note that the initial vowels in some of the verbs of this data set are different from their corresponding reconstructed forms, and different from the usual citation forms, e.g., compare /kwanyesa/ with /-onyesa/, PSA *-Wonyešy-. It is not certain how representative this data is.

/muk/ → mk → mk͡p
/mug/ → mg → mg͡b

Some seriation is evident in the distribution of labialvelarization in Mijikenda. The development of [g͡b] < /gw/ is uniformly attested with little variation in all the dialects, with [k͡p] also attested in all dialects but with considerable variation with [k͡pʷ] ~ [kw] ~ [kuʷV]. The shift /mw/ > [ŋ͡m] is only attested in some dialects, and the most restricted shift of all, /bw/ > [g͡b], is found in Giryama:

Stage 1:	/gw/	>	[g͡b] ~ [g͡bʷ]	Mijikenda
Stage 2:	/kw/	>	[k͡p] ~ [k͡pʷ] ~ [kw]	Mijikenda
Stage 3:	/mw/	>	[ŋ͡m] ~ [ŋ͡mʷ]	Giryama, Duruma...
		>	[ŋ͡mʷ] ~ [mw̃]	Chonyi
		>	[mw̃]	Digo
Stage 4:	/bw/	>	[g͡b]	Giryama

Labialvelarization effectively defines Mijikenda, even though it has marginally spread outside the group. It occurs as an articulatory feature in the non-urban Swahili dialects of the Mombasa area, e.g., Jomvu.[42] Outside Sabaki there are languages in which similar processes are taking place. Thus in Shona, /w/ assimilates to preceding labials or alveolars, but not to velars: /pw/ → [px], /bw/ → [bɣ], /tw/ → [txw], /dw/ → [dɣw], etc., but /kw/ → [kw]. Shambala has /mw/ → [ŋw] (*mwana > ŋwana) (Tucker and Bryan 1957:48), and Nyanja has /mw/ → [mŋʷ] in careful speech and [mŋ] in normal conversational tempo. While these are reminiscent, they are independently based assimilations, distinct and unconnected with the somewhat unique process in Mijikenda.

§9.2. **Glide-deletion: Swahili and Comorian.** Several rule processes operate on consonant-glide sequences. Some we have already discussed in earlier sections. Here we discuss G-deletion, plus several other processes involving CG sequences. Most change of this type characterizes the Swahili dialects (mainly ND) and Comorian. The other languages are conservative and for the most part retain CG sequences. Except for a general glide-deletion rule in at least one variety of Mwiini, deletion only affects glides in specific consonantal environments.[43] Some Mwiini examples showing vowel-lengthening and deletion in word-medial position are:

*mbwa 'dog'	mbaa ~ mbwa	*-lwal- 'be sick'	-laaɬa
*lwembe 'razor'	leembe	*mulungwana 'freeborn'	mɬungaana
*muvwi 'arrow'	mvi	*mwana 'child'	maana

[42]In Vumba and Chifundi (and to a very minor extent in Mtang'ata), PSA *-gwa > /-bwa/ 'fall'. This is reminiscent of labialvelarization in Mijikenda: *-gwa > -gbwa > /bwa/. We have no evidence that this is a more general process in these dialects.

[43]Vianello (pers. comm., 1988) reports two dialects of Swahili spoken in Brava, one with *w-deletion (as recorded by Whiteley 1965) and the other which preserves *w (as recorded by Kisseberth and Abasheikh in their materials).

*nywe 'you (pl.)'	ni	*-nyw- 'drink'	-n̰a
*-vwala 'wear'	-vaɬa	*-twal- 'take, marry'	-taaɬa
*-akwe 'his, her'	-aake	*lufwa 'crack'	lufa
*nswi 'fish'	nsi	*swe 'we'	si
*muji + ni 'in town'	mwiini (= Mwiini language)		

Similar examples also show Mwiini y-deletion, but earlier rules, e.g., causative (§5.3) and palatalization (§4.1.3.2), which are triggered by *y and entail its absorption in the process, have eliminated most cases where G-deletion would otherwise have applied, so examples are limited:

*lupyagilo 'broom'	lpeeɬo
*kivyele 'very old person'	čizeeɬe (*vy > zy > z, §5.1.1)

In certain cases, *y-deletion is bled by other phonological processes. For example, after *l, the glide provided the motivation for the shift of *ly to /ǰ/:

*-lya 'that'	-ǰe (cf. *ly > ɟ in Pokomo)
*-ly- 'eat'	-ǰa

Certain dialects agree in deleting either *w or *y after certain consonants. Swahili and Comorian tend to delete *w after *f/*v and *s:

*lufwa 'crack'	lufa (Mn), ufa (Am,Ung,Ng)
*-fw- 'die'	-fa (Am,Ung,Ng), but -fwa (Vu)
*luvwi 'white hair'	mvi (Am,Ung,Ng)
*-vwal- 'wear'	-vaa (Am,Ung,Ng)
*nswi 'fish'	nsi (Pa,Si,Mw), isi (Ti), mfi (Ng)
*swe 'we'	si (Mw), isi (Ti/Si), sisi (Am,Pa)
	wasi (Nz), sisi (Ng)

W-deletion after *f/*v characterizes ND and SD, but it is not clear whether *w deletes after *s in SD. We have too few reconstructed items with this sequence, and there is considerable reflex variation in the one reliable item. Thus in SD we find both /s/ and /sw/ for *swe: /swiswi/ (Vu), /swiswi ~ siye/ (Mt, Mak), /siye/ (Pe, Tu), but consistently /sw/ for *nswi: /nswi/ (Ung, archaic), /swi/ (Vu, Mak, Mt). The front glide *y is dropped in Swahili (mostly ND) and Comorian in the following:

*lupyagilo 'broom'	fagio (Ung) (*py > fy > f), peleo (Ng), upeo (Am), peo (Nz)
*muvyele 'elder'	mzeeɬe (Mw), mzee (other ND)
*mulyango 'door'	mlango (Am,Ng), mlongo (Nz), nlango (Jo,Ng)
*kyakulya 'food'	čakula (Am), šahula (Ng,Nz,Ma)
*-ly- 'eat'	-la (Si,Pa,Ti,Am,Mv,Ung,Ng,Nz); -lya (Mak,Tu,Pe,Mt)
*-navy- 'wash hands'	-naza (Am) (*vy > zy > z)
*-lovy- 'make wet'	-loza (Ng)

Simple deletion explains most of these examples, but the deletion of *y in *vy sequences in ND and Comorian involves other processes, see §5.1.1 (coronalization), and §5.3 (causative) which entail the absorption of *y.

In Unguja, where *y-deletion shows up in lexical items, these must be loans, since other SD do not delete glides, e.g., *-vyal- 'give birth' > /-vyala/ (Vu) (cf. Ung °-zaa), *-levy- 'make drunk' > /-levya/ (Ung, Mak), *muvyazi 'parent' > /-mvyazi/ (Vu, Mak) (cf. Ung °mzazi); *-ly- 'eat' > -lya (Pe, Tu) (cf. Ung °-la); *kyakulya 'food' > /kyakulya/ (Vu) (cf. Ung °čakula). However, Unguja regularly deletes glides after *f, whereas other SD do not, e.g., /-fyuka/ 'go off on a tangent' < *-pyuka 'deviate' (Pe, etc.), /-fa/ (ND, SD, Com), but -fwa (Vu) < *-fwa 'die'. Mwani also preserves a reflex of *y: /-rya/ 'eat'.

After /h/, Comorian and Mijikenda delete /w/:

*-kwel-	'climb'	-he(y)a (Ng,Nz,Ma) (*kw > hw > h)
*-akwe	'his/her'	-ahe (Ng,Ma)
*mukwe	'in-law'	mha (Nz)
*-twik-	'put up load'	-hika (Gi) (*tw > hw > h)
*-twala	'take, marry'	-hala (Gi)

Here a genetic connection is unlikely, deletion being natural and phonetically motivated. The subgrouping implications of G-deletion are minimal because deletion is involved. It is nevertheless notable that again we find a process shared between ND and Comorian:

ND and Comorian Glide-Deletion

*fw	:	f	*sw	:	s
*vw	:	v	*ly	:	l (in Mwiini: *ly > ǰ)

10.0. PSA Seven Vowels, Long Vowels, and Vowel Processes

§10.1. Proto-Sabaki seven vowels. There is good evidence in support of reconstructing seven vowels for PSA: *į, *i, *e, *a, *o, *u and *ų. Briefly, the argument includes: (a) Both *į and *ų are needed to explain strengthening in Comorian and SD, e.g., /-iba/ (Com, SD) < *-įWa 'steal', /mvuba/ (Com) < *muvųWa 'bellows'; discussion and supporting data can be found in §6. (b) Comorian affrication also requires the [+HV] feature to derive forms such as /mwidzi/ (*mwįvį > mwįzį > mwidzi). In such cases it is clear that we need [+HV] front vowels quite late in the post-PSA period to trigger the shift (for details see §6.3). (c) The spirantizing diversity in Sabaki and Seuta is easier to understand if we have seven vowels rather than five. The Swahili dialects in Sabaki and the Seuta dialects, specifically Shambala, in NEC do not mirror the situation in other languages, which suggests that Bantu Spirantization was not a fully inherited feature of either PSA or PNEC, and was still in operation after the individual PNEC dialects began to diverge. The diversity, and "irregularity," is more easily explained if we can assume seven vowels until quite late in the evolution of PSA. (d) Finally there is evidence that one of the Sabaki languages, Elwana, still attests seven vowels, along with fricatives. Möhlig (1986a, 1986c) records nine vowels, claiming phonemic status for seven. Maddieson and Sands (field notes 1991)

also record at least 10 phonetic vowels. However, when this material is compared with Guthrie's CB, we only find a partial match as far as etymological [+HV] vowels are concerned; for example, the following correspond (the data are Möhlig's): ívu < *ɨvụ 'ashes' (CB *-bʉ́), vúla < *mvụla 'rain' (CB *-bʉ́dà), sîmbɔ < *mfɨmbo 'stick' (CB *-pímbò), mwîwa < *mwɨWa 'thorn' (CB *-yɨ́bà), but others do not: mufóópà < *mufụpa 'bone' (CB *-kʉ́pà), mwéèzɪ < *mwezɨ 'moon, month' (CB *-yédì), -élʊ < *-ɨlu 'black' (CB *-yɨ́dù).[44]

We have therefore set up for PSA both seven vowels and spirants, a system that, as we have seen, has many parallels in modern-day East Highlands languages (§5.4). However, the near-uniform distribution of five vowels in the Sabaki languages today suggests that this system collapsed fairly early, and did not long endure after Bantu Spirantization. That Elwana still retains seven vowels reflects its isolated geographical location and/or conservatism.

§10.2. Long vowels. There are three pieces of data that support PSA long vowels. One is from Mwiini and possibly Elwana, which still attest etymological long vowels; the second is from Comorian, and the third from Mijikenda.[45]

Mwiini has phonological long vowels which derive from different sources. Some arise from the same low-level phonetic processes that give rise to phonetically lengthened vowels in other Bantu languages. Thus, Mwiini attests long vowels, for example, before nasal-consonant clusters, e.g., /muuntu/ 'person' (*muntu); after *CG clusters where the glide has deleted, e.g., /kaa/ 'at, for' (*kwa), /mooši/ 'smoke' (*mwošɨ); where an intervocalic consonant has been deleted, e.g., /peela ~ pela/ 'sweep' (*-pyagila), /nduu ~ ndu/ 'relative' (*ndugu), etc.

Mwiini also has long vowels that have reconstructed CB etynoms with long vowels:

Mwiini	**CB**	**Mwiini**	**CB**
baaba ~ waawa 'father'	*bààbá	-weeka 'put'	*-béék-
-kuuɬa 'extract'	*-kúúd-	-ɬaaɬa 'sleep'	*-dáád-
-ɬeeta 'bring'	*-déét-	maa 'mother'	*máá
maama 'mother'	*mààma	či-nooɬo 'whetstone'	*-nòòd-
taano 'five'	*-táánò	-tuuɬa 'calm down'	*-túúd-
yuuzi 'day before yesterday'	*-jóódɨ		

[44]Möhlig (1986a:277) states: "As to the qualities and quantities of the (Elwana) vowels, only in about 50% of the Bantu vocabulary of Elwana, can we establish regular correspondences with the hypothetical vowel system of Guthrie's Common Proto-Bantu." It should also be noted that there are some discrepancies in transcription between Möhlig 1986a and 1986c; this causes some problems in working out the correspondence relationships.

[45]The evidence for Elwana is not certain. Möhlig (1986a, 1986c) describes phonemic length for Elwana, but the small amount of data given by him does not provide sufficient correspondences with Guthrie's reconstructed long vowels. The one possibility is /sáànɔ/ < *-cáánò 'five', but this item is skewed; CB *c : č in Elwana. Ipu (pers. comm.) also reports length for Pokomo, and Batibo for Standard Swahili has described a limited length contrast (Batibo 1990).

However, there are many examples that have long vowels with short-vowel etymons. We are unable to explain the long vowels in these, and list only a few:

činoofu 'piece of meat'	*-nòkù̱	ikooko 'crust'	*-kókò
-fuura 'swell'	*-fura (EC)	ikoopa 'cup'	*(i)kopa (Port.?)
-piima 'measure'	*-pi̱m-	-vuuta 'pull'	*-dù̱t-

And then there are cases where the etymon has long vowels, but the Mwiini form has short vowels:

iweɬe 'breast'	*-béédè
kešo 'tomorrow'	*-kéékį̱-

While the distribution of length in all cases in Mwiini is unclear, etymological long vowels can only be explained if we assume PSA long vowels which were inherited from pre-PSA stages, as reflected by Guthrie's CB forms. While the rest of Sabaki has lost phonemic length, Mwiini has preserved the feature and capitalized on it by phonologizing phonetic long vowels.

In Comorian, as outlined earlier (§5.1.3, §6.3), PSA *z (< CB *l/__ i̱) has become /dz/ in several environments: (a) when bounded on either side by *i̱ e.g., dziwa 'milk' < *i̱zi̱Wa, cf. hiziwa 'pool' < *kizi̱wa; (b) before *yV sequences in causatives, e.g., -žadza (Ma, Nz) 'fill' (caus.) < *-ji̱jazya, but -ǰaza (Ng); and iii), most important for our analysis here, before *VV and *yV (both CB *VV) sequences:

-dziha (Ng, Nz) 'bury'	*-zi̱i̱ka, CB *-di̱i̱k-
-dzama (Nz) 'sink'	*-zyama, CB *-di̱àm-, but -zama (Ng)
-dzua (Ma) 'raise'	*-zyuka, CB *-di̱ùk-

If we cannot posit long vowels for an older stage of Comorian, we cannot explain Comorian *z-affrication in these environments. Moreover, Comorian must have had long vowels until rather late, since it is only Nzuani and Maore which regularly attest the shift in this environment.

The third piece of evidence comes from the analysis of prosodic phenomena in Mijikenda (and Seuta) (see G. Philippson in §16.2 and §16.6). In a number of apparently disyllabic words, a high tone ought to spread from the penultimate syllable onto the next word, but it does not. The explanation is that although today the H is on the penultimate, at the time when tone-spreading developed the H was in fact on the first mora of a long stem vowel, not on the penultimate vowel, and thus failed to spread on to the next word. At this stage ("Stage 2" of the analysis in §16.6), the Mijikenda languages still had long vowels.

Thus different and independent kinds of evidence from three of the six Sabaki languages requires the reconstruction of long vowels for PSA. And it should be remembered that the length situation in Elwana, although unclear, may also point in the same direction.

§10.3. **Vowel processes**. A number of vowel processes, mostly involving juxtaposition, often of prefix vowels with V-initial nominal and verbal stems, operate in Sabaki. We will not, however, survey or discuss these for a number of reasons: (a) most processes for

which we have evidence are inherited from earlier Bantu stages (this is reflected in our reconstructions, e.g., Class 15 *ku- (C-stems) ~ *kw- (V-stems), Class 7 *ki- (C-stems) ~ *ky- (V-stems), etc.; (b) we do not have sufficient data to deal comprehensively with vowel processes, or the changes involving vowels, other than those already discussed, e.g., 7 > 5 vowel neutralization, and labialvelarization (§9.1); (c) given the near universality of such processes, they have little subgrouping utility; and (d) most synchronic treatments of Sabaki languages deal with vowel-rules, and these can be referred to for detail (see the Bibliography).

When vowel rules appear not to have operated, this is often because they were blocked by etymological consonants which have been deleted, e.g., normally in Tikuu /i/ and /e/ coalesce: /ni-enda-o/ → [nendrao] 'I am going' , /si-end-i/ → [sendri] 'I am not going', but /va-ni-amb-ie/ → [vanambie] 'they told me' (not *[vanambe]). The explanation for the blocked vowel rule in the perfective suffix /-ie/ (< *-ile) is the now (recently) deleted consonant. Other motivations include the avoidance of loss of paradigmatic marking function; note in this particular case that a coalesced perfective /-ie/ (viz., *[-e]) would be homophonous with the subjunctive affix, /-e/ [-e]. The following examples illustrate the failure of the operation of various rules due to a now deleted consonant:

Digo	mu(h)aso (not *mwaso) 'medicine'	< Proto-MK *mu̯aso
Maore	muomo (not *m(w)omo) 'mouth'	< *mulomo
Unguja	kioo (not *čoo) 'mirror'	< *-lol- 'look at'
Unguja	kienge (not *čenge) 'torch'	< *-Wèng- 'become red', but cf. mwenge 'torch')

Diachronically, this blocking is of interest because it is evidence that such C-deletion is relatively later than most Bantu vowel rules. However, there is a lot of irregularity about this blocking process within each language. For example, in Amu palatalization is blocked in /kievu/ 'chin' (*čevu), but vowel-gliding applies in the case of /mwaa/ 'dwarf palm' (cf. mulala El, UP, Gi), mwamu 'sibling in law' (cf. mulamu El, Gi), /mwan̯di/ 'bamboo' (cf. mulandzi Ng), etc. This is similar to the variation noted just above in Unguja (kienge ~ mwenge 'torch'). Vowel rules are often the type which tend to operate, whether phonologically or paradigmatically motivated, whenever the conditions for their operation are created, similar to homorganic nasalization in Class 9/10. This might explain why gliding applies in the above forms, while palatalization does not. The matter requires more study from a synchronic perspective (see Kutik 1983).

§10.4. Non-Class 5 nominals and verbs with V-initial stems. PSA (also CB) has a number of nouns, adjectives, and verbs with vowel-initial stems. The most common vowel involved is CB/PSA *į, and *i, but other vowels also occur. The reflexes of these items have not always retained the vowel. At this point we deal with such items, e.g., *-įkala 'sit, live', *-įkuta 'be satiated', *-įpika 'cook', *-ingila 'enter', *-ingin- 'other', *-ingi 'many', *-imba 'sing', *-įWa 'steal', *-įja 'come' (see Appendix 2 for others). No Sabaki language appears to treat these items with complete consistency, so we have to talk

of "tendencies" rather than absolute processes. Sabaki languages have a general tendency to delete the initial vowel in items originally having three (or more) syllables, but to keep it in items with two original syllables, and apparently for much the same reason that Class 5 monosyllabic nouns tend to keep an initial vowel in the prefix: deletion of the initial vowel would lead to a monosyllabic stem with attendant suprasegmental difficulties. This tendency can be illustrated by Lower Pokomo, where the reflexes of the above are /-kaa/ 'sit', /-kuha/ 'be satiated', /°-mbika/ (°mb) 'cook', /-njia/ 'enter', /-ngine/ 'other' with /i/ deleted, but /-inji/ 'many', /-imba/ 'sing', and /-iwa/ 'steal' where /i/ has not been deleted. The only exception is /-dza/ 'come', where one would expect retention of the vowel. Although there are clear syllable structural similarities between the initial vowel in Class 5 and in this set, there is also a difference; in Class 5 the *į plays a morphological role, which is not so in this set, where one or more other inflectional morphemes normally precede, leading to differential phonological adjustment.

Hence the geographic deletion of Class 5 *į- does not correspond to deletion of this vowel. A reflex of Class 5 *į- before consonant-initial polysyllabic stems is kept only in Elwana, Mwiini, Tikuu, and Siu, having been deleted in other Sabaki languages. For this verb-initial vowel, the two most conservative languages are Digo and Mwani. The former keeps the initial vowel in all the items above except in the reflex of *įkala (maybe a loan from Swahili?); Mwani retains /i/ in all of the items above except /-ja/ 'come' (but the original tone of the stem vowel in this word is kept on the syllable preceding /-ja/). SD, Mwiini, Elwana, Mijikenda, and Comorian largely keep the vowel, but tend to lose it from originally trisyllabic verbs. Pokomo and ND (not Mwiini) delete the vowel more widely, the latter (together with Mijikenda) even in some originally bisyllabic words (e.g., ND -ngi, MK -nji 'many').

In Comorian, before it was lost, the high vowel had the same strengthening effect on eligible consonants as in Class 5 and elsewhere (see §6), e.g., /-ka(y)a/ 'sit', /-kuha/ 'be satiated', /-piha/ 'cook' for the first three items above. In Mijikenda the initial vowel has been widely shifted to some other [-high, -back] vowel. Thus, taking Unguja as representative of other Sabaki, we find, e.g., /-inua/ 'lift', /-imba/ 'sing', /-ingia/ 'enter', /-eupe/ (*-elu+pe) 'white', also /-embamba/ 'narrow', but in Mijikenda (various dialects) /-enula ~ -anula/, /-emba/, /-angia/, /-aruhe/, /-ambamba/. Although all Mijikenda dialects shift the initial vowel extensively, they do not do it consistently from dialect to dialect or word to word. We do not know why this has happened.

§10.5. Mwiini and Comorian vowel-deletion. Mwiini attests vowel-deletion in a number of environments, a process which adds to the divergent appearance of the language: (a) Class 7 /či-/ has the allomorph /š-/ before Ç (e.g., /čiguwo/ 'cloth, rag' versus /škapu/ 'basket', /špande/ 'piece', /šṭana/ 'comb', but /čisima/ 'well', not *šsima, *šsiwa).[46] Similar deletion of /i/ following /s/ and /š/ and preceding Ç also occurs in non-

[46]Very similar forms of palatalization occur in Somali dialects presently or formerly adjacent, (see Nurse 1985b:249-250, Lamberti 1986a:246, 261).

Class 7 cases, e.g., /paŝpo/ 'without' (cf. pasipo Ung), /iŝkiɬo/ 'ear', /sku/ 'day' (cf. siku Ung), /masku/ 'night'. (b) In Class 8 there is /zi- ~ s-/ allomorphy, e.g., /zi-/ with M-stems and Ç-stems, and /s-/ with Ç̥-stems, thus /zili/ 'beds', /ziburi/ 'small sand dunes', /zintu/ 'things', but /spande/ 'pieces', /skapu/ 'baskets', /stenzi/ 'poems'. (c) In Class 15 there is /ku- ~ x-/ allomorphy, e.g., /kubiga/ 'to hit', /kugiiṯa/ 'to pull', /kuɬima/ 'to farm' versus /xpeɬeka/ 'to send', /xtuɬuka/ 'to fall', /xkahaṯa/ 'to hate', /xsooma/ 'to read'. This parallels Class 7 allomorphy and derives from vowel-devoicing, and has apparently been generalized to /u/ following other [+grave] fricatives, e.g., /mafta/ 'oil' (cf. mafuta Ung). (d) Class 11 has the following allomorphy: /lu- ~ l-/, e.g., /luti/ 'stick', /lgoongo/ 'midrib of palm leaf', /limi/ (*lulimi > llimi > limi) 'tongue', /lpeembe/ 'horn', /ltaambi/ 'wick' (PSA *lutambi).

In most of the above examples, vowel-deletion seems to have been triggered by vowel-devoicing in voiceless environments, thus /V/ > V̥ > Ø. This is clear in the case of Class 7, in Class 15 allomorphy, and elsewhere (/sku/ < /siku/). Class 8 is less clear; however, the shift of /zi-/ > /s-/ parallels the process in the singular. There we can assume a vowel devoices if bounded by two voiceless segments before undergoing full deletion; here the appropriate environment is the stem-initial Ç̥, with regressive assimilation spreading the feature [-voice] first to the vowel and then to the consonant before vowel-deletion, e.g., *zipande > zi̥pande > si̥pande > /spande/. This explanation is not adequate for Class 11, where /u/ deletes in voiced and voiceless environments, and for other cases of deletion in the environment of /l/, e.g., /čiβli/ 'shadow' (cf. kivuli Ung), /muβli/ 'man' (cf. muvuli Am), /-uβla/ 'kill' (PSA *-Wulag-), /-elpe/ 'white' (PSA *elu+pe). Here simplification of velarized sequences is likely in play, thus /lu/ or /ul/ > /l/, comparable to /l/-deletion after back vowels in some dialects of American English, e.g., /kawl/ > [kaw] 'cowl'. Devoicing-assimilation does not adequately explain every example of Mwiini vowel-loss, because metathesis or prothesis may underlie apparent deletion in some instances, e.g., /-iŝkiɬa/ 'descend' (cf. -šuka Ung), /-uskuṯila/ 'wash out mouth' (cf. -sukutua Ung), /-ihṯaja/ 'need' (cf. -hitaji Ung), /-ijṯihaḍi/ 'make an effort' (-jitihadi Ung); the latter two clearly do not support a devoicing hypothesis. Vowel-deletion of voiceless and voiced vowels in loans is pervasive where other ND and SD show vowels today, but this may reflect the source rather than a process in Mwiini, e.g., /barḍi/ (cf. baridi 'cold' Ung), /sarkali/ (cf. serikali 'government' Ung), etc.

Mwiini vowel-loss is reminiscent of a situation in Nzuani which (variably?) deletes a high vowel, e.g., /šisima/ [šsima] 'well', /šisiwa/ [šsiwa] 'island', /šikele/ [škele] 'coconut cut in two', /šitrandra/ [štrandra] 'bed'.[47] Whether this is further indication of historical relatedness, either genetic or of other sorts, is not clear. More likely than not, vowel-loss in Nzuani and Mwiini are cases of independent innovation, given the different conditions

[47]While Mwiini *V-loss undoubtedly stems from a devoicing process, there are instances where the proper conditions for devoicing are not met: /šibatsa/ [šbatsa] 'type of basket' and /šilavo/ [šlavo] 'oath'. We assume paradigmatic leveling is responsible for examples such as these.

under which vowel-deletion occurs in the various languages. Another phenomenon in Nzuani follows upon vowel-devoicing; sequences of /si/ before Ç have been deleted, e.g., /kio/ 'ear' (*įsįkilo), uku 'night' (*Wusįku), /mkiri/ 'mosque' (msikiti Ung).[48] The examples point to vowel-devoicing and simplification of devoiced sequences: /usiku/ > [usiku] > [usku] > /uku/.

§11.0. The Phonology of the Noun-class Prefixes

The nominal class prefixes as we have reconstructed them for PSA (see Chap. 4, §2.1.1) will be outlined in this section. The shapes before both C-initial and V-initial stems are given. Although, in most cases, certain synchronic phonological theories would allow for subsuming these variants under one form—e.g., the common alternation of /mu-/ and /mw-/ in many Bantu languages can be handled with a single underlying form and rules which derive the surface form appropriately: /mu/ → mu- / ___ [+C-stem]; /mu/ → mw- /___ [+V-stem]—we have not set up single reconstructed prefix shapes here. We have chosen to handle the alternations lexically rather than deriving the forms with phonological rules, which in nearly all cases were operating at a pre-PSA level and would thus play no role in either defining or subgrouping Sabaki.

Class	C-stem	V-stem	Section	Class	C-stem	V-stem	Section
1	*mu-	*mw-	11.1	2	*Wa-	*W-	11.2
3	*mu-	*mw-	11.1	4	*mi-	*m(i)-	11.3
5	*į-	*įj-	11.4	6	*ma-	*m(a)-	11.5
7	*ki-	*ky-	11.6	8	*vį-	*vy-	11.7
9/10	*N-	*ny-	11.8	11	*lu-	*lw-	11.9
12	*ka-	*k(a)-	11.10	14	*Wu-	*W-	11.11
15	*ku-	*kw-	11.12	16	*pa-	*p-	11.13
17	*ku-	*kw-	11.13	18	*mu-	*mw-	11.13

Certain changes which we have already discussed operate on these forms, e.g., Class 7 *ky- > /č/ by palatalization (§4.1.3.2), Class 8 *vy- > /ž/ (§5.3), nasal-loss and aspiration in Class 9/10 (§8.2.2, §8.4), Class 2 *Wa- > /a-/ (§4.2.1), Class 15 *kw > /kw/ [k͡p]/__V (labialvelarization (§7.1), etc. Other shapes are stable and undergo no change, e.g., Class 12 *ka-. In this section we will discuss those changes whose domain is most obvious in this area of the grammar, e.g., *mu-syncopation (Classes 1 and 3), *m-assimilation (Classes 1 and 3), and changes involving Class 5 nominal stems.

§11.1. Classes 1 and 3: *mu-. We have reconstructed *mu- for PSA Classes 1 and 3 prefixes with C-initial stems (*mw- with V-stems). Two changes affect this form: *mu-syncopation and *m-assimilation. The same changes affect Class 18 *mu-, and sequences of *mu in lexical items, e.g., *-lamuka 'wake up' (Ung -aṃka). The general forms of the correspondences are:

---

[48]We are indebted to G. Philippson for bringing this process to our attention.

*mu- /___ CVCV stems:

mʊ- ~ mu- ...	El (other variation also noted)
mu-	UP
m(u)-	LP, NMK, Com, Chi, Vu (?)
m̩-	SMK, SD, ND including Mw
N̩-	Si, Pa, Am, Jo, Pe, Mn
m(u)- ~ N̩-	Ti

Reports differ for Vumba. Lambert (1957, but based on field work in the 1920s) only shows /m/ < *mu; Bakari (1985:108) reports /mu/.

The following examples are helpful, but not fully reflective of variation and exceptions (for other examples see Appendix 2. Here we mark syllabicity, in Appendix 2 we do not):

*mufuko 'bag'	mʊfuko (El), m̩fuko (LP,Am,Mv,Ung,Vu,Chi,Ng)
*mucaWi 'witch'	mučawi (El), mučaβi (UP), mutsai (Gi,Du), mut̪avi ~ n̪t̪avi (Ti), n̪t̪aʊi (Si), m̩čawi (Ung), mučawi ~ m̩čawi (Chi), n̩sawi (Mn)
*muvyazi̧ 'parent'	m̩vyazi (LP,Vu,Mak), mužazi (Gi), m̩zazi (Ung), m̩vyazi (Vu), m̩zaazi (Mw), m̩ðaði ~ n̩ðaði (Ti), n̩zazi (Jo), m̩dzadze 'mother' (Ng), m̩dzade (Ng)
*mukono 'hand'	mʊkɔ́nɔ ~ mɔ́kɔ́nɔ (El), mukono (UP,LP,Gi,Ch,Ra,Di,Du), m̩kono (Di,Mw,Am,Mv,Ung,Mak,Tu,Chi,Vu,Mt), n̩kono (Ti,Pa,Si,Jo,Pe, Mn), °k^{h}ongo 9/10 (°ng ¬ n) (Vu), mhono (Ng,Nz), mhono ~ muhono (Mh,Ma)
*muntu 'person'	mu̧du ~ móòdu (El), muntu (LP,UP), mut^{h}u (Gi,Ch,Du,Di), m̩t^{h}u ~ n̩t^{h}u (Jo), m̩t^{h}u (Am,Mv,Ung,Vu,Mak,Tu,Mt,Vu), m̩t^{h}u ~ mut^{h}u (Chi), muuntu (Mw), n̩č^{h}u (Pa,Si,Ti), n̩t^{h}u (Jo,Pe), m̩ndru [m̩dru] (Ng), muntru (Nz), m(u)tru (Ma), munu (Mn)
*muti 'tree'	móṯɪ (El), mu̯ri (UP), muhi (LP,Gi,Ra,Di,Du), muti (Mw,Mn), n̩či (Ti,Si,Pa), m̩ti (Am,Mv,Ung), n̩ti (Nga,Jo), m̩ri ~ muri (Chi), m̩ɟiti (Tu), m̩ɟiti/miɟiti ~ uɟiti/miɟiti (Mak), n̩ti/miti (Pe), m̩ri (Vu), m̩ri [m̩dri]/miri (Ng), muri (Mh), mwiri (Nz,Ma)

The vowel-loss affecting [u] in these items is due to a quite natural and widely attested process in Bantu languages. The change is underlyingly assimilatory in nature and founded on the collocation of two labial segments, *m and *u. In assimilating to the nasal element, the syllabicity of [u] is transferred to the nasal, thus preserving syllabicity constraints. In Bantu languages where there is such feature agreement in prefixes there is a tendency for one of the elements to delete or be assimilated by the other. Contrast, for ex-

ample, the behavior of *mu- with *mi-, the Class 4 nominal prefix, which rarely is affected by change (Hinnebusch 1973:21ff.).

Syncopation sets up the proper conditions for a further assimilation in which the syllabic prefix assimilates the features of the following consonant:

*mu- > m̩ > N̩-
[+syllabic]

The final step of this process should not be confused with N-assimilation in Class 9/10 and similar environments, e.g., *ni '1sg subject marker', etc. which is universally attested in Bantu. Although one might wish to argue for a synchronic operation of Class 9/10 N-assimilation and describe both using a conflated rule formalism, the two processes are quite distinct both chronologically and in their end results. M-assimilation is characterized by a homorganic [+syllabic] nasal, N-assimilation by a homorganic [-syllabic] nasal which deletes in some languages, and which fuses in most languages with a following voiced consonant. N-assimilation is pre-PSA, M-assimilation is post-PSA. The only synchronic connection we might want to suppose for the two is where N-assimilation applies to cases of *ni which often synchronically syncopate and assimilate (for examples see Appendix 1 and Nurse 1982b:83-84).

The languages which most regularly attest syncopation are Swahili, Southern Mijikenda (Digo), and Mwani, although individual languages vary somewhat in the details. Of these Tikuu, Siu, Pate, and Amu (ND_3), Jomvu, (Chifundi), Pemba, and Mwani go the final step and add m̩-assimilation, a variable rule. At the other end of the scale are Elwana and Upper Pokomo, which show no syncopation.

Between these are Lower Pokomo, Comorian, and Northern Mijikenda, which have variable syncopation. In Comorian, Lower Pokomo, and Northern Mijikenda (?), the syncopated form [m̩] is normal in regular or rapid discourse, but the full form [mu] can be heard in slower or more formal speech and/or with older speakers. Thus several sources for Comorian phonemicize the Class 1/3 as /mu-/: Rombi (1983) notes Class 1 /mu-/ is realized as [m̩] in rapid discourse; Sibertin-blanc (1980) records both Class 1 and Class 3 examples for Maore with a syncopated prefix, e.g., /m̩ri/ 'tree', /m̩dru/ 'arm'. The analysis was made somewhat more difficult for us in general by normal orthographic practices obtaining in these languages, which do not always correspond to the actual phonetic situation: it appears normal to write the full form in Northern Mijikenda and Comorian, while Lower Pokomo orthographic practice seems to vary between /mu-/ and /m-/.

The contexts for change can be seen by considering these "intermediate" languages. Lower Pokomo, for example, syncopates regularly when the initial-stem consonant is a labial, e.g., *mufupa > /m̩fuɸa/ 'bone', *mufuko > /m̩fuko/ 'bag' versus *mukono > /mukono/. Also, speakers of Lower Pokomo and Giryama pronounce /mu-/ with a somewhat devocalized vowel and somewhat rounded nasal, which suggests that the conditions are there for full syncopation. Both situations give some insight into the evolution of the change: devocalization of the vowel, followed by deletion beginning with labial-initial

polysyllabic stems. The second step is also noted by Bakari (1985:57-58) for older Chifundi speakers, who generally have /mu-/ except with stems beginning with labials, e.g., /mfuβa/ 'bone', /mvuvi/ 'fisherman', /mpaka/ 'boundary', /mboni/ 'castor oil tree', but /mučumba/ 'fiancee', /muremi/ 'leader', /mučawi/ 'witch', etc. Younger speakers have generalized this to other CVCV-stems but not to monosyllabic stems, probably under influence from standard varieties of Swahili, or possibly Digo, where syncopation in both cases is the rule.

In syncopating languages, some contexts are exempt. Mijikenda, Mwiini, and Mwani generally exempt monosyllables, e.g., Southern Mijikenda (Digo) /mutsi/ 'daytime', /munda/ 'field', /muβya/ 'new' Classes 1 and 3, /mufu/ 'hard' Classes 1 and 3, /mut^{h}u/ 'person' (but [mt^{h}u] was also recorded). In the third and final step of the process, this constraint is given up and even /mu-/ with monosyllabic stems is relaxed, as it has already in most Swahili dialects. In Unguja, however, there is an interesting exception wherein /mu-/ is preserved before /w, y, h/:

Unguja

muwa 3 'sugarcane' muwi 3 'mangrove sp.' muwa 3 'palm tree sp.'
muyombo 3 'tree sp.' muhangata 3 'mninga tree' muhogo 3 'cassava'
muhindi 3 'corn plant'

Ngare/Jomvu

muyahudi 1 'Jew' muhunzi 1 'smith'

These were recorded by F. Johnson (1939a) for Unguja, and Lambert (1958a) for Ngare and Jomvu. Historically, *mu-syncopation did not operate before [-voc, -cons] segments. However, such forms are apparently slowly losing ground in favor of the predominant paradigm, as synchronic doublets indicate: /muyombo ~ myombo/, /muyahudi ~ myahudi/, /mhindi ~ muhindi/. A similar phenomenon is noted in Digo, which otherwise syncopates, e.g., /muhaso/ 'medicine'.

Where M-assimilation occurs, it is generally optional in the sense that other variables play a role in its actuation. Eastman (1967) discusses M-assimilation in Amu, Jomvu, and Tikuu. She indicates that in Amu the change is optional and only affects Class 1, it affects both Class 1 and 3 in Tikuu, and in Jomvu is pervasive. Nurse (1982d:81-82) points out that in Amu consonant assimilation is frequent in normal speech, rare in formal speech, and never indicated in writing, e.g., [haŋkuliwaye] < /hamkuliwaye/ 'what's your name?', [ŋkate] < /mkate/ 'bread', [nti] < /mti/ 'tree'.

*Mu-syncopation has wide distribution in eastern and southern Bantu. All NEC except Gogo and Kagulu (Tucker and Bryan 1957) attest the shift. Such distribution normally would indicate early innovation, and inheritance from NEC, but the fact that it cuts right through Sabaki, plus the range of variation where it is attested, indicates that it innovated rather late in Sabaki. Digo is a good case for illustrating the point. Here *mu-syncopation must be very recent. It is regularly attested with polysyllabic stems, but usually not with monosyllabic stems. It is not attested in Giryama, Duruma, or Chonyi, and so post-

dates the breakup of Mijikenda. Native speakers have reported that older people are heard to use nonsyncopated forms. It also postdates Mijikenda/Pokomo *t-lenition, since *mu- does not syncopate before /h/ < *t, e.g., /mu-haso ~ mu-aso/ 'medicine', /muhama ~ muama/ 'millet'. If *mu-syncopation applied earlier than *t-lenition, we would expect to have syncopated forms here as well.

Variation in other dialects reinforces this conclusion. Nurse (1982b:81-82) reports that in ND, *mu-syncopation is the norm, but /mu-/ can still be heard in speech, especially with M-stems. It is common in older writing, and still occasionally used in verse. Specifically for Tikuu, speakers along the mainland coast syncopate more than speakers on northern Pate Island, and younger speakers are less conservative than older ones. *Mu-syncopation is still a functioning Sabaki change, whether internally motivated or as an areal feature under the pressure of standard forms of Swahili.

As to the source of *mu-syncopation, it is likely to have moved in a northerly direction spreading from NEC, probably first to Mwani and SD, where the change is most complete, and then to ND, and more recently to Southern Mijikenda (Digo). Its presence in Comorian may well be convergence, or a shared innovation with SD/ND. If the latter, then it would suggest that Comorian was in close association with SD/ND quite late in the Sabaki period. However, Bell (1970) holds the opinion that, as a very naturally motivated change, *mu-syncopation, where attested, represents instances of parallel yet independent change. Moreover, because *mu-syncopation entails loss it is of minimal value in subgrouping.

§11.2. Class 2: *Wa-

*Wa-/__C-stem	wa-	El, SD, Mn, Com
	ʊa-	UP, Am, Mv, etc. (Bakari 1985)
	va-	Ti
	a-	MK
*Wantu 'people'	wantru (Nz), wat^{h}u (Ung), ʊač^{h}u, (Pa,Si), vač^{h}u (Ti), wandru (Ng), watru (Ma), at^{h}u (Gi)	

Where this morpheme occurs with V-stems, deletion or coalescence of the vowel is usual:

*Wa-/__V-stem	w(a)- ~ ʊ(a)	Sabaki
	a-	MK
*Waana 'children'	wana (Ung,Ng,etc.), ʊana (Am), vana (Ti), ana (Gi)	
*Waivi 'thieves'	wevi (Ung), ʊezi (Am), viði (Ti), aivi (Gi), aevi (Ch), wedzi (Ng,Nz,Ma), etc.	

The form of the prefix here is governed by general processes affecting *W (see §4.2.1.1), and various vowel rules.

§11.3. Class 4: *mi-

*mi-/__C-stems	mi-	Sabaki
*mi-/__V-stem	m(i)-	Sabaki
	ny-	ND
	mi- ~ my- ~ ny-	Jo, Gi (?)
*miezi 'months'	miezi (LP,Gi,Ung), mezi (Vu), nyezi (Am), nyeði (Ti), me(e)zi (Ng,Nz,Ma)	
*mioyo 'hearts'	mioyo ~ myoyo ~ nyoyo (Jo)	

Before V-stems, /mi-/ is quite stable in not gliding, in contrast to the singular, e.g., /mwaka, miaka/ 'year' (LP, Gi, Ung), /mwezi, miezi/ 'month' (LP, Gi, Ung, Ng); however, note /mungu, myungu/ 'God/gods' (Di), /mwewe, myewe/ 'hawk sp.' (Di), /mweko, myeko/ 'spoon' (Di). The form /ny-/ derives from an intermediate *my, as attested by the variation in Jomvu, and is a general process in all ND, e.g., /mwamba, nyamba/ 'rock', /-zinya/ 'extinguish' < *-zimya, /nṯonyi/ 'fisherman' < *-com-i̧- 'pierce' (agentive). According to Stigand (1915) this shift does not occur in Amu with nominals of plants or their derivatives, e.g., /miembe/ 'mango trees', /miavuli/ 'umbrellas'. Also elicited for Giryama (TJH 1972-73 field notes) were a few items of a similar sort, but in variation with /mi-/, e.g., /nyaka ~ miaka/ 'years', /nyezi ~ /miezi/ 'months'. These are likely to be borrowings, rather than remnants of more widely attested change, but as we have not systematically investigated other Mijikenda dialects, we cannot be sure of our conclusions here.

§11.4. Class 5. The behavior of Class 5 prefixes with consonant-initial stems throughout NEC and Sabaki is fairly regular, and easily analyzed by a small set of rules. This is not the situation with monosyllabic (M-) stems or vowel (V-) stems. From language to language throughout NEC we find a diverse and complicated synchronic situation which is the result of the application and interaction of both straightforward morphophonological processes and not so straightforward, analogical restructurings. Analogy is functioning to level out irregularities under paradigmatic pressure from a number of different sources: (a) the overall singular/plural nominal pattern and rules for plural formation in both Classes 5 and 6 and all other classes, (b) the predominant shape of Class 5 polysyllabic nominals, and (c) the verbal concord *li- (with [-HV]). It has been nearly impossible to account for every last detail or exception. It is not possible to distinguish the source of change or variation, to identify the analogical model, or to be certain in all cases about the point, or level, at which something innovated. Analogy can operate in terms of multiple and conflicting models, and independently from one language to the next, as well as begin innovating at some early proto period and carried down to reflex languages in somewhat irregular ways.

An example of analogy is seen in Tikuu's treatment of CB *-bú̧ 'ash(es)', with an identifiable prefix and root, /i-vu, ma-vu/, versus Digo /ivu, ma-ivu/, where the prefix is

reanalyzed as part of the stem. Tikuu still preserves an identifiable, functioning Class 5 /i-/, but Digo does not, and has restructured the form to conform to the predominant C-stem pattern. Note also Unguja /ɟiko, meko/ (V-stem) but Lower Pokomo /dziko, madziko/ 'fireplace'; Nata /įrįoβa/ (V-stem) 'sun' but Giryama /dzuwa, madzuwa/; Maŋanja /dzina, maina/ (V-stem) 'name' but Unguja /ɟina, maɟina/. That these reformation processes are still functioning is indicated, for example, by generational variation, e.g., Lower Pokomo /iǰi, maiǰi ~ iǰi, maǰi/ 'egg', where the younger generation no longer identifies the initial /i-/ as a prefix.

Despite the difficulties, we attempt to sketch a general outline of events in this section. For additional discussion we refer the reader to two earlier and somewhat more detailed, but not necessarily more accurate, efforts at analysis (Hinnebusch 1973 and Hinnebusch, Nurse, and Mould 1981).

§11.4.1. Class 5 C-stems. We have reconstructed the preconsonantal prefix as *į- with C-initial polysyllabic stems. A high front vowel is regularly attested in present-day Sabaki languages only in Elwana, Mwiini, Tikuu, and Siu. The [HV] feature of the vowel is required by the consonantal changes seen in Comorian and Swahili strengthening (§5.4). There is also plenty of comparative evidence outside Sabaki to support this reconstruction. Meeussen (1967:97-99) reconstructs the Bantu nominal Class 5 prefix as *į- ([+HV], the verbal as *di- ([-HV]), and suggests that there is evidence that both derived from *dį-, a hypothesis strongly supported by Guthrie's Comparative Series for this class (Guthrie 1967-71, vol. 3:221-222; C.S. 2204a-d). It is yet unclear at what point *lį- (= Meeussen's and Guthrie's *dį-) lost its consonant in the history of eastern Bantu. There is, however, little evidence for a CV- shape for the polysyllabic C-stem prefix in Sabaki. A single vowel, as a prefix, is the norm for a large number of non-Sabaki eastern Bantu languages, thus indicating that the loss of the consonantal element happened early at some pre-NEC level. Most NEC languages also attest Class 5 *į-deletion, having only Ø prefix with polysyllabic C-stems (e.g., Seuta), but in Sabaki we assume *į-loss is a post-PSA development, given the direct and indirect evidence in widely distributed Sabaki languages, as supported by the following examples:

*į-	i-	El[49], Mw, Ti, Si
	Ø + Strengthening	Com, SD (see §5.3 for Strengthening)
	Ø	Po, MK, Mn, ND except Mw, Ti, Si
	i-	general in Sabaki M-stems (see below); in regular stems in El, Mw, Ti, Si

*įpala 'roof' ipaa (Ti), paa (Am,Pa,Ung), paya (Pe), βaa (Vu), paa/maβaa (Ng), pa/mava 'kitchen', pa/maβa 'house' (Ma)

[49]Our sources for Elwana which distinguish more than 5 phonetic vowels usually record the Class 5 prefix as [I]; Maddieson and Sands do so consistently, and Möhlig usually does so as well.

*ɻtumbo 'belly'	ɻumbo (UP), humbo (LP), ituumbo (Mw), ičumbo (Ti,Si), čumbo (Pa), tumbo (Am,Mv,Jo,Ung,Mt,Pe,Tu,Mak,Mn), rumbo (Chi,Vu), trumbo/marumbo (Ng), trimbo/marimbo (Nz)
*ɻWingu 'cloud'	iβingu ~ iwingu (El), βingu (UP), ʋingu (Pa,Am,Mv), wingu (LP,Chi,Vu,Mt,Pe,Tu,Mak,Mn), ingu (Gi,Ch,Ra, Di,Du), iʋiingu (Mw), iʋingu (Si), ivingu (Ti), bingu/ maingu (Ng), vingu/maingu (Nz), βingu ~ bingu ~ wingu/maingu (Ma)
*ɻWiWi 'rubbish'	biwi (Ung), bii (Ng), ʋiʋi (Am), -wiwi (LP)
*ɻziɻWa 'lake'	ízówa ~ ízówa (El), ziʋa (Am), ziwa (LP,Ung), zia (Gi), ziya (Du), dzia/mazia and idziwa ~ hiziwa 7/8 (Ng), dzia (Nz)

The post-PSA processes needed to derive these shapes are *ɻ-loss, Strengthening, and where *ɻ- is retained, its neutralization to /i-/ (7V > 5V reduction). The following derivations are illustrative:

Tikuu	**Swahili**	**Comorian**	
*ɻtumbo	*ɻWiWi	*ɻziɻWa	PSA
--	ɻbiWi	ɻdziɻWa	Strengthening
itumbo	ibiWi	idziWa	7V > 5V
--	ØbiWi	ØdziWa	*ɻ-loss
ičumbo	biwi	dzia	Other rules
ičumbo	biwi	dzia	Output

There are no subgrouping implications attached to these facts, since we are dealing with retention, in the case of ND, and loss in the others. However, strengthening in SD and Comorian attest to the presence of *ɻ- in those languages at a post-PSA period. This is of interest as it does link part of ND, SD, and Comorian by virtue of retaining a historical form. The loss we see elsewhere could be shared innovation, as well as an areal effect which involved other NEC languages, but because it involves loss its utility for subgrouping is limited.

§11.4.2. Monosyllabic stems. Monosyllabic stems (M-stems) provide further evidence for retention of a vowel prefix (i- < *ɻ) as it is preserved in most Sabaki dialects. Also, and most important, we find evidence here of CV- prefix shapes; these are apparent in the examples below (see Nurse 1982b for others):

*ɻvu 'ash(es)'	ívu (El), ivu 9/10 (Mw), ivu (UP,LP,Gi,Ch,Ra,Du,Di,Ti,Pa, Si,Chi,Mt), ɟivu (Ung,Jo,Tu), ɟivu ~ ivu (Mv,Pe), ipu ~ ibu (Mak), ɟivu ~ °livu (°l → Ø) (Vu), vuu (Ng), vu ~ jifu (Nz), vu (Mh,Ma), rivu (Mn). (CB *-bʉ́ɻ)
*ɻpi 'slap, palm'	iɸi (UP), ihi (Gi), ipi (Mw,Am,Ti,Pa,Si), pi/maβi (Ng), pii/mavii (Ma). (CB *-pí)

*ɟiWe 'stone'	°iɟíwe (°ɟ ¬ z) (El), ǰiwe (UP), iwe ~ ɟiwe (LP), dziwe (Du,Ka, Ra,Di), iʋe ~ ɟiʋe (Am), iwe (Gi,Ji,Ch), iɟiʋe (Mw,Si), ɟiʋe (Jo, Mv), iɟiʋe ~ iʋe (Ti), °ɟiwe (°w ¬ b) (Ung), bwe ~ ɟibwe ~ njiɟibwe (Mak), bwe (Tu), ibwe ~ ɟibwe ~ °iwe ~ °ɟiwe (Pe), ɟibwe (Vu,Mt,Chi), bwe/mawe (Ng,Mh,Nz,Ma), riwe (Mn). (CB *-bùè)
*ɟipu 'boil'	iɸu (UP), ihu (Gi), ipu/mapu (Am,Ti), ipu/maipu ~ mapu (Pa), ɟipu/maɟipu (Ung), ipu (Mt,Mak), liβu (Vu)
*ɟigi 'egg'	igi/magi (UP), iǰi (LP), iǰi/maǰi (Gi,Ra), °ingi/maingi (°n) (Ti), i ~ ii/mai (Am,Pa,Vu), igi ~ gi/magi (Pe), ri (Mn). (CB *-gí)
*ɟifu 'stomach'	ifu (Gi), difu 'bird's crop' (Gi), ifu 'coconut center' (Am), ifu 'cow's stomach' (Vu), difu/ma- 'gizzard' (Mt). (CB *-pù̱)

Two rough generalizations can be made about the data in the above display: the vowel prefix is a predominant feature of monosyllabic stems, and a CV- prefix (li-, ri-, or *ji-) also occurs. CV- prefixes tend to be found in SD and Mwani, e.g., /ɟi-we/, /ɟi-vu/, /ɟi-pu/ (Ung); /li-vu/, /ɟi-bwe/, /li-βu/ (Vu), and /ri-vu/, /ri-we/, /ri-ni/ (Mn). V-, as prefix, occurs elsewhere, e.g., /i-vu/, /i-we/, /i-hu/, /i-ni/, /i-ǰi/, /i-fu/ (all MK), but in each case there are some exceptions, e.g., /i-ni/ (Ung), /dzi-we/ (SMK), /i-ji-we/ (Ti). The analytical problem is the interpretation of this distribution.

A number of hypotheses are possible. We assume that a CV- prefix is a retention from a pre-PSA period, deriving from *lɟ. We know that *ɟ is necessary to explain strengthening in Class 5 environments. We also know that nominal prefixes tend to be maintained with M-stems when they are phonologically deleted elsewhere (e.g., Class 9/10 in Unguja /nṭʰa/ 'wax' versus /tʰaa/ 'lamp'), so it would not be unexpected to find inherited M-stems with CV- reflexes of *lɟi-, unaffected by NEC *l-loss which created PSA *ɟi-CVCV shapes—the great majority of Class 5 nominals. The main problem with this is that M-stems in other NEC languages are fairly uniform in having just /i-/, the few exceptions most likely being loans from Swahili.

Another solution involves analogy. The irregularity we see here, especially but not exclusively in the Swahili dialects, is easier to understand if we further assume that monosyllabic stems and V-stems (see examples below) at an NEC period, or at an early PSA level at least, looked like polysyllabic stems and were of the form /V-CVCV/ (e.g., *ɟi-gana 'hundred'): thus, NEC *ɟi-lɟiWe (or *ɟi-zɟiWe) 'stone'. This would fit in with the analysis of V-stems as well, and has some support in the form of possible remnant items, e.g., Tikuu /iɟiǰʰu/ 'giant', /iɟiwe/ 'stone'. Such reanalysis is analogical in motivation and creates uniform Class 5 nominals. The process is also attested in Saghala, Tuβeta, and Chaga, e.g., Tuβeta /iziso, meso/ 'eye/s', /ijani, mani/ 'leaf/leaves', in conformity to the predominant pattern /i-CVCV/.

The pre-PSA phonological rule that created *į-CVCV stems did not affect the CV prefix with monosyllabic stems, nor the prefix with V-stems, setting up the following Class 5 paradigm (etymological noun-class prefixes are asterisked):

Singular		**Plural**
*V-CVCV	V-prefix + polysyllabic C-stem	ma-CVCV (*V replaced by *ma)
*CV-CV	CV-prefix + monosyllabic stem	ma-*CV-CV
*CV-VCV	CV-prefix + polysyllabic V-stem	ma-*CV-(V)CV

Once the CV- prefixes on V- and M-stems were no longer felt by speakers to be prefixes but part of the stems, the forms were then remodeled after the most frequently occurring pattern, namely, V-CVCV/ma-CVCV singular and plural stems, to give:

Ø-*CV-CV (M-stems) > *V-CVCV / ma-CVCV
Ø-*CV-VCV (V-stems) > *V-CV(V)CV / ma-CV(V)CV

In the case of V-stems, the initial stem vowel often interacted with the prefix vowel and either syncopation or coalescence resulted.

A possible etymon that might explain the Sabaki M-stem forms is *(į)lį-. A form with *l, not *j as for V-stems, more easily accounts for reflexes with /l/ and /r/. A form with *j, or just *į- without a consonant, however, would be appropriate if /l/ and /r/ can be explained by other than a morphophonological mechanism. Consider a derivation with *(į)lį- as the reconstruction:

*įlį- > (zi-?) > ʒi-	Spirantizing (essentially Ung), e.g., ʒi-we
> li-, ri-	7V > 5V (other SD, Mn), e.g., ri-we (Mn), li-βu (Vu)[50]
> i-	*l-loss, etc. (mainly Po, MK, ND), e.g., i-we (Gi)
> Ø	*i-(or į-)-loss (Com), e.g., bwe (Com)

Some of these processes are also needed to derive some NEC forms, e.g., 'stone': /libwe, mabwe/ (Luguru), /dibwe/ (Kwere, Kami, Kutu, Sagara), /jiwe/ (Zaramo), /iwe/ (Doe); 'ashes': /ivu/ (Doe, Kwere), /livu/ (Luguru). Their distribution indicates early innovation in NEC, but not uniform nor simultaneous application in Sabaki in these few M-stems. Their differential behavior is explained by the fact that, even though functioning in NEC, they did not uniformly affect PSA early enough to give us regular synchronic reflexes throughout. PSA already had to have been a dialect cluster. Thus rule reordering and non-application of other rules undoubtedly played a role in the derivational development. For instance, we have to assume that Spirantization, which still operated at an early PSA period, was suspended in the case of those languages which still show a nonspirantized, nonpalatalized reflex, namely all but Unguja.

Several alternatives to the above scenario are also possible:

---

[50]It is likely that Mwani /ri-/ is a loan from neighboring Zone P languages, and thus not helpful in reconstruction.

1. The shapes /li-, ri-/, etc. (not ɟi-) derive from verbal *li- with [-HV]. Pokomo and Mijikenda use a form identical in shape to the verbal form of the Class 5 prefix in marking Class 5 quantifiers, and certain adjectival stems (usually M-stems, but not always):

Pokomo: diɸya 'new', dingine 'other', dyumu 'hard', dyalusi 'black', di- 'verbal prefix'

Giryama: rire 'long', ridzo 'good', rii 'bad', raruhe 'white', ri- 'verbal prefix'

This resembles the sometimes variable use of the verbal prefix in certain Class 5 contexts in Swahili:

Unguja: lenye 'having', lingine ~ ɟingine 'other', leupe ~ ɟeupe (Sacleux 1909) 'white', etc.

Tikuu: lororo 'soft', lembamba 'thin' (Sacleux 1909)

Amu: yororo ~ lororo, yembamba ~ lembamba (Sacleux 1909)

The presence of /li-/ in rural SD (and less likely /ri-/ in Mwani) could well be due to a morphological and/or analogical marking process involving the copying of verbal concord, and not a retention from an earlier proto level.

2. The assumption of a phonological process in which *i̧li̧- yields *i̧ji̧- (possibly via *i̧zi̧-) is what we assume happened in the case of V-stems at some point prior to or early in the development of Sabaki (see §11.4.3). It is not unlike the Comorian shift where CB *d : dz/i̧ ___ i̧ and CB *d : dz/___ i̧V, e.g., /-dziha/ 'bury' (CB *-di̧i̧k-) (see §5.3); except in Unguja the process would be specific to Class 5 environments. However, Unguja /ɟi-/ on M-stems probably has nothing to do with phonology and may simply stem from analogical processes and be modeled after V-stems (see §11.4.3) and/or the regular use of /ɟi-/ to mark augmentatives (see below §11.4.4).

3. Polomé (1967:74-75) suggests *gi- as an analogical model for /ɟi-/ (Ung). Meinhof (1932:118-119) suggests *ɣi̧-. Both are reasonable hypotheses, in that velars often yield palatals before high vowels, but there is no direct comparative evidence that we are aware of that supports either proposal (however, see discussion of augmentatives in §11.4.4).

In Appendix 2 we list Class 5 M-stems with a reconstructed *i̧(CV)- prefix. The CV portion within brackets represents the possibility that some other source and/or mechanism is responsible for synchronic forms, apart from straightforward morphophonological processes. This caveat reflects the indeterminacy of our analysis. For this reason, and the fact that so few forms are involved, we make no claim about subgrouping on the basis of how M-stems behave.

§11.4.3. Vowel-initial stems. In etymological V-stems, as in M-stems, analogy masks etymology. Thus, stems which Guthrie gives as V-stems (his *y-stems, e.g., CB *-yícò 'eye') can be treated as monosyllabic stems, e.g., /ɟiɻo/maɻo/ (Mv). Sometimes the singular prefix appears to follow the pattern of prefixes with M-stems, and other times with V-stems. Reformation of such stems, for the same reasons discussed earlier for M-stems,

makes it difficult to be unequivocally clear about what has happened historically in terms of both the reconstructed shape of the prefix and the phonological processes involved in deriving reflex forms.

We have reconstructed the prefix as *ɟjɟ- for Class 5 V-stems. The correspondences are:

*ɟjɟ- :	ɟi-	Ung, Mv, Pe, Vu, Mt
	ǰi-	UP
	dzi-	MK, Com, (LP)
	izi-	El
	yi- ~ Ø	Mw, Am, Si, Pa
	yi- ~ ži ~ Ø	Ti
	°ri-	Mn
	zi- ~ ri- ~ di-	other SD which can also attest /ɟi-/

*ɟjɟco 'eye' ɪzîčɔ (El), ɟičo (UP,Ung,Vu,Mt), ɟičo ~ zičo (Pe), dzitso (LP, Gi,Ch,Ra,Du,Di,Ng,Mh,Nz,Ma), iʦo (Mw,Am,Si,Pa), yiʦo ~ žiʦo (Ti), ɟiʦo (Mv,Jo,Chi), ɟičo ~ dičo (Mak), dičo (Tu), °liso (Maf), °riso (Mn). (CB *-yɟcò)

*ɟjɟko 'fireplace' meko (El), eko (LP), dziko (Du), iɟiiko/miiko ~ maɟiiko 'kitchen' (Mw), °yeko (Ti), iko (Am), ɟiko (Mv,Ung,Mak), ziko/maziko (Pe), dziho (Ng). CB *yɟkò

*ɟjɟna 'name' ɪzina ~ ízínà (El), ǰina (UP), ɟina (Mv,Jo,Ung,Chi,Vu,Mt,Tu, Mak), dzina (Gi,Du,Ra,Du,Di,Ng,Mh,Nz,Ma), ina (Mw,Pa, Si,Am), yina ~ žina (Ti), ɟina ~ zina (Pe), °zina (°z ¬ j) (Mn). (CB *-yɟnà)

*ɟjɟno 'tooth' dzino (Gi,Ch,Du,Di), yiino (Mw), yenyo (Ti), yeno (Si), ino (Am,Pa), ɟino (Mv,Jo,Ung,Vu,Chi,Mt), ɟino ~ gino (Mak), gino (Tu), zino ~ ɟino ~ gino (Pe), dzinyo (Ng,Mh,Nz,Ma), °rino (Mn, °ri- < Zone P languages), °lino (Mgao, °li- < Zone P). (CB *-yɟnò)

The initial vowel is reconstructed for the same reason we need an initial vowel for the prefix with M-stems, and moreover is supported by direct evidence: /izičo/ 'eye' (El), /iɟiko/ 'fireplace' (Mw). It also triggers strengthening in Comorian: /dzinyo/ 'tooth'. Our reconstruction of *j, however, is problematic. On the one hand, it is well supported by Ung (ɟi- < *jɟ-), Mijikenda (dzi- < *jɟ-), ND (yi- ~ ži- < *jɟ-), and possibly Comorian (dzi- < *ɟjɟ- or *ɟzɟ-). It is uncertain, however, how some forms in other dialects fit this hypothesis. Forms in Mwani, other SD, and Elwana suggest alternative analyses. The Unguja, ND, and Mijikenda forms are derived as follows:

*ịjị	> ịʄị-	Sw (retained as /ʄi-/)
	> dzi-	MK, LP (Sw /ʄ/ : MK /dz/, see §4.2.4)
	> dzi-	Com (§5.1.3)
	> ž > y > Ø	ND (see §4.2.4)
	> z	Pe

Of the Pokomo dialects, only Lower Pokomo participates in this schema, e.g., /dzitso/ 'eye', Upper Pokomo retaining a palatal affricate. Elwana may fit as well, if we assume as we did earlier (§4.2.4.1) that the shift of PSA *j > /z/ (e.g , *-ịja 'come' > -iza) is an Elwana post-PSA event, but the fact remains that Elwana forms (e.g., /izičo/ 'eye') are similar to Seuta and Ruvu forms; this may indicate a retention from a PNEC period with *z. Seuta and Ruvu are reconstructable with *(ị)zị-.

The source of Comorian /dz/ is different from its source in Mijikenda, where any *j shifts to /dz/. In Comorian, PSA *j generally corresponds to /ǰ ~ ž/, depending on dialect, whereas /dz/ derives from PSA *z before *y, *ịV, and between two front [+HV] vowels, e.g., *-zịịka 'bury' > /-dziha/; *ịzịma 'whole 5' > /-dzima/; *ịzịWa 'milk' > /dziwa/ (see §4, also §6.3). Earlier (§6.3) we claimed that this rule also applies to *j between two front [+HV] vowels; we make this claim because of the overall correspondence relationships in Sabaki, namely, /dzitso/ (Com) : /ʄičo/ (Ung) : /yiṭo ~ žiṭo/ (Ti) : /dzitso/ 'eye' (Mk/Po). For Comorian /dz/ to result from a neutralization of *z and *j in these environments is not unusual; nevertheless, there are reasons for questioning it. Comorian /dz/ (PSA *j) may just as easily point to NEC *z, or to an internal Comorian shift in which *z/ị __ ị > /dz/ (see *z-affrication §6.3): NEC *ịzịno > ịdzịno > dzịno > Comorian /dzinyo/ 'tooth' (cf. Sagala, G.39, /idzino, mazino ~ meno/ 'tooth', example from Johnston 1919), or PSA *ịjị-ịno > Comorian *ịzị-ịno > idzi-nyo > /dzinyo/. Either analysis would work, but with different subgrouping implications. The former places Comorian outside of Sabaki, the latter does not. Unfortunately, we do not have good direct evidence to make a choice between the two. However, there is an interesting situation in Pare/Tußeta which might help (see just below).

The forms with /zi, ri, di/ in rural SD also complicate the analysis. Pemba /zi-/, e.g., /zino/ 'tooth', could be a borrowing, or might attest earlier NEC *z, or be a remnant from a general Sabaki shift of *j > /z/ (see Comorian immediately above), whose effects were later undone by other change. Makunduchi /di-/, e.g., /dičo/ 'eye', could be a stray strengthened form of *l or *j. Mwani /ri-/ could reflect similar possibilities: an earlier NEC *lị-, loaning from Bantu P20-30, analogical spreading of *li- verbal concord, etc. It is difficult if not close to impossible to sort it all out, given the apparent restructuring that has taken place. Though we can posit the general outlines of an analysis, we cannot unequivocally explain away the exceptions.

Despite these problems, our best hypothesis remains *ịjị-, because it covers most Class 5 V-stem reflexes for a core of Sabaki languages, with the possible exceptions of Comorian and Elwana, and parts of other dialects. This is an apparent PSA innovation whose immediate source is probably PNEC *(ị)zị-, whose reflex /zi-/ is widely attested

throughout the Seuta and Ruvu branches, e.g., Proto-Seuta/Ruvu *ziso 'eye', *zino 'tooth':

Pre NEC *(ị)lị- > PNEC *(ị)zị- > PSA *ịịị- Class 5 V-stems

(The parentheses around V_1 in the prefix vowel simply indicate a non-etymological source, such as analogy, which may have independently operated at different times; though indeed the V may well be the Bantu pre-prefix.)

The NEC forms derive from a spirantizing change, probably part of Bantu Spirantization, a change which helps distinguish NEC from Western Tanzanian languages which still attest *lị-, e.g., Sukuma /lịịso/ 'eye', /lịịno/ 'tooth'. The shift of PNEC *(ị)zị- to PSA *ịịị- looks very much like strengthening, akin to Comorian/SD strengthening, namely, Sabaki *ị-zịto 5 > /dziro/ 'heavy' where [+continuant] > [+stop] (see §6.3). It is not clear that the two are part of the same historical Sabaki process; nevertheless the similarities are suggestive.

For subgrouping purposes, the shift to *ịịị- does not unequivocally define PSA. We cannot be certain that it is attested in either Elwana or Comorian, as already noted above. The Elwana forms, e.g., /izičo/ 'eye', /izina/ 'name', could either be retentions from NEC *ịzị- or derived from the general shift of PSA *j to Elwana /z/. Comorian Class 5 /dz/ could also stem directly from *z, rather than *j, as we earlier suggested (see §6.3).

Outside Sabaki, *ịịị- is also reconstructible for Pare/Tußeta Class 5 vowel stems. To make this clear, note the following correspondences:

Proto-Pare	So. Pare	Tuβeta	No. Pare	PSA	CB
*z	z	z	ð	*j	*j
*nz	nz	Øz	nz ~ nð	*nj	*nj
*j	j	j	j	*z	*d/__ ị
*nj	nj	nj	nj	*nz	*ng/__ ị

Note the flip-flop (hereafter the "Pare flip-flop rule"): PSA *(n)j corresponds to Pare *(n)z, and PSA *(n)z corresponds to Pare *(n)j, e.g.:

PSA	:	Pare		PSA	:	Pare	
*mwezị		mweji	'month'	*-ịja		-za	'come'
*zị-		ji-	'Class 10 concord'	*-ịjala		-izua	'be full'
*-ziịka		-jika	'bury'	*njala		nzaa	'hunger'
*macozị		masoji	'tears'	*-ganja		-ɣanza	'palm'
*ngonzị		ŋonji	'sheep'				
*-sịnziịla		-šinjia'	'sleep'				

Class 5 V-stems in Pare/Tußeta have the following form: /iziso/ (Tußeta/So. Pare), /iðiθo/ (No. Pare) 'eye'; /iziko/ 'fireplace' (Tußeta); /izina/ (Tußeta/So. Pare), /iðina/ (No. Pare) 'name'. To account for /z/ in these forms only a reconstructed *j is consistent with the correspondences in the above table, thus:

Proto-Pare					
*ɟiɟiso 'eye'	>	iziso (So.Pare/Tußeta)	>	iðiθo (No.Pare)[51]
*ɟiɟina 'name'	>	izina (So.Pare/Tußeta)	>	iðina (No.Pare)
*ɟiɟiko 'fireplace'	>	iziko (Tußeta)	(>	nzukweni (So/No.Pare))

Thus Proto-Pare forms are compatible with the PSA reconstruction of *ɟiɟi- and indicate early genetic affinities between the two:

NEC *ɟiɟi-	>	Pare and PSA *ɟiɟi-	>	izi-	Pare (via flip-flop rule)
NEC *j	>	Pare and PSA *j	>	z	Pare (" " " ")

PSA and Pare shared the shift of NEC *ɟiɟi > ɟiɟi. Shortly thereafter, Pare innovated its flip-flop rule and all *j were shifted to Pare *z. Elsewhere (Chap. 3, §2.3; Ch 5, §2.1.5.1) we conclude that a particularly close relationship between Pare/Tußeta and Sabaki, within the general framework of NEC, cannot be supported on the basis of lexicostatistical evidence alone. The Class 5 evidence, however, indicates a closer genetic relationship between Pare/Tußeta and PSA than other data might indicate. Furthermore, if our assumption that Pare/Tußeta *ɟiɟi- and PSA *ɟiɟi- have the same source, then we have to conclude that the innovation of a palatal in Class 5 contexts is early, possibly pre-dating PSA itself. It further follows from this that, if Elwana and Comorian are truly genetic affines of other Sabaki languages, a fact we do not doubt, whose development postdates the early split of Pare from Sabaki, then it is more likely than not that Elwana and Comorian forms also derive from *ɟiɟi- as well, and not from *z. This is indirect, but important, evidence in support of a uniform analysis for all Sabaki, and helps support the subgrouping utility of the *ɟiɟi- innovation.

§11.4.4. **Augmentative *ji-**. In Unguja the Class 5 V- and M-stem prefix /ɟi-/ is identical to the augmentative prefix /ɟi-/, e.g., /ɟitu/ 'giant', which is also used in diminutives, e.g., /ki-ɟi-tu/. Unguja is somewhat unique in Sabaki in that these morphemes have fallen together. Polomé appeals to analogy to explain the identity here, arguing that a CB prefix *gi- ([-HV]), indicating size, is the source of the allomorphic distribution of Class 5 prefixal marking. This reflects Meinhof (1932:118-119), who derives the Ø reflex of Class 5 polysyllabic stems from *li- and the others from *ɣi- ([+HV]). Hinnebusch (1973) (also see Hinnebusch, Nurse, and Mould 1981) attempts a uniform treatment for Class 5 and augmentatives, and argues for a single etymological source from a pre-PSA *liɟi-. This may have been a misguided attempt, as the evidence is good that PSA *ɟiɟi- (< PNEC *(ɟ)ziɟi- < pre-NEC *(ɟ)liɟi-) is distinct from the augmentative which we reconstruct as PSA *ji-, where the vowel is [-HV].

Sabaki augmentatives are diverse in their formation, sometimes closely following Class 5 marking as in Unguja, sometimes not. The irregularity here is a beacon indicating

[51]Hinnebusch field notes (1972/73) for Tuβeta give /iɟiso/ 'eye'; other sources give /iziso/. /iɟiso/ may be a borrowing from Swahili, or may represent an archaism, supporting the reconstruction; note that /iɟani, mani/ 'leaf/leaves', attested in Tuβeta, No. Pare, and So. Pare, also does not fit the pattern.

post-PSA reanalysis, and confusion between the two functions. In augmentatives, Giryama uses /ri-/ with polysyllabic stems, e.g., /ri-tsoka, mi-tsoka/ 5/4 'big axe', and /dzi-/ with M-stems, e.g., /ri-dzi-tswa, mi-dzi-tswa/ 5/4 'huge head'. Giryama also marks certain adjectival stems with /ri-/ as well: /ri-re/ 'long', /ri-dzo/ 'good', /ri-ša/ 'new', /r-aruhe/ 'white'. Digo uses /dzi-/, e.g., /dzi-mu-hi, ma-dzi-hi/ 5/6 'big tree', /dzi-ndata, ma-dzi-ndata/ 'big stick', as does Duruma, e.g., /dzi-hi/ 'large tree , /dzi-dzi-hi/ 'huge tree'. In both languages /dzi-/ is cognate with the Swahili augmentative /ɟi-/, but there is a lack of conformity between languages. In Giryama, the use of /ri-/, deriving from verbal *li-, with certain adjectives and quantifiers is probably part of the use of the same form. Comorian marks its augmentatives with /ǰi-/: /ǰondo/ (Nz) 'terrible trouble' (see /-onda/ 'make poor', /šondra/ 'wound'; *kilonda 'wound'), /ǰitro/ (Nz) 'valley', /ǰito/ (Ng) 'large river' (see *muto 'river'), /ǰivi/ (Ng, Nz) 'heavy wooden door' (see CB *-yi̧bi̧), /ǰindru/ (Ng) 'giant' (see *muntu), /ǰimbo, maimbo/ 'song, singing' (Ng, Nz) (see *lwimbo), /ǰindru, maindru ~ maǰindru/ (Ng) 'large object'; /ǰiba/ (Ng) 'large bone', (see šiba 'bone'); /ǰosi/ 'smoke' (see Nz mosi ~ moši). The Comorian data is the best Sabaki evidence that Class 5 forms and augmentatives are etymologically distinct. The Comorian augmentative does not show the same changes found in Class 5, namely, *i̧ji̧- > /dzi-/, nor does it participate in other regular Comorian phonological changes, namely, PSA *j > Nzuani /ž/, e.g., /-žaya/ 'be full' (*-i̧jala). Only Comorian preserves the historical distinction; it is in the rest of Sabaki that the augmentative *ji- gets mixed up with other Class 5 marking functions.

This confusion of Class 5 vowel stems with *i̧ji̧- and augmentative *ji- probably began after the Sabaki seven vowels reduced to five. That *ji- augmentatives and Class 5 vowel stems were once distinct is further supported by morphological evidence. Augmentatives were formed with Class 4 plurals, while Class 5 V-stems had Class 6 plurals. Class 4a augmentatives are still seen in Giryama, e.g., /ri-dzi-tswa, mi-dzi-tswa/ 5a/4a 'huge head', with remnant traces in the rural dialects of Swahili, e.g., Pemba: /zino, mi-zino/ 'tooth/teeth', /ɟibwa, miɟibwa/ 5/4a 'large dog' (Whiteley 1958); Makunduchi: /mɟiti, miɟiti/ 3a/4a 'trees'; the latter most likely stems from an older augmentative pairing Class 5/4 (Whiteley 1959).

§11.5. Class 6: *ma-. PSA *ma- before C-stems is stable. Before V-stems, various vowel rules operate, as illustrated in the data below. We will not discuss these changes.

*ma-/__C-stem	:	ma- Sabaki
*matunda 'fruit'		matunda (SD, ND), matuunḍa (Mw), mahunda (Di), marunda (Ng,Vu), etc.
*ma-/__V-stem	:	m(a)- Sabaki
*maiko 'fireplaces'		miiko (Mw), meko (rest of Swahili), maeko (UP)
*maino 'teeth'		meno (SD,Po,MK), manyo (Ng,Nz,Ma)

§11.6. Class 7: *ki-. For C-stems we have:

*ki-/__C-stems	ki-	El, UP, Gi, SD, ND, Mn
	ki- ~ či	LP
	ky-	Du
	či-	Di, Du, Ch
	či- ~ š-	Mw
	i- ~ hi-	Ng
	ši-	Nz, Ma
*kicwa 'head'	kečwa ~ kíčwa (El), čitswa ~ kitswa (LP), kitswa (Gi), čitswa (Ch,Ra,Di), kyitswa (Du), čiṯa (Mw), kiṯwa (Ti,Am,Si,Pa, Mv), čiṯwa ~ kiṯwa (Chi), kičwa (UP,Ung,Vu,Mt,Pe,Tu,Mak, Ung,Maf), hitswa (Ng), šitswa (Mh,Nz,Ma), čiṯwa (Chi), kiswa (Mn)	
*kikapu 'basket'	čikaɸu (LP), kyikaβu (Du), čikaβu (Di), škapu 7/8 (Mw), kikapu (Am,Ung), °ikapu/zi- (°k/p ¬ h/β) 7/8 'type of basket from Madagascar' (Ng)	
*kintu 'thing'	kintu (LP), kit^{h}u (Gi), kyit^{h}u (Du), čiint^{h}u (Mw), k^{h}iču (Ti,Pa, Si), k^{h}itu (Am), kit^{h}u ~ k^{h}itu (Mv,Ung), k^{h}itu (Nga), čitu (Jo), kitu (Pe), hindru (Ng), šintru (Nz), šitru (Ma), kinu (Mn)	
*kisima 'well'	kisima (El,LP,Gi,Am,Ung), kyisima (Du), čisima (Mw), isima (Ng), šisima [ššima] (Nz), šisima (Ma)	

Lower Pokomo variably shows /ki- ~ či-/. The palatalized form is used by older speakers; /ki-/, under the influence of Swahili, is used by younger speakers (H. Ipu, pers. comm.). Tikuu, Amu, and Siu in some contexts also show variability (Nurse 1982b:84).

Mwiini attests /či-/ before C-stems with an initial voiced consonant, e.g., /čilima/ 'small hill', /čiguwo/ 'cloth, rag', /činema/ 'cinema'; before Ç-stems we find /š-/, e.g., /škapu/ 'basket', /špande/ 'piece', /šṯana/ 'comb', /šiniko/ 'lid' (but /čisima/ 'well', /čisiwa/ 'island', not *ššima, *ššiwa). Before monosyllabic stems, regardless of the voicing feature of the initial consonant, neither the shift of *či > /ši/ nor vowel-deletion is seen with monosyllabic stems, e.g., /čifo/ 'mortality', /čike/ 'feminine', etc. (see §10.5 for further discussion).

The Mwiini shift of *ki- > či- > /š-/ is similar to events in Comorian. Comorian dialects have a general shift in which *ki- > /ši-/; Ngazija adds a further step, namely, *ki- > ši- > /hi-/; and Nzuani deletes the vowel, e.g., /šisima/ [ššima] 'well', /šisiwa/ [ššiwa] 'island', /šikele/ [škele] 'coconut cut in two', /šitrandra/ [štrandra] 'bed'. The similarities here between Mwiini and Lower Pokomo, and Mwiini and Ngazija, are probably due to different underlying motivations and offer nothing for subgrouping.

In Ngazija, the normal realization of the Class 7 prefix is /i-/ before most polysyllabic C-stems, and /hi-/ usually with M-stems (TJH field notes 1984, Sibertin-blanc 1980);

Sacleux (Chamanga and Gueunier 1979) records only /hi-/ in both contexts. On the surface these correspondences appear to be a product of Comorian *k-lenition, viz., *k > /h/, but the circumstances of the shift in Nzuani/Maore, which have /ši-/, and the attestation of /š-/ with V-stems in all of Comorian, suggest a different possibility: Class 7 *ki- (Sabaki) > /ši-/ (Comorian) > /hi-/ (Ngazija) > /i-/ (Ngazija). This series of changes has to be ordered before Comorian *k-lenition, otherwise we would uniformly find Class 7 /hi-/; contrast this with Comorian Class 15, where *k-lenition is responsible for its reflexes.

For V-stems we find:

*ky-/__V-stems:	ky-	El, UP, Du, Vu, Pe
	č-	LP, ND, other SD, Gi, Di, Mn
	š-	Com
*kyula 'frog'	kyola (El), kyula (UP,Du), kyura (Vu), kyura ~ šura (Pe), čuya (LP), čulwa (Ch), čula (Gi), čura (Ung,Mv), čuuɬa (Mw), čwačwa ~ čuačua (Am,Ti,Si,Pa), shwa-toto (Ma)	
*kyakulya 'food'	kyakuɟa (UP), kyakulya (Mak,Pe,Vu), kyakura (Du), čakuɟa (LP), čakuǰa (Mw), čakula (Am,Ung), šahula (Ng,Nz,Ma), čakurya (Gi,Mn)	
*kyala 'finger'	kyaala (El), kyala (Du), čala (Gi,Maf,Mn), čaaɬa (Mw), šaya (Ng), ša (Nz)	

The changes operating here are identical to those applying generally to sequences of *kyV, thus *kyV > kyV > čV > šV, and require no further explication (see §4.1.3.2).

§11.7. Class 8: *vị-

*vị-/__C-stem:	vi-	Sabaki	zi-	ND_2, Com
	ði-	Ti, Pa, Si	zi- ~ s-	Mw
*vịntu 'things'	vit^{h}u (Ung), zit^{h}u (ND), ðič^{h}u (Ti), zintu (Mw), zindru (Ng), ziintru (Nz), zitru (Ma)			

Changes involving this prefix form are general. The shift of *v > z/___i is discussed in §5.1.1, the ND shift of /z/ > /ð/ in §5.1.5, and the Mwiini shapes in §10.5.

For V-stems:

*vị-/__V-stem :	vy-	Sabaki	ž-	Gi
	z-	ND_2, Com	ð-	Ti
*vyakulya 'foods'	viakuɟa (UP), vyakuɟa (LP), žakurya (Gi), vyakurya (Di, Du,Ch), vyakula (Ung), zakula (Am), ðakula (Ti), zahula (Ng,Nz,Ma)			

The changes that underlie the correspondences here have been discussed elsewhere: Comorian/ND_2 *v > z/___ ị (§5.1.1); Giryama *vy- > ž (§5.3).

§11.8. Class 9/10: *N-. The Class 9/10 prefix in PSA participates in a number of changes which involve the initial consonant of the stem to which it is prefixed. These changes are discussed in great detail in §8; also see Herbert (1986). The prefix itself, whenever it surfaces in synchronic grammars, is realized as a stressless, prenasalized homorganic nasal. In monosyllabic environments it can also surface as a stressed nasal, as in Swahili. Our ignorance of Sabaki suprasegmental systems precludes our saying more about similar marking functions in other languages.

*N-/__C̥-stems:	N-	Sabaki
	Ø	El

*nguWo 'clothing'	guo (El), nguo (LP,Gi,Du,Am,Si,Mv,Ung,Ng,Nz,Mn), nguvo (Ti), nguʋo (Mw)
*mbwa 'dog'	bóa (El), mbwa (LP,Am,Si,Pa,Chi,Mak,Ng,Mh,Nz,Ma), mbaa ~ mbwa (Mw), imbwa ~ mbwa (Ti), imbwa ~ umbwa (Mn). (CB *-búa)

*N-/__C̥-stems:	N-	Ng, Nz, Mn
	N- [N̥]	Po, Mw
	N- [N̥] > Ø	MK, SD, ND (not Po, Mw)
	Ø	Ma, El

*mpula 'nose'	bula ~ bóla (El), mpula [m̥pʰula] (UP), mpuya [m̥pʰuya] (LP), pʰula (Gi,Ch,Ra,Du); pʰura (Di), mpuɬa [m̥pʰuɬa] (Mw), pʰua (Ti,Si,Pa,Am,Jo,Mv,Jo,Vu,Chi,Mt,Pe,Ung,Mak), pʰuya (Tu), mbua (Ng,Mh), mpua (Nz), pua (Ma), mula (Mn).

The devoicing process, effects of which are still directly apparent in Pokomo and Mwiini, and which underlies N-deletion in Mijikenda and in the Swahili dialect continua, is discussed in §8.2.2. The deletion in Elwana is part of a language-specific process in which *NC̥ sequences become voiced and the nasal subsequently deletes (§8.1.1). There is no evidence that deletion in Maore is connected with devoicing in other Sabaki languages. In Mwani sequences of *NC̥ become /N/, e.g., /mula/ < *mpula; here the reflex nasal presumably derives from *N and the oral stop deletes (*mpula > mmula > mØula > /mula/). With monosyllabic stems the facts are different, e.g., Swahili does not delete the prefix in such environments (see §8.4).

For stems with initial fricatives, the correspondence relationships are different. Generally deletion is the norm in Sabaki if the following fricative is voiceless, the exceptions being Lower Pokomo and Comorian (Ng, Nz), where the nasal has been preserved (for full details, cf §8.3). Before voiced fricatives, deletion is less widely attested.

For V-stems we have reconstructed *ny- [ɲ], a widely attested Bantu form of the prefix with vowel-initial stems; for the most part this is stable and undergoes little change except as noted:

*ny-/___V-stems :	ny-	PSA
	n̪-	Ti
	ny- ~ n̪- ~ n-	Mw
*nyota 'star'	nyoha (LP), nyota (Am,Jo,Mv,Ung,Mt,Pe,Tu,Mak), nyoča (Pa,Si), nyora (Vu,Ng,Mh,Nz,Ma), n̪oča (Ti), noota (Mw)	
*nyumba 'house'	nyóba (El), nyumba (LP,Gi,Du,Am,Mv,Jo,Ung,Vu,Pe,Ng,Nz, Mn), nuumba ~ n̪uumba (Mw), n̪umba (Ti)	
*nyuni 'bird	nyɔ́ɔ̀ni (El), nyuni (UP,LP,Gi,Ch,Ra,Du,Di,Am,Si,Pa,Jo,Mv, Chi,Mt,Mn), nyuni (obs) (Mv,Ung), nyunyi (Mw), n̪on̪i (Ti), nuni ~ nyunyi (Ng), nyunyi (Mh), nyunyi ~ nyuni (Nz), nyuni (Ma)	

Refer to §7.1 for more details on the Tikuu and Mwiini dentalization.

§11.9. Class 11: *lu-. Correspondences for the Class 11 nominal prefix with polysyllabic and monosyllabic stems are:

*lu- :	lu-	El, UP, MK, Mn
	l- (lu-)	Mw
	ru-	Di
	yu-	LP
	u-	Am, Ung, Com
	u- (~ vu-)	Ti
*lukope 'eyebrow'	lukoɸe (UP), yukoɸe (LP), lukohe (Gi), ukope (Am,Mv,Ung), lkope (Mw), lukope (Mn), uhoβe (Ng)	
*lupembe 'horn'	rupembe (Digo), lpeembe (Mw)	
*luti 'staff, stick'	lut̪i (El), yuhi (LP), uti (Am,Ung), luti (Mw), uri (Ng)	
*luto 'thick fluid'	luro̥ 'clear honey' (UP), luho (Gi), uto (Am), učo ~ vučo (Ti,Si), uto 'oil, thick oily fluid' (Ung), utro 'mold' (Nz)	

Changes to this prefix for the most part follow changes affecting *l elsewhere (see §4.2.2); *l before *u in non-Class 11 contexts deletes in Swahili, Comorian, and Lower Pokomo, thus the shapes here are derived from the general shift affecting *l overall, with the exception of Pokomo. In Lower Pokomo, *l generally shifts to /y/ which deletes before /u/ and in other environments, except in Class 11 where /y/ is retained, e.g., *-luma > /-uma/ 'bite' versus *lukope > /yukoɸe/ 'eyelash'.

*l-deletion is a late areal change in Sabaki that started only after the differentiation of the Sabaki dialects. Part of the evidence for this is the behavior of Mwiini vis-à-vis the rest of Swahili. Mwiini preserves *l both in specific contexts in the general lexicon and in Class 11, where the vowel of the prefix is deleted rather than *l (except in M-stems, where a CV form of the prefix is preserved; see §10.5). Mwiini had to have already been distinct

as a dialect, if not geographically separated, when *l-deletion began to operate, otherwise we would expect Mwiini forms to be congruent with other Swahili dialects.

Where *l-deletion has occurred in Sabaki languages, Class 11 has merged with Class 14, which has also lost its etymological consonant (see §11.11; Chap. 4, §2.2.3). The two are usually kept paradigmatically distinct by their plurals: Class 11 nouns usually have Class 10 plurals, and when plurals are possible, Class 14 nouns have Class 6 forms. In data for Tikuu, reported by Bakari (1985), we see a further collapse of distinction, where presumably formerly /u-/ marked Class 11 nouns are attracting the Class 14 noun marker: /vulimi, ndrimi/ 'tongue', /vudove, ndove/ 'fingernails' (both etymological Class 11/10 nouns; see /vudongo/ 14 'mud, clay'). This would seem to be a recent phenomenon in Tikuu. Other sources for Tikuu neither report on this nor record it (e.g., Nurse 1982b; Sacleux 1909, 1939; Stigand 1915). It is also not certain what model is responsible, since the vast majority of Class 14 nouns in Tikuu, as reported by the sources and our own field work, are only marked with /u-/; the exceptions are monosyllabics, e.g., /vu-yi/ 'porridge' (cf. /u-ɟi/ Ung, /u-i/ Am), /vu-so/ 'face' (cf. /u-so/ Ung), and stray lexical items, e.g., /vu-ongo/ 'clay' < *Wudongo versus /u-vongo/ 'earth'. The only synchronic Class 11 example we have noted with /vu-/ is /u-vuča, mbuča/ 'bow' (Sacleux 1939/41). However, it has been reanalyzed as part of the stem, and moreover is from an earlier Class 14 item, thus *Wu-ta 14 > vu-ča > /u-vuča, mbuča/ 11/10; see /uta, nyuta/ 11/10 (Ung).

For V-stems:

*lw-/__V-stem	lw-	El, MK, Mn
	l-	Mw
	lw - ~ yw-	UP
	ǰw-	LP
	Øw-	Com, rest of Sw
*lwala 'fingernail'	lwala (El), lwala ~ ywaya (UP), ǰwaa (LP)	
*lwembe 'razor'	lwembe (Gi,Mn), ǰwembe (LP), leembe (Mw), wembe (Am,Ung,Nz,Ng)	

Swahili and Comorian are consistent in attesting *l-loss here, just as they do in other contexts. The other dialects preserve a consonantal reflex: Mwiini deletes the glide, Upper Pokomo appropriately attests /lw- ~ yw-/ from *lw-. The Lower Pokomo correspondence of *lw- : /ǰw-/ is part of a general process specific to Lower Pokomo in which all sequences of *lw are realized as /ǰw/, e.g., *-lwana 'fight' : /-ǰwana/. The distribution of the reflexes of *lu- (C-stems) in Pokomo suggests that *lu- > /lu-/ (as in El, UP) > /yu-/ (as in LP), thus Lower Pokomo /ǰw-/ would seem logically to derive from a similar seriation: *lw- > lw (upper UP) > yw- (lower UP) > /ǰw-/ (LP); this is supported by other data in Lower Pokomo where *y > /ǰ/ in intervocalic environments, e.g., *-Wuya 'return' > /-uǰa/,

*lucaya 'cheek' > /yučaǰa/, *mbegu > mbeu (*g-loss) > mbeyu (G-insertion) > mbeǰu.[52] This continuum, with retention in the north ranging to innovation in the south, matches a similar one involving *t, namely, *t > /ɽ/ (UP) > /h/ (LP).

§11.10. Class 12: *ka-. For a justification and discussion of this PSA class, see Chap. 4, §2.2.1. There is no convincing evidence for Class 13.

§11.11. Class 14: *Wu-

*Wu-/___C-stem :	u-	Sabaki (including Ti)
	vu-	Ti (M-stems)
*Wuso 'face'	oso (El), vuso (Ti), uso (rest of Sabaki)	
*Wuntu 'humanity'	unt^{h}u (LP), ut^{h}u (Gi,Am,Ung), undru (Ng), untru (Nz)	
*WucaWi 'witchcraft'	učawi (El,Ung), učaβi (UP), utsai (Gi), uṯaʋi (Am,Mv), uṯavi (Ti), učwai (Ng), usawi (Mn)	

For V-stems:

*w-/___V-stem : w-	Sabaki
*w-ivu 'laziness'	wivu (most Sabaki)
*w-įvį 'theft'	wivi (most Sabaki), ʋiði (Pa,Si,Ti), widzi (Ng)

The reconstructed prefix *w for Class 14 V-stems derives from the earlier deletion of the consonantal portion of the prefix: *Wu- + vowel-stem > /u-/ + V-stem > /w-/ + V-stem. While /w/ from *W is deleted in Mijikenda, approximants from vowel sources are not. However, in some languages /w/, regardless of source, is assimilated when the following stem vowel is /o/, e.g., /ongo/ (Gi) < *WuWongo 'brain'; /oa/ (El), /oga/ (Gi) < *w-oga 'fear'. For further discussion see §4.2.1.2.

§11.12. Class 15: *ku-

*ku-/___C-stem :	ku-	Sabaki
	hu-, u-	Com
	ku- ~ x-	Mw (see §10.5)
*ku-įpika 'cook'	ku-pika (most Sabaki), hupika (Ng,Nz), upika (Nz), xpika (Mw)	

The Comorian shift of *k > h is part of the general lenition of *p, *t, *k (see §4.1). The further shift of /hu-/ > /u-/ is specifically Comorian and appears to be synchronically active, as indicated by variable descriptions. Thus Sacleux (Chamanga and Gueunier 1979) gives /hu- ~ hw-/; Sibertin-blanc (1980) reports /hu- ~ u-/ for Ngazija, Nzuani and Maore, but Rombi (1983) only marks infinitives with /u-/. The circumstances under which deletion occurs are not certain.

---

[52]Lower Pokomo regularly preserves *g, thus /mbeu/, which is the immediate source of LP /mbeǰu/, must be borrowed from a neighboring Sabaki language which attests the word in this form, e.g., Tikuu, Pate, Siu, and Amu.

For V-stems:

*ku-/__V-stem :	kw-	Sabaki
	kw- [k͡p(ʷ)-]	MK
	hu-, hw-	Com
*kw-enda 'to go'	kw-enda (Ung, etc.), hw-enda ~ hu-endra (Ng), hu-endra (Nz), kw-enda [k͡pʷ-enda] (Gi), k-enda (Mw)	

Labialvelarization (*kw- > /kw-/ [k͡p(ʷ)]) affects Mijikenda infinitives variably. For example, in Giryama very few labialvelarized infinitives were recorded where the initial vowel of the stem was [-low], in contrast to Chonyi, where [k͡p] is more common before /i/ and /e/. Although Giryama does not altogether avoid labialvelarization of the prefix in vowel environments, it appears to utilize certain strategies to avoid them. So /ku-/ tends to be realized before /i, e/ as [kuʷi] and [kuʷe], and the initial vowel frequently deletes before /a/, resulting in variation, e.g., /kw-ambia/ [k͡pambia ~ kumbia] 'to tell' (see §9.1 for more examples and discussion).

In Mwiini, the prefix vowel deletes before V-stems, with compensatory lengthening of the stem vowel, e.g., /ku-atuɬa/ > [kaatuɬa] 'to split' (cf. kwatua Ung), /ku-ondoka/ > [koondoka] 'to wake up', /ku-oɬa/ > [kooɬa] 'to write', /ku-umba/ > [kuumba] 'to create', /ku-uɬa/ > [kuuɬa] 'to buy', /ku-enda/ > [keenda] 'to go', /ku-inenda/ > [kiineenda] 'to walk'. Vowel-lengthening in Mwiini is complex and due to a number of language-specific mechanisms (see §10).

§11.13. Classes 16, 17, and 18: *pa-, *ku-, and *mu-. Except in Mwani, the locative classes are synchronically active only in marking various grammatical relations (see Chap. 4, §2.2.2) and as such they undergo regular phonological changes, e.g., *p-lenition (*p > h, etc.), labiovelarization (*kw > [k͡pʷ], etc.), syncopation (*mu > m̩), etc.

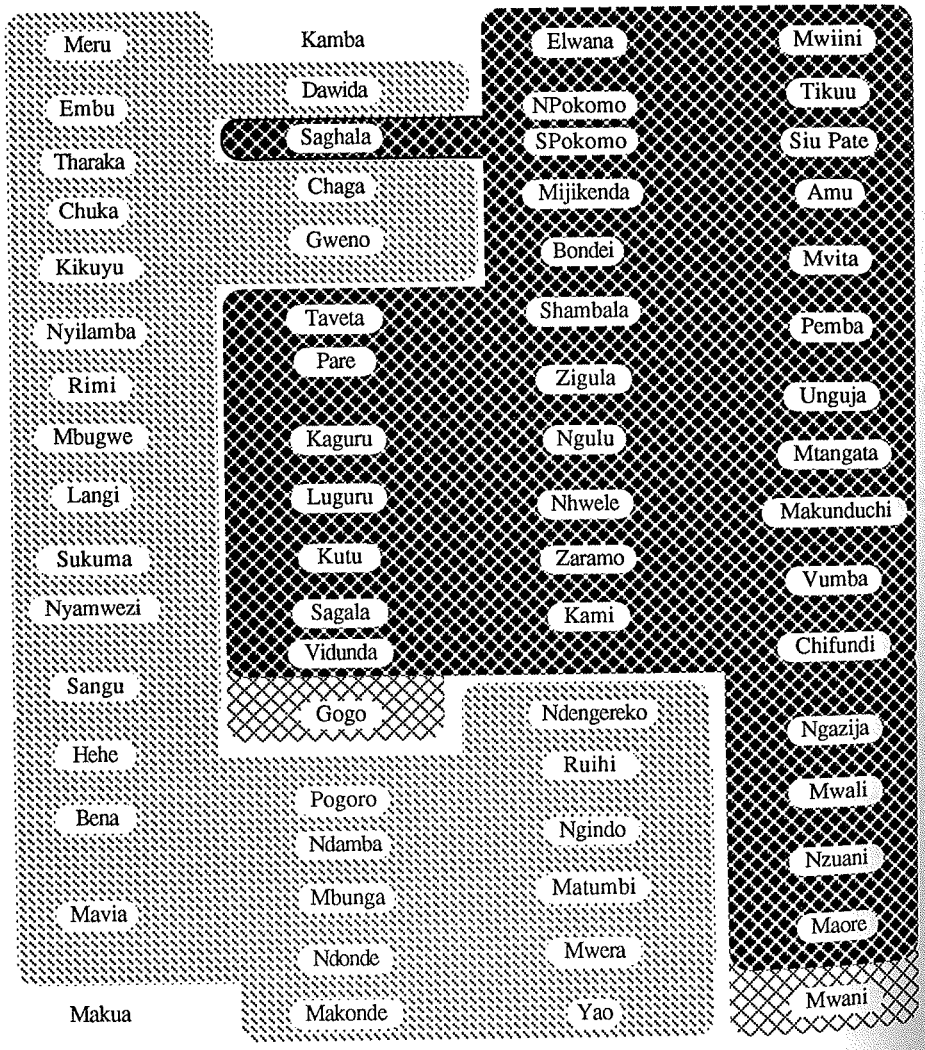
§12.0. NEC: Its Phonological Definition

NEC as a dialect cluster is difficult to define phonologically for various reasons (see Hinnebusch, Nurse, and Mould 1981). There is only one clear NEC phonological innovation (see §12.1). We therefore must use other criteria, largely involving a combination of features. Some of these are widely attested by NEC and other groups, but nevertheless are useful in defining other parts of NEC's boundary. Some are innovations which are found in adjacent non-NEC languages, thus further defining boundaries. Our approach here is largely, but not completely, typological. The following section outlines briefly the types of features that can be used to define the group.

§12.1. An NEC innovation. NEC languages are virtually uniquely defined by a shift in which the Class 5 nominal prefix with vowel stems is attested as either *j(i̧)- or *z(i̧)-, both probably derived ultimately from *i̧zi̧- (Chart 19). This contrasts with surrounding languages which preserve an older state of affairs, namely reflexes of CB *di̧- (or *di) with /l/ or /r/ (Hinnebusch, Nurse, and Mould 1981). Thus, for example, NEC is distinguished

Chart 19

NEC *jį- ~ *zį- vs. *lį Class 5 VS



Key:

*jį- ~ *zį Cl. 5 Vowel Stems

*lį- ~ *li- Cl. 5 Vowel Stems

NEC exceptions (Gogo and Mwani attest /ri-/

Non-NEC exceptions: Saghala, Kamba, Makua (see Ch. 2, §12.1)

from Central Kenyan to the northwest, which attests /ri-/ (CB *di-), and similarly for other surrounding groups. Kamba is an exception in attesting a spirantized reflex, /sy/, but this is an idiosyncratic Kamba development wherein /s/ corresponds to CB *j, CB *gụ, CB *gịV, CB *di/___Cl. 5, etc. (Nurse 1979a:424-425), and is not derivable via Bantu Spirantization. Daiso, which is spoken within NEC territory, is another exception. It is similar to Kamba in this respect, and clearly behaves as it does by virtue of its Central Kenyan connections. Saghala, which otherwise subgroups with Daßida (Hinnebusch and Nurse 1981), falls in with NEC in attesting /z(i)-/, but this is a late borrowing from NEC, as are other aspects of Saghala. Two major NEC exceptions are Mwani and Gogo with attestations of /ri-/. In the former, this is likely to be a borrowed feature from P20 languages; its source in Gogo is unclear.

§12.2. **Non-unique NEC features.** In this section we outline phonological characteristics in a typological fashion. These are features which are not necessarily unique to NEC, but do serve to demarcate boundaries with other groups (for details, see Chap. 1, §3.1 and 3.2.1). Also important are innovations in languages outside the group. We discuss below several features of both sorts which help to define NEC boundaries (see Chap. 1, §3.2.1, for membership of the language groups discussed in this section).

§12.2.1. NEC, Central Kenya, and West Tanzania

1. Bantu Spirantization (§5.2 and Chart 20). This, attested in NEC, provides an isogloss to distinguish it from Central Kenya (CK), which has not undergone this change at all, and from West Tanzania (WT), where it is only partly attested in a few languages, and therefore must be recent.

2. Reduction of 7 to 5 vowels (§10 and Chart 21). This is a late post-PNEC shift which has affected all NEC languages, except possibly Elwana. The isogloss sets NEC (5 vowels) apart from Central Kenya and West Tanzanian (7 vowels—except Sumbwa, a West Tanzania dialect with 5 vowels).

3. Neutralization of inherited V length (§10 and Chart 22). All NEC languages (except Mwiini, and apparently Elwana), together with a number of other eastern Bantu languages, attest this shift. Central Kenyan and West Tanzania are conservative in retaining vowel distinctions based on length.

4. Shifts affecting *N̥Ç clusters (§8.2 and Chart 23). Part of Central Kenya Bantu has *N̥Ç > NÇ, thus adding to the distinctiveness of part of Central Kenya vis-à-vis northern NEC. Elwana, the only NEC language to attest this shift, may do so as a consequence of contact with Central Kenyan languages which do. The instance of Comorian voicing of *N̥Ç is internal to the group.

§12.2.2. **NEC, and Chaga/Taita.** Chaga/Taita shares many features with NEC languages (e.g., Spirantization, 7-Vowel reduction, *VV-neutralization), but can be distinguished from them by a combination of retentions and innovations:

1. Shifts affecting *NC̥ clusters (§8.2 and Chart 23). All of the Chaga/Taita group have voiced *NC̥ clusters. The only NEC languages in this area to do so are Tuβeta and North Pare, while South Pare has shared in the Sabaki and Seuta shift of *NC̥ > NC̥^{h}.

2. Bantu Spirantization (§5.2 and Chart 20). A specific fact of this shift, namely, *l > /r/ (or derivatives of /r/) before [+HV], is an innovation which helps define Chaga/Taita as a group, and supplies evidence that spirantization or parts thereof in Chaga/Taita was independent of Spirantization in NEC (see Hinnebusch, Nurse, and Mould 1981, and Hinnebusch and Nurse 1981). Saghala, which otherwise looks like a NEC language, also shares partially in this shift and is distinguished from NEC by this isogloss.

3. Spirant-devoicing (Chart 24). This is a shift in which fricatives, which were derived from CB *b, *d, *g by Spirantization, are devoiced, and then in some cases weaken to /h/, or are even deleted (Hinnebusch, Nurse, and Mould 1981). It is widely attested in a number of areas in eastern Bantu, including a large group of languages south of NEC (see §12.2.3 below). Its occurrence in Chaga and Gweno helps define these against NEC. The Taita languages, Daßida and Saghala, which pattern with Chaga and Gweno on other grounds, are also therefore distinct (see Hinnebusch and Nurse 1981).[53] Some southern NEC languages, however, attest this shift (e.g., Luguru, Kaguru, Kami, and Kutu).

§12.2.3. NEC's southern boundary. NEC shares a number of features (e.g., Spirantization and 7-5 vowel reduction) with other adjacent eastern Bantu groups south and west of the Ruvuma River: Rufiji-Ruvuma, Kilombero, and Southern Highlands. To varying degrees these languages also pattern with NEC. There appear to be particular similarities with Kilombero and Southern Highlands (see Chap. 1, §3.1, and Nurse 1988e). Distinctiveness is created, however, by sets of retentions and innovations in both:

1. *p-lenition (§4.1.1.1 and Chart 25) separates most of NEC, in which it is attested, from all the southern groups, plus parts of West Tanzanian which are conservative in not having undergone this shift. The one outstanding exception is the Swahili group (with some exceptions), which does not share *p-lenition with the rest of NEC. So the isogloss is only partially useful.

2. Spirant-devoicing (Chart 24). This is also only partially useful in defining NEC. Languages south of the Ruvuma (plus a few NEC languages) attest spirant-devoicing. Its wide attestation in Southern Highlands, Rufiji-Ruvuma, and Nyakyusa, and in a few languages in adjacent southern NEC (Kaguru, Luguru), is a classic example of an innovation areally radiating to contact languages on the periphery. While overlapping part of NEC, the shift is still useful, in that most of NEC does not attest the shift.

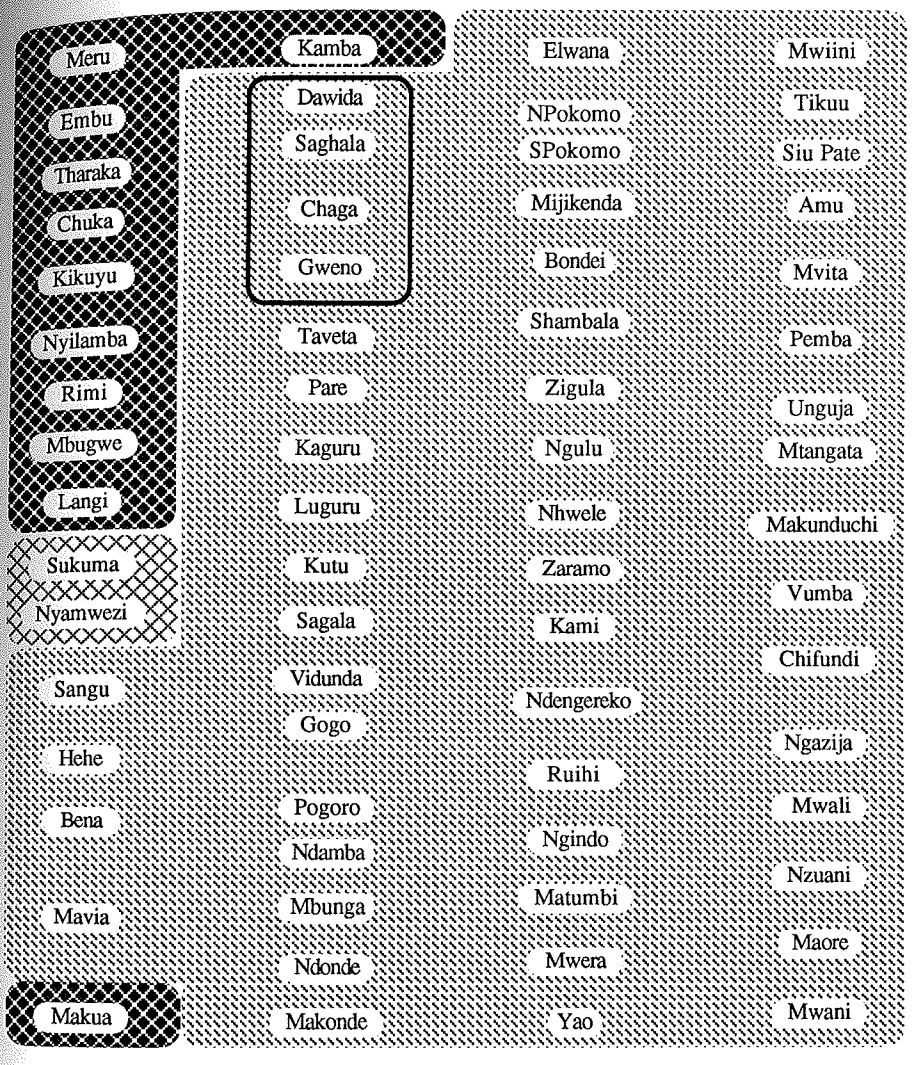
3. Spirant-weakening/loss (Chart 24). This follows on spirant-devoicing and leads to the shift of voiceless spirants to /h/ and in some cases full deletion (Hinnebusch, Nurse, and Mould 1981). It correlates with spirant-devoicing on NEC's southern borders but does not include any NEC language, and thus is helpful in negatively defining NEC's southern boundary.

---

[53]One Chaga dialect appears to have voiced fricatives (see Nurse 1979a: 393-394).

Chart 20

NEC Bantu Spirantization



Key:

Spirantizion not attested

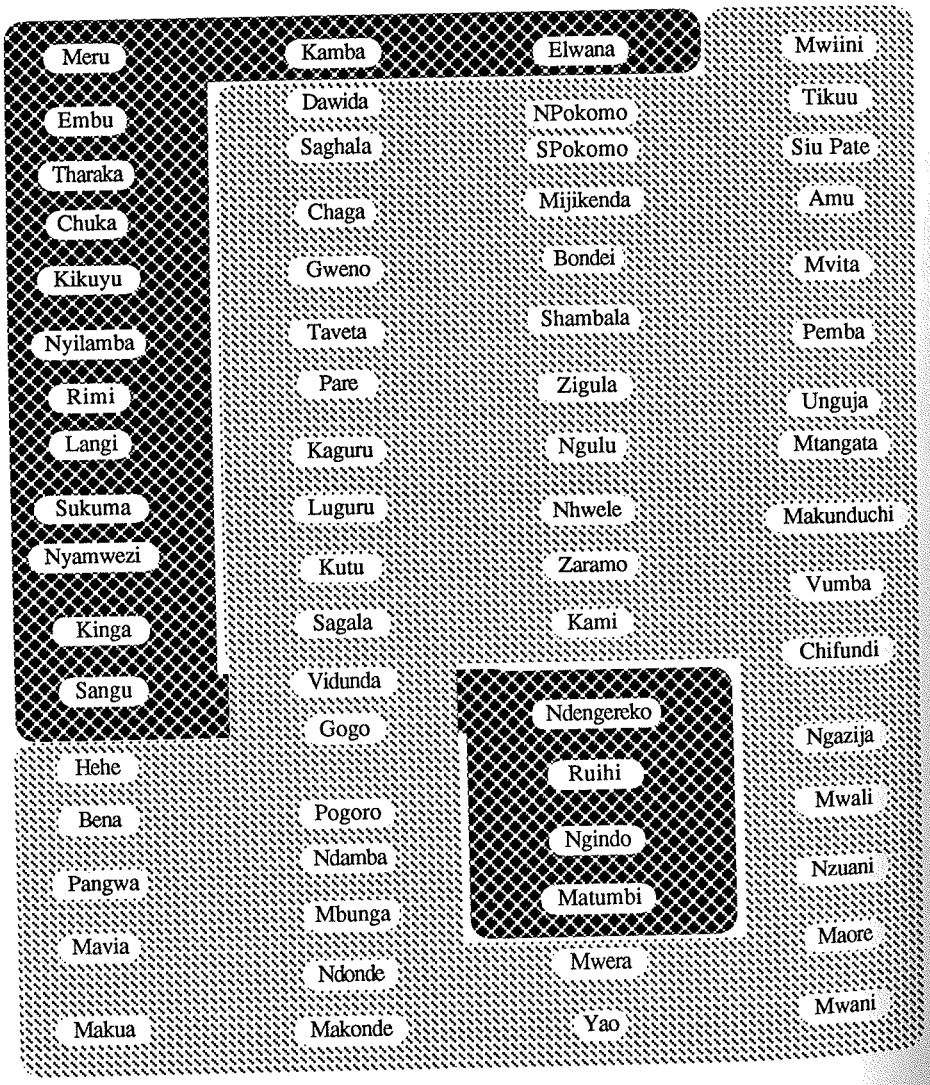
Spirantizion Partially attested

Spirantization attested

*l : r / ___[+HV] (Saghala *l : r / ___ μ)

Chart 21

NEC 7 > 5 Vowels



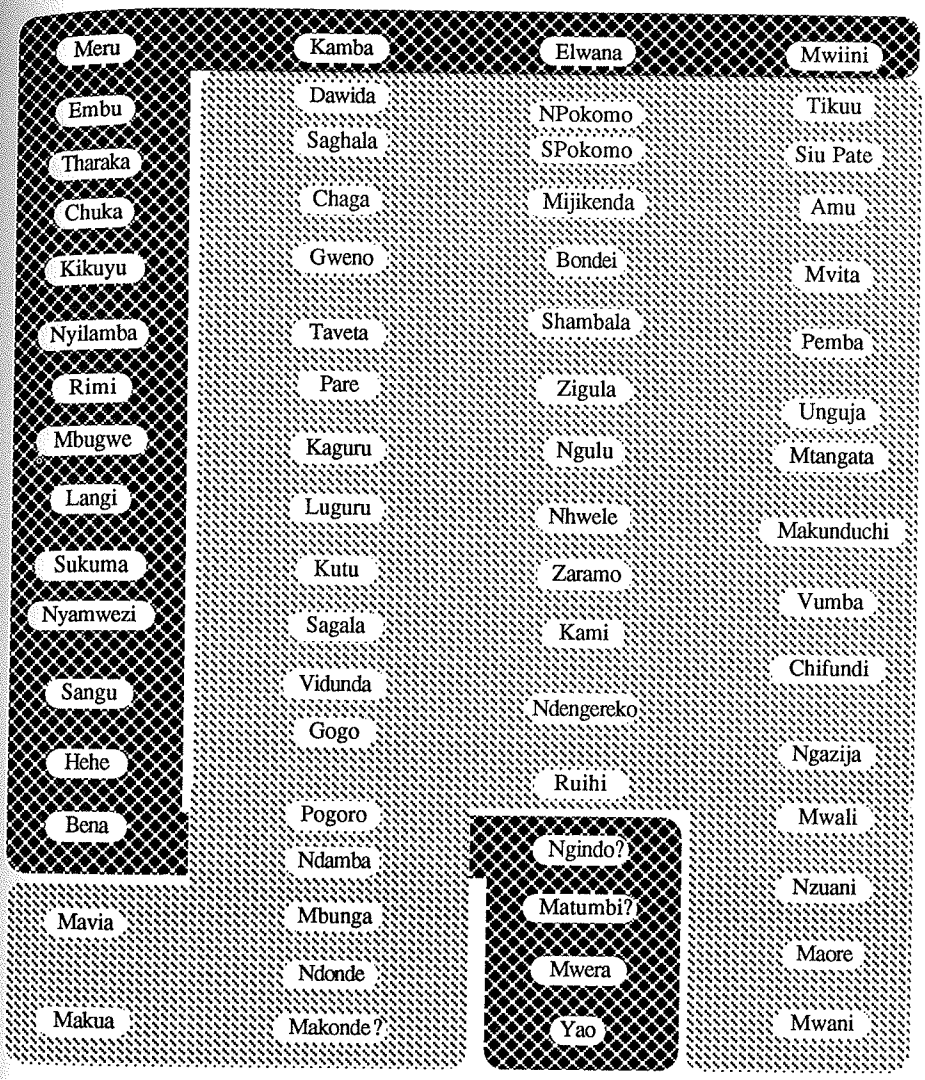
Key:

7 Vowels

5 Vowels

Chart 22

NEC VV > V



Key:

VV attested

VV > V

Chart 23

NEC Voiceless *NC̥

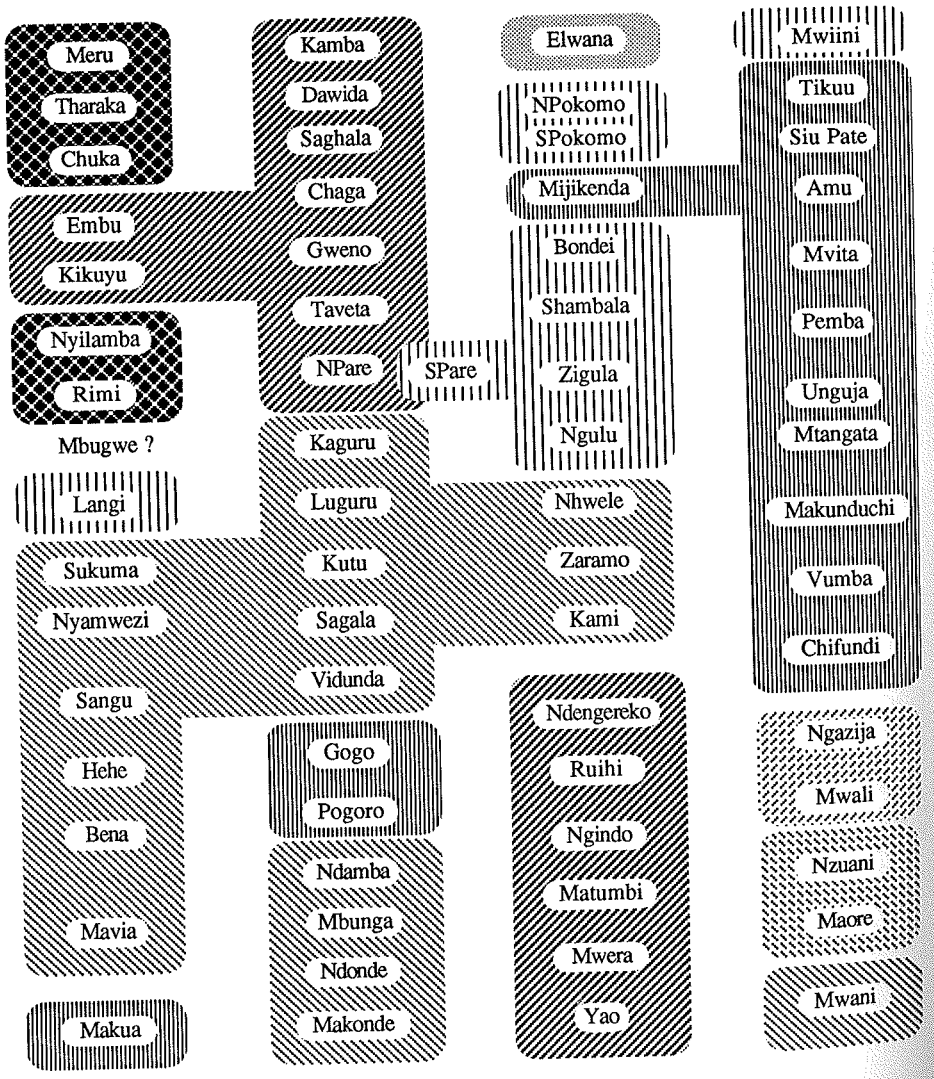
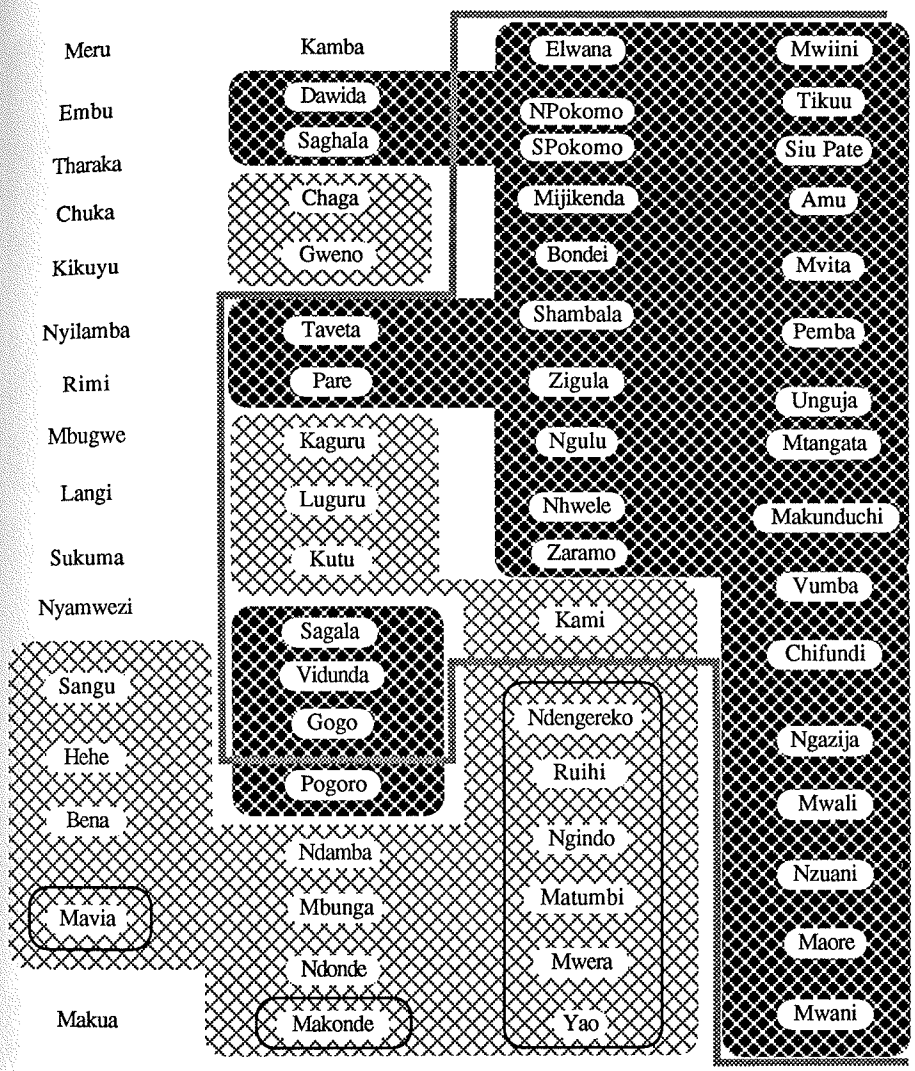


Chart 24

NEC Spirant devoicing



Key:

Voiced spirants < CB *b, *d, *g / ___ [+HV]

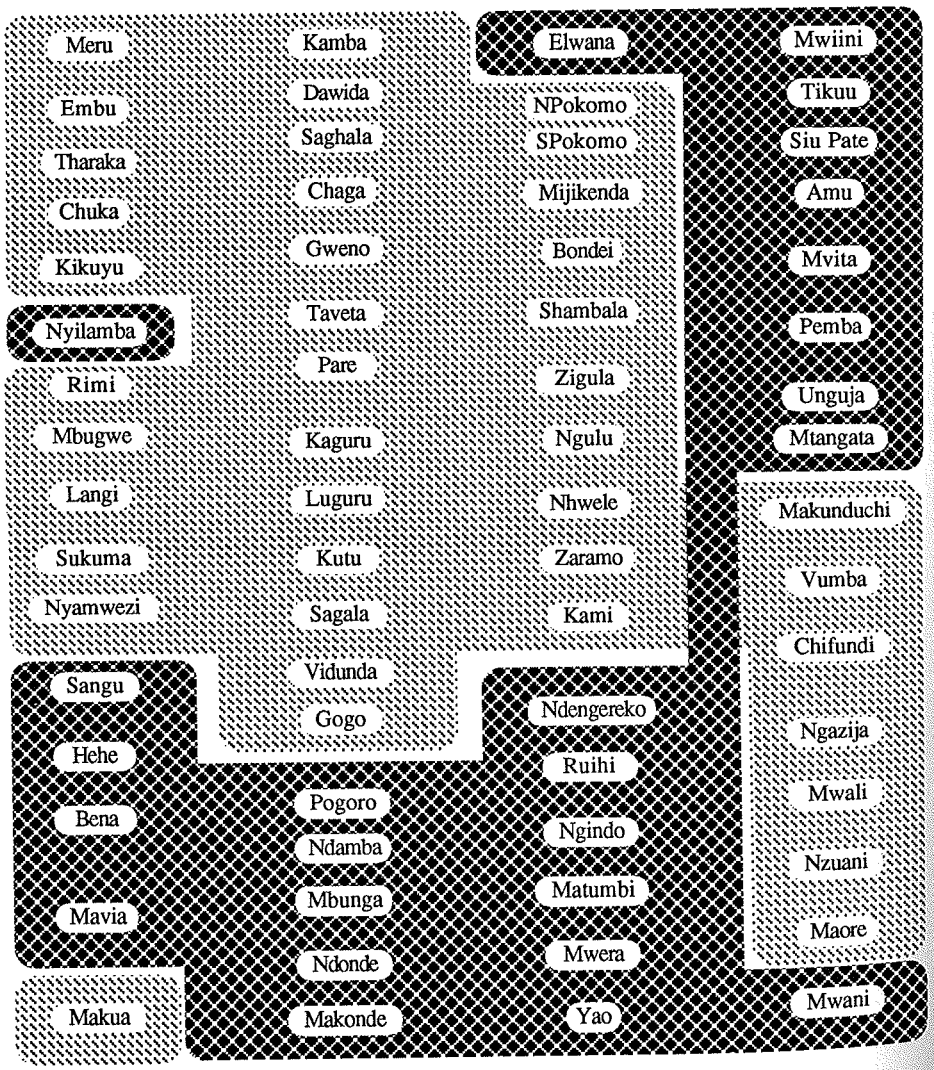
Voiceless spriants < CB *b, *d, *g / ___ [+HV]

Voiceless spirants ~ /h/ ~ Ø < CB *b, *d, *g / ___ [+HV]

NEC Languages

Chart 25

NEC *p-lenition



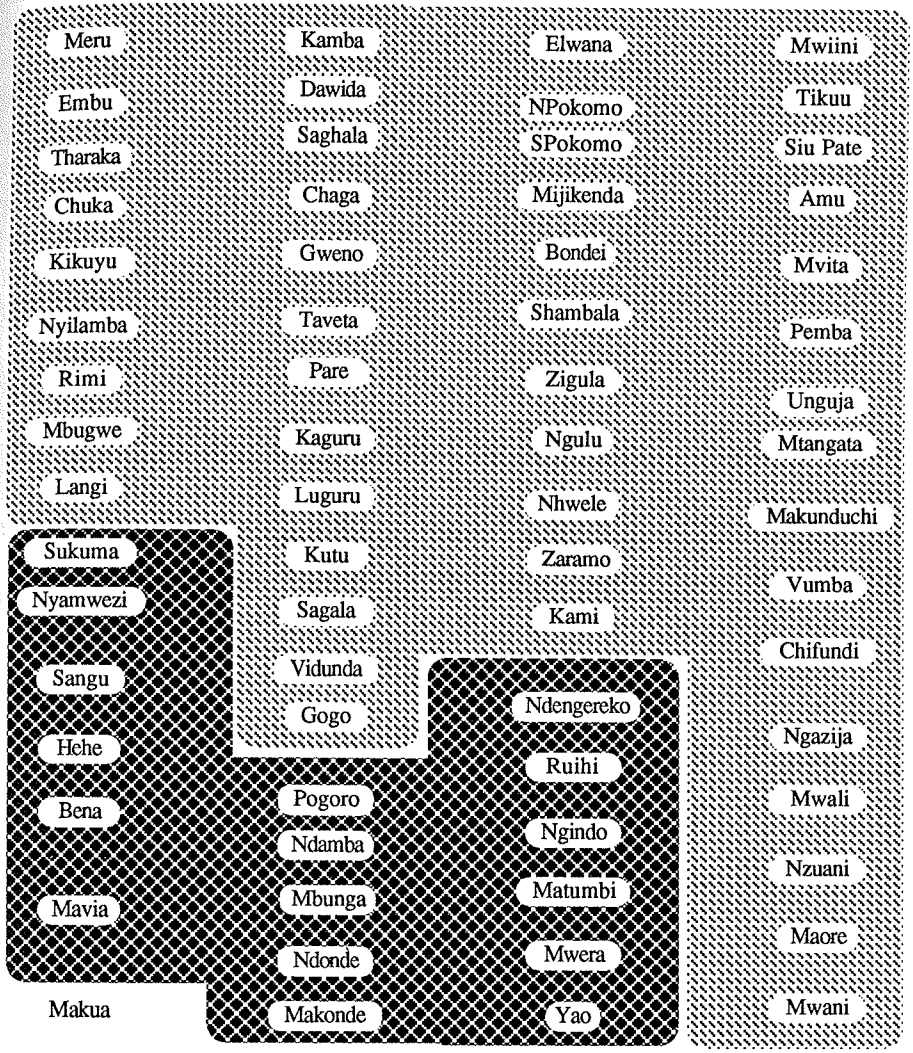
Key:

*p : p

*p-lenition (*p : β, v, ɸ, h, Ø)

Chart 26

NEC Class 5 *l-loss/____CVCV



Key:

Reflex of Cl. 5 *l/___CVCV stems retained

Cl. 5 *l-loss attested

4. *l-loss in Class 5 CVCV-nominals (§11.4.1 and Chart 26). *l-loss in Class 5 nominals is a pan-Bantu change. In eastern Bantu languages, only Rufiji-Ruvuma, Kilombero, and Southern Highlands preserve a reflex in this environment (Hinnebusch, Nurse, and Mould 1981), thus adding to the set of typological features (but not innovations) defining NEC on its southern boundaries.

5. *VV-neutralization (Chart 22). This feature only partially helps, in that the Southern Highlands group preserves the distinction between CB *VV and *V, along with Yao, Mwera, Ngindo (?), and Matumbi (?). Some non-NEC languages along its southern periphery, e.g., Ndengereko and Rufiji, which are otherwise distinct when we consider other criteria, share this shift with NEC.

6. Shifts affecting *NC̥ clusters (Chart 23). Only some groups south of NEC demonstrate differential behavior here. Southern Highlands falls in with Ruvu (*NC̥ > N^h/N), but Rufiji-Ruvuma attests NC̥-voicing: *NC̥ > NC̥, as does part of Comorian.

7. Retention of 7 vowels (Chart 21). Parts of Rufiji-Ruvuma and Southern Highlands retain 7 vowels, thus adding to the distinctiveness at this point along the boundary.

The Kilombero group, which lies south of the Ruvuma, is often conservative in respect to some of the events which are found in Southern Highlands and Rufiji-Ruvuma. It thus sometimes resembles NEC more than surrounding languages; it has replaced tone by stress, and attests nasal-devoicing in Pogoro (Nurse 1979a), two innovations which are found in Sabaki. It remains conservative in still attesting a CV reflex of CB *di- in Class 5 CVCV nominals; NEC is innovative for this feature.

§13.0. NEC, Sabaki, and Interrelationships

This section discusses those features which internally subgroup NEC phonologically. We can define PSA and Pare versus the rest of NEC by the shift of NEC *i̧zi̧ > PSA *iji̧- (Class 5 vowel stems; §11.4.3). Pare became phonologically distinct by innovating a further shift whereby Class 5 *iji̧ > /zi-/. There is no other Sabaki phonological innovation, properly speaking, that can define PSA vis-à-vis the rest of NEC. There are, however, important distributional facts that are helpful in making distinctions.

1. Shifts affecting *NC̥ clusters (§8.2 and Chart 23). A major isogloss cutting through NEC reflects the different ways in which NEC treats CB *NC̥ sequences. In a subset of languages (Swahili, Mijikenda, Pokomo, and Seuta), nasal-devoicing, and in most cases nasal-deletion, has occurred, namely, CB *NC̥ > N̥C̥^{h} > C̥^{h} (see §14.8 below). The other group of NEC languages, specifically Ruvu, rather than deleting the nasal portion of the cluster deletes the voiceless stop, sometimes leaving a trace in the form of aspiration: CB *NC̥ > N^h ~ N ~ (N̥).[53] Neither of these changes is diagnostic for any NEC

[53]Phonologically, /N^h/ may be read as /N/ or / N̥ /. Proper phonological analysis of the affected languages has never been done; it is unclear whether any language ends up with a phonological voiceless nasal. Guthrie, for example, gives N̥ for Sukuma/Nyamwezi, but phonologically this could be /N/ [N̥] (=[N^h]. For a recent study of Sukuma voiceless nasals, see Maddieson 1991.

group, but both provide an isogloss separating Sabaki/Seuta from Ruvu, with some exceptions, e.g., Gogo, which is diverse.

2. CB *c (§4.1.4 and Chart 27). The reflexes of CB *c in NEC languages provide further definition for Sabaki vis-a-vis other NEC languages. Sabaki is conservative in retaining [-continuant] reflexes for CB *c (/č, ts, ʈ/), while the rest of NEC has fricatives corresponding to CB *c (/s, š, θ, etc./). Non-Sabaki NEC falls in with the rest of eastern Bantu in this regard. We earlier discussed the possibility that these distributional facts perhaps indicate that CB *c : /č, ts, ʈ/ is innovatory for Sabaki and not retentive. Were we unequivocally able to accept this analysis, we would in fact have a Sabaki innovation. Nevertheless, for whatever reason, Sabaki does stand out against the rest of NEC and surrounding non-NEC Bantu languages by behaving differently. Mwani, a Sabaki language, however, no doubt under the influence of P20 languages, attests /s/ < *c.

3. CB *j (§4.2.4 and Chart 28). The problems of interpreting CB *j are similar to those for CB *c. Sabaki either attests [-continuant] reflexes of CB *j (namely, /dz, ʄ, etc./) or reflexes that are clearly derivatives of a palatal affricate or stop (namely, Maore/Nzuani /ž/ < Comorian *ǰ; Mwiini, Amu, and other ND /ž ~ y ~ Ø/ < PSA *ʄ); an exception is Elwana, which has /z/, whose source is uncertain. The rest of NEC, with some exceptions (e.g., Saghala), attests a fricative. This separates Sabaki from Pare/Tußeta and Seuta/Ruvu. Isoglosses for CB *j, however, do cut through the eastern Bantu area, so are not unique to NEC and thus not very helpful for subgrouping.

4. *nc (§8.2.3 and Chart 29). Sabaki retains stops or affricates as reflexes of *nc, whereas most of the rest of NEC innovates. Sabaki has CB *nc : /(n)č^{h}, (n)ts^{h}, etc./ versus Seuta and Ruvu CB *nc : /(n)s, š, θ/, Rufiji and Pogoro CB *nc : /s, h, Ø/, Guthrie's P20 CB *nc : /s, ny, h/. Sabaki here is distinct by virtue of innovations in adjacent groups.

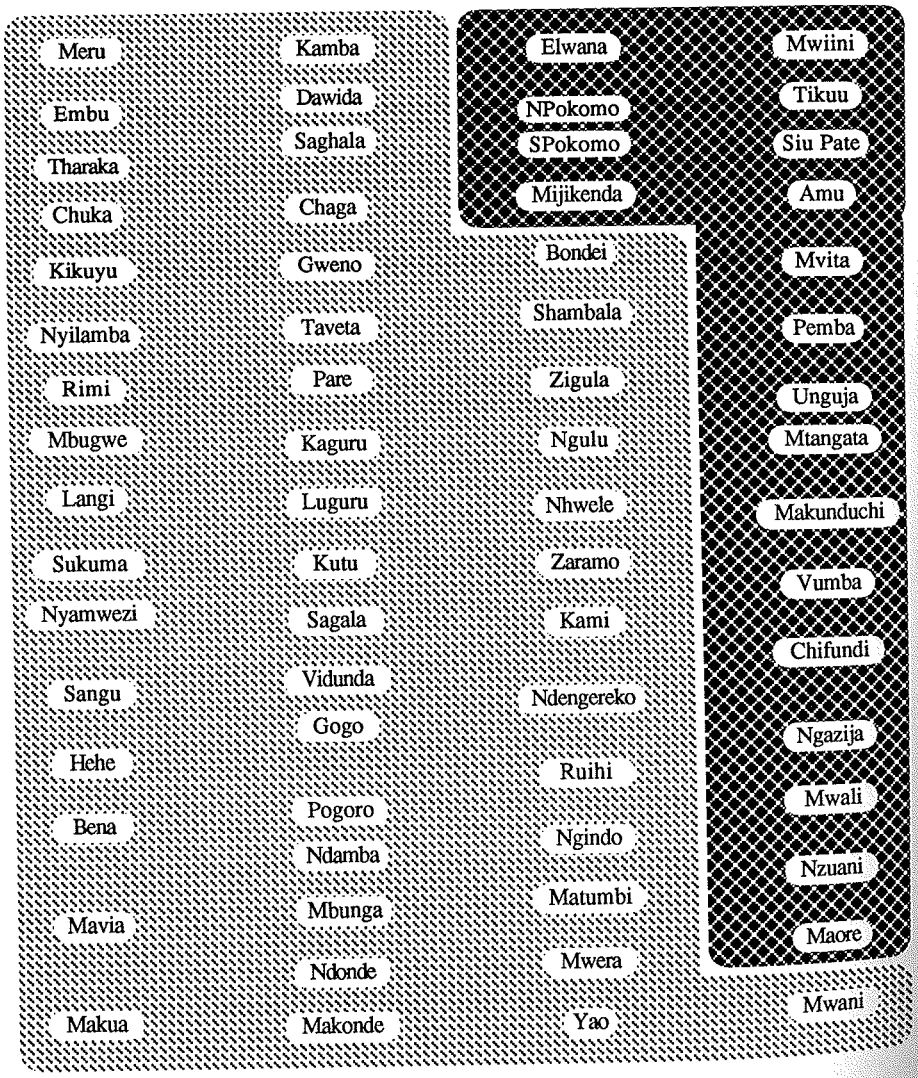
5. *nj (§8.1.4 and Chart 30). This is a less helpful set of changes, in that Sabaki is conservative along with others in eastern Bantu which preserve a nasal + palatal affricate or derivative for *nj. Sabaki has CB *nj : /nǰ, ndz, nɖ, ǰ/ versus CB *nj : /s/ in Seuta and *nj : /nz/ in Ruvu (but /nj/ in Kaguru). Innovations outside Sabaki provide the definition. Some adjacent groups are conservative, matching Sabaki: Rufiji and Kilombero have CB *nj : nǰ.

6. Dahl's Law (Bennett 1967, Davy and Nurse 1982, Nurse 1987). This is a dissimilation process whose most general form is that, in a sequence of two voiceless obstruents in adjacent syllables, the first becomes voiced (e.g. *-tátù > -datu 'three'). It is restricted to Bantu languages spoken in or on the fringes of East Africa. Within East Africa, it is synchronically or diachronically attested in Central Kenya, the Lacustrine languages, West Tanzania, Chaga-Taita, and Southern Highlands, with traces in Kilombero. There are no, or only scattered, traces of it in the other languages of southern Tanzania. In NEC there are many lexicalized traces in Seuta, Ruvu, and Pare/Tuβeta, which suggest that it once played a more active role. In Sabaki, these traces are very few:

CB *-kùùta 'shout'	Swahili -guta
CB *-pèta 'twist'	Swahili -peta 'twist', -betabeta 'curve (as a road)'

Chart 27

NEC *c



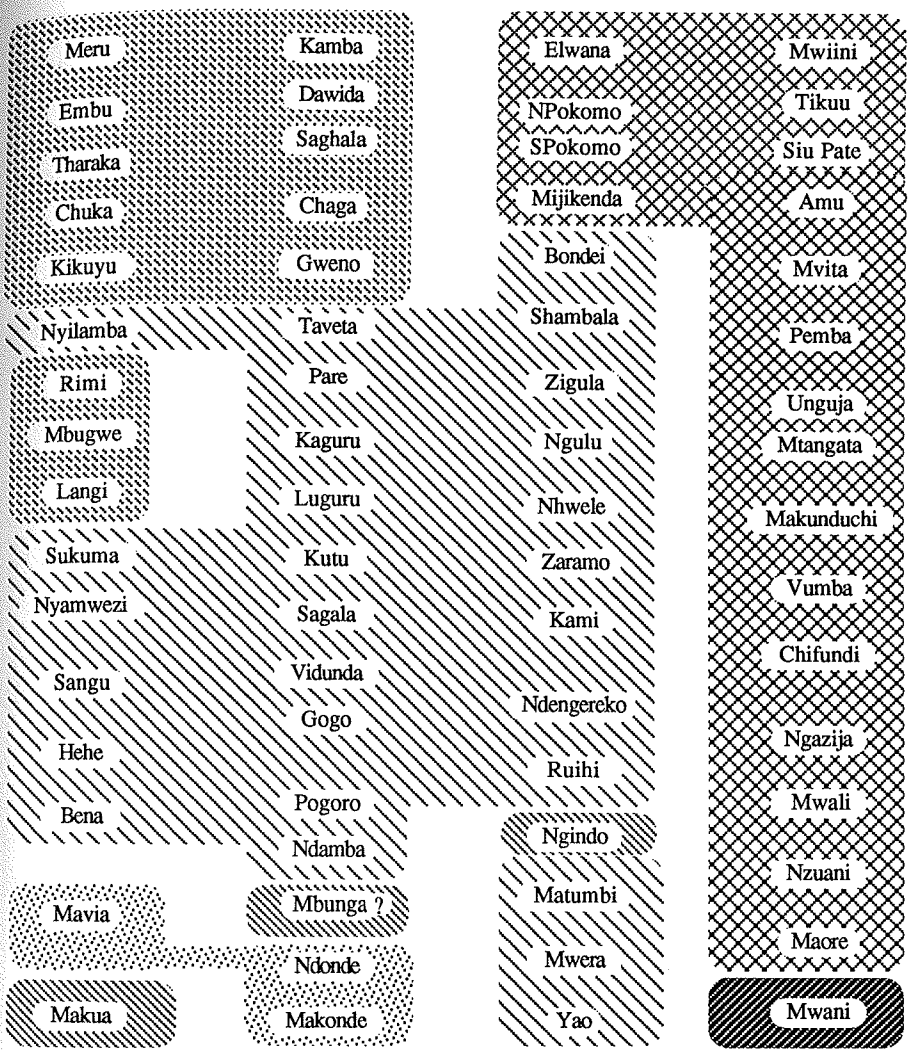
Key:

*c > č ~ ts ~ t̪

*c > s ~ θ ~ h ~ Ø

Chart 28

NEC *j



Key:

CB *j : *j > ǰ (~ ž) ~ ɟ (~ ž ~ y ~ Ø) ~ z Elwana

CB *j : *j > j ~ c ~ š ~ s

CB *j : ɟ ~ z

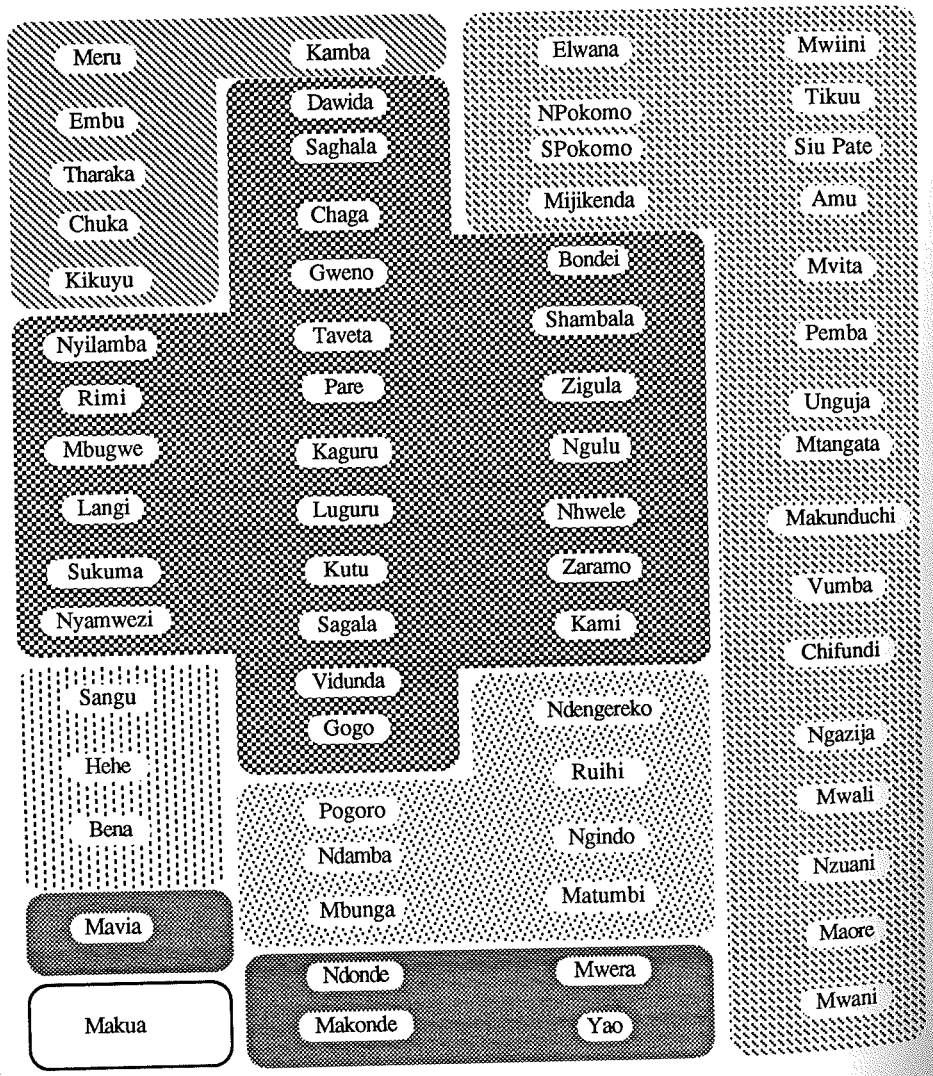
CB *j : *z ~ s ~ š ~ ts ~ Ø

CB *j : d

CB *j : c

Chart 29

NEC *nc

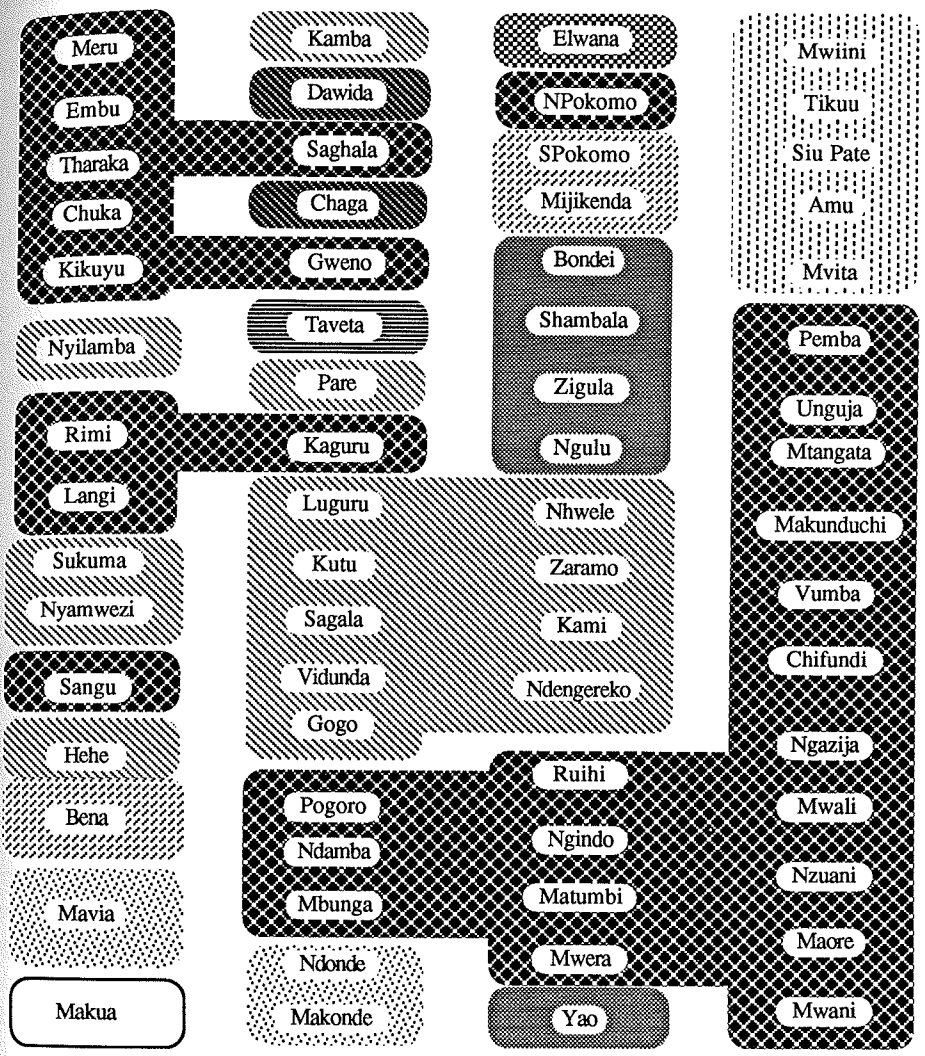


Key:

- *nc : (n)č ~ (n)ts
- *nc : (n)s ~ θ ~ š
- *nc : s ~ h ~ Ø
- *nc : ny ~ some (n)s
- *nc : s ~ ny ~ h
- *nc : $t\theta^h$
- *nc > nj ~ ns

Chart 30

NEC *nj



Key:

*nj : nǰ
*nj : ndz
*nj : nd̠
*nj : ǰ
*nj : nz
*nj : s
*nj : nð
*nj : nj ~ č ~ š
*nj : tθ
*nj : nd

That the traces of Dahl's Law are so minimal in Sabaki, but so widespread in other NEC languages, could be explained as a genetic and/or diffusional process. We could posit active Dahl's Law for PNEC, lost early in PSA, so leaving only slight traces, but kept till later in the other NEC languages, leaving more traces. A parallel exists in the Central Kenya languages, most of which have active synchronic forms of Dahl's Law, deriving presumably from Proto-Central Kenya, but Kamba has lost it entirely, keeping only a few relics.

Alternatively, we could claim that Dahl's Law was never part of PNEC, and that the forms found today in non-Sabaki NEC are the result of post-PNEC diffusion from one of the neighboring groups. Against this hypothesis we must count the contemporary distribution of Dahl's Law: if it results from post-PNEC diffusion over the last 1,500 years or so, why has it diffused so widely into Pare, Seuta, and Ruvu, but hardly at all into any Sabaki language, whose communities have never formed a hard and fast boundary with their neighbors, and how have traces got into Sabaki languages not in contact with non-Sabaki neighbors?

In view of this, we tend to favor the genetic explanation, which would imply that PSA early lost an active form of Dahl's Law, while early non-Sabaki languages kept it longer. However, we have to say that we cannot definitively prove this explanation.

§14.0. Internal Sabaki Interrelationships

In this section we call attention to those phonological features which serve to internally subgroup Sabaki. Our focus here is Swahili, although some discussion will perforce concern other groupings.

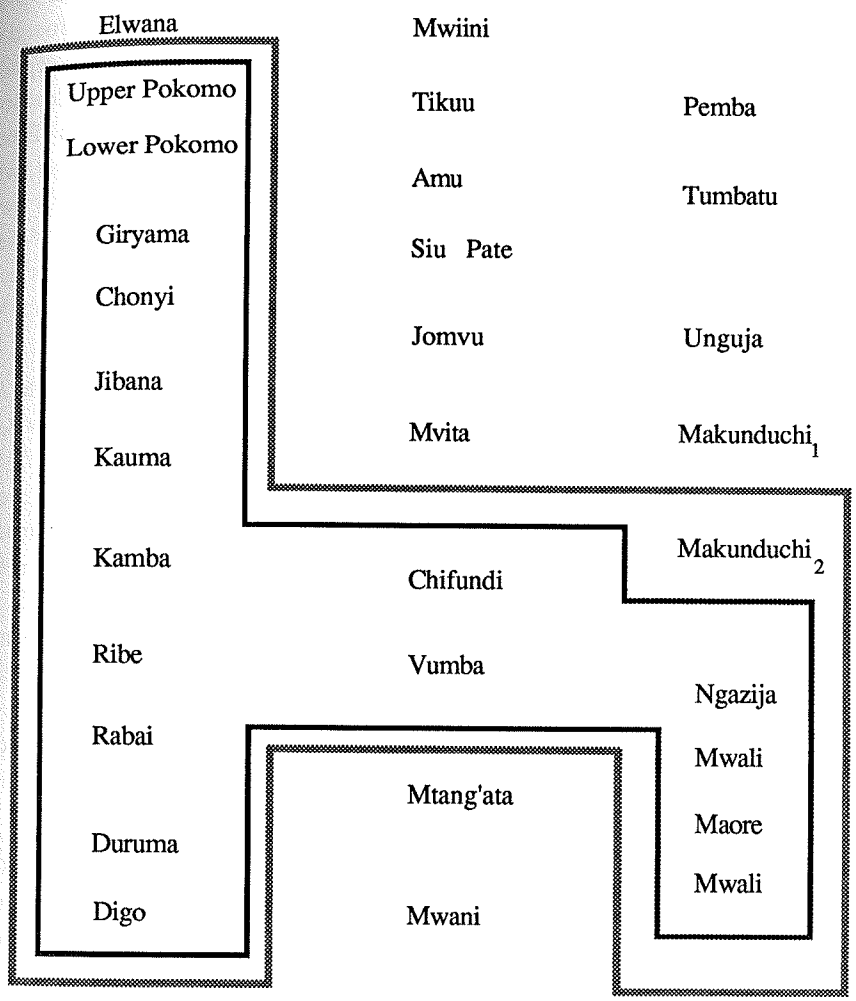
§14.1. *p-lenition (§4.1.1 and Chart 31). Pokomo, Mijikenda, Comorian, and three Swahili dialects, Makunduchi, Vumba, and Chifundi, attest *p-lenition. The remainder, the bulk of the Swahili dialects (including Mwiini), Mwani, and Elwana, are conservative.

*p-lenition is a widely attested Bantu change (Chart 25). Interpretation of the distribution of *p-lenition in East Highlands Bantu indicates either an areal change that spread late from south to north, reaching the Sabaki dialects while the early PSA community was still a close-knit dialect cluster, but with part of the community distinct enough, either geographically or socially, to rule out complete coverage. This interpretation is reinforced by the diversity of reflexes in Sabaki dialects today, and the fact that most of the Swahili group, including Mwiini, plus Elwana and Mwani, preserve *p.

Alternatively, while the phonetic preconditions of lenition were shared quite early, *p-lenition may have been a late internal Sabaki innovation in a already diverse dialect continuum. The different reflexes in the different dialects must surely have arisen after the individuation of the various dialects, otherwise we should expect more uniformity. Nevertheless, *p-lenition is an innovation that sets Pokomo, Mijikenda, and Comorian against Swahili, except for the already-noted exceptions. A serious problem arises in deciding why the isogloss splits Swahili as it does. Vumba, Chifundi, and Makunduchi have a lenis

Chart 31

Subgrouping *p/t Lenition



Key:

*p-lenition

*t-lenition

reflex of *p, the rest of Swahili does not. Several explanations are possible; these are detailed in Chap. 5, §13.6.

§14.2. *t-lenition (§4.1.2 and Chart 31). Lenition of *t is more localized than *p-lenition, probably coming as a specific Sabaki response to *p-lenition. It is rarely found outside the Sabaki group, and for that matter rarely occurs in eastern Bantu (one exception is the Chaga group, where *t : r ~ h ~ x ~ d, etc.). Its isogloss coincides with *p-lenition in negatively defining Swahili, with the notable exceptions of Vumba and Chifundi. Elwana, and Mwani, both on the periphery, are also conservative. Makunduchi, however, which attests *p-lenition, does not attest *t-lenition. If a borrowing hypothesis is valid for explaining *p-lenition in Vumba and Chifundi, then it would be valid here as well. Whether borrowing, as the ultimate reason for these cross-cutting isoglosses, is confirmed or not by other independent evidence, lenition of both *p and *t serves to demarcate Pokomo, Mijikenda, and Comorian vis-à-vis Swahili.

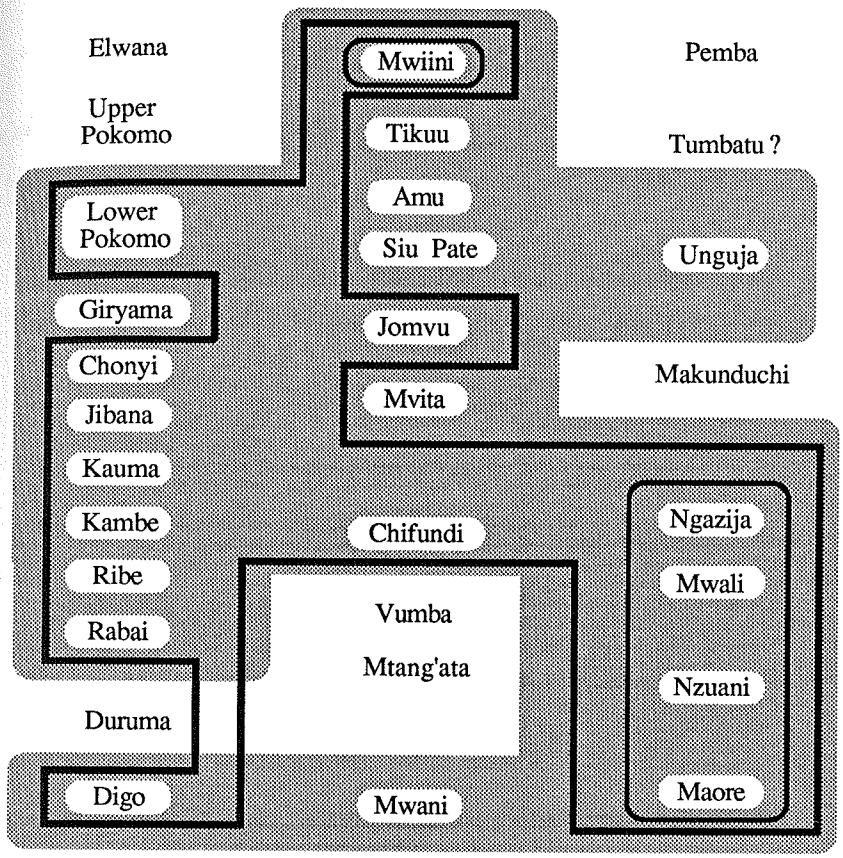
§14.3. *k-palatalization (§4.1.3.2 and Chart 32). The shift of *ky/__V-stems to /č/ is widespread throughout Sabaki. Only a few languages have not participated: Elwana, Upper Pokomo, Pemba, Vumba, and Mtang'ata; these are all conservative in retaining /ky/ before V-stems. The extension of the change to any *k before *i is restricted, however, to part of Mijikenda, Mwiini, Jomvu, Chifundi, and all of Comorian. Palatalization before *e is even more restricted. Whether Comorian is part of this depends on the analysis of facts, namely, that the shift of Sabaki *k > š > (/h/) had the intermediate shift of *k > /č/. The apparent subgrouping of Mijikenda and Comorian here is similar to findings for lenition and changes involving Sabaki *c. Mwiini also falls within this isogloss and also gives rise to /š/ in Class 7 environments, but whether this has any connection with Comorian /š/ is uncertain. As for the rest of Swahili, Jomvu and Chifundi fall in with Mijikenda and Comorian while Pemba, Makunduchi, Tumbatu?, Vumba, and Mtang'ata are conservative in retaining /ky/ in vowel stems.

§14.4. *c-dentalization (§4.1.4.1 and Chart 33). Comorian, Lower Pokomo, and Mijikenda share the shift of *c > /ts/. ND can also be subgrouped with these three, assuming the first of the two possible seriations outlined in §4.1.4.1 is correct, namely, *c > /ts/ [t̪s] > t̪). In any case, the shared fronting of *c, or dentality, of the resulting reflexes (/ts/ [t̪s] in Lower Pokomo/Mijikenda/Comorian, and /t̪/ in ND), and the seriation we have proposed, point to connections between ND and Lower Pokomo/Mijikenda/Comorian.

The isogloss has the effect of not merely setting SD apart from Mijikenda, Lower Pokomo, and Comorian, but significantly, it splits ND from most of SD, which remains conservative for this change and for the others we have so far discussed. At a later period, if we are correct in assuming that /t̪/ derives from *c via an intermediate /ts/, ND becomes distinct from Sabaki mainland dialects in shifting *ts (from earlier *c) to /t̪/. The same reflex subdivides the Swahili continuum: SD, which is conservative in retaining /č/ (: *c), versus ND, which attests /t̪/ (: *c). Elwana and Mwani maintain their distinctiveness in not participating in this innovation.

Subgrouping *ky and *ki

Chart 32



Key

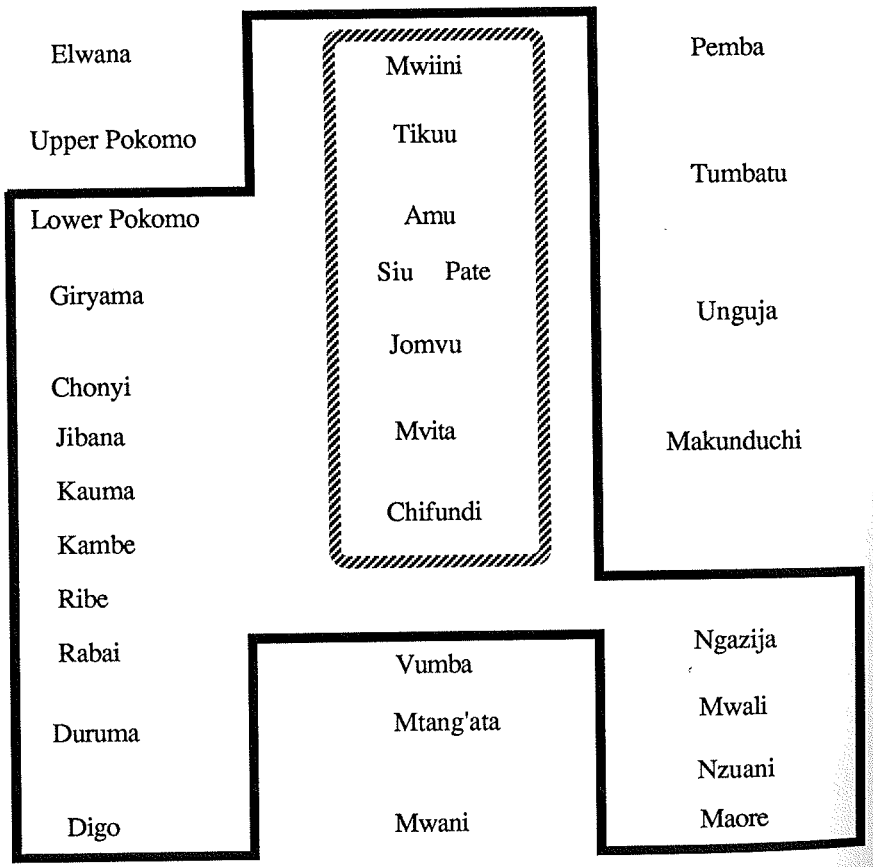
*ky > č ~ š /___V (e.g. *kyakulya, *-kya, etc.; languages unmarked attest /ky/)

*k > č ~ š /___i

Com *c > š ~ h ; Mwiini /č/ ~ /š/)

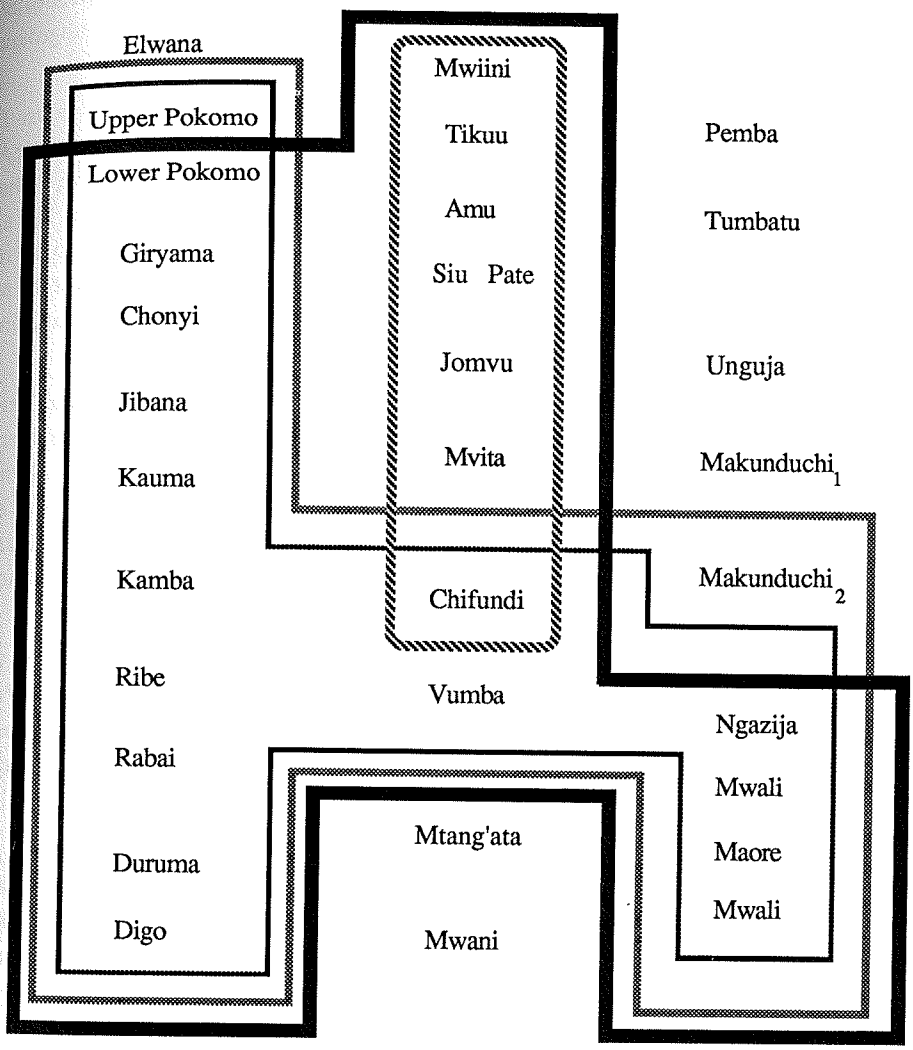
Chart 33

Subgrouping *c-dentalization



Subgrouping *p/t & *c

Chart 34



Key

*p-lenition

*t-lenition

*c : [-pal, +dent] reflex (= ts ~ t̪)

*c : t̪

Chart 34 demonstrates the combined effect of lenition and *c-dentalization. Of interest here is the behavior of Chifundi and Vumba (§15.10 below), which do not pattern consistently with other groups. The most striking fact here is the conservatism of most of SD, including Unguja. Likewise, Elwana and Mwani do not fit, nor can they be subgrouped with SD, other than sharing earlier Sabaki features.

§14.5. Sabaki W/___[-round] (§4.2.1.1 and Chart 35). The value of this feature for subgrouping rests on analysis, and depends on whether Sabaki *W was /ʊ/ [+dental] or /w/ [-dental], a question we cannot unequivocally resolve. The surface facts, however, do help with internal subgrouping, even though we cannot identify the innovative feature. ND and Upper Pokomo (Elwana?) attest a dental articulation of Sabaki *W, in opposition to the rest of Sabaki, where a nondentalized /w/ or Ø are attested. Mijikenda is set off against the rest in showing complete loss of *W, an event partially shared by Comorian, but again, since loss is involved, we have a less reliable indicator of relatedness.

§14.6. *W-loss/___[+round] (§4.2.1.2). Since most of Sabaki, the exceptions being Tikuu and Mwiini, has lost *W before [+round] vowels, we have no subgrouping evidence here. Because of the presence of a reflex of *W in this environment in Tikuu and Mwiini, we have one small bit of evidence that *W here was a PSA feature—albeit one that was lost early—and that loss before other vowels in Mijikenda and Comorian may be related, as rule-simplification and extension to nonround environments. Thus loss in Comorian is similar to loss in Mijikenda.

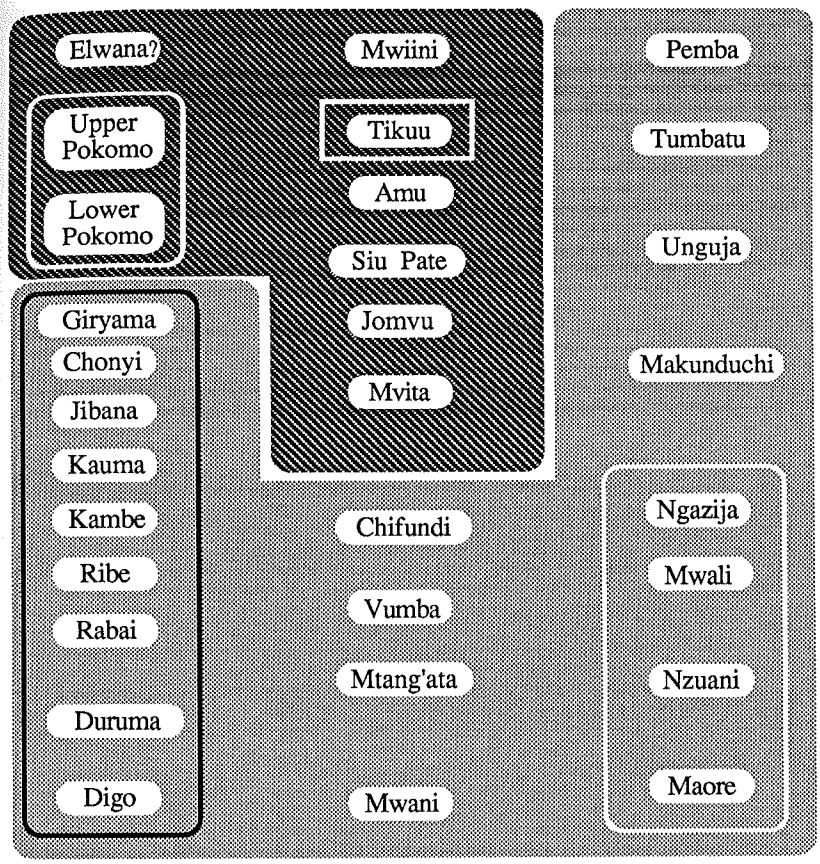
§14.7. *l-loss (§4.2.2 and Chart 36). There is some evidence that *l-loss in Sabaki was preceded by the shift of *l > y, and that this shift preceded in stages beginning before *u, a stage we still see operating in Pokomo. Though loss is ultimately involved, the intermediate shift does subgroup Swahili and Comorian, the exception being Mwiini, which is conservative. *l-loss has also occurred in Class 11; the facts here match those for general *l-loss.

§14.8. *g-loss (§4.2.3 and Chart 37). Though loss is not a reliable defining feature for subgrouping purposes, we want to note the fact that in both Comorian and ND (Tikuu, Siu, Pate, Amu), PSA *g has been lost in a lot of lexis, in contrast with SD and Mijikenda, which are conservative for the most part. In light of other shared features, this may be significant and indicate genetic connections. However, the further fact that Mwani, Elwana, and Mwiini attest regular loss of *g makes no sense in terms of any other patterning that we have noted. Here we may well be looking at independent innovation leading to loss.

§14.9. Sabaki *j (§4.2.4 and Chart 38). Swahili as a group can be characterized by the attestation of the palatal stop, /ɟ/, or a reflex clearly derivable from /ɟ/, as the reflex of CB *j. However, it is not clear whether this feature is the result of an innovation or a retention, because we are not certain of the phonetic value of CB *j. If it were certain that CB *j was

Chart 35

Subgrouping *W/___[-round]



Key

*W : w

*W : υ

*W : w > Ø

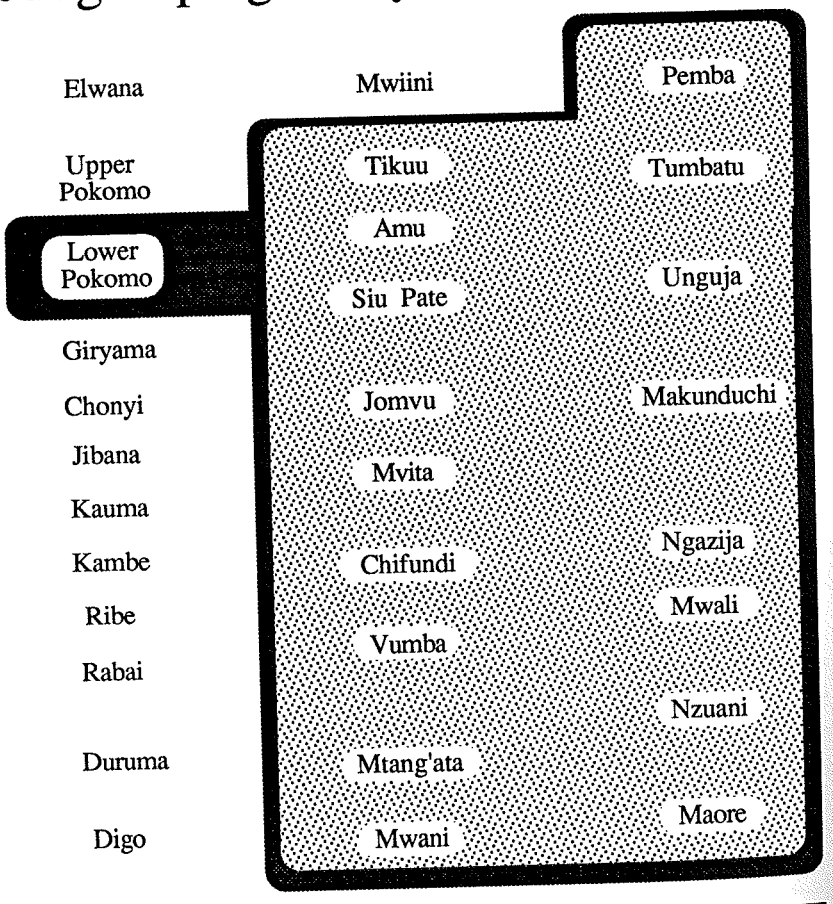
*W : w ~ Ø ~ (h)

*W : υ ~ β ~ v

*W : υ > v

Chart 36

Subgrouping *l > y/___u



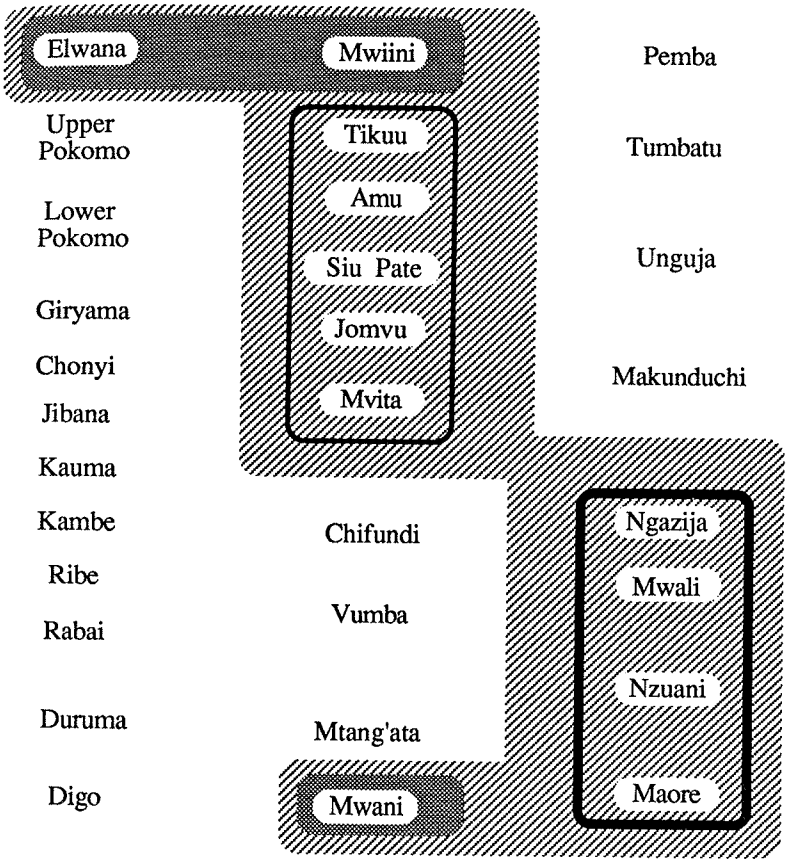
Key

*l > y/___u [-stress] (e.g., Classs 11 *lu- /___C-stems)

*y > Ø/___u [-stress]

Subgrouping *g-loss

Chart 37



Key

*g-loss attested

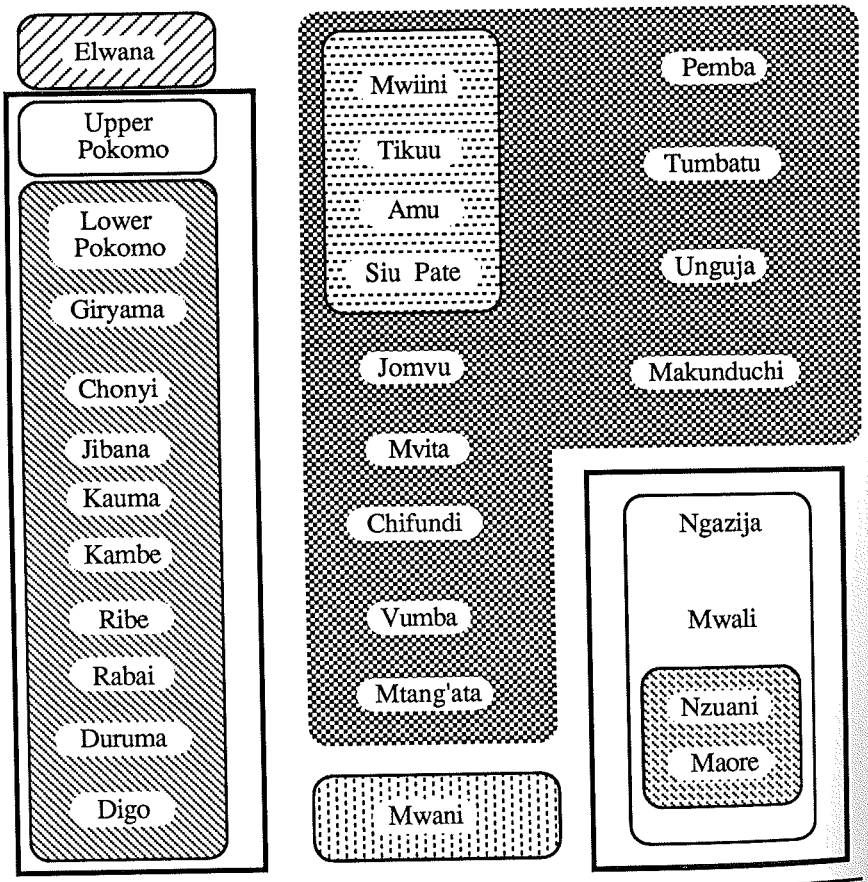
*g-loss regular

*g-loss (irregular)

*g-loss (some exceptions)

Chart 38

Subgrouping CB *j : *[+affricate] : * ʄ



Key

*j > dz

*j > *ʄ > ʄ

*j > *ʄ > ʄ ~ ž ~ y ~ Ø

*j > *ǰ > ǰ

*j : [+affricate]

*j > *ǰ > ž

*j > z

*j > ʄ

an affricate, namely, /ǰ/, then we have an innovation (CB *j : > ǰ > Swahili ɟ). Nonetheless, geographic distributions are helpful. It is only the Swahili group and Mwani (Schadeberg, pers. comm.) that attests CB *j : ɟ. Comorian, Mijikenda, and Pokomo have an affricate corresponding to CB *j: /dz/ in Lower Pokomo and Mijikenda, /ǰ/ in Upper Pokomo. In Comorian, presumably /ž/ in Nzuani/Maore derives from a Proto-Comorian affricate *ǰ, as does Ngazija /ǰ/, the main evidence for positing an affricate in Comorian. Elwana has CB *j : z, a correspondence which is unique for Sabaki but not for other NEC languages. Mwani conserves /ɟ/. But it may be significant that ND_2 (Tikuu, Siu, Pate, Amu), Mwiini, and Comorian all attest lenition of CB *j: ND *ɟ > ž > y > (Ø) and CB *j : Comorian *ǰ > /ž/ (in two dialects). An alternative interpretation of the Comorian lenition of CB *j : /ž/ may indicate evidence for an earlier Comorian *ɟ rather than *ǰ. The former is more likely to weaken to [ž]. This would require the assumption that while part of Comorian weakened, the other strengthened, CB *j > *ɟ > ǰ (Ng). This would then support further Swahili and Comorian connections. However, at most, all we want to say here is that the only clear subgrouping we get is Swahili and Mwani (with reconstructible *ɟ) versus the rest of Sabaki.

§14.10. Bantu Spirantization

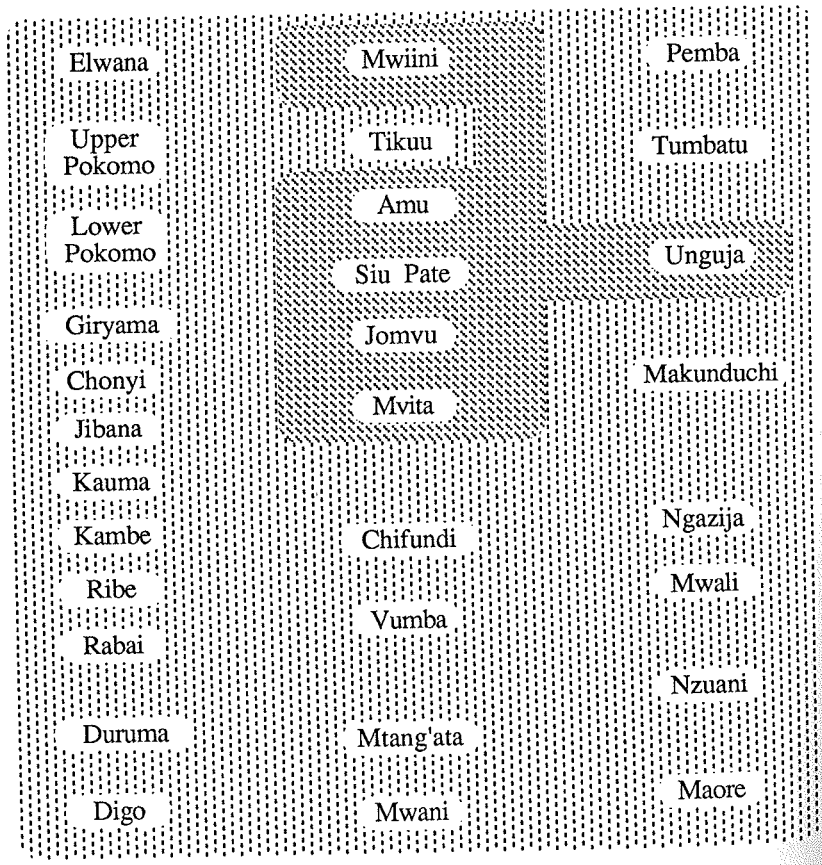
§14.10.1. Spirantization (CB *k : š/___*i̯) (§5.1.4 and Chart 39). PNEC, like most other East African Bantu groups except Central Kenya and West Tanzania, underwent spirantization before the [+HV] vowels. The most obvious result of this was the appearance of fricatives (§5.1) in all NEC languages. There is, however, one respect in which the reflexes of Spirantization in Sabaki differ from the rest of NEC. We have argued that in early PSA, as part of Bantu Spirantization, CB *k became š/___ i̯, and that the resulting palatal later merged with /s/ (< *t/___ i̯) in most Sabaki languages. The clear exceptions, where the palatal is kept, are Unguja and all of ND, except Tikuu. The other SD have a rather mixed situation, some lexical items having the palatal, others the alveolar.

For subgrouping purposes, however, we cannot make too much of this, because the same set of arguments which set up *š for Sabaki could be used at the PNEC level instead, and so on up the genealogical tree. Nevertheless, the distribution of the reflexes /š/ and /s/ do have the effect of setting off much of Swahili, with the exceptions noted, against the rest of Sabaki. ND and Unguja retain a palatalized reflex of original *š; the rest innovate by merging *š and *s. That Unguja shares this feature with ND may reflect the same motivation and source that explains the large amount of ND lexis in Unguja (Chap. 3, §4.7).

§14.10.2. Coronalization (PSA *f/v > s/z : CB *p,b/___*i̯) (§5.1.1, 5.3 and Chart 40). Mwiini, Amu, Tikuu, Siu, Pate (ND_2), and Comorian share the articulatory shift in which PSA *f/v > s/z before the high front vowel. This includes the inherited verbal causatives in which PSA *fy/vy > s/z. This innovation not only links ND_2 and Comorian, but also differentiates part of ND from SD (already seen in §13.3) and the rest of Sabaki. Because SD and the rest of Sabaki are retentive here, this isogloss does not serve

Chart 39

Subgrouping Spirantization *k/___*į



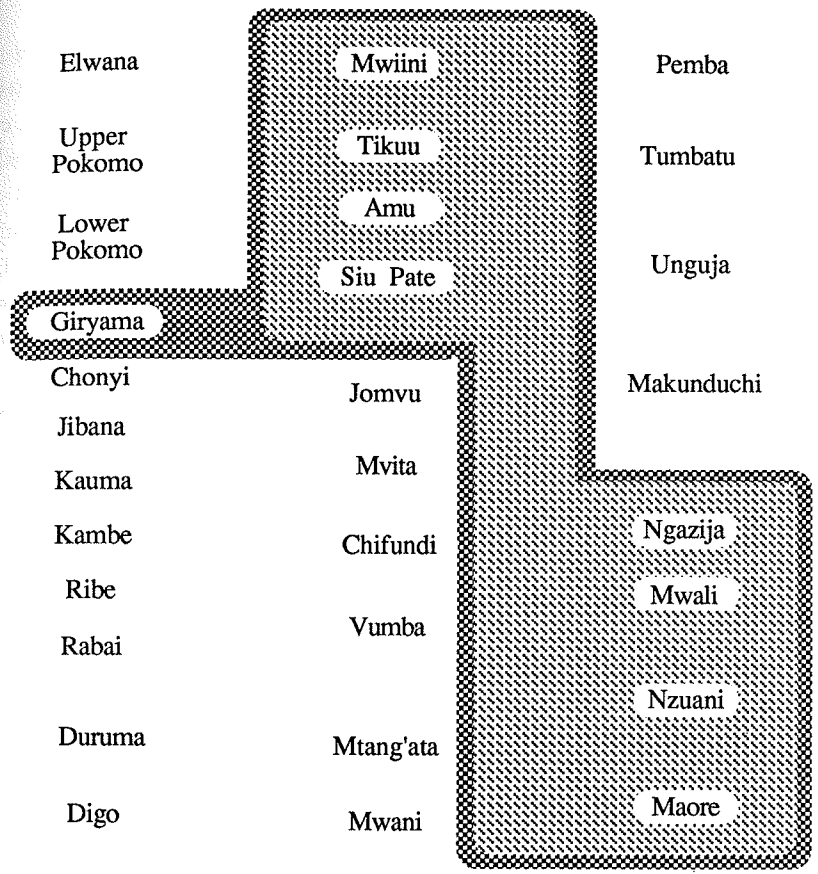
Key

*k : š / ___ *į CB *nkíngò, *-kéékį-, *-yókì

*k : s / ___ *į CB *nkíngò, *-kéékį-, *-yókì; *t : s / ___ *į

Chart 40

Subgrouping *f/v-coronalization



Key

CB *pį/bį : PSA *fį/vį > si/zi

PSA Causatives *fy/vy > š ~ s / ž ~ z ~ dz

to subgroup them, other than negatively. PSA causatives *fy and *vy show the same innovation and are shared by the same languages, including Giryama. The attestation in Giryama is probably a case of convergence.

§14.11. Strengthening (§6 and Chart 41). Strengthening is a complex multi-staged set of changes, partly phonologically, partly morphologically, conditioned. In its basic form it sets up stop and nonstop alternations in Class 5/6 (e.g., /β ~ w, ɓ ~ Ø, ɗ ~ l/, etc.). The conditioning environment is *i̧, which also governs the presence of stops in lexical items that otherwise would attest weakening or even deletion (e.g., -iba (SD) versus -iwa (Am, Ti, Si, Pa, LP, Mn), -ia (MK)). While the total body of evidence for this phenomenon presents challenges for analysis, with a residue of doubt about its extent in NEC and Sabaki, it is clear that both SD and Comorian share this shift. The specifics are quite diverse between the two, SD showing just the barest outlines of a once-rich set of alternations and Comorian a systematically active synchronic set of alternations.

Strengthening is the only phonological innovation linking SD and Comorian (but see §14.17 below). It is also the only case where the Swahili continuum is subdistinguished by innovation in SD. Of interest is the cross-classifying nature of the various isoglosses which Swahili (or parts of Swahili) and Comorian share (e.g., strengthening, *f/v-coronalization, *nd-retroflexion, behavior of *g) in contrast with this link between SD and Comorian. Of note, as well, is that Mvita, otherwise congruent with ND, shares in the shift.

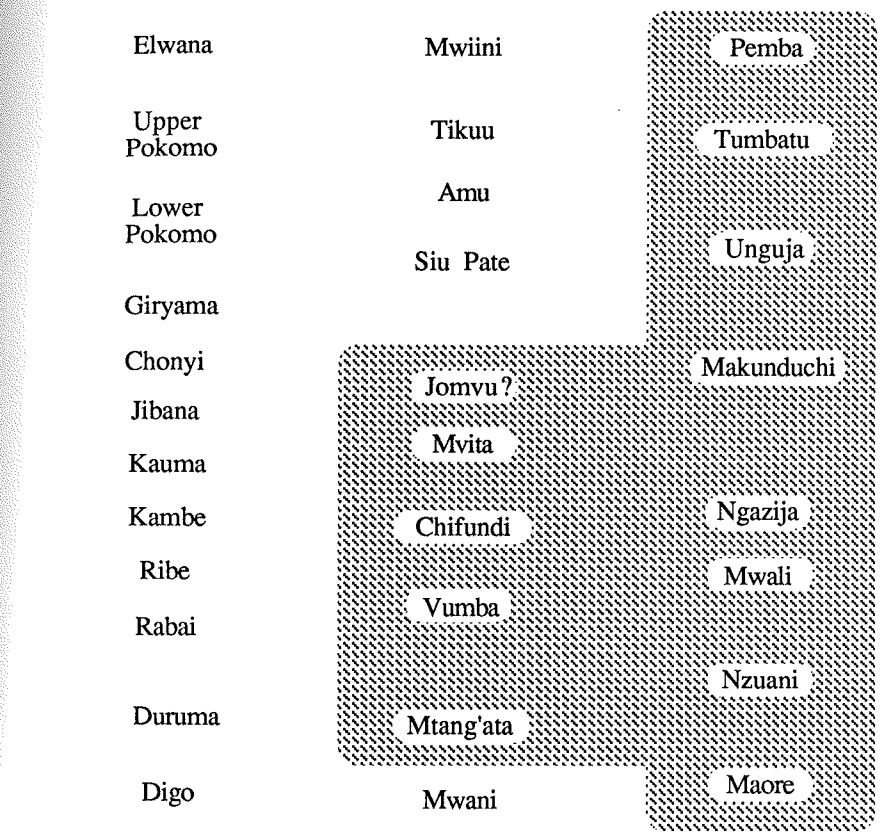
§14.12. Nasal-depalatalization (§7.1 and Chart 42). The palatal nasal *ny has regularly shifted to /n/ before *w, e.g., *-nywa > /-nwa/). Though our database only includes one lexical item attesting this phenomenon, it is shared by ND, Elwana, Mijikenda, and Comorian, and excludes SD and Mwani. Because this isogloss is based on a single lexical item, not much diagnostic value can be attached to it.

§14.13. Nasal-retroflexion: *nd > ndr (§ 8.1.2 and Chart 43). ND (Mwiini, Tikuu, Siu, Pate, Amu, Mvita) and part of the Comorian continuum share the retroflexion of *nd (e.g., /ndani/ [nd^{r}ani]). The lack of complete coverage here suggests that this is a late low-level phonetic rule; but given the other shared features between ND and Comorian, we view this as a possible historically shared feature, though one which may reflect contacts between some ND speakers and some Comorian speakers after the migration of pre-Comorians from the mainland.

§14.14. Nasal-palatal clusters (§8.1.4, 8.2.3 and Chart 44). Proto-Sabaki *nc and *nj become /(n)ts/ or /(n)t̪/ and /ndz/ or /nd̪/ respectively. The underlying shift here is from a palatal articulation to a dental one, shared by Lower Pokomo, Mijikenda, ND, Chifundi, and Comorian in contrast to SD, Elwana, Upper Pokomo, and Mwani, which are conservative. Further subgrouping is evident with Mwiini and the rest of ND, including Chifundi, which share the shift of *nc and *nj > /(n)t̪/ and /nd̪/. The effect is to add to the isoglosses distinguishing ND from SD and from the rest of Sabaki (see §14.3 and 14.11).

Chart 41

Subgrouping Strengthening

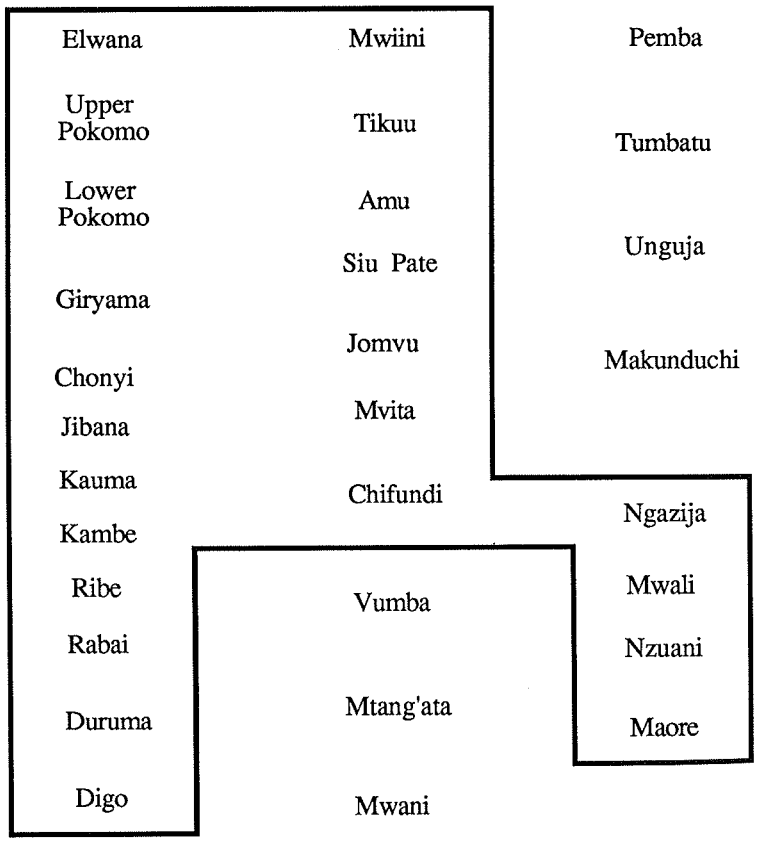


Key

Strengthening; cf. *-ịWa, *mwịWa, *ịWe ~ *jịịWe

Subgrouping *ny > n/___w

Chart 42



Key

——— *ny > n/___w

Chart 43

Subgrouping *nd Retroflexion

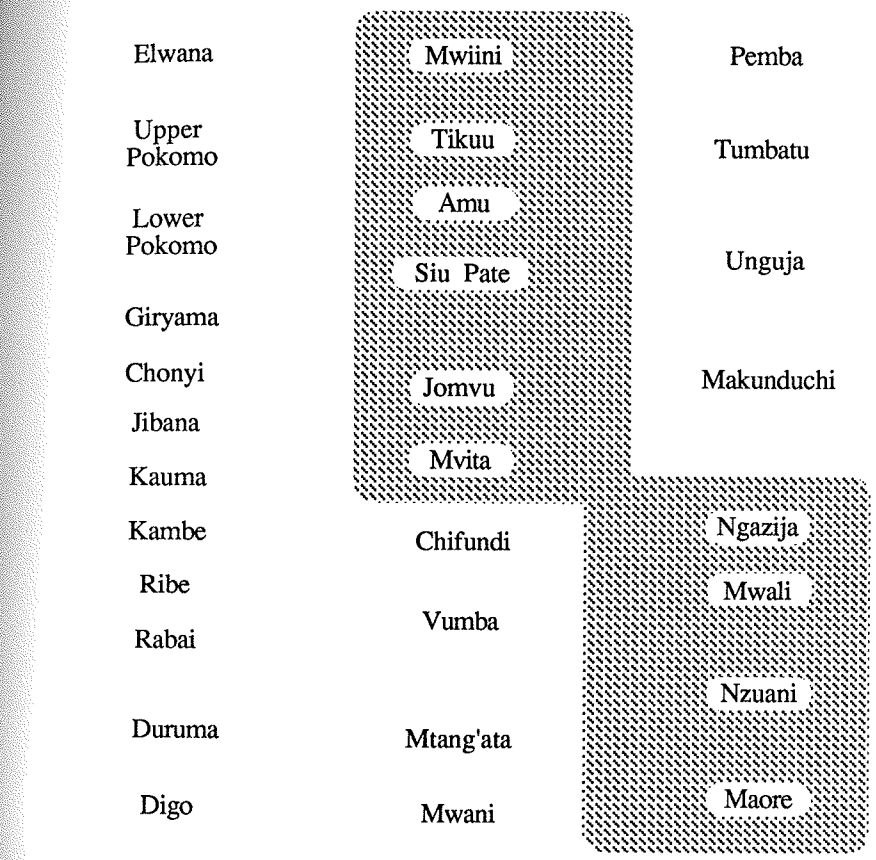
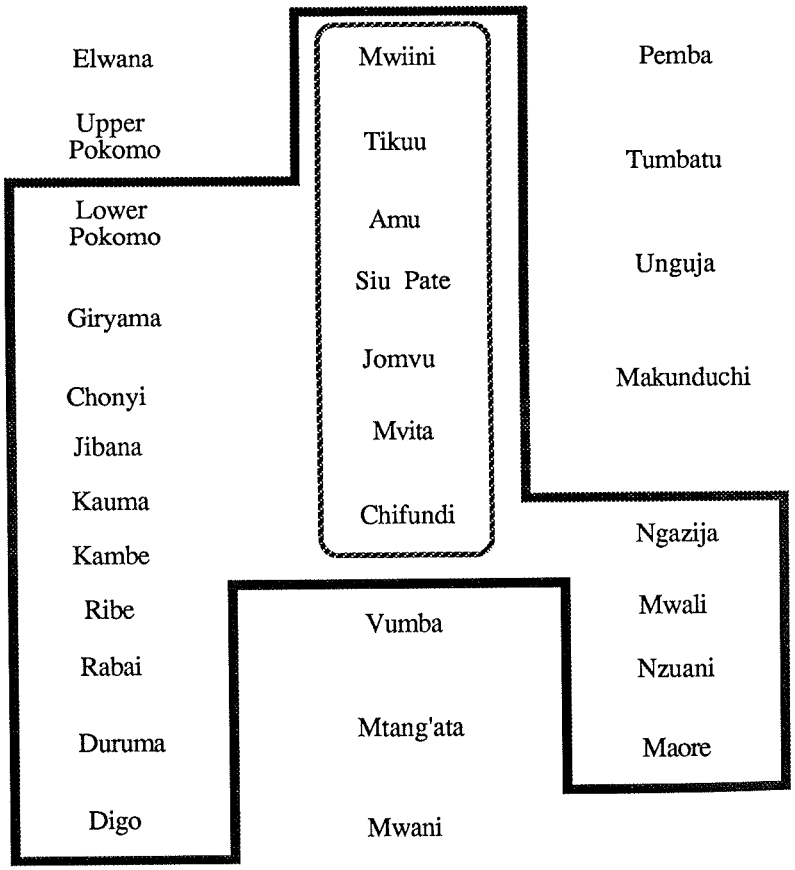


Chart 44

Subgrouping *nc and *nj



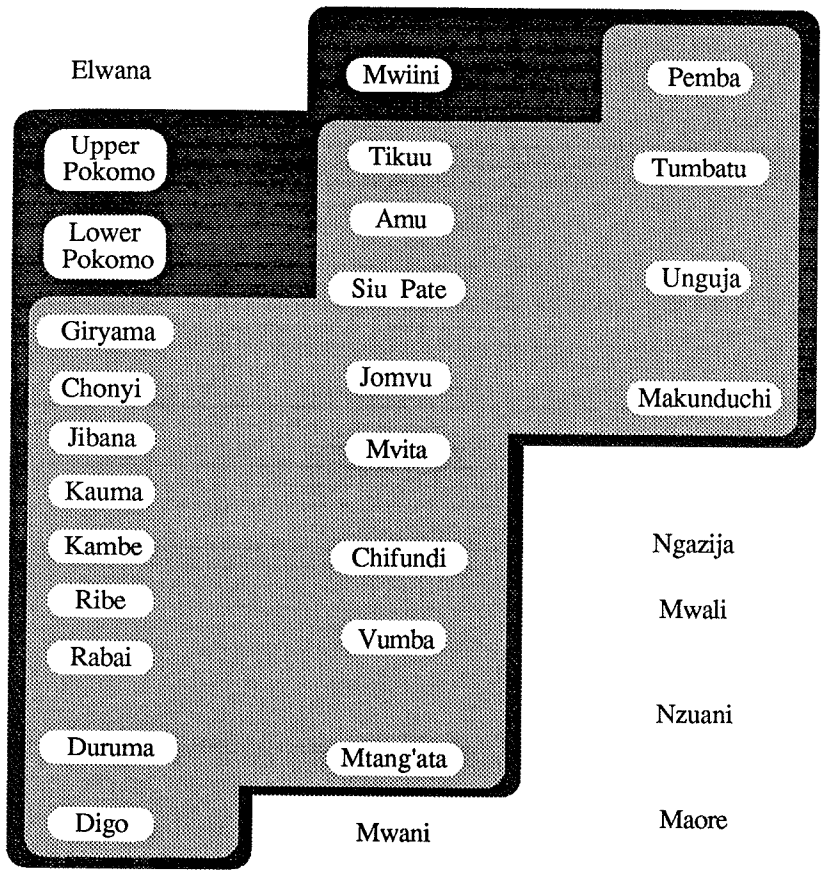
Key

*nc > (n)ts
*nj > ndz
([+palatal] > [+dental, -palatal])

*nc : (n)t̪
*nj : nd̪

Chart 45

Subgrouping N-devoicing/Aspiration



Key

*NC̥ > NC̥^{h}

*NC̥^{h} > C̥^{h}

§14.15. Nasal-devoicing/aspiration (§8.2.2 and Chart 45). PSA *NC̥ underwent a two-stage shift in which the nasal was first devoiced with the development of concomitant post-consonantal aspiration, followed by the loss of the highly marked voiceless nasal. The first stage is widely attested in NEC, specifically in Seuta and Southern Pare, and outside of NEC in Langi (F33), and Pogoro (G51), so it does not throw any special light on Sabaki interrelationships, or subgrouping, except negatively, insofar as Elwana is again isolated.

Comorian is also excluded, because the facts there do not synchronically mesh with this reconstruction, even though Maore does delete the nasal in such clusters. Ngazija, which otherwise voices postnasal stops, synchronically only shows nasal-deletion of an underlying nasal after pause (e.g., /nkuhu/ → [kuhu]/#___, and [ŋkuhu] /V___). The specific behavior of Comorian *NC̥ suggests that nasal-devoicing/aspiration is a late areal seriation that spread after Comorian separated from other languages. This is also supported by the linguistic analysis: rule-ordering indicates that nasal-loss must postdate *p/t-lenition, a feature that Comorian does share with Sabaki mainland languages.

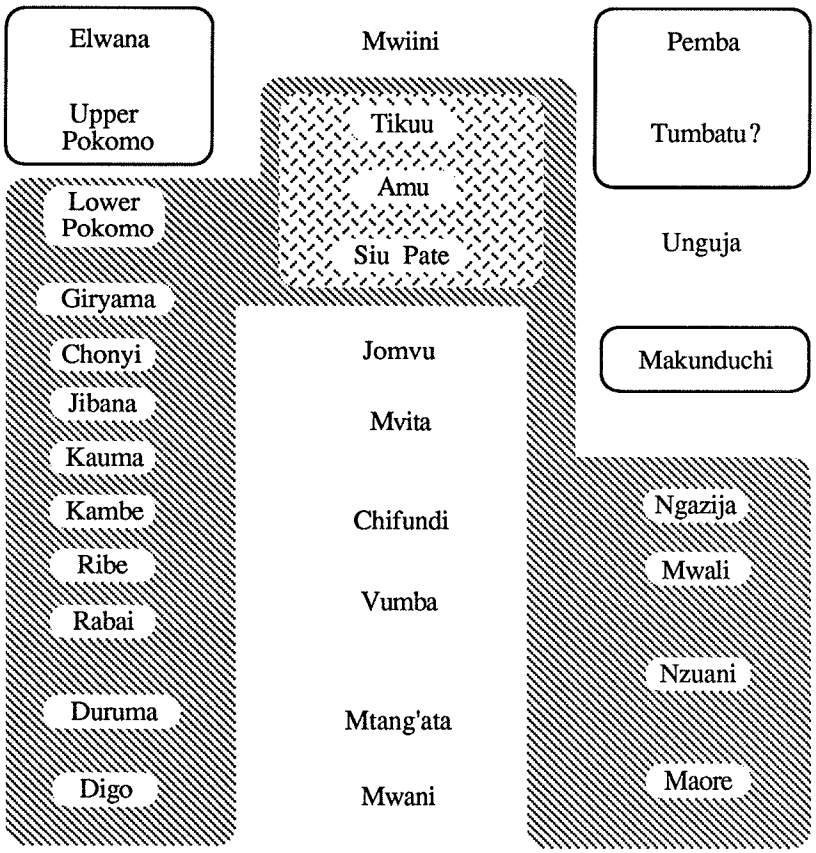
The second stage of this change, nasal loss (*NC̥^{h} > C̥^{h}), encompasses Mijikenda and all of Swahili, but not Pokomo and Mwiini, dialects on the periphery which still attest direct reflexes of Stage 1, namely, /NC̥/ [NC̥^{h}]. Although Stage 2 involves loss of a marked feature, it is nonetheless interesting to examine the distribution involved, because it is one of the few instances where all of Swahili (but not Mwiini) shares an isogloss with Sabaki mainland dialects. The fact that the nasal is still attested in older ND literature, e.g., *wantu* 'people', *yonte* 'all (Class 4)', *ntano* 'five (Class 10)' (see Knappert 1968, 1979), suggests that this took place in ND only in the last three or four centuries. Its innovation probably centered among prestigious Swahili speakers and spread to both rural Sabaki and Seuta speakers, in much the same way that animate concord is a Swahili innovation which spread to languages on the Swahili periphery (Wald 1975).

§14.16. Nasal-fricative clusters (§8.3 and Chart 46). Most of Sabaki has lost the nasal element of *ns and *mf sequences, with a few languages still retaining original PSA forms, such relics serving as our main evidence supporting the reconstruction of a nasal in these environments. Nothing of subgrouping interest hinges on these observations. For voiced nasal-fricative clusters, the picture is more complicated.

The behavior of *nz in Mijikenda, Comorian, Lower Pokomo, and ND is similar to that of *nj (see Chart 44), except that *nz has shifted to /nd̪/ only in Tikuu, Siu, Pate, and Amu (ND_3), thus helping to distinguish ND_3 from everything else. There is a scatter of attestation of loss in Pokomo, and parts of SD, but nothing that patterns in any clear way. The only apparent subgrouping, in the broadest sense of the term, is that most of SD, Chifundi, and the adjacent parts of ND (the Mombasa dialects) are conservative. Loss of the nasal before /v/ also tells us nothing; there is no obvious pattern, and it would appear that each case is one of independent loss.

Chart 46

Subgrouping *nz



Key

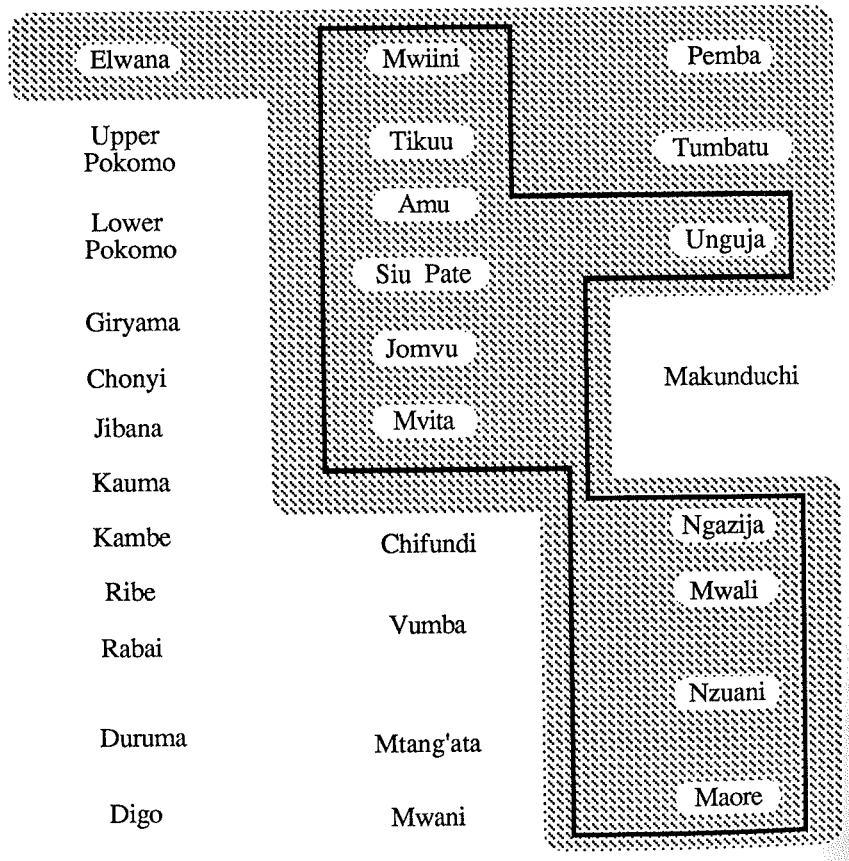
*nz > /ndz/ versus /nz/ [ndz] attested elsewhere

*nz > ndz > nd̪

*nz > Øz

Chart 47

Subgrouping Glide-deletion



Key

*-fwa > -fa

*lya > -la (-ǰa Mwiini, ɟa LP)

§14.17. Glide-deletion (§9.2 and Chart 47). Glide-deletion in Sabaki, in which *y and *w following certain consonants delete, e.g., *-fwa > /-fa/ 'die' and *-lya > /-la/ 'eat', is found in Comorian and some Swahili dialects. Though such glide-deletion or absorption involves loss of a segment, and is a possibly independently motivated loss, it is notably a feature shared by Comorian and ND, plus Unguja, and in some cases rural Zanzibari dialects. Elwana attests loss of *w after *f, thus falling in with the others, one of the few phonological changes it shares with other Sabaki languages along with the shift of *ny > n/___w (§14.12). None of this is highly significant nor diagnostic for Elwana, or, for that matter, any of Sabaki, since glide-deletion of this sort is a common NEC feature. Most of Seuta and Ruvu show deletion of *w after *f. However, *y-deletion after *l would appear to be a post-PSA shift; most of Seuta and Ruvu palatalize sequences of *ly, e.g., *-lya 'eat' > /-ja/. Mwiini (*ly > /ǰ/) and Pokomo (*ly > /ʄ/) also attest this, but we see no evidence of specific Seuta/Ruvu connections with them because of this.

§14.18. *mu-syncopation (§11.1). This is widespread in Bantu and NEC. It is not helpful in subgrouping, as it involves loss. The evidence indicates that it is an areal feature, which has probably been spreading from south to north relatively recently. It is attested in its full form in ND, Southern Mijikenda, Mwani, and most SD, and in an intermediate form in Northern Mijikenda, Lower Pokomo, Comorian, Chifundi, and Vumba.

§15.0. Subgrouping Summary

This section gives potential subgrouping affiliations based on phonological innovation, and loss (with the usual caveat). The list below should be viewed as suggestive of historical relatedness, a function of a mix of genetic and areal causation, of innovation and retention. Any subgrouping conclusions have to be judged against this measure. Thus for Swahili, while we can list features that distinguish it from other Sabaki languages, we have to recognize in what way they do so. In most instances the same features that "define" Swahili are also PSA or PNEC features. In some respects Swahili stands out only because surrounding languages have innovated. For example, Swahili (also Elwana and Mwani), has /p/ in contrast to a lenis labial in Pokomo, Mijikenda, and Comorian. This is a retention, and Swahili looks different because it has not shared in events that are responsible for the situation in the others. Of course, this is an important fact in developing some sense of the history of the language. Why, for example, did Swahili preserve *p; what does it reflect about the early Swahili community?

The following lists interpret the contents of this chapter. These can be viewed in several ways; our historical interpretation is set out in Chapter 5. These lists are intended only as a guide and should be used with the associated charts, with our reservations about specific points of the analysis, stated earlier, clearly in mind. What follows sets out only the most salient subgroupings. Note that neither Elwana nor Mwani figure in any important ways in the subsets below (§15.1-15.11). Like Swahili (mainly Unguja) they retain PSA or PNEC features, or are shifted in other directions through contact with different sets of neighbors, e.g., Mwani with Makua and other Zone P languages.

§15.1. Pokomo/Mijikenda/Comorian

*p-lenition	Charts 31 and 1 (also Chifundi, Vumba, Makunduchi)
*t-lenition	Charts 31 and 2 (also Chifundi, Vumba)

§15.2. Lower Pokomo/Mijikenda/Comorian

*k > [+palatal]/___i	Chart 18 (and 5) (not Giryama; includes Mwiini, Jomvu, Chifundi)
*(n)c > (n)ts	Charts 33 (and 6) and 44 (and 18) (see §15.2)
*nj > ndz	Chart 44 (and 11) (see §15.2)
*nz > ndz	Chart 46

This grouping is based on surface reflexes attested in the respective languages, in contrast to the subgrouping in §15.3, which is based on feature identity of the reflexes involved. Thus ND is excluded here but included below.

§15.3. Lower Pokomo/Mijikenda/Comorian/ND_2

*(n)c > [+dent, -pal]	Charts 33 and 44 (/(n)ts/ in Mijikenda, Lower Pokomo, and Comorian; /(n)t̪/ (via *(n)ts?) in ND, also Chifundi (see Charts 6 and 18))
*nj > [+dent, -pal]	Chart 44 (/ndz/ in Mijikenda, Lower Pokomo, and Comorian; /nd̪/ (via *ndz?) in ND, also Chifundi, see Chart 16)
*nz > ndz > nd̪	Chart 46 (ND_3 (not Mwiini); /ndz/ in Mijikenda, Lower Pokomo, Comorian; /nd̪/ in ND_3)
*ɲ > n/___w	Chart 42 (also Elwana, Chifundi)

§15.4. Comorian/ND (Mwiini, Tikuu, Amu, Siu, Pate)

*f/v-coronalization	Chart 40 (Giryama has a similar process /___ i̯V, see §5.3)
*fy/vy > *sy/zy	Chart 40
*nd > ndr	Chart 43 (includes Mvita and Jomvu, but in Comorian only Nzuani/ Maore)
Attested *g-loss	Chart 37 (also Elwana, Mwiini, and Mwani, where *g-loss is most regular)
Glide-deletion	Chart 47 (also Elwana and some SD, not all possible glides in Mwiini, Elwana, SD)

§15.5. Comorian/SD

Strengthening	Chart 41 (occasional/irregular attestations in other Sabaki and NEC)

§15.6. Comorian/Mijikenda

CB *j : [+affricate]	Chart 38

§15.7. Swahili/Pokomo/Mijikenda

*N̥Ç > N̥Ç^{h} > (Ç^{h})	Chart 45 (Pokomo and Mwiini attest / N̥Ç^{h} /, Mijikenda and Swahili attest /Ç^{h}/

§15.8. Swahili/Mijikenda

*N > Ø/____vl. stop	Chart 45 (not Mwiini, which preserves *N before voiceless stops)

§15.9. Swahili (ND and SD) and Mwani

CB *j : ɟ	Chart 38 (Swahili innovation if CB *j = affricate, or fricative; ND feature if ND /y/, /ž/, Ø < *ɟ)

§15.10. Swahili, ND versus SD (mainly ND innovations, or typological features)

*(n)c : (n)t̪ (ND) : (n)č (SD)	Charts 44, 6 and 18 (Chifundi attests /t̪/)
*nj : nd̪ (ND) : nj (SD)	Chart 44 (Chifundi attests /nd̪/)
*j : ž ~ y ~ Ø (ND$_2$) : ɟ (SD)	Chart 38
*W [+dent] (ND) vs. [-dent] (SD)	Charts 35 and 7 (Pokomo also [+dental])
š (ND) : s (SD) < CB *k/__ i̯	Chart 39 (exceptions: Tikuu /s/, Unguja /š/)
*g-loss (ND) : *g-retention (SD)	Chart 37

§15.11. Swahili (ND and SD)/Comorian

*fw > /f/, *ly > /l/	Chart 47 (versus some SD, Pokomo, Mjikenda)
*l > y (including Cl. 11)	Charts 36 (includes Lower Pokomo)
*l > Ø (including Cl. 11)	Charts 36

These all involve loss and thus are not reliable for subgrouping.

§15.12. The phonological definition of Swahili. We have only one feature among the myriad discussed in this chapter that can be called an innovation (§15.9). It defines Swahili phonologically. Swahili can be further defined typologically by its retention of features that contrast with innovations in other groups, in the same way that we distinguished NEC from neighboring languages earlier. Thus, Swahili doesn't undergo lenition, except for a small subpart of SD, whereas Pokomo, Mijikenda, and Comorian do (Chart 34). Nevertheless, it is interesting that what unites the various subparts of Swahili adds up to considerably less than the diversity that separates them, either because parts share in innovations with non-Swahili languages (e.g., strengthening in SD and Comorian, but not in ND), or because parts innovate separately (e.g., ND /(n)t̪/ versus SD /(n)č/) (see §15.10). The lack of definition here is parallel with that noted in our discussion of the lexical definition of Swahili (see Chap. 3).

§15.13. Mwiini, a Swahili dialect. Linguists have not always agreed about the status of Mwiini. The lexicostatistic analysis in Chapter 3 (§2.3) shows that Mwiini's highest percentages are with ND$_3$ dialects, but in an alternative analysis Mwiini shows less affinity to Swahili than does Mwani (Chap. 3, §2.3.1). Because lexicostatistical results for Mwiini

do not unequivocally define it as Swahili, we review here the phonological features, both genetic and typological, that define Mwiini as Swahili, and specifically as ND:

1. *c-dentalization *(n)c : (n)t̪ (Chart 33)
2. PSA *W : ʊ (Chart 35)
3. CB *j : *ɟ (Chart 38)
4. š < CB *k / ___ i̯ (Chart 39)
5. *f/*v-coronalization (Chart 40)
6. *nd-retroflexion (Chart 43)
7. *nj > nɖ (Chart 44)

It is significant that Mwiini shares the putative Swahili innovation of CB *j : *ɟ. The other features link Mwiini with ND. The simplest explanation for this sharing is that Mwiini is Swahili, rather than a Sabaki dialect on the order of Mijikenda or Pokomo, which, because of its proximity to ND, changed through contact in the direction of ND. Its divergent features, a mix of retentions and innovations, e.g., preservation of *l in unstressed environments and of *nz (Chart 36 and Chart 46), palatalization of *k before *i and the morphophonemics of Cl. 7/8 (§11.6-7), and preservation of the nasal voiceless aspirated stops (Chart 45) developed after it split from ND.

§15.14. Vumba and Chifundi. As already noted (Chap. 1, §1.2.1), the status of Chifundi as ND or SD is open to interpretation (see Chap. 5, §7.6). Vumba and Chifundi (and, to a lesser extent, also one Makunduchi dialect) behave unlike other Swahili dialects. Here we set out relevant phonological isoglosses that will clarify what is happening in these two dialects (see also Chap. 5, §13.6):

Mijikenda-like features

1. Chifundi, and Vumba (and Makunduchi) share *p-lenition (Chart 34); Chifundi and Vumba share *t-lenition (Chart 34). These features form part of the definition of Mijikenda, Pokomo, and Comorian, and are otherwise unattested in Swahili.
2. Chifundi and Vumba have some lexicalized traces of labialvelarization (e.g., -bwa < -gwa 'fall'), as does Mtang'ata. But they do not have this as a productive process, such as characterizes Mijikenda, thus such traces most likely stem from nongenetic sources, as they do in Mvita.
3. Neither Chifundi, Vumba, nor Makunduchi shows signs of glide-deletion whereas a major portion of Swahili (mainly ND) does (Chart 47); neither does Mijikenda or Pokomo. But glide-deletion is somewhat variable in SD in general.
4. Chifundi attests *ɲ > n/___w (Chart 42); so do Mijikenda, Comorian, and Pokomo. But so also does ND, with which Chifundi shares other isoglosses.[55]

[55]This isogloss is based on a single lexical item (*-nywa 'drink') and is therefore problematic.

5. Chifundi has palatalization of *k and *g before a front vowel, as have (Southern) Mijikenda, Lower Pokomo, Comorian, and Jomvu (Chart 32).

Swahili features (innovations and retentions)

1. Mijikenda, Lower Pokomo, and Comorian have shifted *(n)c and *nj/nz to (n)ts and ndz, respectively. Vumba and Makunduchi retain the PSA/SD stage here (/č/ and /ǰ/) (Chart 44); Chifundi shows predominantly ND dentalization (*(n)c/nj > (n)t̪/nd̪), except, like Mvita, it keeps /nz/ < *nz (Chart 46).
2. In Class 5 V-stems, Chifundi, Vumba, and Makunduchi do not show the /dzi-/ characteristic of Mijikenda, Lower Pokomo, and Comorian, but have the Swahili situation, with derivatives of PSW *ǰ.
3. Chifundi and Vumba have undergone, and Makunduchi is undergoing, loss of *l. This characterizes Swahili and Comorian, not Mijikenda. This is also true of *l in the Class 11 prefix.
4. Chifundi and Vumba consistently retain the PSA *į of vowel-initial verb-stems (e.g., *-įjua 'know'), as does Makunduchi in some items. Mijikenda has consistently shifted this to different vowels.
5. Mijikenda has lost PSA *W in all nonprenasalized environments. Although Chifundi and Vumba have a few borrowed items with *W deleted (e.g., -ili 'two'), in the vast majority of items it is kept.

Nondiagnostic features. The following are only a sample:

1. Chifundi, Vumba, and Makunduchi share nasal-devoicing/aspiration, and nasal-loss with both Swahili and Mijikenda (Chart 12)
2. All three share some form of *mu-syncopation with Swahili and Mijikenda.
3. All three have *k > s/___ į. The only Sabaki dialects to consistently have /š/ in this context are Mwiini, Siu, Pate, Amu, Jomvu, Mvita, and Unguja. So, in this respect, Chifundi, Vumba, and Makunduchi behave like Mijikenda and the rest of SD.

Both languages share features with both Swahili and neighboring Mijikenda dialects, and in a sense are transitional dialects, but this does not address the question of genetic affiliation. If phonology were our only set of criteria, it might be difficult to classify Vumba and Chifundi. For example, we can't unequivocally identify an innovation that uniquely defines Swahili as Swahili. The closest we come is the correspondence CB *j : ǰ, where /ǰ/ is a non-affricated palatal. If we assume that this specific articulation (non-affrication) is an innovating feature for Swahili, then Vumba and Chifundi fall clearly into the Swahili subgroup. If Swahili /ǰ/, on the other hand, represents Sabaki articulation of the palatal, then we don't have an innovation, but a PSA retention. Practically speaking, we can resolve the issue by relying on surface distribution and counts of typological features, and conclude that Vumba and Chifundi are Swahili dialects because they have more Swahili-like features (= PSA-like features in some cases, even NEC features in other cases) than they do Miji-

kenda-like features, and that they have the latter features because of contact mechanisms operating. This solution, of course, is arbitrary, unless we can assign chronological priority to specific features, but this is most difficult to do in a non-arbitrary manner. How do we weight lenition, for example, against any Swahili-like feature that Chifundi and Vumba might also attest? We have more specific proposals in Chap. 5, §13.6.

Tone (and Stress) in Sabaki

by Gérard Philippson

§16.0. Introduction

In order to reconstruct the evolution of tone systems in Sabaki satisfactorily, we need to assess the validity of the available data. Apart from the Swahili dialects which have only demarcative stress on the penult and are of no use for this section, reliable information on the prosodic systems of Sabaki languages is limited to Comorian (primarily Nzuani, then Mwali and Ngazija), Mwiini, Mwani, Mijikenda (especially Digo, with some much more restricted, but vital, data, on Giryama), and rather less satisfactorily Lower Pokomo. Of Elwana we know next to nothing.[56]

Fortunately, we can also draw on our rather considerable knowledge of the Seuta languages and of Pare, all closely related to Sabaki within NEC; the Shambala tone system is well known since Röhl's (1911) pioneering study, with a fresh theoretical perspective by Odden (1982, also 1985), while Zigua (Kisseberth and Kenstowicz 1988, Kisseberth and Mochiwa 1988) and Bondei (G. Philippson, forthcoming) have been recently studied. East Ruvu languages have lost all prosodic distinctiveness, as has Swahili; while the West Ruvu languages, which have not (Kagulu, Lugulu, probably Sagala), are barely known tonally, and the better known Gogo is too remote to shed much light on Sabaki. Pare still maintains a rather conservative approximation to the ancestral eastern Bantu system, and is useful for reference purposes. The other languages of the area, Chaga and the Central Kenya lan-

[56]The Elwana data provided by Maddieson and Sands arrived too late to take into account here. There is some evidence that certain SD dialects (Makunduchi) might also have a partly distinctive accent.

guages, are distant genetically and of a quite different tonal type. We will first briefly sketch the prosodic systems of the individual Sabaki languages, and then attempt a diachronic reconstruction.

Some initial clarification of terms is necessary. "Pitch" used on its own refers to phonetics, and is self-explanatory. "Tone" here means the distinctive use of pitch, *where more than one contrast in pitch can be found underlyingly within one morpheme.* "Pitch-accent" or just "accent" means distinctive but restricted use of pitch, with no more than one pitch contrast per morpheme—in some cases not more than one such per word. "Stress" means prominence, generally of loudness or length, but even of pitch *if the language does not otherwise make use of pitch contrasts,* as in Swahili, for instance, where the prominent syllable in a word is both longer and higher in pitch than other syllables. In other languages, however, like Mijikenda or Bondei, where the penultimate syllable is always quite prominent, this prominence is due to length only, pitch contrasts being assigned distinctively, as we will see.[57] The same can be said of Lower Pokomo, where this prominence is however rather less audible.

All Sabaki languages with distinctive pitch contrasts can be said to represent pitch-accent systems, as defined above. These languages can be considered as having only one underlying high (H) per morpheme; or in other terms, one syllable—at most—per morpheme bears a mark. (In the framework of Halle and Vergnaud (1987) that we follow here, adapting it slightly in accordance with eastern Bantu facts, a mark can be symbolized by a lexical line 1 asterisk, i.e., assigned underlyingly before the application of any rule constructing metrical constituents.) The same is true of Seuta (although Shambala presents some complications) which belongs to a prosodic continuum with Sabaki, as will be seen below. Pare, on the other hand, is a "pure" tone language, where certain morphemes can have two underlying H. Comparison of disyllabic noun stems makes the situation clear: thus Pare distinguishes in all contexts /mbújì/ 'goat' and /nkúkú/ 'chicken'. Sabaki and Seuta languages do not differentiate the patterns of these two items, but treat them alike as members of one and the same prosodic class; cf. Shambala /mbúzí/ and /ngúkú/, Bondei /mbuzi ída/ 'that goat' and /nguku ída/ 'that chicken', Nzuani /mbúzi/ and /nkúhu/, Pokomo /mbûzi/ and /nkûku/.[58]

Although this is to a certain extent true of most eastern Bantu languages, the prosodic systems of Sabaki and Seuta make reference to an unusual degree to the syntactic context (corresponding more or less to Selkirk's (1980) definition of the "phonological phrase"), to the extent that, in Comorian and Mijikenda at least (as in Bondei and Zigua), the pitch contour of most items cannot be established in isolation, but only by reference to surrounding items. This is not true of Pare, nor of Shambala. Consider:

---

[57]The penult can also receive H pitch due to accent assignation rules. Also, there is a final cadence when both penult and final are not accented, which makes the final sound rather lower in pitch than the penult and previous syllables.

[58]Bondei nominals cannot usually be distinguished by their tonal pattern in isolation.

Nzuani	múntru 'person'	mwána 'child'
	muntru mzúri 'good person'	mwana ḿzuri 'good child'
Digo	nyama 'meat'	nyumba 'house'
	nyama yangú 'my meat'	nyumba yâ̂ngu 'my house'

Compare with Mwiini:

naxsoomá 'I am reading'	naxsomaa čuwó 'I am reading a book'
naxsóoma 'he is reading'	naxsomaa čúwo 'he's reading a book'

Now we consider each system in turn.

§16.1. Pokomo. The Pokomo system seems to be the simplest, which may be due to its being little known, and also the most divergent of all Sabaki systems—that much is relatively clear.[59] There are three accent groups in disyllabic stems, e.g.:

Pokomo (a) mašóká 'axes' (b) magego 'teeth' (c) máwee 'breasts'

For (c), the pattern is realized as such in isolation, but in context we would have /máwée/; in general, Hs always go in pairs in Pokomo, but a HHL pattern is always reduced to HLL before pause.

The situation for longer stems is not clear, but judging from what happens in verbs in the infinitive, it seems the tripartite pattern really means (a) final H, (b) no H, and (c) penult H. For instance, although the verb /kúona/ 'to see' has a H on the prefix (in context also on the root syllable), the derivative /kuónea/ has it on the root syllable and, in context, also on the penult, thus /kuónéa/. Since in /máwee/ and /kúona/ above, where a H is realized on the prefix, it is clear that the determinant must be the first stem syllable (otherwise we would expect the other nouns in (a) and (b), or a verb such as /kubiga/ 'to hit', to have a H prefix), in the same way a H antepenult must be determined by the penult. It thus seems that H pitches in Pokomo always appear in pairs, one on the affected syllable itself and the other one on the previous syllable, so that the lexical forms of the above words can be written: /mašoká, magego, mawée, kuóna, etc/. If this generalization is proved true (it is certainly true in verbs) and the nonfinal H is always on the penult, even in cases such as /kuonéa/ where its underlying origin is clearly from the root syllable, then the most convincing account would be to construct right-headed unbounded metrical constituents in all words with a nonfinal H, with the proviso that the final syllable is always nonmetrical. For words with a final H, the same construction rule holds, except that the final syllable is marked and so is not subject to extra-metricality. This approach implies that certain words will not construct any constituents at all, which appears to be a necessary conclusion in light of the Comorian data, dealt with below. Thus:

[59] Another exception is Mwani, which does not seem to be part of the Sabaki spectrum at all (see §16.4).

(.	.	*)				(.	*)	
(*	*	*)	*	*		(*	*)	
ma	šo	ka	ma	ge	go	ma	we	e

Phonetic realization rules will then apply to interpret the accent mark as a H pitch, then associate it to the previous syllable as well, delete the final H of a HHL pattern before pause, etc. The double association rule is post-lexical, since it regularly applies across word (and phrase?) boundaries, e.g., /ngombé ntahu/ 'three cows', which in context would be /ngombé ntáhu/, where the final H on the noun comes from the first stem syllable of the adjective. In case the marked syllable happens to be word-initial and the word stands before pause, the H is realized as a fall on the marked syllable itself (e.g., wêe 'breast'). Conversely, another H deletion rule will apply to the H of a marked final syllable when the word stands before another word whose first syllable bears a H:

míhi míhahu 'three trees', but míhí miwíí 'two trees' (lexical form /mihí/)

On the other hand, a penult H will not delete under the same conditions:

nyímí nkúu 'big tongues', not *nyimí nkúu (lexical form /nyímí/)

Furthermore, a "doubled H" can replace a final lexical H: /kítswá cúmu/ 'hard head' (lexical forms: /kitswá/ and /cúmu/), where the final H of /kítswá/ has been disassociated according to the "final H deletion rule," only to be replaced by the double association (across word boundaries) of the following H. All this suggests that most H association rules in Pokomo seem to be post-lexical, and this should be formalized in a complete account of the language (for which we do not have enough information to attempt here).

§16.2. Mijikenda. Accent in Mijikenda is more complex and must be studied in a syntactic context. Digo disyllabic noun stems, as in Pokomo, exhibit a threefold division, but this is not readily apparent in isolation, where only two patterns are found:

Digo (a) ŋombe 'cow' (b) kandé 'food'

In other words, disyllabic nouns appear with either a H on the final, or else with no H at all. But in context, it can be seen that category (a) in fact subdivides into two parts:

Digo kuona ŋombé 'to see a cow' and kuona nyûmba 'to see a house'

where the verb /kuona/, itself realized without a H, puts a H on the final syllable of /ŋombe/, but a falling contour on the penult of /nyumba/. In the same context, words of category (b) surface as /kuona kándé/.[60]

Thus pitch realizations depend on the underlying accent of neighboring words. The case of polysyllabic words is not entirely clear due to the paucity of data, but there seem to be three main categories only:

---

[60]We ignore phonetic adjustments here, mostly the fall of every sentence-final H and the rise on a L penult before final H.

(a) nguruwe	'pig'	kuona nguruwé	'to see a pig'
(b) čikarangó	'frying-pan'	kuona čikárángó	'to see a frying-pan'
(c) gunguhi	'bed-leg'	kuona gungûhi	'to see a bed-leg'

The fact that the number of categories is limited to three strongly suggests an analysis in terms of accented syllables, although the rather straightforward metrical pattern of Pokomo does not seem to apply here. Words of category (a) appear to be unmarked, since they offer no resistance to a H coming from the previous word (that the final H of /kuona nguruwé/ is due to the verb can be seen by considering the phrase /kugura nguruwe/ 'to buy a pig', where no such H appears). Words of category (b) are certainly marked, since they always have a H in isolation; and furthermore, the precise location of the mark can be ascertained from the examination of verbs with extended stems, where the only element liable to inherent marking is the verb root: extended verb bases divide into those with no H and those with final H, e.g., /kugurira/ 'to buy for' and /kuonyesá/ 'to show'. In fact, the noun /čikarangó/ given above derives from /kukarangá/ 'to fry'. In sum, it seems almost certain that the accented (marked) syllable in such words is the root, i.e., the first stem syllable, as in /čikårango/.

Now we know that /nguruwe/ is not marked and that /čikårango/ is marked on the first syllable. The remaining group, that represented by /gunguhi/, must then be marked on some other syllable, either the penult or the final. Taking it to be the penult allows us an important generalization, namely, that a H from a previous word will always land on the marked syllable of the following word, or, in case this word has no such syllable, on the final, thus:

kuồna čikårango	>	kuona čikárangó
kuồna gungûhi	>	kuona gungúhi
kuồna nguruwe	>	kuona nguruwé

In the first example, the final H comes from the root syllable. There are obvious differences between the first two forms and those above. When a H is placed on a word already containing one, the intervening syllables assimilate in pitch. Also, a H landing on a penult followed by a L is realized as falling, so /čikárangó/ > /čikárángó/, /gungúhi/ > /gungûhi/, etc. It is thus apparent that there are no finally accented words in Digo, since they would be indistinguishable from unmarked words. What is really interesting is that when the *penult* of a word bears the mark, H cannot appear on the word itself. On the other hand, if the antepenult or a previous syllable bears the mark, the H will appear on the final; so we can deduce that H needs at least two syllables after the marked syllable to associate. The general principles of pitch association in Digo are thus the following:

1. The last H of a word associates to the final syllable, provided there are at least two syllables after the mark; otherwise it remains unassociated.

2. Any previous H will associate to the next marked syllable; if there are several syllables within a word bearing a H after this, all intervening syllables will appear as H.

Thus, in the case of disyllables, category (b) should not exist if the H cannot associate, unless two syllables are present after the marked syllable. And indeed such stems are anomalous and correspond for the most part to CB stems of the shape -CVVCV, so that at an earlier stage of Mijikenda, when long vowels were phonemic, the count must have been by mora and not by syllable.[61] Synchronically, the exceptional character of these stems must be indicated in the lexicon.

Giryama, although little known, appears to be just like Digo, with the exception that the last H of a word does not move right to the *final* but to the *penult,* e.g., /kuona kirónda/ 'to see a wound', but /kutsola kironda/ 'to choose a wound'. In the same way, a previous H does not move to the next marked syllable, but to the one before it:

kuona máfuha 'to see oil' (underlying /mafůha/)

Again, exceptional disyllabic noun stems do not bear a H on the final but on the penult: /kutsola kázi/ 'to choose work' (compared with /kuoná ! kázi/). As can be seen, two consecutive H belonging to two different constituents are separated by a downstep in Giryama, but apparently not in Digo.

In sum, Digo has undergone a shift by one syllable to the right, as compared with Giryama (and Bondei and Zigua, which are practically the same prosodically). Due to this difference, Giryama has a *fourfold* division of disyllabic stems (including the exceptional group). The fact appears clearly in context:

kuona kirónda mačéro 'to see a wound tomorrow'

kuona muhóho má!čéro 'to see a child tomorrow'

The word /muhoho/ 'child', undistinguishable from kironda when it stands in final position, displaces a H on the following word, and so must have the underlying representation /muhohǒ/. Digo has eliminated this category.

A strange characteristic of Mijikenda is that, contrary to the situation obtaining in neighboring languages but common in Southern Bantu, voiced (nonprenasalized) obstruents block H pitch. We do not have space here to consider this phenomenon, which is rather common in the languages of the world, but unique—to the best of our knowledge—in East Africa.

§16.3. Comorian. Comorian appears to stand apart from the two previous languages in that no word may surface without a H, and noun forms may never have more than one H (with the qualifications given below). So in Nzuani, for example, disyllabic items in isolation are of two types only: (a) /dágo/ 'house' and (b) /nyumbá/ 'house'. Nevertheless, not all category (a) words behave identically when suffixed with the locative /-ni/: /dagóni/ 'in the house', but /gidzoní/ 'in the market' (in isolation, /gídzo/ 'market'). In some words the H moves to the next syllable, whereas in others it goes right to the final syllable.

[61]Exactly the same situation obtains in Seuta.

In Nzuani and Mwali, this type of difference in noun stems plays a very small role in the language, being restricted almost entirely to the locative (in /-ni/ or /-ju/; for further details, see Philippson 1988). In Ngazija, however, this difference between the two categories appears in several other contexts, e.g., /wajeni watíti/ 'small strangers', /wana wátiti/ 'small children'. In isolation, the two noun stems have the H on the penult: /wajéni/, /wána/. This contrast does not obtain in the other two dialects, e.g., in Mwali, which is otherwise almost indistinguishable from Ngazija, both segmentally and tonally, one would have /wajeni wátiti/ and /wana wátiti/.

The behavior of nouns such as /wajéni/ in Ngazija, which do not influence the pitch contour of other words in the phrase, whereas words such as /wána/ exert a "pull-back" effect on a following H, can best be accounted for if one takes these words as being non-marked underlyingly, and receiving their H (always penult) by a default rule. In Ngazija, this default rule applies only at phrase level (so post-lexically), whereas in the other two dialects it applies at word-level, i.e., lexically. For this reason, in Nzuani and Mwali it is impossible to know whether a noun is inherently marked on the penult or has just received a default accent, unless it can be put in the locative, as /gidzoní/ above. In Ngazija, on the other hand, the problem can be readily solved with the aid of any adjective whose own pattern is known.

This behavior is not limited to nouns but, in Ngazija and Mwali at least, also occurs in verbs, e.g.:

Mwali	nitsotsímba	'I'll dig'	and	ngamá 'hole'
but	nitsotsimba ngamá	'I'll dig a hole'		

where the fact that the H of the complement noun has not moved indicates that the previous word itself is not marked. Compare in another tense:

Mwali	karitsímbi	'we never dig'
	karitsimbi ngáma	'we never dig holes'

Here the H of the complement has been pulled back, showing the influence of a previous marked syllable. (In Nzuani, on the other hand, there is no apparent difference: /nitsotsimba ngáma/ and /karitsimbi ngáma/, showing that Nzuani also treats default accent in verbs as a lexical rule, whereas Mwali does it only for nouns, and Ngazija not at all).

A complete formal treatment would require more space than is appropriate here (see Philippson 1988 and forthcoming for a more extended treatment), but suffice it to note that:

1. Comorian is amenable to a metrical analysis in terms of *obligatory* unbounded feet (Halle and Vergnaud 1987:69-74) whose heads are underlyingly accented syllables.

2. These feet must be constructed on *two different metrical planes* (ibid.: 1987:66): on the first they are left-headed, e.g., in Nzuani:

*			*			*			*		
*	*	*	*	*		(*	*	*)	(*	*)	
wa	na	wa	ti	ti	>	wa	na	wa	ti	ti	'small children'

Then, line 0 (= the bottom line) of asterisks is copied on the second plane, *but not line 1*:

```
(*   *   *)  (*   *)
wa   na  wa   ti   ti
```

On this new plane, the feet are now identified as *right-headed,* and new line 1 asterisks are placed accordingly:

```
          *         *
(*   *   *)  (*   *)
wa   na  wa   ti   ti
```

There are two asterisks on line 1, but only one syllable is identified as accented on the surface, as shown above (wana wátiti). To obtain this result, we must first construct a left-headed unbounded constituent on line 1 and identify its head on a third line (line 2):

```
          *
          (*        *)
(*   *   *)  (*   *)
wa   na  wa   ti   ti
```

Then lines 1 and 2 must be conflated, which explains the nonappearance of a secondary accented syllable on the final (Halle and Vergnaud 1987:52):

```
          *
          (*        )
(*   *   *)  *    *
wa   na  wa   ti   ti
```

3. If there are no asterisks in the underlying representation, no constituents will be constructed on plane 1. Instead, a default constituent will be constructed on plane 2, according to the specific headedness instruction of that plane, namely right-headed; interestingly, the final syllable of the string will not be included, being *extra-metrical,* e.g., (Ngazija):

```
              *
             (*)
(*   *   *   *)
wa  ndru wa  dzi  ro   'heavy people'
```

Indeed, final syllables can always be considered extra-metrical in Comorian, unless they are themselves marked with an underlying line 1 asterisk. So that a more faithful rendering of the constituents of /wana wátiti/ might be:

```
*              *
(*   *     *)  (*)
wa   na    wa  ti   ti    'small children'
```

Comorian is thus not very different from languages such as Lithuanian, Classical Greek, Somali, or, even more closely, Tokyo Japanese, where basically metrical constituents are interpreted in terms of pitch. In Comorian, the line 1 constituents are not defined in terms of individual words but of prosodic phrases; high pitch is assigned to the accented syllable and spreads to the following syllables until the end of the prosodic phrase is reached (before pause the last two syllables are lowered, unless bearing the line 2 asterisk), so that a longer phrase for instance would have the following pitch pattern (underlyingly accented syllables are underlined):

Nzuani arenge biríká lá máží dribwávu 'he took a large water basin'

§16.4. Mwani. Mwani prosodic features are rather different from what we have seen above, insofar as its H assignment patterns are restricted entirely to word-level, and so somewhat simpler. Basically, the first syllable of (nominal and verbal) stems bears an underlying line 1 asterisk: right-headed unbounded constituents are constructed from left to right, and final syllables are extra-metrical. A H pitch is then associated with each syllable bearing a line 1 asterisk, e.g., /kukósomóla/ 'to cough':

```
      *                          *     *
*     *     *     *          (*   *)  (*   *)
ku    ko    so    mo    la  >  ku    ko    so    mo    la
```

In case the two H are adjacent, the second is deleted (malúlumi 'tongues'):

```
      *                      *     *
*     *     *            (*   *)  (*)
mu    lu    lu    mi  >  ma    lu    lu    mi
```

There are several complications that we will not treat here in their entirety. The main ones are:

1. In nouns, but not in verbs, the H cannot appear on the initial syllable, but is shifted one syllable to the right:

lulúmi/malúlumi 'tongue'
kongólo/makóngolo 'bone

unless the next syllable is the final one, in which case the H stays on the initial. If the stem is monosyllabic, the H will appear on the prefix:

sóka/masóka 'axe' kífyu/vífyu 'knife'

Some stems, mostly loans, have a line 1 asterisk on the penult, which is then always stressed:

sidádi/masidádi 'city'
masikíni 'poor'

Others have epenthetic vowels which do not count in the assignment:

ám(u)ri 'order' mét(u)ru 'meter'

2. Verbs can have a H on the first syllable (= prefix) if this has an underlying line 1 asterisk (níkubáli 'I agree'):

```
*                       *       *
*   *   *               (*)  (*  *)
ni  ku  ba  li   >      ni   ku  ba  li
```

Verbs in the subjunctive and imperative have their H on the final syllable. This is accounted for by attributing an underlying line 1 asterisk to this syllable, which thus ceases to be extra-metrical; the first stem syllable is not underlyingly accented in these tenses:

afungiriwé 'let him be tied down' lamukaní 'get up (pl.)!'

Other tenses either have one H regularly on the penult, and can then be considered as entirely without line 1 asterisks, the only constituent being a default one:

akiperekíwa 'he was brought (narrative)' atijengíre 'we didn't build'

or else the H is on the second stem syllable, with another one on the penult (if the two are adjacent, the second one is of course deleted); if the stem is mono- or disyllabic, the H will remain on the penult. What is particularly interesting is that "stem" here must include the object prefix (OP). This is reminiscent of languages such as Matuumbi (see Odden 1988):

aalombéka	'she's not to be married (neg. imperfect)'
aunisíngana	'you won't find me (id.)'
nilawíre	'I who went out (rel. perfect)'
arisínjire	'the one who cut himself (id.)'
asiláwe	'let him not go away (neg. subj.)'
asimuúlaye	'without killing him (id./neg. narrative)'

A possible analysis would be as follows. The first stem syllable, including OP, gets a line 1 asterisk, and *so does the final syllable,* which should produce a configuration such as:

```
          *       *
(*   *    *)  (*  *)
a    a    lo  mbe ka
```

Since no forms in Mwani ever surface with a H on both the final and a previous syllable, a rule must be posited which deletes the first constituent and redefines the second as left-headed. We will not try to formalize the rule here, but it might use the device of constructing separate metrical planes, as we saw in Comorian.[62] The reason for this apparently devious proposal is that subjunctive forms, which *are* finally accented, as we saw above, seem to have the pattern in question (the data unfortunately contain an apparently contradictory form):

---

[62]Indeed, the grammar of Comorian does seem to contain a similar rule reorganizing constituents the last of which is final (see Philippson 1988).

wapereké 'let them bring' wanipéreke 'let them bring me'

Furthermore, when there is no reason whatsoever to suspect a final accent (as in the infinitive, for instance), then the H surfaces on the *first* stem syllable, including OP:

kukútuma 'to send you' kutíperekéra 'to bring to us'

3. Finally, the case of the past is rather intricate; the majority of forms seem to correspond to the pattern just mentioned:

wampa 'they gave him' wapongóriwa 'they were born'
kanisíngana 'he found me' rivunjíka 'it broke'

However, forms with a Class 1 prefix (including 'I' and 'you' for some speakers, but only 'you' for others), and no OP *have no H at all* —the only forms in the language to behave that way, e.g.:

kapongola 'she delivered'
kuuka 'you went'
nisinja múti (~ nisínja) 'I cut a tree'

It is difficult to account for this exceptional behavior in the framework followed here. Either the forms in question have no line 1 asterisk and constitute an exception to the rule constructing default constituents, or they have an underlying final asterisk on line 1 (which would be consistent in our analysis with their behavior when preceded by an OP, see above), but exceptionally delete it unless another accented syllable is present in the form. This accented syllable can than be attributed either to the OP or to the subject prefix of classes other than Class 1. This at least is quite in keeping with the evidence of other Eastern Bantu languages, even quite remote ones such as Chaga, which generally tonally contrast the exceptional Class 1 /ka-/ prefix of the perfective (or narrative as the case may be).

§16.5. Mwiini. Finally, Mwiini presents certain curious features within a system by and large quite simple, and probably well on its way to the general Swahili pattern of penultimate stress only.

1. Nominals, including adjectives, without exception receive a H on the penult when standing at the end of a prosodic phrase (P-phrase):

aɬama yaa mvúɬa ni mawíingu 'the sign of rain is clouds'
wáantu wó̱te 'all people'
wapiši wíingi hangamiza mt̪úuzi 'too many cooks spoil the broth'

A monosyllabic noun or adjective is always accented on its only syllable:

nt̪í nii nkávu 'the earth is dry'
múuntu wene n̪umbaa né 'the man saw four houses'

The line 0 constituents might be considered unbounded and right-headed, with an extra-metric final, but the intersection of accent placement and length makes another solution more likely. Vowels can be underlyingly long or short in Mwiini, the only language

we have so far considered which has kept this inherited contrast on the surface. But additionally, not only can length contrasts only appear on the penult or antepenult of a P-phrase, not only is length reduced in antepenult position if the penult itself is heavy (i.e., has a branching rhyme), but conversely, a short final vowel will be lengthened if it stands in P-phrase antepenult or penult position (unless of course followed by a penult with a branching rhyme):

mwáana 'child'	mwana wa síimba 'child of lion'
mwana waa sába 'seventh child'	mwana wa píili 'second child'
mwáana ṯinzilee n̪áma kaa císu	'the child cut the meat with a knife'
cf. also nṯii nkávu 'dry earth', but	nṯí nkávu 'the earth is dry'

In the last two examples, the first form is within a P-phrase, lengthening applies, and there is only one accent. In the second, there are two P-phrases (indeed, two intonational phrases, as we will see shortly), lengthening does not apply, and there are of course two accents.

The interaction of accent and length within P-phrases makes it more likely that the feet of Mwiini are actually binary and right-headed. Underlying length can then only surface if the syllable in question belongs to the last foot of the P-phrase, either coinciding with the accented syllable or immediately preceding it. Furthermore, if the first syllable of the last foot coincides with a word boundary, the vowel of that syllable will be lengthened. The situation can be handled by constructing constituents on two different metrical planes: on the first are constructed right-headed binary feet from right to left, the final syllable being extra-metrical; line 1 constituents are right-headed and unbounded, and lines 1 and 2 are conflated, getting rid of secondary accents.[63] The same line 0 constituents are copied on the second plane, but there syllables with underlying branching rhymes get a line 1 accent; the first syllable of the last foot also receives a line 1 asterisk if it is at the same time word-final; then all line 1 asterisks are deleted except the rightmost one, and all CVV- syllables not bearing an asterisk are reduced to CV-. For example:

somanii cúwo 'read (pl) a book'

foot-construction:

```
(*   *)  (*   *)
so   ma   ni   cu   wo
```

underlying length:

```
(*   *
(*   *)  (*   *)
so   ma   ni   cu   wo
```

[63]It is in fact not quite clear whether secondary accents exist in Mwiini. If they do, then line conflation does not occur.

lengthening:

```
*   *   *
(*  *)  (*  *)
so  ma  ni  cu  wo
```

line 1 constituent construction:

```
        *
(*  *  *)
(*  *)  (*  *)
so  ma  ni  cu  wo
```

output:

somanii cúwo

As constituent construction occurs within P-phrases, it is useful to consider what syntactic structures P-phrases actually consist of. The main examples are to be found among the following: noun + most elements of a NP (but not -o̱e 'all'), verb + one complement, either Goal or Beneficiary, but not both, see:

mwáana mwan̪dikilile mwaalímu xáa̱i 'the child wrote the teacher a letter'

Furthermore, negative verbs don't form P-phrases with their complements:

čaɬa čimóoyi hačúβli n̪táwa 'one finger does not kill a louse'

Writing about Haya where a similar situation obtains, Byarushengo and Hyman (1984:72) attribute it to the intrinsically focused character of negative verbs.

2. Nothing has been said so far specifically about accent in verb forms. Verbs, as nouns, are underlyingly unaccented in most cases. Nevertheless some morphemes induce a behavior rather different from what we have seen so far. Most striking is the fact that this new accent pattern is not expressed at P-phrase level, but at a higher syntactic level that we will call the intonational phrase or I-phrase (after Nespor and Vogel 1982). The I-phrase may encompass several P-phrases, and does not interfere with the previously assigned pattern of any but the final one—basically replacing penult by final accent, but leaving length assignment intact. See the following examples:

mwan̪dikilile mwaalímu xáa̱ti	'he wrote the teacher a letter'
nimwan̪dikilile mwaalímu xaa̱tí	'I wrote the teacher a letter'
mpeɬe mwáana čibúuku	'he gave the child a book'
nimpeɬe mwáana čibuukú	'I gave the child a book'
naxsulaa kúǰa	'he wants to eat'
naxsulaa kuǰá	'I want to eat'
naxsula kuǰaa n̪áma	'he wants to eat meat'
naxsula kuǰaa n̪amá	'I want to eat meat'
naxsula kuǰa n̪ama yaa mbúzi	'he wants to eat goat meat'
naxsula kuǰa n̪ama yaa mbuzí	'I want to eat goat meat'

múke ʈiɬanziɬee ɳáma kaa čísu	'the woman cut the meat with a knife'
ɳama yaa múke ʈiɬanziɬo kaa čisú	'the meat the woman cut with a knife'
múuntu ikusíɬe	'the man is satiated'
muuntu ikusiɬó hamwíiwi mwenyee nɖáɬa	'the satiated man does not know the hungry one'
mkono waa wé huxaɖíri ku'uʈiinɖá ubúuse	'the hand you can't cut, kiss it'

This "final accent" pattern is triggered by 1st- and 2nd-person subject prefixes (both sg and pl) of two tenses: the present progressive /-naku-/ and the perfect /-iɬe/. Furthermore, it also appears with all subject prefixes in the /-ka-/ tense and all relative forms, whether or not they differ segmentally from their nonrelative counterparts. In all these the accent will be placed on the final syllable of the last word of the I-phrase, the latter covering the maximal string of constituents forming the VP whose head is the verb in question, whether main or embedded. Another verb form, the negative imperative /si-...-e/, also takes this pattern, but it never forms an I-phrase, nor indeed a P-phrase, with a complement:

sisoomé čibuku ičíje 'don't read that book'

The evidence that I-phrase accent is assigned by an entirely different mechanism from the one assigning P-phrase accent is that the former has no relationship at all with length, whereas we have seen that length surfaces in accordance with the constituents generated at P-phrase level. The parameters for I-phrase accent are thus unbounded and right-headed, with extra-metricality entirely determined by the verb form. I-phrase accent assignment overrides previous P-phrase accent on the last word, which can easily be accounted for by the following rule: Delete a line 2 asterisk if it is immediately followed by another line 2 asterisk. If the verb form assigns penult accent, however, or if the last word of the I-phrase is monosyllabic, no change takes place vis-à-vis previous P-phrase accent, e.g.:

múuntu wene ɳumbaa né	'the man saw four houses'
muntu wene ɳumbaa né	'the man who saw four houses'

§16.6. The prosodic system of PSA. At the prosodic level, Mwani has little in common with other members of Sabaki, since it consistently displays *first stem syllable accent,* a feature conspicuously absent from the rest of Sabaki, but well attested in some Rufiji-Ruvuma languages, particularly in Mozambican dialects of Makua. On the other hand, Mwani seems to have the penult lengthening typical of Mijikenda and Swahili (but also of numerous other Bantu languages in eastern and southern Africa) and also apparently constructs default feet on the penult, which does not seem to be the case in the southern languages. Due to the problematic nature of the evidence, we omit Mwani from further discussion.

We assume that the original eastern Bantu tone system was very similar to that exhibited today by Pare, from which diverged the Proto-Central Kenya and Proto-Chaga systems, on the one hand, and PNEC and Proto-Western Tanzania on the other. The latter two share at least one important innovation, namely, the reduction of the contrast between

*-CV́CV̀ and *-CV́CV́ stems, thus taking the first definite step towards changing from a tone ("several H per morpheme") to an accent ("one H per morpheme") system. Central Kenya and Chaga have not made this step.[64] Since this phenomenon is common to both NEC and Western Tanzania, it can safely be assumed for PSA.

Another process which must be assumed for PSA, since it is common, in one form or other, in Seuta and Western Tanzania,[65] is the process of tone assimilation, or tone spread, by which a H assimilates (or spreads onto) a following L syllable. This, which in languages like Langi or Remi is limited to one such syllable, seems to have applied iteratively in PSA, somewhat in the manner apparent in Shambala today; see the following examples:

Pare	kuvóna nyumbá 'to see a house'
Shambala	kuóná nyú!mbá
Giryama	kuona nyúmba
Nzwani	kuona nyúmba

The Shambala situation helps to shed light on what takes place in Mijikenda and Comorian. It can be seen that the H linked to the first stem syllable spreads to all nonH syllables following (the Pare form is a good equivalent of underlying tones in Shambala). In the other two languages, the situation emerging is plain: the only H to appear is the one which in Shambala was final in the first constituent. It is thus likely that the surface realization in an earlier stage of those two languages (also Bondei and Zigua) was similar to the Shambala form, and that the whole initial string of H, except the last, was dissimilated. It would then appear reasonable that PSA should have started with a system rather like that of Shambala.

This being so, the case of Pokomo appears paradoxical, since this spreading process is not much in evidence there, and if anything, H seems to associate *to the left* (as seen above). Thus the phrase above would be

Pokomo	kúǫ́na nyúmbá́

where the marked syllables are underlined. Nevertheless, there is no necessary contradiction, if we assume—as we must because of its absence in any neighboring languages—that this leftward spreading is a recent development in Pokomo. Leaving it aside, the only other generalization we felt we could safely make in §16.1. above was that Pokomo constructs *word-internal* right-headed unbounded constituents. Since these are just the type of constituents we find in Mijikenda and Comorian as well, except that in these two languages the

---

[64]Daßida, genetically a part of the Chaga-Taita group, has undergone a reduction in distinctiveness, but quite a different one, since it neutralizes the -CV̀CV́ and -CV̀CV̀ contrast. This is reminiscent of what happens in Interlacustrine languages, but distance and genetic remoteness preclude any possible connection. The prosodic evidence for Saghala, whose own genetic position is controversial, is confused, but it appears to be closest to MK.

[65]It seems not to appear in Nilyamba, but our evidence is meager.

constituents are constructed at phrase level,[66] we will admit that the spreading process in PSA was at first restricted to within word-boundaries, as indeed it still is in Shambala, provided the spreading H is followed within the word by at least two nonH syllables. Compare with the above the following example:

Shambala kunibánángía nyumbá 'to destroy a house for me'

Here the H spreads only to nonfinal syllables within the word, not affecting the following word. In order to maintain the hypothesis, otherwise well motivated, of a close relationship between Pokomo and Mijikenda, it is thus probably best to accept that H-spreading in PSA was so restricted. The development of spreading within the phrase would thus belong to a stage after the separation of Pokomo, Mijikenda, and Comorian, although one is left with the unpleasant fact that a very similar development has occurred not only in these last two but also in Seuta, similarities so close that they can hardly be due to chance. It should be noted that in Mijikenda, although phrase-initial (or single) H sequences are dissimilated as seen above, this does not apply to following ones. If the initial H sequence is short enough, the pitch pattern of the whole phrase will look quite like Shambala, thus:

Giryama	ná!óná má!fúhá má!njí dzana	'I saw a lot of oil yesterday'
Shambala	ná!á!óná má!vútá! mángí	'I had seen a lot of oil'

Areal influences might be invoked in the case of Digo, which is adjacent to Seuta, but Giryama is quite remote geographically, and furthermore separated by a stretch of arid steppe. Besides, areal influences, at least modern ones, are less than likely in the case of Comorian.

We will suggest a model which tries to bypass these difficulties. We start by assuming for PSA and Proto-Seuta a prosodic system where H spreads word-internally to all following nonH syllables except the final one (interestingly enough, the southern dialect of Pare exhibits such a process, albeit restricted to verb stems). For reasons difficult to determine, Pokomo stopped at this stage (Stage 1), thus remaining to the best of my knowledge the only East African Bantu language to have accent restricted to word-level with line conflation but no default constituents. In other terms, words may at most have one H but can have none; other languages in East Africa with line conflation construct default constituents, like Comorian or Kinga; languages allowing words with no H also generally allow words to surface with several H, although these may be restricted to verb forms, as opposed to nouns. It is possible that, as available data on Pokomo accent are scarce, further research might modify this impression, but won't change the fact that Pokomo developed differently from its nearest relatives.[67]

---

[66]The same analysis would apply to Shambala, insofar as it is possible to view it metrically—which we find attractive, but will not develop here.

[67]One might be tempted to look for a solution in the direction of Cushitic. What is known of the Southern Oromo accent system does not look at all close to the situation in Pokomo. It is in fact simpler underlyingly, but definitely involves several phrase-level accent rules, which, we argue, Pokomo lacks, unlike the other Sabaki accent languages.

Leaving Pokomo aside to develop eventually, at an undetermined date, its leftward association rule, we can visualize the process in the other Sabaki and Seuta languages as follows. The spreading rule began to apply within the P-phrase, probably already distinguished at this stage (Stage 2) by some kind of penult prominence, although not necessarily similar to the one exhibited by contemporary Mijikenda and Bondei—all the more so since vowel length was presumably phonemic then. The evidence for phonemic length at this stage derives from the facts on Shambala quoted above, where H spreads across word boundaries only if penultimate. A number of disyllables are apparent exceptions to the rule, e.g.:

Shambala	nkánde	'food'	šáno	'five'
	mée	'milk'	kukwéa	'to go up'
	kuvyáa	'to give birth'	etc.	

All these words have a -CV́VCV̀ pattern and do not spread their H onto the following word, just as if they were trisyllables. The obvious reason is that they all derive from Bantu stems of the shape *-CV́V̀CV̀, and H-spread must have started by counting morae rather than syllables, so length was still significant at the beginning of the process. Practically the same words appear as exceptions in Mijikenda and Bondei, where no H appears normally on disyllabic stems in isolation, since three syllables are needed, but where a number of words nevertheless associate a H to the penult (Bondei, Giryama) or the final (Digo), e.g.:

Bondei	mêe	'milk'	šâno	'five'
Digo	kandé	'food'	kuvyará	'to bear children'

Although Bondei and Mijikenda have lost the rule counting syllables (or morae) for H assignment at phrase-level, e.g.:

Digo	kuona nyûmba	'to see a house'
	kubananga nyûmba	'to destroy a house'

contrasted to

Shambala	kuóná nyú!mbá	but	kubánánga nyumbá

they still retain at word-level traces of the former vowel-length contrast.

Assuming then that the stage immediately following that of PSA would have been very similar to Shambala today, the next step (Stage 3) was for Bondei, Zigula, and the Sabaki languages, minus Pokomo, to stop counting syllables and apply the constituent construction rule across the board (within P-phrases, of course, as is still the case today). There again the existence of P-phrase penult prominence might well have played a role. Mijikenda is still by and large at this stage, with Digo having later, and probably quite recently, developed a rightward shift by one syllable. Comorian, on the other hand, went boldly forward and completed the move to a neat accent system by constructing default feet (on the penult, it should be noted) and getting rid of multiple H in a phrase by applying line

conflation.[68] Although penult prominence is not particularly audible in most cases in Comorian, this is doubtless due to the intonation pattern which puts a string of H from the accented syllable, sometimes right to the end of the phrase.

How do we fit Swahili into this pattern? There is no reason to assume that PSW did not share in all the developments above, right down to Stage 3. If we take Mwiini as representing the stage just before the abandonment of all prosodic distinctiveness, which the other Swahili dialects—certainly ND—once went through, does this stage directly follow Stage 3, or could it be placed down the line from the Comorian pattern (Stage 4)? Unfortunately, there seems to be no evidence bearing on this point, since intermediate stages are missing for Swahili (perhaps the clue might be found in some dialect of Rural Zanzibar?). What all Swahili dialects clearly have is penult prominence, but we have argued that this probably existed in PSA as well, and thus cannot be considered an innovation. Nevertheless, it seems probable that a Stage 3 language would not go directly to complete loss of prosodic distinctiveness without first having developed default feet, which would then make it more Comorian-like. The alternative would be a language contrasting, say, words with penult H versus words with no H; such a language is not inconceivable, but, to the best of our knowledge, does not exist in Bantu (East) Africa.

[68]This account is based on the assumption that Nzuani is overall the most conservative, although at word-level Ngazija better reflects the Stage 3 situation. It is not possible to discuss the complexity of the facts here, but it can be assumed that the loss of this word-level distinctiveness is very recent in Nzuani. The restriction on the number of feet per line 1 constituent found in Ngazija and Mwali we take to be a simplification of the Nzuani pattern (for more facts, see Philippson forthcoming).