languages

MDPI

Article

PI-Effects in South Bantu: Consonant Changes Due to a Preceding Front Close Vowel

Jeffrey Wills

Ukrainian Catholic University, 79000 Lviv, Ukraine; wills@ucu.edu.ua

Abstract: An important set of sound changes affected the South Bantu languages through the impact of front vowels on following consonants, most notably under the form of the class 5 nominal prefix *i-. These consonant changes are well known, but their extent has been underestimated, as the substantial data in this paper show. There is not even a standard name for these changes, which are here called "Preceding-I effects". This paper offers a detailed study of the relevant conditioning factor, calling attention to the understudied category of hiatus resolution in the history of Bantu languages. Although the reflexes in individual languages vary and levelling often reduced the number of surviving examples, indications of systematic PI-effects in all the subgroups of the South Bantu branch contrast with other Bantu branches and suggest a common conditioning factor was present in Proto-South-Bantu.

Keywords: South Bantu languages; Eastern-Bantu; sound change; Preceding-I effects; progressive palatalization; Nyanjic; Shonic; sandhi; morphophonology

Academic Editors: Koen Bostoen and Sara Pacchiarotti

Received: 19 August 2024
Revised: 2 November 2024
Accepted: 4 January 2025
Published: 27 January 2025

Citation: Wills, J. (2025). PI-Effects in South Bantu: Consonant Changes Due to a Preceding Front Close Vowel. *Languages*, *10*(2), 23. https://doi.org/10.3390/languages10020023

1. Introduction

Over the last decades, the phylogeny of the 400+ Bantu languages has been intensively studied and its major branches clarified. As a result, now is a propitious time to reconstruct the linguistic features of those intermediate branches; for example, large ones like Eastern Bantu (with over 200 languages) and smaller ones like its daughter South Bantu (with about 30 languages).[1] Proto-Eastern Bantu (PEB) has been reconstructed with a time depth of about 2500–3000 years, and was at that time generally located in the western Great Lakes area; the time depth of reconstruction is perhaps 2000 years for Proto-South Bantu (PSB), a branch established by lexicon-based phylogenies (Koile et al., 2022; Grollemund et al., 2015).

An important set of sound changes affected the South Bantu languages, and a few other Eastern Bantu languages, mainly through the impact of the class 5 nominal prefix *i-. For example, the class 5 word for 'ember, charcoal' has a different initial consonant (C_1) from the class 6 stem in these singular/plural pairs: Chewa (N31) *khala/ma-kala*, Manyika (S13) *gara/ma-kara*, and Tsonga (S53) *khálá/mà-kálá*. More fully, Table 1 shows the difference in South Bantu languages between unconditioned outcomes of PEB *k in *kám 'squeeze' (BLR 1689) and the unusual reflexes of PEB *k in the class 5 noun *i-kómì 'ten' (BLR 2027).[2] Examples throughout this paper are given with the conventional orthographies of those languages.[3]

In all these languages, it is clear that the singular reflexes of PEB *k in the class 5 form 'ten' are different from what we would expect in an unconditioned environment, seen both in 'squeeze' and in the cl. 6 plural 'tens'.

Table 1. Unusual South Bantu reflexes of Proto-Eastern-Bantu *k.

Language	*kám 'squeeze'	*ì-kómì 'ten'	*mà-kómì 'tens'
N31 Chewa	**k**àmà	**kh**úmì	mà-**k**úmì
P31 Makhuwa	ama	ni-**kh**úmi	—
S12 Zezuru	**k**áma	**g**úmí	mà-**k**úmí
S21 Venda	**h**áma	**f**úmí	mà-**h**úmí
S42 Zulu	**k**háma	í-**sh**ùmì	ámá-**sh**ùmì
S53 Tsonga	**k**áma	**kh**úmè	mà-**k**úmé
S62 Gitonga	**g**àmá	lì-**kh**ùmì	mà-**g**ùmì

However, these special reflexes of *k are not restricted to class 5, since they also occur in a verb like *jíkʊt 'be sated' (BLR 3445), e.g., Chewa ***kh****utira*, Zezuru ***g****úta*, Venda ***f****úra*, or a pronominal like *jèká 'only, alone' (BLR 3293), e.g., Chewa *-e**kh**a*, Zezuru *-e**g**a*, Gitonga *-e**kh**a*. The environment can also not just be connected to the PEB phoneme *k, because the relevant phonological changes in South Bantu languages are a wide system that includes reflexes of all the consonants *p, *t, *k, *b, *d, and *g. This is not the place to give examples of all those consonants for 30 languages, but we see in (1) that, as a result of these changes in class 5, unvoiced stops in Chewa become aspirated and those in Manyika become implosives, and in both languages voiced consonants become implosives (written *b* and *d* in the conventional orthographies).

(1) Consonant changes in cl. 5 vs. unconditioned cl. 6 forms in Chewa and Manyika

a. Chewa (N31)					
*i-p > ph vs. *p > p	*phéwà/mà-péwà*	5/6	'shoulder'	<	*pègà
*i-t > th vs. *t > t	*thákò/mà-tákò*	5/6	'buttock'	<	*tákò
*i-k > kh vs. *k > k	*khálà/mà-kálà*	5/6	'ember'	<	*kádà
*i-b > ɓ vs. *b > w	*bérè/mà-wérè*	5/6	'breast'	<	*béèdè
*i-d > ɗ vs. *d > l/r	*dèwù/mà-lùwà*	5/6	'flower'	<	*dòbà
b. Manyika (S13)					
*i-p > ɓ vs. *p > p	*bángá/mà-pángá*	5/6	'knife'	<	*pángà
*i-t > ɗ vs. *t > t	*dákó/mà-tákó*	5/6	'buttock'	<	*tákò
*i-k > g vs. k > k	*gàrá/mà-kàrá*	5/6	'ember'	<	*kádà
*i-b > ɓ vs. *b > w	*bùwè* —	5	'spider'	<	*bòbì
*i-d > ɗ vs. *d > r	*dòtà/mà-dòtà*	5/6	'ashes'	<	*dòtà
	ròtà/mà-ròtà	5/6	'ashes'	<	*dòtà

Already, we can make some important observations:

(a) Because the special reflexes are seen in nouns, verbs, and pronouns, we are clearly discussing a phonological change and not just a morphological change of class 5.

(b) In certain South Bantu subgroups, this change became morphophonemic, distinguishing class 5 singulars (with change) from class 6 plurals (without change). In languages like Chewa, the morphophonemic rules continue to be productive, even extending to borrowings, e.g., *thímàtì/mà-tímàtì* 'tomato(es)'.

(c) In many South Bantu languages, however, analogical levelling significantly removed singular and plural differences, e.g., in the Manyika doublet for 'ashes' in (1) and the Zulu form for 'tens' in Table 1, thus greatly reducing the data set and potentially obscuring earlier patterns.

(d) The full sets of reflexes for each language subgroup are presented in Section 3, but the major outcomes are quite varied: aspiration (e.g., Chewa, Makhuwa, Copi, Gitonga), palatalization (e.g., Sotho-Tswana), or strengthening (the impact on voiced consonants in most languages).

(e) The cause of the phonological change is not obvious from the BLR reconstructions we saw: *ì-kómì 'ten', *jíkʊt 'be sated', *jèká 'only, alone'. However, if we remove *j (following Wills, 2022) and if we improve some vowel reconstructions, we can likely clarify the relevant conditioning environment.

The general phenomenon of an unusual C_1 for class 5 or certain verbs has been known since the beginning of Bantu studies (Bleek, 1862, p. 162). Changes were pointed out by Meinhof (1899, 1901) on Sotho and Venda, and then Eiselen (1924) presented the basic comparative data a century ago. Some changes are recorded by Guthrie (1967) in his inventory lists under *yį clusters (equivalent to BLR *ji) and unusual features of class 5 are discussed in many grammars of South Bantu languages as well as in a few languages in other branches of Bantu, e.g., Ganda (JE15), Comorian (G44), and Lega (D25).[4] So, these consonant shifts are hardly unreported, but they are underestimated. First, they are thought of as occasional and language specific because the extent of them has not been realized. Second, they are seen as the consequence of a change that is not well defined, if at all. The result is that there is not even a standard name for these consonant changes, which here are called "Preceding-I effects" (PI-effects), with capital "I" representing a front close or near-close vowel in certain conditions. Third, the phylogenetic implications have not been discussed. Beyond the 30 languages in South Bantu, these consonant changes are fairly rare, with fewer than a dozen other Bantu languages providing much evidence.

The goals of this paper are (1) to compile the substantial South Bantu data for these changes (File S1) and present the relevant reflexes, (2) to demonstrate that the relevant conditioning factor is more complicated than the generally assumed front close vowel, (3) to highlight the varying factors that affected the final results of these sound changes, and (4) to discuss what, if any, shared innovation accounts for these effects across South Bantu. The paper is mostly aimed at those familiar with Bantu diachronic linguistics. However, it is hoped that the discussion of conditions will also appeal to phonologists interested in the effects that a preceding front vowel or diphthong can have on a consonant, and that historical linguists will also appreciate the ways once-regular sound changes can become obscured over time. A change conditioned in part by sandhi settings can be hard to recover, especially when analogical levelling is subsequently applied.

Because the unifying factor to the changes is the conditioning environment, we begin in Section 2 by studying the structures with preceding *i or *ɪ where PI-effects are seen. In Section 3, we examine the reflexes themselves and consider the possible phonetic paths to them. Section 4 reviews the analogical responses have often limited our data for PI-effects in class 5 nouns. Finally, in Section 5, we look at whether this is a shared innovation of the South Bantu branch of languages.

Unless otherwise noted, all reconstructions for Proto-Eastern-Bantu (PEB) in this paper are the author's, usually the same as those in *Bantu Lexical Reconstructions 3* (BLR), so BLR numbers are given for reference. Note, however, that reconstructed roots are written here without a preceding or following hyphen and without *j, following Wills (2022). Also note the use of *s instead of BLR's *c, since its reflex is a sibilant in the vast majority of Eastern Bantu languages, thus following the suggestion of Meeussen (1967, p. 83), "Instead of /c/, one might just as well use the symbol /s/". Although some scholars tend to cite all BLR reconstructions as PB, many of those reconstructions only have evidence from certain zones and subgroups and should be labelled as such. By restricting ourselves to PEB, we avoid the need to constantly distinguish which reconstructions are PB and which are only PEB. Certain reconstructions can only be made for a more limited group of languages and are labelled as such, e.g., PSB. An extensive supplement (File S1: South Bantu Data for PI-effects) summarizes the changes in each language with numerous examples, as well as the relevant sources and traditional orthographic conventions of the languages.

South Bantu Languages

We have not begun to look at all the evidence yet, but Table 1 and the lists in (1) show that we are dealing with a set of phonological changes that have both breadth (involving over 30 languages) and depth (affecting every oral consonant in many of them). These changes were a major event in many languages, affecting many consonants, so not easily borrowed as a phonological system. Rather they indicate common conditions in the South Bantu branch which individual languages later developed in different ways. The languages that we discuss are most of Guthrie zones NPS; they are displayed in Figure 1 and listed in (2), with more details in File S1.

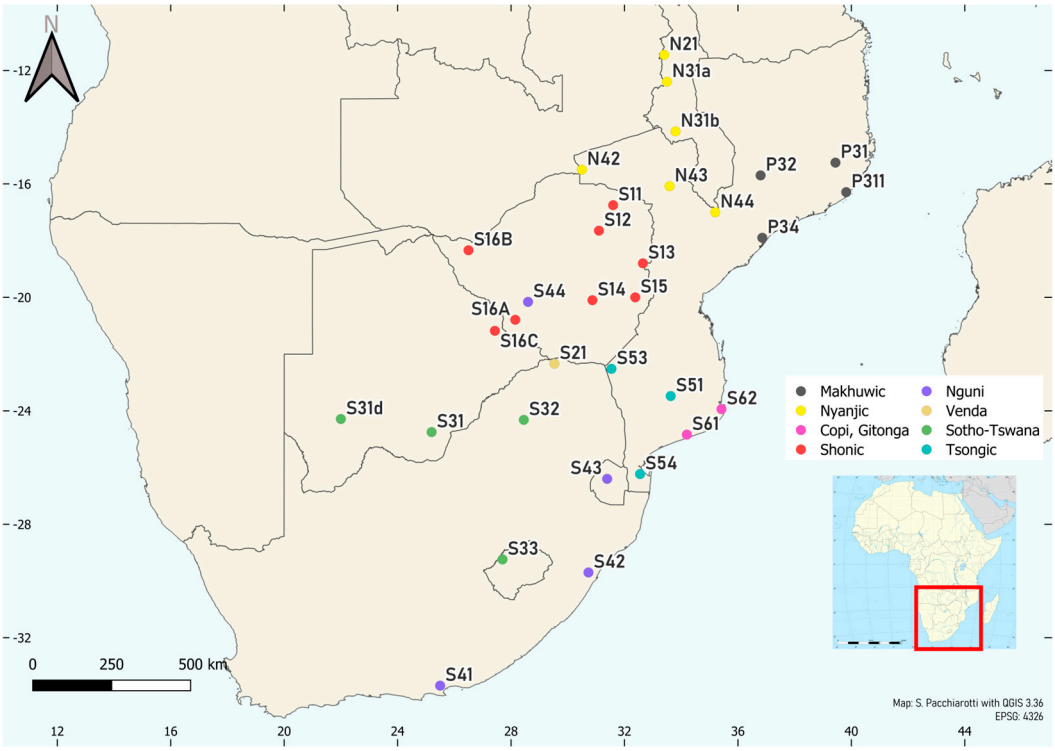

Figure 1. South Bantu languages.

(2) South Bantu subgroups and languages[5]

Nyanjic[6]: N20 Tumbuka, N31a Nyanja, N31b Chewa, N42 Chikunda, N43 Nyungwe, N44 Sena

Makhuwic (P30): P31 Makhuwa, P311 Ekoti, P32 Lomwe, P34 Cuwabo

Shonic (S10)[7]:

Central Shonic (S11 Korekore = Ko, S12 Zezuru = Z, S13 Manyika = M, S14 Karanga = K)

Eastern Shonic (S15 Ndau = Nd)

Western Shonic (S16 Tjikalanga = Tj, Ikalanga = Ik, Nambya = Na)

Venda (S21)

Sotho-Tswana (S30): S31 Tswana, S311 Kgalagadi, S32 Northern Sotho, S33 Southern Sotho

Nguni (S40): S41 Xhosa, S42 Zulu, S43 Swati, S44 Zimbabwean Ndebele

Tsongic (S50)[8]: S51 Tswa, S53 Changana-Tsonga, S54 Ronga

Copi (S61)

Gitonga (S62)

I use the term "South Bantu" for this branch to distinguish it from the term "Southern Bantu" conventionally used for the narrower group of all languages of Zone S (Doke, 1954;

van der Spuy, 1990; Gunnink et al., 2022) and the term "South-eastern zone" for the yet narrower languages of Zones S20–60 (Doke, 1954; Cole, 1959; relabeled as "Southern" in Janson, 1991). "Southwestern Bantu" in Angola and neighboring countries is obviously a different branch.

2. The Conditioning Environments for PI-Effects

In addition to the position after the class 5 prefix, there are several other environments which show the same series of reflexes different from those of unconditioned consonants. In all four environments of PI-effects, the consonant undergoing change is preceded by *i or *ɪ, as exemplified in (3).

(3) PEB environments showing PI-effects
- a. Stems with internal vowel + *i/ɪ
  - *dàip 'be long' > Zezuru *rèba*
  - *áító 'our' (1pl. poss.) > Chewa *-áthù*, Zezuru *-èdù*, S. Sotho *-eso*
- b. Stems with initial *i/í
  - *ìt-ɪd 'pour (out)' > Chewa *thìla*, Venda *shèla*
  - *ít-(an) 'call' > Ekoti *iitha*, Tjikalanga *dána*
- c. Verb stems preceded by the reflexive pronoun *i
  - *í-kʊt 'become satiated' > Zezuru *gúta*, Venda *fúra*
- d. Nouns with class 5 nominal prefix *i
  - *i-tákò buttock' > Chewa *thákò*, Zezuru *dákó*
  - and other nouns illustrated in (1) above

PI-effects can happen inside a morpheme or across morpheme boundaries (so called *sandhi* discussed in Section 2.5). We begin with examples of morpheme-internal changes because those roots (although fewer) often have reflexes across most languages. However, the most common source for PI-effects is the final category: class 5 nouns where the nominal prefix *i was at the morpheme boundary, causing changes to the stem's initial consonant.

We now review each of these environments in more detail and then consider the phonological process uniting them (Section 2.5).

*2.1. Stems with Internal Vowel + *i/í*

PI-effects inside a morpheme are easily seen in the reflexes of the 1pl. possessive pronoun PB *áító 'our' (Meeussen, 1967, p. 107) which are grouped in (4) by their consonant and vowel changes.

(4) Southern Bantu reflexes of *áító 'our' and unconditioned *t

a with PI-effects		
Chewa	*-athu*	*t > t
Venda	*-ashu*	*t > r
Gitonga, Copi	*-athu*	*t > r
e with PI-effects		
all Shonic varieties	*-edu*	*t > t
S. Sotho (collective)	*-eso*	*t > r
N. Sotho	*-gêšo*[9]	*t > r
Ekoti	*-eɬthu*	*t > r
e without PI-effects		
Zulu	*-ethu*	*t > th
Tsonga	*-erhu*	*t > rh

Note the conspicuous lack of a two-vowel reflex (either in hiatus or in a diphthong or a long vowel) in any modern language. That is because South Bantu undergoes synizesis (i.e., monophthongization with shortening) reducing all PEB long vowels and diphthongs to short vowels, as discussed in Section 2.5. Thus, the *i of the proto-form never survives as an independent segment. The specific choice of vowels with consonant mutation is the same in South Bantu forms of **bàiy* 'work wood' (BLR 8930), except that *a* rather than *e* is seen in Zulu *bàza*.

Another stem connected with pronominal functions, PSB *ika 'only, alone' (BLR *jèká 3293), has a similar structure and also exhibits PI-effects.[10] The initial *i* in the stem combines with the vowels of preceding markers to create *aika/oika, e.g., Ntomba (C35a) *kaíka* and Gikuyu (E51) *-oika*, so it is not surprising that we regularly see consonant changes in the South Bantu languages, where the root is used to form exclusive pronouns in different persons or classes, as in (5).

(5) Reflexes of PSB *ika 'only, alone' and unconditioned *k

a. Aspiration		
Chewa	*-ekha/-okha*	*k > k
Nyungwe and Sena	*-ekha/-okha*	*k > k
Makhuwa	*-eekh-*	*k > ø
Lomwe	*-ekha-*	*k > ø
Copi	*-ekha/-okha*	*k > k
Gitonga	*-ekha/-okha*	*k > ɣ
b. Voicing		
Central Shonic	*-ega/-oga*	*k > k
Ikalanga	*-ega/-oga*	*k > k
c. Strengthening or preservation		
Cuwabo	*-eká/-oká*	*k > ø

The reflexes in (4) and (5) give us a good survey of the typical PI-effects for *t and *k across South Bantu: aspiration in most languages, palatalization in S20–30, voicing in Shonic, strengthening or preservation in Cuwabo. We return to this variety of effects in Section 3.

Proto-Bantu noun and verb roots are typically reconstructed with the shape CVC, and sometimes with long vowels, but rarely with two different vowels inside the root, and if so, they are usually with a high vowel as V_1, but almost never as V_2. However, Meeussen (1967, p. 88) gives "four examples of a special type -CV-VC-, in which the extension -VC- is a suffix that must be present, so that the radical is never used alone". One of these is *tó-ık 'put load on head; give to carry' (BLR 3029) with PI-effect voicing in various Shonic varieties, e.g., Manyika and Zezuru *twíga*. Other examples involving impositive *-ık are *bí-ık 'put away; bury, lay eggs' (BLR 200) > Karanga *víga* 'bury', Tjikalanga *bígà* 'put, store', Nyungwe *ikha* 'put, store'; *tó-ık 'burn, tr' (BLR 4857) > Manyika, Zezuru *túguka* 'burst into flames'; and *té-ek/*tá-ık 'put; place on fire' (BLR 2831) > Karanga, Zezuru *téga* 'set up firmly (pot on hearth)'.

There are also a few adjectives with CVV structures from which verbs were derived by the suffix *p, thus creating the necessary environment for PI-effects, as in (6).

(6) Examples of PI-effects before the verbal suffix *p

(a) *dàì adj 'long' (BLR 3705) >> *dàì-p 'be long' (BLR 784)
Zezuru *rèba*, Ekoti *lepa*, Copi *làpha*, Gitonga *làpha*

(b) *bíì adj 'bad' (BLR 5841) >> *béep 'tell lies' (BLR 156)
Karanga both *nyèba* and *nyèpa* 'lie'

(c) *pì adj 'black' (BLR 6406) >> *pìip 'be black' (BLR 2584)
Shonic varieties both *svìba* and *svìpa* 'be dark, black, dirty'

In Ekoti, the unconditioned reflex of *p is /v/, but in *lepa* 'be long' we see the PI-reflex /p/ instead, just as the Zezuru form *rèba* shows voicing in contrast to Korekore *rèpa*. However, PEB also derived new adjectives by means of the deverbative *-u suffix, and that high vowel suffix triggered Bantu Spirantization, which we see instead of voicing, e.g., *dàìp-ú 'long' adj. > Zezuru *rèf-ú*, Kalanga *lèf-ú*; *pìip-ú 'glossy black' > Manyika *svìf-ú*.

In general, there were few opportunities in PB for the environment inside a morpheme to create a PI-effect. By the time of PEB, we only see a few more CVVC roots with *i or *ɪ as the second vowel, e.g., *tòìgà 9/10 'giraffe' (BLR 3028) and *bàiy 'work wood'. However, there are only a handful of these.

2.2. Roots with Initial *i or *ɪ

Because early Bantu noun and verb roots regularly had CV- or V- prefixes, roots which begin with *i or ɪ almost always formed words with vowel + *i/ɪ in the middle of a word. This structure is not so different from the category just discussed, especially regarding nouns which have fixed sg/pl prefixes. For that reason, the PI-effects are regular and predictable.

There are not many PEB nouns with initial *i or *ɪ before a stop, but a few of them have South Bantu reflexes which show PI-effects inside the stem. The word *ìkò 'ladle' (BLR 3443, classes 3/4, 11/10) is a good example of a noun where a prefix vowel could create adjacent vowels which triggered loss of *i* and changes to the subsequent consonant. This word has PI-effect reflexes through most of South Bantu, as seen in (7).

(7) PI-reflexes of *ìkò 'ladle, spoon' and unconditioned *k

Nyanjic: Nyungwe *lu-kho* 11, Chewa *chì-khò* 7 'cup'	*k > k
Shonic: Korekore *rù-gò* 11, Ikalanga *lù-gò* 11	*k > k
Venda *lù-fò* 11	*k > h
Tswana *lo-so*/*din-tsho* 11/10 & *le-sô* 5	*k > χ
Copi *m'-khò*/*mì-khò* 3/4	*k > k
Makhuwa *ni-ikho* 5, Cuwabo *mú-kò* 3	*k > ø

Other PB nouns with frequent PI-effects inside the stem are *íkò 5/6 'fireplace, country' and *ísò 5/6 'eye', both discussed in Section 3.

Apart from reflexives, verbs with vowel-initial roots are also present in many Bantu languages, and some with initial *i/í have been reconstructed for PB, as seen in (8).[11] The unconditioned reflexes are given in parentheses.

(8) PI-effects in verbal stems reconstructed with initial *i or *ɪ

a. *ìp-ɪk 'cook, boil' (BLR 3496)
   Chewa *phìkà*, Nyungwe *phika* (*p > p)
   E.-C. Shonic (all) *bìka*, Nambya *bhìká*, Ikalanga *bhìkà* (*p > p)
   Cuwabo *píyà* (*p > v)

b. *ìp-ʊd 'take (food or pot) off the fire' (BLR 3500)
   Chewa *phùlà*, Nyungwe *phula* (*p > p)
   E.-C. Shonic (all) *búra*, Nambya *bhùlá*, Ikalanga *bhùlà* (*p > p)
   S. Sotho *tshola* 'dish up food' (*p > ɸ)

c. *ít(-an) 'call' (BLR 3379, 3508)
   Ekoti *iitha* (*t > r)
   Karanga & Zezuru *dána*, Kalanga (both) *dánà* (*t > t)
   Gitonga *tháná* (*t > r)

d. *ìt(-ɪd) 'pour, pour out' (BLR 3503, 3504)
Makhuwa *iittha* (*t > r)
Chewa *thìlà*, Nyungwe *thira*, C. Shonic *dìra*, Ikalanga *dìlà* (*t > t)
Venda *shèla*, Tswana *tshela* (*t > r)
e. *ìkad 'dwell, be, sit, stay' (BLR 3441)[12]
Chewa *khàlà*, Nyungwe *khala*, C. Shonic *gàra*, Ikalanga *gàlà* (*k > k)
Xhosa & Zulu *hlàla* (*k > kh)
Copi *khala* (*k > k), Gitonga *khala* (*k > ɣ)
Makhuwa *khala*, Cuwabo *-kala* (*k > ø)
f. *íb 'steal' (BLR 3387)
Chewa *bà*, Nyungwe *ba* (*b > w/ø)
E.-C. Shonic *bá*, Nambya & Ikalanga *kwíbà* (*b > v/β/ʋ)

Further verb stems are *ìpag 'kill' (BLR 3494), *íbɪd 'sink' (BLR 3397), *ítab(-ɪd) 'answer' (BLR 3381, 3509, 6031), *íd 'get dark' (BLR 6142), and *idìmà 'darkness' (BLR 3411).

Because of the various possible hiatus resolutions at the initial morpheme boundary, each verb and each language hold a separate story. So, in addition to these forms, sometimes languages also have forms which do not show consonant changes. This is particularly true when the initial vowel is preserved, a common reflex in Nguni languages, e.g., Zulu *ephùla* & *ophùla* 'take off the fire', Swati *ètsa* 'pour out' and *éba* 'steal', but occasionally elsewhere, e.g., Chewa *(y)ìtànà* 'call', Makhuwa *iiya* 'steal'.

2.3. Verbs with Reflexive Prefixes

Another category showing PI-effects comprises verb stems preceded by the reflexive pronoun or prefix *i-. Since the reflexive immediately precedes the verb stem, it is in the appropriate position to influence the stem's initial consonant.[13] Whether *i- should be reconstructed for the reflexive at the PB stage or not (Meeussen, 1967, p. 109; Polak, 1983), there is good reason to reconstruct it for the PEB stage (Guthrie, 1967, C.S. 2238; Wills, 2022, p. 85). However, it was frequently replaced or combined with other object markers (Polak, 1983; Dom, 2024). Hence, there were limited time frames in which the reflexive in *i- could have operated to cause PI-effects on the initial consonant of a verb stem. In many languages, any reflexive consonant variant has been levelled out in analogy with that of the basic verb form.

The most well-attested example of an apparently reflexive verb in South Bantu is *í-kʊt 'become satiated', with the examples we saw in the introduction.[14] Otherwise, not many reflexive verbs can be reconstructed for early South Bantu. First, to 'bear offspring' (< *í-bad 'carry a child' BLR 3391) is attested in the northern part, e.g., Chewa *bàla* and Zezuru *bára*. Throughout the region, we find reflexes of *ì-bed-ɪk (BLR 3394) 'bear (child); carry (child) on back', e.g., Nyungwe *bereka*, Tswana *belega*, Swati *béleka*. From *í-tʊk 'be frightened' (BLR 3511), we see Karanga *dúka*, Venda *shùwa*, Kgalagadi *tjhóxa*, and with vowel retention in Zulu *ethuka*. For South Bantu, we can reconstruct *í-pat 'hold' based on PEB *pát 'hold' (BLR 2414), although variants in other zones (*bát 'seize, hold' zones AC, BLR 8483; *píat 'seize', zones ABS, BLR 2556) suggest that the reflexive effects must have been earlier. Forms include C. Shonic *báta*, Tswana-Sotho *tshwara*; the vowel in Zulu & Xhosa *phátha* may also suggest a prefix (Rycroft, 1980).

Inside languages, there are not many examples of reflexives with contrasting initial consonants due to PI-effects, but there are some examples, as seen in (9).

(9) Examples of reflexives with PI-effects in contrast to forms without them
a. *pís 'hide, cover' (BLR 2563): Makhuwa provides both a transitive verb *o-vitha* 'hide (tr.)' (with the usual reflex of *p > v) as well as reflexive *w-i-pitha* 'hide oneself' (with *I-p > p) (Schadeberg, 1999, pp. 386–387).
b. *gìd 'abstain' (BLR 1394): there are a several C. Shonic forms but notice the contrast in Karanga between *ìra* 'abstain from (tr.)' and *zìra* 'abstain from food (intr.)', where the old reflexive apparently gave the intransitive the PI-effect.
c. *ítan 'call, name' (BLR 3380, 3508): Gitonga has the expected *ráná* 'call (someone)' and *tháná* 'respond when called; know one another' with PI-effect.

2.4. Class 5

The class 5 nominal prefix *i- is particularly useful for studying PI-effects, because it gives the largest set of examples, with a wide range of consonants in the dozens of nouns in this class, as we have seen above. However, the pairing of class 5 singulars with class 6 plurals, which did not undergo these changes, naturally led to levelling in one direction or the other, which reduced the total number, as discussed in Section 4.

While the PB class 5 concord marker *dɪ- has been preserved throughout South Bantu with some phonological adjustments, e.g., C. Shonic and Tsongic *ri-*, Makhuwa *ni-*, the corresponding nominal prefix has evolved in varied ways. The class 5 nominal prefix in *i- was unusual for its vowel-initial shape, which made it a candidate not only for development into an off-glide in certain sandhi contexts, but also for replacement. For class 5 prefix variations throughout Bantu, see Kamba Muzenga (1988). In the word for 'ten', some languages have generalized the pattern in which the prefix has been lost with change of consonant, e.g., Zezuru *gúmí*, while some have generalized the pattern in which the prefix stays vocalic, e.g., Zulu *i-shumi*, and some have replaced the nominal prefix with the concord or perhaps combined them, e.g., Kgalagadi *le-shome*. Different languages went different directions, and some languages have multiple patterns, as seen in (10). PI-effects can be seen in each of these scenarios but to varying degrees.

(10) Scenarios for the class 5 prefix and augment
a. *Loss of *i-*: Nyanjic, Shonic, Venda, Tsongic. In these cases, consonant changes were very regular, compensating for the loss of the prefix by marking the stem itself.
b. *Preservation/restoration of *i-*: Nguni. Consonant changes are rarely visible in Nguni, perhaps because in all cases the prefix was maintained or restored by analogy. Eventually, a new prefix structure incorporating the concord also developed in monosyllables, e.g., Zulu *i-tshe* & *ili-tshe* 'stone'.
c. *Replacement by or combination with the concord *dɪ-*: To some extent in almost all groups except Shonic and Tsongic. In this range of options, it is not always easy to trace the full story. Does the current prefix derive from the historic combination of augment and prefix (*dɪ-i-) or does it derive from pattern adding the concord after the consonant change, or from pattern a replacement of the old prefix by the concord?

The first two scenarios reflect the expected sandhi patterns for *i-: (1) in the middle of a phrase, after a preceding vowel, the *i undergoes a change accompanied by PI-effects; (2) at the beginning of a phrase, the prefix preserves its vocalic shape. Each of these two categories has levelled in one direction or the other. The third scenario has a variety of stories in it, some of which only affect a few nouns in a given language. The Sotho-Tswana languages at present have no augment and only a standardized single class 5 prefix *lɪ-* but a limited number of class 5 nouns can be used without a prefix or suffix as adverbs of place.

In short, the combination of augment + nominal prefix provided a steady phonological environment for transforming the *i with change of the initial stem consonant (*dɪ-i-CVC > *dɪ-I-CVC > *dɪ-**C̲**VC). The triple vowels in cl. 5 *dɪ-i-iso 'eye' may explain the varying singular/plural reflexes in Swati *lîsó/émêhlò* and Gitonga *lìsó/màhó*, which have PI-effects only in the singular even though the *i of the root precedes the consonant in both singular and plural. However, it is also possible that the unaugmented form caused the change and was then supplemented with a new prefix (**i*-CVC > **C̲**VC > *lɪ*-**C̲**VC). Both preservation and supplementation are possible scenarios and could even work together.

Separate from these scenarios are the "Z forms" of vowel-initial roots, e.g., Zezuru *zísó/mésó* 'eye', discussed in Section 3.3, where the initial consonants are often mistaken as descendants of the concord but actually develop from the nominal prefix *i-.

2.5. What Is the Phonological Process Causing PI-Effects?

To generalize about these four categories, we can say that PI-effects are seen in environments where *i or *ɪ could be preceded by a vowel within the same morpheme or concluding the previous morpheme. What is it about those environments that leads to PI-effects: the phonetic character of the front vowel or the fact that it is preceded by a vowel, or both?

Because of the well-known phenomenon of Bantu Spirantization, in which consonants are affected by immediately subsequent close (level-1) [+ATR] vowels but not near-close (level-2) [-ATR] vowels,[15] scholars have tended to assume that PI-effects are likewise due to a preceding close *i vowel (e.g., Hyman, 2003, p. 56; Nurse & Hinnebusch, 1993, p. 139). Since the effects of this preceding front vowel are similar to those caused by nasals in some languages, it was suggested a century ago that *i exerted on the following consonant the same influence as a nasal (Eiselen, 1924) or that *i first induced the insertion of a nasal which consequently exerted its influence on the following consonant (Bourquin, 1932–1933). There is much to be said for more research on the parallels between the presence and impact of preceding nasals and *i* in Eastern Bantu, which are both lost in some languages, but in the most recent examination of that question, Creissels (1999, pp. 329–331) rejected the nasalization solutions by Eiselen and Bourquin. Looking carefully at the Tswana evidence, he argued for some type of rule with "the context *≠ i̯ _ (or perhaps *yi̯ _ or * i̯ _ with the condition that no consonant precedes immediately i̯)". For the mechanism, he cautiously proposed an insertion rule *i̯CV > i̯Ci̯V, which would account for the Tswana data,[16] although he understood that might not work elsewhere, and in fact many other languages have different reflexes for PI-effects than effects from following vowels or glides.

However, we can recognize several problems with the focus on the close vowel:

(1) PI-effects occur not only after close *i but apparently also after near-close *ɪ, as we have seen, e.g., *ít-(an) 'call' or some verbs with *-ɪk. It is likely that there are more instances after near-close *ɪ, since Bantu reconstructions are biased by the assumption that only close vowels cause consonant changes. In this case, besides near-close *jít (BLR 3379) attested throughout Bantu, there is also a reconstruction of close *jít (BLR 3507) for the two zones G and S, simply for the sake of explaining the consonant changes—no actual forms with close vowels from seven-vowel languages are cited by Guthrie (1967, p. 4.190). Likewise, all six of Guthrie's modern reflexes for C.S. 2098 *yi̯tuk 'become startled' have vowels that could historically be near-close, but he reconstructs a close vowel for the sake of the consonant change. Accordingly, all the reconstructed roots listed in Section 2.3 should be reviewed and studied for their attested vowels.

(2) If the very close level of the preceding vowel were the trigger, we would expect to sometimes see these effects on C_2 in numerous CVC roots, just as we see Bantu Spi-

rantization regularly on C_1. However, we never do. Rather, these changes only happen in CViC or CV#iC environments, when front vowels are in environments where they are (or could be) preceded by another vowel, i.e., in an environment of potential diphthong formation.

(3) PI-effects are different in important ways from Bantu Spirantization. Their difference in reflexes could perhaps be explained by some difference between the BS process of turbulence in the release of a consonant before a close vowel and whatever phonetic process would be involved in PI-effects after a close vowel. However, there is also a major difference in the fact that the vowel is inevitably preserved in Bantu Spirantization of CVC syllables, but the vowel often disappears in the PI-process. This loss or absorption is much more like some Bantu "consonant + glide" changes (Hyman, 2003, p. 55) which happen when both close or near-close vowels become glides. Also troubling is the difference between the consistent regularity of Bantu Spirantization before close vowels (even across morpheme boundaries) and the inconsistent occurrence of PI-effects. Something besides, or in addition to, the character of the vowel seems to be involved.

The cross-linguistic evidence does not limit the options. Although consonant changes due to preceding front vowels or glides are less common than following triggers, Bateman (2007) found a dozen modern languages with progressive palatalization where the trigger precedes the target. Triggers include various front vowels as well as the palatal glide, isolated or in a diphthong. Diachronic examples include the progressive palatalization of Common Slavic (Wandl, 2020) and Old English forms like *iċ* "I" and *dīċ* "ditch, dike" after /i/ or *weġ* ("way") after /e/.

For all these reasons, in Bantu we must consider not only the character of the vowel, but also the segmental context. PI-effects only happen in CVIC or CV#IC environments, i.e., when front vowels are in environments where they are (or could be) preceded by another vowel, i.e., in an environment of potential diphthong formation. Inside roots like *dàìp 'be long' or a hypothetical inflected form like *báìkòté 'let them be satiated' (my reconstruction), this diphthong potential is obvious. However, diphthongization can also occur across a word boundary, i.e., sandhi (Myers, 2020; Kosch, 2003), which would make an off-glide of the initial *i* of a class 5 noun.

The avoidance of the symbols *y* and *w* by BLR hides the fact that diphthongization was likely at early stages of Bantu languages. To some extent, on-glides are well known in the cases of Ci or Cɪ roots and the causative *-i suffix, where both the relevant consonant changes and the modern orthographic use of *y* and *w* make them clear. Off-glides, however, are almost never indicated by Bantu orthographies, although there is an increasing number of scholars noting them in speech, including across word boundaries, e.g., de Haas (1988) on Kikuyu (E51), Overton (2018) on Simbiti (JE431), Mtenje-Mkochi (2018) on Cindali (M301), and Mhute (2016) on Shona. Discussing adjacent vowels in Kikuyu when two words are placed next to each other to form a phrase, Mugane (1997, p. 7) notes, "in careful speech, there is no diphthong formation, but when uttered in the context of normal speech we get diphthongs". In general, more languages should be studied for the phonetics of *ai* and *au*, as well as whether the frequent coalescence of *ai > e involved a stage of diphthongization, especially in languages without long vowels today.

An analysis of PI-effects that includes diphthongization with off-glide formation has several advantages. First, this approach would explain the fact that PI-effects are only seen in CViC or CV#iC environments, i.e., where diphthongization is possible. Second, it would explain why front vowels of varying closure are possible triggers since they will all yield palatal glides after another vowel. Third, the phonetics of certain PI-reflexes are easily explained by the development of a palatal glide at some stage, e.g., tone-bearing geminate

consonants in Ganda (Meeussen & Tucker, 1955; Myers et al., 2019). Fourth, the adaptability of a glide would explain the variability in segmental outcomes: an off-glide could assimilate to a following consonant (gemination) or be absorbed in a consonant change or assimilate or coalesce with the preceding vowel. For example, the word reconstructed as *cáijà 'old man; man' (BLR 9282 for Zone J) yields Tooro (JE12) *-sáíjà*, Ganda (JE15) *-sá'jja*, Soga (JE16) *-sààḍa*, Bwisi (JE102) *-sàásâ* with varied reflexes: preservation of *ij*, gemination of consonant, change of consonant, lengthening of vowel. Because this diversity of processes all took place in fairly closely related languages, a broad trigger in needed which can involve both vowels and consonants. The diversity of South Bantu forms, discussed in the next section, also make it clear that, whatever the starting point, there were a number of different phonetic paths to the modern reflexes.

The diphthongization that is seen in some modern Bantu languages could also have occurred at earlier Bantu stages, both inside words and between words. Since PEB only had open syllables, any initial front vowel would always follow another vowel, except at the beginning of a phrase. We can see some typical V + VC combinations in PB phrases reconstructed by Meeussen (1967, pp. 117–118, preserving his orthography):

*miedé ítátu 'three knives'
*miedé gíá-baána ítátu 'the child's three knives'

Not surprisingly, the numeral "three" is a word which shows PI-effects in various languages (and different hiatus resolution patterns for *e-i and *a-i). Meeussen even provides some phrases which show the varying environments that the class 5 prefix might encounter:

*i̯pía dídimá bantu 'the garden which people cultivate'
*mu-i̯pía dídimá bantu 'in the garden which people cultivate'
*muntu judima i̯pía 'a person who cultivates (his) garden'

These varying environments (with *i̯* phrase initial, inside an inflected form, and following a vowel) are ripe for the sort of language-specific variations and levelling which we see with PI-effects in South Bantu.

However, where are those diphthongs in South Bantu languages today? As we saw in Section 2.1, all the reflexes of *áító 'our' have either short *a* or *e* today. However, it is not only diphthongs that are absent, but also long vowels. The general South Bantu loss of long vowels and two-mora vowel sequences required *ai to undergo synizesis (i.e., monophthongization with shortening), with tonal consequences as well. It is possible that this pressure towards single-mora syllables is somehow connected to the development of the second-vowel into a component that united with the following consonant to cause PI-effects.

This synizesis in South Bantu can also explain the curious loss of the class 5 prefix *i- in Nyanjic and Shonic. The awkward anomaly of a prefix-less class could be explained if the inherited class marker had become part of a long vowel or diphthong that then had to be simplified. There is evidence for this from the Shonic varieties which have some sandhi at word boundaries as well as very strong PI-effects. In Shonic, historical prefixes have been described as "a latent initial vowel, whose phonological influence is discernible in certain words preceding nouns and on a number of other occasions" (Doke, 1954, p. 49; other languages in de Blois (1970)). For example, C. Shonic *na 'with, by' coalesced with the historical class 5 prefix *i to form *ne-* in *ne-banga* 'with a knife' vs. *na-mapanga* 'with knives' (Fortune, 1985, p. 1.12; Harford, 1997). The wide scope of these rules is seen in verb phrases where the terminal vowel *a coalesces with the historical prefix of the following class 5 noun, e.g., Karanga *ndakavanike g̯umbo* 'I am broken as to the leg' < *ndakavanik-a i-kumbo (Fortune, 2004, pp. 9, 91). Thus, it appears that the historical *-a#i- had become tau-

tosyllabic and then reduced from two morae to one mora, yielding *e*. Most varieties have *ndakavanika gumbo*, replacing the *e* with *a* by analogy to the isolated verb form. Whether the development *-a#i- > *e* went through a diphthongal stage of *ay* or not, it seems likely that the subsequent synizesis is what severed the class 5 prefix from its noun. If we recognize that the Shonic class 5 prefix developed in this way *-a#i-C > *-e **C*** (synizesis with mutation), rather than simply being lost, we can consider it a regular development, parallel to that which we saw in *dàip 'be long' > Zezuru *rèba*, *áító 'our' > Zezuru *-èdù*.

Synizesis with consonant mutation might also be relevant for some reflexes of the PEB augment in Shonic languages. Shonic has lost the augment almost everywhere but there are traces of it in sandhi contexts, e.g., Ndau *baso muriro* 'make a fire' < *basa u-mu-riro. Possible cases of consonant effects are seen in C. Shonic which has both cl. 11 *ru-* and *rwu-* (e.g., *r(w)ùkúní* 'piece of firewood', *r(w)ùókó* 'arm') and class 13 *t(w)u*, e.g., *t(w)ùvànhù* 'small people' (Fivaz, 1970, pp. 59–60). These post-consonantal glides could be due to the old augments in *ʊ- in these classes and would be evidence for the impact of a preceding near-close ʊ; however, it is also possible they are due to some influence of the demonstratives *urwo*, *utwo* in those classes where the glide is regular < *u-ru-o, *u-tu-o. What one might call "preceding-U effects" are also seen in Men and Kom (Grassfields Ring languages), where labialization, velarization, and labio-palatalization of initial consonants are caused by the historic **u-* of class 3 and class 8 prefixes. "After palatalisation is completed, there is a tendency for the triggering spiranticising semi-vowel to disappear" (Kießling, 2010, p. 9).

In sum, the conditioning factor for PI-effects may not be (only) the level or character of the front vowels involved but rather the position of *i/ɪ after another vowel, resulting in diphthongization and its consequences. Just as Bantu Spirantization is connected to the reduction of 7 > 5 vowels in those languages, it is possible that PI-effects in SB are connected to the reduction of long vowels and diphthongs in its particular languages. That need not be the case in all instances of PI-effects since they also occur in languages which maintain long vowels, as we just saw in Zone JE10 where the reflexes of *aiC all maintain two morae: *aij*, *aaC*, and *aCC* (tone-bearing consonant in Ganda).

The influence of the class 5 prefix on consonants was termed "palatalization" by Meinhof (1899, 1901) based on Sotho and Venda, but then it was called "vocalization" by Doke (1931, 1954) based on Shonic and that term continued in South Africa, although it was recognized that it was also unsuitable (Westphal, 1946, pp. 29–31). Rather than focusing on the result, the term "Preceding-I effect" employed here is useful as a cross-linguistic term that can describe any phonetic process somehow conditioned by a preceding close or near-close front vowel.

3. Outcomes of PI-Effects in South Bantu

So far, we have examined the conditioning environments in which preceding *i or *ɪ triggered a set of consonant changes. In this section, we address the reflexes themselves, first looking at the outcomes of the major PEB unvoiced and voiced consonants, then at the special "Z-effects" related to *g and vowel-initial stems.

3.1. Subgroup Reflexes and Apparent Processes

The PI-effects in South Bantu follow certain patterns but vary in detail by language and by consonant. The data are summarized for voiceless stops in Table 2 and for voiced consonants in Table 3 but see supplementary File S1 for details and variants. For comparison and contrast with nasal effects, reflexes of *N-p are given (effects on *N-t and *N-k are correspondingly similar).[17]

Table 2. PI-effects of voiceless stops in South Bantu.

Proto-Eastern-Bantu	***p**	***t**	***k**	***I-p**	***I-t**	***I-k**	***N-p**
aspiration							
N20 Tumbuka	p	t	k	ph	th	kh	mph
N31,42–44 Nyanjic	p	t	k	ph	th	kh	mph
S61 Copi	h	r	k	ph	th	kh	ph
S62 Gitonga	β	r	ɣ	ph	th	kh	ph
P31 Makhuwa	v	r	∅	ph	t̪h	kh	ph
P34 Cuwabo	v	r	∅	p	t̪ ?	k	p
voicing or implosion							
S10–15 E.-C. Shonic	p	t	k	ɓ	ɗ	g	mh
S16 W. Shonic	p	t	k	b	d, d̪	g	ph
palatalization							
S31 Tswana	h	r	χ	ts^h	s, (t)ʃ	s, ʃ	ph
S32–33 N.-S. Sotho	ɸ	r	χ, h	ts^hw	ʃ	s, tʃ	ph
mixed or unclear							
S21 Venda	ɸ	r	h	f	ʃ	s, f	ph
S40 Xhosa, Zulu	ph	th	kh	?	s ?	hɬ, ʃ	mp
S53 Tsonga	h	r	k	ph ?	tʃ	kh	mh

Table 3. PI-effects of voiced consonants in South Bantu.

Proto-Eastern-Bantu	***b/β**	***d/l**	***I-b**	***I-d**	***N-b**
strengthening or preservation					
N20 Tumbuka	β	l	b	d	mb
N31,42–44 Nyanjic	∅, β/w	l/r	ɓ	ɗ	mb
S10–15 E-C. Shonic	v/β/w	l/r	ɓ	ɗ ?	mb
S16 W. Shonic	ʊ, β	l	b	d ?	mb
S40 Xhosa, Zulu	ɓ	l	ɓ	d ?	mb
S62 Gitonga	ʊ	l	ɓ	ɗ	mb
S61 Copi	ʊ	l/d	p	t	mb
P34 Cuwabo	ø	l	b	d̪	b
P31 Makhuwa	ø+	l	p	t̪	p
palatalization or mixed					
S21 Venda	β	l	v	dz	mb
S31 Tswana	b	l	ts+	ts	p’
S32–33 N.-S. Sotho	b	l	ts	tʃ	p’
S53 Tsonga	ʊ	l/r	b	ts	mb

For unvoiced stops, we can group the outcomes as reflecting aspiration, palatalization, and voicing or implosion. It is important to remember that sometimes there were probably several steps and processes to reach a given outcome.

Aspiration or strengthening: A very common outcome in South Bantu is aspiration, seen fully in Tumbuka, Nyanjic, Copi, Gitonga, and Makhuwa-Lomwe, as well as partially in Tsonga. For the Zone N languages, the aspiration appears to be a first and only step, which contrasts with the unconditioned preservation of *p*, *t*, *k*. The other languages were affected by extensive lenition (e.g., *t > *r*), and it is unlikely that the lenited forms became aspirates, so the PI-effect process must have started very early before full lenition. If there was early aspiration in P30, Cuwabo must have then undergone general de-aspiration (also seen in pre-nasalized combinations). It is possible, however, that the first step was some preservation/strengthening which prevented *I-p, *I-t, *I-k from leniting; cf. Lega (D25) *ì-támà/mà-rámà* 5/6 ‘cheek’, and the modern Comorian (G44) outcomes *p*, *tr*, *k* rather than unconditioned *β*, *r*, *h*. In that case, the unaspirated forms in Cuwabo would represent the first step of the P30 group and then the aspiration in Makhuwa and Lomwe would be later.

Palatalization: Other common outcomes are spirants and affricates in Sotho-Tswana and the closely related Venda, which must be due to palatalization at some stage. Palatalization is hardly surprising in environments like ours adjacent to front vowels and is also partially seen in Nguni and Tsongic. Commenting on Tswana, Creissels (1999, p. 331) notes, "the crucial observation is that, in all the correspondences listed above, the form of class 5 would be perfectly regular … if an additional *i̯ were present immediately after the initial consonant of the stem". This very process can be seen transparently in the Eastern Bantu language Nilyamba (F31): *tʸíma/ma-tíma* 'liver', *lʸùkà/mà-lùkà* 'leaf', *ʃʊnsó/ma-sʊnsó* 'udder' (Yukawa, 1989). Elsewhere, Rundi (JD62) also shows palatalization as a PI-effect. To reach Nguni *hɬ*, one sequence is for *I-k > *s early (as in S20–30), which would then feed into the areal change *s > hɬ seen in S30–50. South Bantu reflexes of *I-n with a palatal nasal would also fit into this category.

Voicing or implosion: The W. Shonic reflexes *b*, *d*, *g* (from which the E.C. Shonic implosives likely derive) perhaps reflect assimilation of voicing to the preceding element. Voicing is also seen in Aka of Bayanga (C10): *bàpá/mà-pàpá* 5/6 'wing'; *dòngà/mà-tòngà* 5/6 'double oyster' (Grégoire, 2003, p. 362). A possible mechanism is seen in Hendo (C73), where an allophone of the vowel /i/ in the class 5 prefix lengthens voiced consonants but appears as an unreleased voiced palatal implosive before voiceless consonants, e.g., [ddemba] 'body' but [i̯pEka] 'hand' (Demolin et al., 1999). Assimilation of the inherited consonant to this new implosive element would explain the E.C. Shonic forms.

Gemination: The gemination of almost all phonemes is a well-known PI-effect in Ganda (JE15) (Meeussen & Tucker, 1955), and *t, *d, and *s are geminated in Tetela (C71) (Kadima, 1969, p. 40). In South Bantu, it is only seen as a late effect in Makhuwa varieties Ikorovere and Enahara. However, the assimilation of a glide or other consonantal element to a consonant might have been a first step in some cases, after which the gemination was simplified leaving what looks like a preservation or aspiration scenario.

In South Bantu, the unconditioned reflexes of the PEB voiced consonants *b and *d are almost all continuants, which is why they are often reconstructed as continuants, e.g., *β and *l, rather than stops at the PSB stage and earlier. However, the PI-reflexes are mostly voiced stops (or implosives), an effect seen also in Comorian. The phonetic mechanism for the strengthening or possibly preservation of *b/β and *d/l after *I needs further exploration, but possibly a path through gemination could play a role; cf. Tetela (C71) *ddèmbà/à-lèmbà* 'body'. Subsequent devoicing is seen in Makhuwa and Copi. The affrication that we see again in Sotho-Tswana, and for *I-d in Venda and Tsonga (cf. *I-t), is easy to explain through palatalization. Because *g is lenited to a glide or is lost in South Bantu, it has a special series discussed in Section 3.3.

In sum, these reflexes show several different PI-effects: mostly aspiration of unvoiced stops, strengthening of voiced consonants and palatalization in certain subgroups. If these languages had a shared innovation, it would only have been at the first step before other developments. Perhaps the best that can be said is that all these languages and subgroups (even Nguni, discussed in Section 5) show enough PI-reflexes that we can reconstruct the frequent existence of initial conditions (activated I + consonant) at the PSB stage. South Bantu subgroups surely varied in how long these conditions continued to operate. For example, a longer life of these conditions in Shonic is suggested by the preservation of sandhi phenomena, the high frequency of PI-effects, and the morphophonemic impact on secondary input.

*3.2. PI-Effects on *s and *n*

A preservation effect is also seen for *I-s, and palatalization for *I-n. The class 5 word *ísò 'eye' is a good example of a noun where the prefix *i can create a double vowel with ef-

fects on the subsequent consonant, in this case *s (also commonly written *c; cf. Meeussen, 1967, p. 83). In some languages the PI-effect on *s is palatalization in some form, i.e., *ʃ*, *sy*, *si*, but in several South Bantu languages the role is preservation (11). Especially noteworthy are the examples from Nguni (languages a-c), which generally have few remaining PI-effects.

(11) PI-reflex es of the root *ísò 'eye'
a. Xhosa *iliso/amehlo*, Zulu *iso/amehlo*
b. Swati *lîsó/émêhlò*
c. N.Lala, S. Lala *liso* & *lihlo* (Zungu, 1999, p. 44)
d. Gitonga *lìsó/màhó*

The unconditioned Nguni reflex of *s is the lateral fricative *hl* and that is indeed what we see in the plural forms (van der Vlugt, 2023; see also van der Vlugt & Gunnink this volume). However, the PI-effect on the class 5 singular prevented the original **s* from eliding (or reversed whatever lenition had started). Of course, this difference in class 5/6 forms is natural fodder for analogy leading to the adoption of the lateral fricative in the singular also. Zulu apparently has also had *ihlo*, and the more common form in Zimbabwean Ndebele is now *ílìhlò/ámèhlò*, as also in Southern Ndebele.

In other South Bantu languages, Venda apparently has a PI-reflex (*I-s > s) that is different from the unconditioned reflex (*s > ʂ), e.g., *sámbá/mà-ʂámbá* 'word of abuse, insult'. Possibly also Nyanjic languages (e.g., Chewa *tsàmbà/mà-sàmbà* 'leaf') and Copi *di-tshiku/ma-siku* 'day', where in both languages unconditioned *s > s.

Nasals are rather infrequent at the beginning of Bantu roots, but there are two roots with medial *n that seem to show PI-effects in several languages. Reflexes of the root *ínò 'tooth' occasionally show palatalizations of *n*: Xhosa *izinyo/amenyo*, Zulu *izinyo/amazinyo*, Swati *lítînyo*, Manyika *zínyó* & *ìnyó/manyo*. So, also Ganda (JE15) *erinnyo* and Comorian (G44) *dzinyo/manyo*, both languages with a high rate of PI-effects. For the 2pl. possessive 'your', Meeussen (1967) reconstructed PB *áiɲú, thus incorporating a palatalization, but very few languages support an early palatal nasal, so better is *ínú, cf. Guthrie (1967) C.S. 2074 *yɨ́nú and Babaev (2008) *inu(e) for the independent pronoun. In South Bantu, languages with a palatal nasal are C. Shonic *-ényú*, Makhuwa *-ínyú*, and Cuwabo *-ènyù*. Sporadic palatalizations of this sort are not extraordinary, but unlike the context of Karanga *mu-pínyí* 3 'handle of tool' (*pínì BLR 2519), there is no overt *i* in C. Shonic *-ényú*. Rather, the clear PI-effect on the 1pl. possessive *áító 'our' (Section 2.1) makes a similar case for a PI-effect in the 2pl plausible.

Otherwise, instances of relevant palatalization in nasal roots are found only occasionally, e.g., in C. Shonic reflexes of extensions of *ín 'dip' (BLR 3460): *ín-ık 'dip' (BLR 3461) > *nyíka* (also Chewa *nyìkà*); *ín-ʊd 'take out of water' > *nyúra*; *ín-ʊk 'be taken out of water' > *nyùka*. Apparently, also *ìn-am 'bend over' (BLR 3465) > S. Sotho *inama* 'bow, stoop', but Chewa *nyamula* 'pick up', *nyamuka* 'rise up, set off'.

3.3. "Z-Reflexes"—Mutations of Initial Glides

So far, we have omitted a discussion of PI-effects on *g because one of the most well-attested changes in PSB concerns the loss or lenition of PEB *g (except after nasals), yielding either ø or glides. It turns out that glides also arose at the beginning of vowel-initial roots and together these two sets of glides had similar PI-effects which we will call "Z-reflexes".

Let us begin with the lenition of *g: for example, *gèd 'flow' (BLR 1347) becomes Shonic (MZ) *yèra*, Tswana *elela*, Xhosa *èla*. The loss of **g* could lead to vowel coalescence or formation of a glide from a preceding high vowel, e.g., **mà-gàdí* 6 'blood' (BLR 1299) > Tswana *màdí* 6 or **dò-gòdí* 11 'string' (BLR 1417) > C. Shonic *rw-òdzí* 11. However, in other

cases, as the result of *g, we see the glide *y* and occasionally *w* next to back vowels, either as a remnant of lenition, or more likely as a hiatus-filler. For example, the well-attested PB **gènd* 'walk, go' (BLR 1362) yields forms with and without a glide: Tjikalanga *(y)ènda*, Chewa *(y)èndà*. In short, we should be prepared that PSB *g can have reflexes of either *ø* or *y*, depending on local environments inside the same language.

It is also likely that there were often hiatus-filling glides at the front of vowel-initial nouns in early Bantu (Wills, 2022), just as we see today in Songe (L23) *e-yiso*/*ma-yi*so 'eye'. The fact is that orthographies often do not record glides, so it would be easy to think that Tswana, a language which has undergone numerous elisions, somehow keeps hiatus in *le-ino [liíná]* 'tooth'. However, Krüger and Snyman (1988, pp. 112–114) devote three detailed pages with a diagram explaining the glide-insertion rules of Tswana including the insertion of the palatal glide [IPA j] in any "succession of front vowels". For glide epenthesis in Malawian Tonga, see Mkochi (2019).

Since roots with initial *g and roots with initial vowels could both end up with initial glides, the unusual initial consonants of *ínò 'tooth' in Karanga *zínó*/*mènó* 5/6 and Chewa *dzínò*/*mànó* are the same as for *gànà 5 'hundred' (BLR 1314) in Karanga *zànà* and Chewa *dzànà*. At first, one might think that this correspondence (Karanga *z* = Chewa *dz*) has an origin in the class 5 marker *dɪ, but that does not work. For comparison, unconditioned *d > *r*, *l* in these two languages and *dí-a 'eat' (BLR 944) > Karanga *dyá*, Chewa *dyà* & *dìà*. Even if one could merge the augment and nominal prefix (*dɪ-i) and expect Bantu Spirantization, the pattern of Karanga *z* = Chewa *dz* does not match the results of Bantu Spirantization, e.g.,. *dím 'extinguish' (BLR 1046) > Karanga *dzíma*, Chewa *zìmá*, or *dì-am 'sink' (BLR 1008) > Karanga *dzàma*, Chewa *zama*.

Rather, Karanga *z* = Chewa *dz* seems to be the result of a strengthened glide, in other words, the PI-effect on **y*. Evidence comes from a number of word-internal contexts, e.g., **égam* 'lean against' (BLR 3291) usually has a glide in it, e.g., Swati & Zulu *eyáma*, but the initial front vowel can also create a structure suitable for PI-effects, as we see in Ganda *eggama* 'shelter' and Chewa *èdzàmà* 'lean against'. The noun *tòìgà 9/10 'giraffe' (BLR 3028) also has gemination in Ganda *en-tugga* with PI-effects in Shonic (KMZ) *twìzà*, Ikalanga *khwìzhà*, Nambya *ìnhwìzhà*, and Venda *ṱhùḓà* 9, which nicely shows the absorption of the glide into the changed consonant.

For class 5 vowel-initial nouns, in the appropriate sandhi conditions, the path of strengthening might be *ì-ínò 'tooth' > *i-(y)ino > *I-yino > Karanga *zínó*, Zulu *i-zinyo*. It is also possible that that the class 5 prefix became reanalyzed as part of the stem and an additional prefix was assigned, e.g., for *àdà 5/6 'trash heap' (BLR 1557) there might have been the path *ì-àdà > *yada > *i-yada > *I-yada > Chewa *dzàlà*, Zulu *í-zàlà*. This last lexeme gives us enough forms from which we can list the reflexes of *I-y for representatives of all the South Bantu groups in Table 4. Also included is another word with a vowel-initial stem *íkò 5/6 'fireplace, country' (BLR 3444) and a word with *g-initial stem *gòdò 5 'evening, yesterday' (BLR 1485).[18]

Because these reflexes from glides are consistent across many words, we should reconstruct some phoneme at the PSB stage, apparently a voiced spirant or affricate, for which I have used the general symbol *Z.[19] To make clear that we need to distinguish this series from the results of Bantu Spirantization, for contrast I have added the reflexes of *di, which are clearly a different series from *Z. The words in brackets show PI-effects inside the stem *íkò, rather than at the beginning of the word—presumably depending on PI-effect or glide epenthesis: *i-iko > *i-**k**o or *i-(y)iko. Notice that a given word shows one change or the other but not both.

Table 4. Z-reflexes from glides contrasted with reflexes of *di.

Language	*Z reflex	'trash heap'	'fireplace, country'	'evening, yesterday'[20]	Spirant
PEB	*iy	*ì-(y)àdà	*ì-(y)íkò	*ì-gòdò 5	*di
PSB	*Z	*(ì)-Zàdà	*(ì)-Zíkò	*(ì)-Zòdò	*dzi
Chewa (N31)	dz	dzàlà	dzíkó	dzùlò	*zi*
C. Shonic (S10)	z	zàràròtà (KZ)	zíkó (M)	zùrò (KM)	*dzi*
Ikalanga (S16)	zh	-zhàlànlòtà	—	zhùlò	*dzi*
Venda (S21)	ḓ	ḓàlà-ḓàlà	—	—	*dzi*
S. Sotho (S33)	tl	le-tlalakala	[ifo]	—	*di*
Zulu (S42)	z	í-zàlà	í-zîkò	í-zòlò	*zi*
Tsonga (S53)	t	tàlà	tíkó	tòlò	*ti*
Copi (S61)	t	dì-tàlà	dì-tìkó [cìkhó]	-tulo	*ti*
GiTonga (S62)	t	lì-tàlà	lì-tìgó	-tulo	*dzi*
Makhuwa (P31)[21]	s	n-sala	[-iikho]	n-suri	*ri*

In the next sub-sections, we consider some more examples of the two phonological contexts that merge in South Bantu with parallel Z-reflexes in the various languages: stems with initial vowels and stems with initial *g.

3.3.1. Z-Reflexes on Stems with Initial Vowels

Many words with initial vowels adopted Z-reflexes, but not uniformly, which is not surprising due to a variable feature like glide epenthesis. We have seen that in general Nyanjic and Shonic languages show PI-effects more extensively than others, and that is also the case with Z-reflexes, as we see in these other examples of common vowel-initial words (12).

(12) More vowel-initial words with Z-reflexes

a. *óbà 5 'sun' (BLR 1614): Tumbuka *zuŵa*, Chewa *dzùwá*, C. Shonic (KKoMZ) *zúvá*, Nambya *ì-zhùbá*, Venda *ḓúvhá*, Cuwabo *ń-zùwá*.

b. *óì 3/4, 5/6 'voice; word' (BLR 1612): Shonic (MZ) *ì-zwí* 5, Cuwabo *ń-zù*, Zulu *í-zwí*, Ronga *rhì-tó*—but without Z in Chewa *lì-ú*.

c. *ánì 5/6 'leaf; grass' (BLR 1567): Chewa *dzánì*/*mà-yánì & mà-ánì*; C. Shonic (MZ) *zánhí* and *zání* (Ko); Ikalanga *zhánì*.

The most famous nouns with initial *i in Bantu are the universal triad of *ísò 'eye', *ínò 'tooth', and *ínà name'—all in classes 5 and 6 (BLR 3405, 3472, 3464). Their status as very common words has allowed them to preserve irregular features, but that does not mean their vowel-initial structure has not been a challenge for each language. Since their class 5 forms often have an initial consonant, let us begin by recalling the evidence that their stems begin with a vowel. First, their structure in many Bantu languages is a predictable one for vowel-initial nouns with class 5/6 prefixes *dɪ-i-/a-ma, e.g., 'eye': Manyanga (H16b) *diisu*/*meeso*, Luvale (K14) *liso*/*meso* (Wills, 2022). Secondly, it is basically only in Eastern Bantu where there is an irregular initial consonant (which led to earlier reconstructions with initial *j), e.g., Unguja Swahili *djicho*/*macho* 'eye' or Zezuru *zínó*/*mènó* 'tooth'—but even in these languages it is conspicuous that the plural has no signs of a *j. In fact, Guthrie (1967, C.S. 2030, 2077) lists the word for 'eye' in over seventy languages, and in only two does he give a consonant other than *m* before the vowel-initial root in the plural.

Even if one could create a magical consonant that disappears in the plural of these words, there is another problem for a purely phonological rule—languages are not consistent about whether they have some consonant at the beginning of the stem in these words.

Sometimes yes, sometimes no, without a set pattern, although there is a gradient: generally 'eye' is least likely to have an additional consonant and 'name' has one about half the time. For example, Songe (L23) *e-yiso/ma-yiso* 'eye' but *e-fina* 'name'; Zulu *ili-so/amehlo* 'eye' but *i-zinyo/ama-zinyo* 'tooth'; Central Kongo H16b *di-inu/meeno* 'tooth' but *zina* 'name'; and Bemba (M42) *ili-ino/ameno* 'tooth' but *i-fina* 'name'. Obviously, each language had more than one prefix or hiatus-resolution strategy and different lexemes evolved in different ways. When we come to South Bantu, we also see a mixture of initial sounds in these special class 5 nouns (Table 5).

Table 5. Class 5 vowel-initial roots in South Bantu.

Language	*Z reflex	'eye'	'tooth'	'name'
Proto-Eastern-Bantu		*í(y)ísò	*í(y)ínò	*í(y)ínà
Chewa (N31)	dz	diso	**dzínò**	**dzínà**
Karanga (S14)	z	**zísó**	**zínó**	**zítá**
Ikalanga (S16)	zh	**zhíshó**	**zhínó**	(zíná)
Venda (S21)	ḓ	*íṱó, líṱó*	*ḽínó*	(dzíná)[22]
Tswana (S31)	tl+	*lè-ítlhó*	*lèínó*	*lèíná*
Zulu (S42)	z	*ísó*	**ízînyo**	—
Tsonga (S53)	t	**tihlo**	**tínyó**	—
Copi (S61)	t	*dìsó*	*dìnó*	*dìná*, **tìná**
Gitonga (S62)	t	*liso*	—	*lìná*
Makhuwa (P31)	s	*ni-itho*	*ni-ino*	**n-sina**
Cuwabo (P34)	z	*ní-ǹtó*	*ní-ìnó*	**ń-zíná**

The South Bantu reflexes on this table derive from three different morphophonemic stories, reflected in three types of print:

1. *italics* = vowel merger of *(dí-)ì-ísò etc. without any initial consonant changes.
2. **bold** = the Z-reflexes that we usually find with vowel-initial roots.
3. box = PI-reflex of *I-d; presumably the vowels in *dí-ì-ísò elided to *diso, then the class prefix *i was added again.

The multiple phenomena involved made it hard for scholars to formulate a single generalization or uniform reconstruction to accommodate all these forms. One can sympathize with Guthrie's tendency to multiple reconstructions of words, e.g., 'name' is both C.S. 831 *gíná and C.S. 2068 *yíná. However, in the end, we have to recognize that there are other processes besides the phonological at play. A second observation is that these basic lexemes were slow to generalize the Z-reflexes that came to apply to most vowel-initial nouns. That is not to say languages were completely conservative with these nouns, since they obviously adopted a variety of strategies, including adapting prefixes and augments for the forms in italics. For example, the new morphologically regular Venda *ḽíṱó* 'eye' shows that the wonderful relic *íṱó* is in the process of renewal. However, in some sense, we are seeing various strata of linguistic history which these frequently used lexemes were slower to erase than other nouns did.

Adding plural forms would complicate the table even further but it would also demonstrate the speaker's challenge of maintaining archaisms. For any language, one advantage of developing an initial consonant in class 5 nouns was that it avoided the problem of a different coalescence in classes 5 and 6, e.g., creating the new *z* in Zulu *i-zinyo/ama-zinyo* 'tooth' avoided the irregular pattern the language has in *ili-so/amehlo* 'eye'. Central Shonic has older forms for 'eyes' like *mèsó* (KKoZ) and *màsó* (M), but it is not surprising that those same varieties have also developed *mà-zísó* (KMZ), a regular plural based on the singular. The avoidance of these irregular class 5 forms with mismatches in the plural was a

reason for the renewal of certain basic words with vowel-initial roots. Besides morphological adjustments, there are even replacements as indicated by the dashes on Table 5. This is another type of story, also indicative of the problematic structures of these words: for the noun 'name' most Nguni languages now have forms of *gama* or *bizo* (<verb 'name'), as do Tsongic languages (Tsonga *vító*). Copi *dìnó* 'tooth', recorded a century ago, now only survives dialectally as *di-dino* (R. Bailey, p.c.), otherwise it has been replaced by *di-kwasa*. Likewise, the words for 'sun', 'hand' and 'egg' would usually be some of the most stable words in a language, but in South Bantu there are a number of replacements or class shifts, possibly due to their phonological structure.

In sum, vowel-initial stems required Bantu languages to employ hiatus-resolution strategies, of which we have seen varying examples of glide formation, vowel assimilation and vowel-coalescence. In the long history of these languages, glide epenthesis, which led to *Z, was clearly also part of the story, but in some words, it developed faster than in others, with morphological factors coming into play.

In addition to the PEB consonants discussed above, Eastern Bantu languages consistently show some sort of stop, spirant, or affricate in a limited set of lexemes including 'water', 'know', 'come', and 'be full' (BLR 3433, 6209, 3425, 3429). South Bantu languages are no exception, where the phonemes in question are identical with the Z-reflexes we have been discussing, e.g., Chewa *mà-dzí* 'water', *dzìwà* 'know', *dza* 'come', *dzàlà* 'be full'. For this series of reflexes, scholars have traditionally reconstructed some ancestor at the Proto-Bantu stage with the symbols **j* or **y* or some combinations, e.g., **yij* (Guthrie), **jij* (BLR), and most recently **iy* (Wills, 2022). In other words, something similar to *Z already existed in the language at the PEB stage, and the development of *I-y > *Z was a merger with it at the PSB stage. Although initially limited, the number of lexemes with this *Z phoneme expanded both from PI-effects and from merging results of spirantization in some languages.

3.3.2. Z-Reflexes on Stems with *g

In order to study PI-effects on *g, we need examples of words with initial or internal *ig as well as class 5 nouns with proto-forms in *g. We have already seen the Z-reflexes in *gòdò 5 'yesterday', *tòìgà 9/10 'giraffe'. A few more show the pattern:

- *gànzà 'hand, palm of hand' (BLR 1329) > Chewa *dzànjà*/*mànjà* 5/6 'hand'. As expected, the class 5 form has a Z-reflex and the class 6 plural shows total loss of any initial consonant.
- *gègò 5/6 'molar tooth' (BLR 1355) > Karanga, Zezuru *zèyò* 5, Ikalanga *zèù*. As expected, the second velar has become *y* (by lenition or insertion), but the strengthening at the beginning has resulted in /z/.
- *gòdò 5 'sky, top' (BLR 1486) > Zulu *í-zùlù*, Swati *lí-tùlù*, Venda *ḽí-ṱà-ḓúlú*, Tsonga *tìlò*, Makhuwa *o-sulu* 14, C. Shonic *ù-zúrù* 'upwards'.

The lenition of *g > y > ø in South Bantu in intervocalic and word-initial position created allomorphs for words beginning with *g: usually initial *ø but post-nasal *ng especially in verbs, e.g., *gàb 'divide' (BLR 1274) > Tswana *aba* 'divide' but *kabô* 9 'division' (*ng > Tswana *k*). However, the rest of the phonological system still had contrasts from *b/*mb, *d/*nd etc. so it was easy enough to restore *g in some verb forms. Even more so, by analogy any vowel-initial word in Sotho-Tswana could be assigned to the series of historic *g reflexes (*N-g, *I-g) for reflexives, e.g., *agêla* 'build for' > *i-kagêla* 'build for oneself'. The South Bantu lenition of *g is part of what led some Bantu scholars to reconstruct certain stems with initial *g instead of initial vowel, e.g., Meinhof and van Warmelo (1932) began many class 5 stems with *ɣ (the graphic predecessor of *g) and Guthrie (1967) has a doublet series in *g for some of these words: C.S. 831 *gína and C.S. 2068 *yínà 'name' and C.S. 828 *gíkò and C.S. 2056 *yìkò 'fireplace'.

In summary, in contrast to the fortition of other voiced consonants (cf. typical *I-b > *b*/*ɓ*, *I-d > *d*/*ɗ*), we do not see velars as reflexes of *I-g. Rather for *I-g, we see Z-reflexes: dental or alveolar fricatives and affricates, e.g., Zezuru *z*, Zulu *z*, Kalanga *ʒ*, Chewa *dz*, which then reduce to stops in some languages (Venda *d̠*), sometimes with devoicing (Gitonga *t*). Accordingly, the PI-effects for *g seem to have arisen at the stage after the lenition of PEB *g > SB *y.

4. Challenges for PI-Effects in Classes 5/6: Levelling and Extension

The analysis of PI-effects is complicated by challenges and limitations in the dataset. Since PI-effects are phonological changes, they can be expected to be regular inside invariable roots where the phonological environment is fixed (Sections 2.1 and 2.2). However, we have observed that there were not more than 20 such roots with an eligible internal structure. On the other hand, PEB had dozens of class 5 nouns which were the most numerous candidates for PI-effects—but the resulting allomorphy in gender 5/6 was then vulnerable to levelling processes. This push for levelling in the gender 5/6 happens because a change might have been triggered in the class 5 singular (prefix *i-) but not in the class 6 plural (prefix *ma-). Levelling is not an unimportant language change and deserves explicit attention from Bantu historical linguists, not only because it can alter patterns in the dataset, but also because differing amounts of levelling indicate differing patterns of associations made by speakers.

Across languages, we see three patterns of activity with the gender 5/6 that we discuss:

1. Levelling of the difference in C_1 in favor of the singular or plural.
2. Preservation of the difference in C_1 between classes 5 and 6.
3. Morphophonemic extension of a pattern of difference to new words.

Thus, depending on the amount of gender 5/6 levelling, South Bantu languages vary between those languages which have greatly levelled the difference (and so preserve some PI-effects inside stems, but with rare class 5 examples) and those which have preserved the difference (with numerous examples of PI-effects in class 5). As a result, although all these modern languages descend from a PEB or PSB stage where apparently the initial i- in class 5 forms could sometimes or regularly cause consonant changes, we now see languages which look very dissimilar in their PI-effect patterns.

4.1. Levelling of the Difference in C_1 in Gender 5/6

The allomorphy of the singular class 5 and plural class 6 forms was a difference which was frequently levelled, either by generalizing the singular stem to the plural or vice versa.

Levelling in general: We have several reasons to think there has been substantial levelling in classes 5 and 6. In cases where we have records of different periods, we can even document the levelling process. Older printed sources sometimes show more singular/plural differences than newer ones, e.g., Lanham (1955) reports Gitonga *li-thambo*/*ma-rambo* 'bone', but Amaral et al. (2007) only *li-rambo*/*ma-rambo* generalizing the unconditioned reflex. Likewise, we can see the levelling process taking place generationally as younger speakers of Gitonga have fewer singular/plural stem differences than older speakers do, e.g., for the plural of *lì-phàdzá* 'hoe' older speakers say *mà-vbàdzá*, but the younger say *mà-phàdzá* generalizing the PI-effect. It would be useful to search early bibles and other sources for signs of levelling in all these languages. Finally, we have variations in subgroup data: for the plural of *dúndú* 5 'chest', Karanga has *mà-túndú* but Manyika *mà-dúndú* has taken on the PI-effect of the singular. Likewise, it is reasonable to think that Ronga (with a stem difference now perhaps only in *khúme*/*mà-kúmé* 'ten') once had as many differences as closely related Tsonga (which has about five).

Another sign of levelling is free variation between allomorphs. In the Sotho-Tswana languages, the few surviving instances of a distinctive class 5 singular are conspicuous, so it is not surprising to see a new palatalized plural based on it, e.g., Northern Sotho *le-šágo* 'buttock' has interchangeable the older plural *ma-rago* (unconditioned) and *ma-šago* where "the modern plural is derived by analogy of the palatalized singular form" (Mokgokong, 1966, p. 69). The Tswana dictionary by Matumo (1993) has numerous examples of multiple singular and plurals, some due to dialectal variation. Likewise, Cole's (1955) grammar of Tswana has entries like this for 'cheek': *le-sama*/*ma-rama* 5/6 with the expected PI-reflex in the singular only, but also singular *le-tshama*, *le-šama*, *le-tšhama*, *le-rama* and "*ma-sama* etc." for the plural. In a multi-dialectal environment, a speaker may assume that an instance of singular-plural stem variation is a mixture of dialects rather than an internal variation—which provides yet another reason for levelling.

Plural to singular: The fact that some languages have fewer remaining PI-effects than others is most likely a testimony to the power of levelling from the unconditioned plural. Especially in cases where there are related verbs or noun forms in other classes, the class 5 form is more vulnerable, because its innovative C_1 is distinct from a whole network of unaffected forms. For example, in Zezuru *búró*/*mà-púró* 5/6 'big threshing stick', the PI-reflex *b* in the class 5 form is different not only from the *p* in the plural, but also different from the verb *púra* 'thresh' and the agent *mù-púrí* 'thresher'. Accordingly, it is not surprising to also find the form *mù-púró* 3/4 'threshing stick', which looks like a replacement for *búró* because instrumentals in *-o* are typically in class 5.

Singular to plural: In the word for 'skin', Nyungwe has begun to level in favor of the aspirated PI-effect in the singular stem: for *khanda* 5, speakers accept both the original plural *ma-kanda* 6 and the new *ma-khanda*. There is good reason to think that Nyungwe has been generalizing the aspiration on a large scale. In Nyungwe, there is no inherited source for initial */kh/* except from PI-effects and almost fifty class 5 nouns in the Nyungwe dictionary of Martins (1991) begin with */kh/*. However, only eight of them have plurals in */mak/*. Likewise, Nyungwe also has almost fifty class 5 nouns beginning with */ph/*, of which only eight have plurals in */map/*, and over sixty class 5 nouns beginning with */th/*, with fifteen having plurals in */mat/*. For the list, as well as examples of levelling from Tumbuka and Tsongic, see supplementary File S1.

4.2. Preservation of the Difference in C_1 in Gender 5/6

Despite the tendency to levelling—which is indeed the major pattern—there are many instances which preserve the difference in C_1 in class 5 and 6 stems. Those are most obvious in languages like Chewa, Zezuru and Gitonga. However, the fewer preservations in other languages are also important: they are often old and basic vocabulary with clear etymologies. In fact, even a few basic words with systematic PI-effects in a language create a prima facie case that PI-effects were once a common pattern in the language. How else to explain these forms? Their rarity is not hard to explain, since levelling so easily reduced the original difference. Also, many newer words simply did not undergo the changes because the environment that caused PI-effects was lost or the sound change ceased to operate. In most languages, the presence of a phonologically variable class 5 prefix did not last. The class 5 nominal prefix **i-* was at risk for several reasons mentioned above and in many SB languages the class 5 prefix system was reshaped into *li-*, *i(li)-* or some variant. Once a new prefix arose of CV shape, the environment for this sound change disappeared and new words were not subject to it.

In other words, in languages where PI-effects were not morphologized (that is to say, in most languages), the presence of that effect in a given noun is a likely indicator of the

antiquity of that noun in the language. For nouns without established PB or PEB sources, this sound change suggests the need to explore their word histories more carefully.

4.3. Morphophonemic Extension of the Difference to New Words

In most South Bantu languages, analogy produced levelling to remove singular and plural stem differences, but differences can not only be maintained but even extended. For example, in the Nyanjic (N20, 30, 42–44) and Shonic (S10) groups, the distinctive PI-reflexes became morphophonemic, marking class 5 forms in contrast to the class 6 stem. In this way, the pattern of class 5 aspiration in Nyanjic and voicing in Shonic was morphologized, extending itself to borrowings, or to the output of later phonological rules, or by producing doublets.

Extending to other nouns: In Shonic, the PI-effects on both *p and *b result in some form of *b*, but the majority of class 5 nouns with initial *b* arise from *i-p. For Ikalanga, Mathangwane (1999, p. 71) counts 30 class 5 nouns with *b*/*p* alternation but only 8 without; 35 with *d̠*/*t̠* alternation and only 5 without; 54 with *g*/*k* alternations and only 11 without. In fact, the voicing pattern became so dominant that it is not surprising to see it extended to new nouns, e.g., in a sort of hypercorrection the *b*/*p* alternation was introduced into Korekore *bèrè*/*mà-pèrè* 5/6 'maize cob' < *bèdé 'millet' (BLR 122).

Extending to borrowings. The fact that class 5 had no prefix in some languages made it a good class for adapting foreign words with unusual initial consonant. This was particularly true for fruits and vegetables, a traditional sub-category of gender 5/6. Examples of PI-effects introduced in these cases include Chewa *thímàtì*/*mà-tímàtì* 'tomato', Gitonga *li-dhalandza*/*ma-lalandza* 'orange' (Portuguese *laranja*), *li-dimwa*/*ma-limwa* 'tangerine-like fruit' (Portuguese *limão*).

Using secondary input: In addition to the environments at the stage of PSB, we see PI-effects operating on later input that developed inside the history of descendant languages. In some cases, this may indicate a continuation of the original process, but most are likely to be independent. An example of continuation is found in a Shonic reflex of *gìd 'abstain'. Because *g is lost by early South Bantu, we see the predictable Karanga *-ìra* and Zezuru *-èra*, but that loss also creates a new vowel-initial stem suitable for a PI-effect, as we see in Manyika *-dà*, which has lost the vowel and voiced the consonant.

One change that operated in various ways in various languages and therefore is assumed to be a later stage is the palatalization of velars, e.g., *kek 'to cut' > C. Shonic *chéka*. However, the class 5 instrumental of this is *jékó*/*mà-chékó* 'sickle', which is the result of applying what became the class 5 voicing rule in Shonic languages and was even extended to a new plural *mà-jékó*. If it were early, we would see *I-k > *g with different results. In Nyanjic languages, the pattern of a distinctive C_1 in through aspiration was also applied in these cases. The word for 'testicle' *kèndè 5/6 (BLR 1775)[23] palatalized and affricated to *chèndè /k/ > /tʃ/. Then, at this later stage, the aspiration rule was applied, /tʃ/ > /tʃ^{h}/, to produce Chewa *tchèndè*/*mà-chèndè*.

Likewise, the Bantu Spirantization (*ku > *fu*) seen in *kúpà 'bone' > Common Shonic and Nyanjic *(p)fúpá*/*mà-(p)fúpá* is regular, but the initial voicing of the affricate seen in the alternative Zezuru *bvúpá*/*mà-pfúpá* is rare and looks later, especially since the Karanga equivalent only means 'big bone' (with cl. 5 used for augmentative). A similar analysis holds for the voicing seen in the class 5 form *dzvìpá* of C. Shonic *svìpá* 'glossy black' (<*pìip).

Incorporation of class prefixes also opened the opportunity for new class 5 treatments. The usual PB *tóì 15/6 'ear' (BLR 3030) with class 15 prefix *ku has been reanalyzed as *kʊtʊ 5/6 in a number of Eastern Bantu languages, including Nyanjic where aspiration is then applied: *khútú*/*mà-kútú*. Likewise, the stem *bosh(w)e* for words meaning 'left hand' in various classes in E.-C. Shonic must be a reanalysis of a class 14 prefix *bʊ + *ósó 'left

arm; left side' (BLR 7360). The voicing alternation in Ndau *bòshwé/mà-pòshwé* must be a late backformation.

In general, PI-effects are a classic case where a phonological change to one part of a paradigm can have the consequence of developing into a morphophonemic pattern. In both Nyanjic and Shonic languages, PI-effects were eventually reinterpreted as markers of class 5, and in the cases when that marker did not or could not apply, e.g., loss of the prefix without PI-effect, the morphological system had to find alternative markers to indicate class 5. This search for alternative markers is one of the factors that led to the use of the class 5 concord as a nominal marker in several languages.

5. Are PI-Effects a Shared Inheritance of the South Bantu Language Branch?

We have seen PI-effects throughout the South Bantu languages, but to varying extents and with varying reflexes. Are these changes attributable to a shared inheritance?

In some ways, it is premature to discuss this question without a full study of PI-effects throughout Bantu, but leaving aside Z-reflexes, my current research indicates that there are only about a dozen other languages that have substantial examples of PI-effects, and a few more with a couple of examples. In other words, fewer than 5% of other Bantu languages exhibit PI-effects and they are almost all small groups or single languages (as we noted with Ganda). In fact, there is a recent independent case within the Makhuwa dialects where, after a vowel-initial *i*, Ikorovere has geminated the following consonant, e.g., *bín 'dance' > *iina* in other dialects, but *inna* in Ikorovere (see File S1 under Makhuwa). In contrast, the concentration of 30 languages with PI-effects throughout the South Bantu region justifies at least raising the question of a shared inheritance.

The South Bantu languages we have been focusing on in the NPS zones are most of the languages of Malawi, Mozambique, and the countries to their south. As the final leg of the Bantu expansion, they are geographically contiguous and partially isolated, with the Indian Ocean to the east and south, and the Kalahari to the west. It is not surprising that they have also been demonstrated as a linguistic branch in several ways.

The Bantu zones of Doke (1945) were mainly geographical but he also tried to identify "areas characterised by uniform or similar linguistic phenomena" (1). His zones 5(b) and (6) covered the South Bantu languages, but also included Yao and Nyakyusa. Guthrie (1967) provided Topogram 15 ("The loss of yị in reflexes of yịC", i.e., PI-effects), which basically colors the South Bantu languages in green, although he omits N40 and S30. Our current list of South Bantu languages was then explicitly grouped by Ehret (1999) under the name of the "Kusi" group on the basis of several "stem-morpheme innovations" (semantic shifts or neologisms) and a narrow sound change.[24] Since then a range of evidence for the South Bantu clade has accumulated:

(a) *Lexicon*: In the most detailed phylogenetic trees of Bantu languages based on lexicon, these languages form a distinct node. The South Bantu clade is supported by the phylogenies of Grollemund et al. (2015) and Koile et al. (2022) with 100% and 99%, respectively, of their posterior samples.

(b) *Bantu Spirantization Exceptions*: In his study of Agent Noun Spirantization (ANS), Koen Bostoen shows that the various regional results of Bantu Spirantization indicate historical isoglosses that accompany or follow other isoglosses. In particular, Bostoen (2008, p. 310) rejects an areal approach: "it is also not plausible that the overall distribution of BS could be the result of a wave-like spread across language boundaries". In particular, he notes that certain South Bantu languages of zones N, P, and S (which he calls 'southern East Bantu') (1) lack ANS and (2) do not attest Bantu Spirantization before the perfect suffix. Bostoen (2008, p. 345) concludes that these geographically

> clustered languages, at an early stage of their common history, "must have broken away from their northern relatives before BS ever became a significant change across morpheme boundaries".

These isoglosses distinguish a branch of South Bantu languages with a common history of expansion in a specific and contiguous geographical region. The consonant shifts presented in this paper share some useful comparisons with the limitation of Bantu Spirantization in South Bantu. First, both can be seen as isoglosses because they are effects seen in other Bantu subgroups but not in neighboring ones. Both sound changes could also create problematic allomorphs, but whereas signs of ANS "were wiped out under analogical pressure" (Bostoen, 2008, p. 338) at an early common stage of South Bantu, the analogical levelling of class 5/6 differences continue into the present in these languages. This is not to say that PI-effects are unique to South Bantu any more than the lack of ANS is. Given that hundreds of Bantu language communities inherited the same phonemes and morphemes, it can be expected that certain changes would occur more than once in various times and places.

The concentration of significant PI-effects in these same South Bantu languages (in contrast to very few languages elsewhere in Bantu) is conspicuous. Since the current reflexes are sometimes quite different in South Bantu subgroups, we can exclude later borrowing as a general explanation. In fact, one of the useful aspects of the differing phonological changes of PI-effects is that they can be used to identify occasional borrowing of specific words and to trace earlier stages of contact: this is most obvious in the differences between older Venda forms and those due to contact with Shonic varieties (see File S1 under Venda).[25] Nor does independent innovation by a dozen South Bantu language subgroups seem likely for a relatively rare phenomenon.

To demonstrate a linguistic isogloss of PI-effects, besides being an unusual phenomenon, two characteristics are needed: (1) a demonstration that all varieties within the isogloss have or had the feature—and these changes are indeed widespread throughout the South Bantu branch, affecting consonants in each subgroup in both nouns and verbs. (2) a boundary that separates varieties with the feature from those without. As to phylogenetic boundaries, the nearest branches to South Bantu, both on the family tree and geographically, are the zone G and P languages of southern Tanzania and northern Mozambique, which do not show PI-effects. Touching South Bantu on the northwestern side is the Botatwe group (M60) which also lacks this system of consonant changes, although a few examples in Tonga (M64) show unusual voicing like that in the adjacent Shonic languages.

Despite the phylogenetic case for a South Bantu isogloss for PI-effects, there are two major difficulties. First, although there are hundreds of examples of PI-effects compiled in File S1, their distribution in South Bantu is uneven. In particular, the Nguni group provides the fewest number, less than ten, if Z-reflexes are excluded. The likely reason is that they have maintained or generalized vocalic reflexes of *i/*ɪ:

- Class 5 nouns: Nguni languages are the only South Bantu languages other than Nambya to maintain the original nominal prefix *i in class 5. In addition, the usual process of levelling could have reduced consonant changes in singular forms in favor of the plural.
- Vowel-initial verbs: More than other groups, Nguni preserves initial vowels in verb stems, e.g., *íb 'steal' > Zulu *èbà*, Swati *éba*; *ìp-ʊd 'take off the fire' > Zulu *ephùla* & *ophùla*; *ìt-(ɪd) 'pour out' > Swati *étsa*; *í-tʊk 'be frightened, startled' > Zulu *ethuka*.

Rather, the question might be: how can we explain the presence of the PI-effects that remain in Nguni except through an early change? Some of the examples look old and solid. For example, in Nguni languages unconditioned *k > kh, but we have several cases where *k has different reflexes in the usual context of PI-effects:

(13) PI-reflexes of *k in Nguni

a. *I-k > hl [ɬ]

*ìk 'come (or go) down' (BLR 3436) > *(i)hlà* (Xh), *èhla* (Zu) 'go down'

*ìkad 'dwell; be; sit; stay' (BLR 3441) > *hlàla* (XhZu), *hlàlà* (Sw) 'sit'

*i-kómbè 5/6 'shoulder blade' (BLR 1922) > *í-hlômbè* 5/6 (Zu), *lí-hlômbè* (Sw)

b. *I-k > s, ʃ/_ u

*í-kʊt 'be satiated' (BLR 3445) > *(e)-sútha* (Zu), *é-sútsa* (Sw), *hlútha* (Xh)

*i-kómì 'ten' (BLR 2027) > *í-shùmì*/*ámá-shùmì* 5/6 (Zu), *lí-shûmì* (Sw)

One might propose an early *I-k > Proto Nguni *s which would be a feeder for standard Proto-Nguni *s > *hl* [ɬ], but with a different pattern for *I-kʊ. As a result, Nguni *hl* is the reflex of both PEB *s and *I-k (and even *I-kʊ in Xhosa). However, could the Nguni PI-effects be borrowed from Sotho? Perhaps 'ten' is a trade word (cf. N. Sotho *le-sóme*/*ma-sóme* 5/6, S. Sotho *le-shome*/*ma-shome*), but the others are core vocabulary with different sounds in the two language groups. For 'be satisfied', we have N. Sotho *khora* and S. Sotho *kgora*. In the end, it is hard to explain away the Nguni examples, even if they are few, so it seems easier to attribute the small number to levelling of the class 5 nouns, and perhaps phonological merger, e.g., we could reconstruct *I-p > ph (as in Tsonga) which merged with the usual Nguni *p > ph.

A second problem for the case of a shared South Bantu innovation is the variety of reflexes for unvoiced stops in the different subgroups: aspiration, palatalization, and implosion. Although some of these groups have been separated for many years and have undergone several sound shifts in that time, it is still not easy to describe an initial shared consonant innovation which could then generate this variety of outcomes. Perhaps further phonetic research will uncover more paths like the unreleased voiced palatal implosive seen in Hendo. At this stage, the most we can say is that the South Bantu isogloss is based on the shared conditions for consonant changes, perhaps connected to the synizesis of PEB long vowels and diphthongs. The phonological variation seen in the modern languages would then be the classic pattern of an early distinction which later underwent further changes. The variation in quantity of results, with some languages treating the reflexes in a morphophonemic way and others levelling against them, is also a pattern consistent with a change that ceased to operate long ago.

6. Conclusions

The series of consonant changes that we have examined in over 30 South Bantu languages arose from the effects of a preceding *i or *ɪ in CVIC or CV#IC environments. These PI-effects are especially visible in class 5 nouns where the nominal prefix *i was somehow activated to cause a mutation of the initial consonants of the noun stem. However, there are also a number of vowel-initial verbs where sandhi and word-internal contexts could create the environment which then affected the consonant, e.g., *ìp-ɪk 'cook, boil' > Chewa *phìkà*. Glide epenthesis, which often happened at the beginning of vowel-initial roots, led to a series of sound changes that we have called Z-effects.

The amount of data varies by subgroup. Some South Bantu subgroups (Nyanjic, Shonic, Gitonga) offer plentiful data from which the phonological pattern of this change can be reconstructed. In these languages, PI-reflexes even developed a morphophonemic character in class 5 with C_1 reflexes different from those in class 6. However, in other languages, changes affecting class 5 nouns were mostly leveled by analogy to class 6. Even when a language has limited data, a few good examples of consonant mutations are enough to demonstrate that a sound change once took place.

The phenomenon of PI-effects is not limited to the South Bantu clade, but it is by far the largest and most concentrated area with such changes among Bantu languages. The

combination of a somewhat uncommon trigger and well-defined geographical boundaries make it unlikely to have happened in so many related languages independently. However, the wide variety of contemporary outcomes (aspiration, palatalization, voicing) makes it difficult to characterize all the consonant changes as shared. Rather, what seems to have distinguished the early South Bantu branch from other branches was a certain set of rules of synizesis and glide epenthesis, i.e., the hiatus-resolution conditions which fostered consonant changes. The question of whether these conditions were an innovation, or a preservation of PEB conditions lost elsewhere, awaits comparative study.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/languages10020023/s1, File S1: South Bantu Data for PI-Effects. The following references are cited in File S1: Alves (1939), Bachetti (2006), Bailey (1995), Baumbach (1974), Baumbach (1987), Bawman (1949), Bennett (2019), Bostoen (2005), Cabrá (2012), Chebanne and Schmidt (2010), Creissels (n.d.), Crosby (1966), de Matos (1974), Dickens (1984), Dickens (1986), Dimande and Chimbutane (2022), Doke et al. (2014), dos Santos (1941), dos Santos (1949), Fortune (1955), Gowlett (1967), Guérois (2015), Guma (1971), Jones-Phillipson (1972), Kalinde (2018), Kishindo and Lipenga (2005), Kisseberth (2003), Kriel (1988), Krüger (2006), Lukusa and Monaka (2008), Malepe (1966), Marques et al. (2017), Mbatha (2006), Moreira (1924), Moreno (1988), Mtenje (2001), Ngulani (2017), Ngunga and Faquir (2012), Olivier (2009), Ponelis (1973), Poulos and Louwrens (1994), Prinsloo (1992), Rycroft (1981), Sapitwa Digital Library (2023), Schadeberg and Mucanheia (2000), Shrum (2017), Shrum (2018), SIL Moçambique (2010a), SIL Moçambique (2010b), Simbe (2004), Sitoe (2011), Sitoe and Ngunga (2000), Sitoe et al. (2008), Smyth and Matthews (1902), Tshabe (2006), Turner (1952), Vail (1971), van der Merwe and Schapera (1943), van der Wal (2009), van Warmelo (1971), van Warmelo (1989), von Staden (1973), Wentzel (1974), Wentzel (1983), Wentzel and Muloiwa (1982), Wilder et al. (1915).

Funding: This research received no external funding.

Institutional Review Board Statement: A historical study based on already published materials did not require ethical approval.

Informed Consent Statement: The only speakers consulted were themselves scholars who gave their responses in written form. Accordingly, consent was implied through the scholars' voluntary participation and written responses.

Data Availability Statement: Data are contained within the article and supplementary File S1.

Acknowledgments: Thanks to the editors and anonymous reviewers for many valuable exchanges, to Thilo Schadeberg and Koen Bostoen for their interest and encouragement, to Sara Pacchiarotti for producing the map, to Gérard Philippson and the participants of Bantu V for their comments, to Andrew Garrett, Rozenn Guérois, and colleagues of the Orikunda project for their teamwork and bibliographic support, and especially to friends and scholars who have shared their knowledge of various languages: Pita Alfandega, Richard Bailey, Tshidzanani Malaba, Winfred Mkochi, Reggemore Marongedze, Lovemore Mutonga, Mthulisi Ncube, Dion Nkomo, Sóstenes Rego, Arigo Saraiva, and Emmanuel Sithole.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

1, 2, 3	noun class 1, 2, 3
BLR	*Bantu Lexical Reconstructions 3* (Bastin et al., 2002)
C.S.	Comparative Series, in Guthrie (1967)
PB	Proto-Bantu
PEB	Proto-Eastern Bantu
PSB	Proto-South Bantu

Shonic (Ko K M Z Nd Ik Tj Na)	Korekore, Karanga, Manyika, Zezuru, Ndau, Ikalanga, Tjikalanga, Nambya
N42, P30, S21	Bantu zones following the Guthrie system

Notes

1 I use the term "Eastern Bantu" for all languages in zones EFGJMNPS (as well as K21 and many in zone D), i.e., node 10 on the Bantu time consensus tree in Grollemund et al. (2015), cf. Koile et al. (2022). For our purposes, that is more than enough material for stable reconstructions, although occasional reference will be made to the earlier stage called Proto-Bantu (PB), i.e., node 2 or earlier on the Bantu tree. South Bantu consists of clades 16 and 17 on the tree of Koile et al. (2022).

2 Unless otherwise noted, all reconstructions in this paper are mine for Proto-Eastern-Bantu (PEB), based on those in *Bantu Lexical Reconstructions 3* (BLR). Note, however, that reconstructed roots are written here without a preceding or following hyphen and without **j*, following Wills (2022).

3 Forms from modern languages follow the conventional orthographies as listed for each language in supplementary File S1, with sources for the cited forms. For simplicity, verb stems are cited without initial hyphens, but hyphens are maintained for incomplete forms of pronominal stems or parts of compounds.

4 Particular attention has been given to the effects in class 5 by Kamba Muzenga (1988) and on individual languages by Creissels (1999) on Tswana, Schadeberg (1999) on Makhuwa, Nurse and Hinnebusch (1993) on Comorian, and Meeussen and Tucker (1955) on Ganda.

5 Group names are formed by replacing final *-a* or *-e* of a major variety or location with the suffix *-ic* (as also Güldemann, 2018, p. 84). They are used merely for convenience and in no way is any particular variety considered more important or older than other varieties. A fuller list of relevant languages in these groups can be found in Maho (2009). For S62 Gitonga is used rather than Tonga to prevent confusion with Zambian-Zimbabwean Tonga (M64) and Malawian Tonga (N15).

6 This group was first called the "Senna Cluster" by Torrend (1891). Werner (1919) called the group "Nyanja" and labelled Sena and Tete (Nyungwe) as dialects of Nyanja (p. 42: "Nyanja extends, more or less, from the north end of Lake Nyasa to the Zambezi and is closely cognate—if not identical virtually with the main speech of Southern Rhodesia"). It is ironic that Watkins (1937), whose book made the variety Chewa famous, considered Chichewa as "only a variant" of the Nyanja language. For a fuller history, see Paas (2024). It is quite possible that Malawian Tonga (N15) is also part of this group but more research is needed to sort out the significant impact of contact from Tumbuka.

7 This varied group was called "Karanga" by Torrend (1891) after the largest variety. In 1928, Clement Doke was hired by the Rhodesian administrators to formulate a standard written language for the eastern part of the colony. The term "Shona" was only popularized after that, which led to a suppression of the names of the local varieties and historical diversity that covers over 150,000 m^2. The resulting Standard Shona was meant to be an amalgam of varieties but it has increasingly come under the influence of Zezuru, the variety around the capital city. In fact, current Standard Shona might better be described as a variety of Zezuru than vice versa. In Harare, Manyika students are sometimes mocked by other students for their forms and Manyika *w* (for Zezuru *v*) is sometimes marked wrong on school exams. If by the term "Shona" scholars mean Zezuru, it is best to say so explicitly, just as one would not usually use the group term "Nguni" to cite a specific Zulu form. Hannan (1984) makes it easy to cite Central Shonic forms by their particular varieties if they are not common to the group—his standard abbreviations are given above, to which I have added Eastern and Western varieties. Ndau, Kalanga and Nambya speakers do not like being described as "Shona", but at present "Shonic" seems the easiest name for the group.

8 The group is usually called "Tsonga" but that is also the name of a major variety and it is useful to be able to distinguish the two levels.

9 The *g* in Northern Sotho is epenthetic as a hiatus-filler.

10 Data for this root was collected by Angenot-Bastin (1977). Forms in *-nka* are found in at least 6 zones and *-(V)ke* or *-(V)ka* in even more, and some form of *-ike* or *-ika* can be found in over a dozen languages.

11 Unfortunately, the reconstruction of such roots was sometimes obscured by linguists who regularized CVC root structures with initial consonants everywhere, e.g., **j* or **y* in Guthrie (1967) or **j* in BLR. For reasons to remove these consonants, see Wills (2022).

12 This very old root (attested in all Bantu zones) has been treated as a unitary root, or has been derived from an extension to **ìk* 'come (or go) down' (BLR 3436): Xhosa *(i)hlà*, Zulu *èhla*. Botne (1991) argues that the initial vowel is part of a valency-changing prefix, but its low tone distinguishes it from the reflexive **í*.

13 See Marlo (2014) for cases in which the reflexive is followed by the 1sg object pronoun which can also interact phonologically with a following root-initial consonant. But the primacy of the reflexive as the immediate preverbal element in PB is seen by the fact that in some Northwest Bantu languages the reflexive is the only object marker that precedes the stem (Polak, 1986, p. 374). For more on the reflexive in Bantu languages, see also Botne (1991) and Marlo (2015).

[14] In fact, this stem so consistently shows PI-effects that one might think the reflexive marker was already incorporated into the verb at an earlier stage. But one data point without a PI-effect shows that either the simple verb and the reflexive both co-existed to a late stage, or the glide version was levelled out: Kgalagadi has the expected change in *-tjhóra* but Tswana has *-kgora* with the nasal replacement effect on reflexives (see the Tswana section of the Supplementary File S1).

[15] For the various descriptions of level-1 Proto-Bantu vowels and the changes caused by tense, [+ATR], or superclose vowels, see Janson (2007), Bostoen (2008).

[16] For earlier proposals of a metathesis of ịCV > CịV in the Sotho languages, see Meinhof and van Warmelo (1932, p. 67) and Mann (1973).

[17] For South Bantu consonant reflexes in general, as well as possiible relationships among subgroups, see van der Spuy (1990), Janson (1991), Gunnink et al. (2022). For Tswana: Creissels (2007).

[18] In addition to the examples of vowel-initial roots and *g-initial proto-roots, both of which often could become *y-initial, let me note that there is one additional root which arguably had an inherited glide: *yóòtà 5, 9 'thirst' (BLR 7055). The class 9 nasal form *nyótá* is the only one in Nyanjic or Shonic (all varieties) and has been adapted into Tswana and S. Sotho *le-nyora* 5, but Venda *ḍòrà*, Tsonga *tórhá* and Copi *dì-tòrá* have the expected Z-reflexes; the *d* in Gitonga *lì-dòrá* is irregular.

[19] As it is generally assumed that the *t* in Nguni Tekela varieties comes from *z, as seen in Zunda varieties, I also assume that the *t* of S50–60 comes from an older *z. Cuwabo also has a *z* in these reflexes.

[20] Some of these now mean the 'day before yesterday'.

[21] The Z-reflex in Cuwabo (P34) is *z*: *ńzíná* 'name', *nzílò* 'yesterday, *ńzánà* 'day before yesterday'.

[22] This is yet another type of story: borrowing. Venda *dz-* probably adapts an older version of Kalanga (Tj, Ik) *zíná*, cf. Venda *dzèú* 'molar' from Ikalanga *zèù*, Tjikalanga *zewu* (with tell-tale *-u*).

[23] In Chewa, the word can mean 'penis' in the singular and 'male genitals' on the plural. Note also *chi-sende* 7 'scrotum of animals'.

[24] The sound change that Ehret (1999, pp. 54–55) proposed (**g* > **j/i_*), if corrected in some ways, is an example of the Z-changes discussed in Section 3.3.

[25] A conspicuous example of contact is seen in the trade word *gànà 5 'hundred', where we would expect the S50–60 languages to have the shape **tana*, but instead we see Tsonga *dzana*, Copi *di-dzana* and GiTonga *li-zana*. These forms cannot be due to influence from the Nguni languages (which have *khulu*) or the Sotho-Tswana languages (which have *le-kgolo*). Rather, these irregular forms are likely borrowings of the Karanga *zànà* from a time when the rulers of Great Zimbabwe dominated the Limpopo trade routes who perhaps had direct impact or just on the Tsonga (who then heavily influenced the coastal population).

References

Alves, A. (1939). *Dicionário português-chisena e chisena-português*. Casa Portuguesa.

Amaral, A. B., Sara, A. J. L., & Eugénio, F. N. (2007). *Dicionârio de português-gitonga, gitonga-português: E compêndio gramatical*. Câmara Municipal de Oeiras.

Angenot-Bastin, Y. (1977). Les mots pour 'tout' et 'seul' dans les langues bantoues. *Africana Linguistica*, *7*(90), 1–56. [CrossRef]

Babaev, K. V. (2008). Reconstructing benue-congo person marking I. proto-bantoid. *Journal of West African Languages*, *XXXV*(1–2), 131–190.

Bachetti, C. (2006). *Gramática da língua ronga*. Paulinas.

Bailey, R. A. (1995). Issues in the phonology and orthography of Chopi (ciCopi S 61). In A. Traill, R. Vossen, & M. M. A. Biesele (Eds.), *The complete linguist: Papers in memory of patrick J. dickens* (pp. 135–170). Rüdiger Köppe Verlag.

Bastin, Y., Coupez, A., Mumba, E., & Schadeberg, T. C. (Eds.). (2002). *Reconstructions lexicales bantoues 3/Bantu lexical reconstructions 3*. Musée Royal de l'Afrique Centrale. Available online: https://www.africamuseum.be/en/research/discover/human_sciences/culture_society/blr (accessed on 15 July 2024).

Bateman, N. (2007). *A crosslinguistic investigation of palatalization* [Ph.D. dissertation, University of California-San Diego].

Baumbach, E. J. M. (1974). *Introduction to the speech sounds and speech sound changes of Tsonga*. J.L. van Schaik.

Baumbach, E. J. M. (1987). *Analytical tsonga grammar*. (Studia originalia, 3). University of South Africa.

Bawman, E. (1949). Gramática lomué. *Boletim da Soc. de Geografia de Lisboa*, *67*(41–82), 177–221.

Bennett, W. G. (2019). "Backwards" sibilant palatalization in a variety of Setswana. In E. Clem, P. Jenks, & H. Sande (Eds.), *Theory and description in African linguistics: Selected papers from the 47th annual conference on african linguistics* (pp. 41–61). Language Science Press.

Bleek, W. (1862). *A comparative grammar of South African languages*. Trübner.

Bostoen, K. (2005). *Des mots et des pots en bantou: Une approche linguistique de l'histoire de la céramique en Afrique*. (Schriften zur Afrikanistik/Research in African Studies 9). Peter Lang.

Bostoen, K. (2008). Bantu spirantization: Morphologization, lexicalization and historical classification. *Diachronica*, *25*(3), 299–356. [CrossRef]

Botne, R. (1991). Variation and word formation in Proto-Bantu: The case of *-*YlKAD*-. *Afrika und Übersee*, *74*, 247–268.

Bourquin, W. (1932–1933). Entstehung von nasalen durch den einfluss von *i* im Bantu. *Zeitschrift für Eingeborenen-Sprachen*, *23*, 195–202.
Cabrá, M. M. (2012). *Gramática de gitonga (guitonga): Língua falada na província de inhambane, moçambique*. Tipografia ABC.
Chebanne, A. M., & Schmidt, D. (2010). *Kalanga: Summary grammar*. (CASAS monograph, 75). Centre for Advanced Studies of African Society.
Cole, D. T. (1955). *An introduction to tswana grammar*. Longmans, Green & Co.
Cole, D. T. (1959). Doke's classification of Bantu languages. *African Studies*, *18*(4), 197–213. [CrossRef]
Creissels, D. (n.d.). *Setswana-French dictionary* [unpublished; author at Université Lumière (Lyon 2)].
Creissels, D. (1999). Remarks on the sound correspondences between Proto-Bantu and Tswana (S.31), with particular attention to problems involving *j (or *y), *i and sequences *NC. In J.-M. Hombert, & L. M. Hyman (Eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (pp. 297–334). CSLI (Center for the Study of Language and Information) Publications.
Creissels, D. (2007). L'influence des voyelles sur les évolutions des consonnes en tswana (S31). *Africana Linguistica*, *13*, 33–52. [CrossRef]
Crosby, J. (1966). *The morphology of the substantive in Ncheu* [Bachelor's thesis, University of the Witwatersrand].
de Blois, K. F. (1970). The augment in Bantu languages. *Africana Linguistica*, *4*, 85–165. [CrossRef]
de Haas, W. G. (1988). *A formal theory of vowel coalescence: A case study of ancient Greek*. (Publications in Language Sciences, 30). Walter de Gruyter.
de Matos, A. V. (1974). *Dicionário português-macua*. Junta de Investigações Científicas do Ultramar.
Demolin, D., Ngonga-Ke-Mbembe, H., & Soquet, A. (1999, August 1–7). *Phonetic characteristics of an unexploded palatal implosive in Hendo*. 14th International Congress of Phonetic Sciences (pp. 1047–1057), San Francisco, CA, USA.
Dickens, P. (1984). The history of so-called strengthening in Tswana. *Journal of African Languages and Linguistics*, *6*, 97–125. [CrossRef]
Dickens, P. (1986). *Qhalaxarzi phonology* [Master's dissertation, University of the Witwatersrand].
Dimande, E. M., & Chimbutane, F. (2022). Reexaminando o prefixo nominal da classe 5 na língua ronga. *Domínios de Lingu@gem*, *16*(2), 870–899. [CrossRef]
Doke, C. M. (1931). *A comparative study of Shona phonetics*. University of the Witwatersrand Press.
Doke, C. M. (1945). *Bantu: Modern grammatical, phonetical, and lexicological studies*. Percy Lund, Humphries & Co.; International Inst. of African Languages and Cultures (IIALC).
Doke, C. M. (1954). *The Southern Bantu languages*. (Handbook of African Languages 4). Oxford University Press.
Doke, C. M., Vilakazi, B. W., Malcolm, D. M., & Khumalo, M. (2014). *English-Isizulu/Isizulu-English dictionary* (4th ed.). Wits University Press.
Dom, S. (2024). Reflexive Morphology in the Kikongo Language Cluster: Variation and Diachrony. *Languages*, *9*, 113. [CrossRef]
dos Santos, L. F. (1941). *Gramática da língua chope*. Imprensa Nacional de Moçambique.
dos Santos, L. F. (1949). *Dicionário português-chope e chope-português*. Imprensa Nacional de Moçambique.
Ehret, C. (1999). Subclassifying Bantu: The evidence of stem morpheme innovations. In J.-M. Hombert, & L. M. Hyman (Eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (pp. 43–147). CSLI (Center for the Study of Language and Information) Publications.
Eiselen, W. (1924). Die Veränderung der Konsonanten durch ein vorhergehendes *i* in den Bantusprachen. *Zeitschrift für Eingeborenen-Sprachen*, *14*(2), 81–153.
Fivaz, D. (1970). *Shona morphophonemics and morphosyntax* [Ph.D. dissertation, Department of African Languages, University of the Witwatersrand].
Fortune, G. (1955). *An analytical grammar of Shona*. Longmans, Green & Co.
Fortune, G. (1985). *Shona grammatical constructions* (3rd ed., 2 vols.). Mercury Press.
Fortune, G. (Ed.). (2004). *Essays on shona dialects*. (ALLEX project). University of Zimbabwe.
Gowlett, D. F. (1967). *Morphology of the verb in Lozi* [Master's thesis, University of the Witwatersrand].
Grégoire, H. C. (2003). The Bantu languages of the forest. In D. Nurse, & G. Philippson (Eds.), *The Bantu languages* (pp. 349–370). Routledge.
Grollemund, R., Branford, S., Bostoen, K., Meade, A., Venditti, C., & Pagel, M. (2015). Bantu expansion shows that habitat alters the route and pace of human dispersals. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(43), 13296–13301. [CrossRef] [PubMed]
Guérois, R. (2015). *A grammar of cuwabo (mozambique, Bantu P34)* [Ph.D. dissertation, Université Lumière Lyon 2].
Guma, S. M. (1971). *An outline structure of Southern sotho*. Shuter & Shooter.
Gunnink, H., Chousou-Polydouri, N., & Bostoen, K. (2022). Divergence and contact in Southern Bantu language and population history: A new phylogeny in cross-disciplinary perspective. *Language Dynamics and Change*, *2022*, 1–58. [CrossRef]
Guthrie, M. (1967). *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages* (Vols. 1–4). Gregg Press.
Güldemann, T. (2018). Historical linguistics and genealogical language classification in Africa. In T. Güldemann (Ed.), *African languages and linguistics* (pp. 58–444). DeGruyter Mouton.

Hannan, M. (Ed.). (1984). *Standard shona dictionary* (3rd ed.). College Press in conjunction with the Literature Bureau.
Harford, C. (1997). When two vowels go walking: Vowel coalescence in Shona. *Zambezia*, *24*, 69–85.
Hyman, L. M. (2003). Segmental phonology. In D. Nurse, & G. Philippson (Eds.), *The Bantu languages* (pp. 42–58). Routledge.
Janson, T. (1991). Southern Bantu and Makua. *Sprache und Geschichte in Africa*, *12/13*, 63–106.
Janson, T. (2007). Bantu spirantisation as an areal change. *Africana Linguistica*, *13*, 79–116. [CrossRef]
Jones-Phillipson, R. (1972). *Affinities between venda and other Southern Bantu languages* [Ph.D. dissertation, University of London].
Kadima, M. (1969). *Le systeme des classes en bantou* [Ph.D. dissertation, University of Leuven].
Kalinde, P. (2018). *Ellomwe-english vocabulary: Emihavani and ekokholani dialects*. (SIL Language and Culture Documentation and Description 39). SIL International.
Kamba Muzenga, J.-G. (1988). Comportement du préfixe nominal de classe 5 en bantou. *Annales Aequatoria*, *9*, 89–131.
Kießling, R. (2010). Infix genesis and incipient initial consonant mutations in some lesser known Benue-Congo languages. In B. Armin, C. El Mogharbel, & K. Himstedt (Eds.), *Form und struktur in der sprache—Festschrift für elmar ternes* (pp. 187–220). Gunter Narr.
Kishindo, P. J., & Lipenga, A. L. (2005). *Parlons citumbuka: Langue et culture du malawi et de la zambie*. L'Harmattan.
Kisseberth, C. (2003). Makhuwa (P.30). In D. Nurse, & G. Philippson (Eds.), *The Bantu languages* (pp. 546–565). Routledge.
Koile, E., Greenhill, S. J., Blasi, D. E., Bouckaert, R., & Gray, R. D. (2022). Phylogeographic analysis of the Bantu language expansion supports a rainforest route. *Proceedings of the National Academy of Sciences*, *119*(32), e2112853119. [CrossRef] [PubMed]
Kosch, I. M. (2003). Sandhi under the spotlight. *South African Journal of African Languages*, *23*(3), 144–153. [CrossRef]
Kriel, T. J. (1988). *Popular Northern Sotho dictionary: Northern Sotho—English—Northern Sotho* (3rd ed.). J L van Schaik.
Krüger, C. J. H. (2006). *Introduction to the morphology of Setswana*. (LINCOM studies in African linguistic, 69). LINCOM Europa.
Krüger, C. J. H., & Snyman, J. W. (1988). *The sound system of Setswana*. Via Afrika.
Lanham, L. W. (1955). *A study of gitonga of inhambane*. (Bantu Linguistic Studies, I). Witwatersrand University Press.
Lukusa, S. T. M., & Monaka, K. C. (2008). *Shekgalagari grammar: A descriptive analysis of the language and its vocabulary*. CASAS.
Maho, J. F. (2009). *NUGL online: The online version of the new updated guthrie list, a referential classification of the Bantu languages*. Available online: https://brill.com/fileasset/downloads_products/35125_Bantu-New-updated-Guthrie-List.pdf (accessed on 4 June 2009).
Malepe, A. T. (1966). *A dialect-geographical survey of the phonology of the central, eastern and southern dialects of Tswana* [Master's thesis, University of South Africa (UNISA)].
Mann, M. (1973). Sound-correspondences and sound-shifts. *African Language Studies*, *14*, 26–35.
Marlo, M. R. (2014). Exceptional patterns of object marking in Bantu. *Studies in African Linguistics*, *43*, 85–123. [CrossRef]
Marlo, M. R. (2015). Exceptional Properties of the Reflexive in Bantu Languages. *Nordic Journal of African Studies*, *24*(1), 1–22.
Marques, O. F. F., Paulo, V. M., Vareiro, E. J. W., & Acácio, J. H. (2017). *Lomwe—English dictionary*. [Webonary.org. SIL International]. Available online: https://www.webonary.org/lomwe/ (accessed on 31 July 2024).
Martins, M. d. A. (1991). *Elementos da língua nyungwe (tete, mozambique): Gramaticá e dicionário nyungwe-português-nyungwe*. Missionários Combonianos.
Mathangwane, J. T. (1999). *Ikalanga phonetics and phonology: A synchronic and diachronic study*. (Stanford Monographs in African Language). Center for the Study of Language & Information.
Matumo, Z. I. (1993). *Setswana-english-setswana dictionary* (4th ed.). Macmillan Botswana.
Mbatha, M. (2006). *Isichazamazwi sesizulu*. New Dawn.
Meeussen, A. E. (1967). Bantu Grammatical Reconstructions. *Africana Linguistica*, *3*, 79–121. [CrossRef]
Meeussen, A. E., & Tucker, A. N. (1955). Les phonèmes du ganda et du bantou commun. *Africa: Journal of the International African Institute*, *25*(2), 170–180. [CrossRef]
Meinhof, C. (1899). *Grundriß einer lautlehre der bantusprachen*. Dietrich Reimer.
Meinhof, C. (1901). Das Tshivenḓa: Linguistische studie. *Zeitschrift der deutschen Morgenländischen Gesellschaft*, *55*(4), 607–682.
Meinhof, C., & van Warmelo, N. J. (1932). *Introduction to the phonology of the Bantu languages*. D. Reimer & E. Vohsen.
Mhute, I. (2016). A look into the concept of diphthongisation in Bantu: A case of shona. *Journal of Culture, Society and Development*, *17*, 33–35.
Mkochi, W. (2019). Glide formation is not motivated by onset requirement in Malawian Tonga, glide epenthesis is. *South African Journal of African Languages*, *39*(3), 273–280. [CrossRef]
Mokgokong, P. C. (1966). *A dialect geographical survey of the phonology of the Northern Sotho area* [Master's thesis, University of South Africa (UNISA)].
Moreira, A. (1924). *Practical grammatical notes of the Sena language*. Anthropos.
Moreno, A. (1988). *Nambya dictionary*. Mambo Press.
Mtenje, A. (2001). *Database of "Chewa dictionary"*. Comparative Bantu Online Dictionary.
Mtenje-Mkochi, A. A. (2018). 'Repair' strategies of vowel sequences in Cindali. *South African Journal of African Languages*, *38*(3), 327–335. [CrossRef]
Mugane, J. (1997). *A Paradigmatic Grammar of Gĩkũyũ*. (Stanford Monographs in African Languages). CSLI Publications.

Myers, S. (2020). An Acoustic Study of Sandhi Vowel Hiatus in Luganda. *Language and Speech*, *63*(3), 506–525. [CrossRef] [PubMed]
Myers, S., Namyalo, S., & Kiriggwajjo, A. (2019). F_0 Timing and Tone Contrasts in Luganda. *Phonetica*, *76*, 55–81. [CrossRef] [PubMed]
Ngulani, M. (2017). *English-Kalanga dictionary*. Yhelane Publ.
Ngunga, A., & Faquir, O. G. (Eds.). (2012). *Padronização da ortografia de línguas moçambicanas: Relatorio do III seminário*. (As nossas linguas #3). Centro de Estudos Africanos.
Nurse, D., & Hinnebusch, T. J. (1993). *Swahili and Sabaki: A linguistic history*. (University of California Publications in Linguistics 121). University of California Press.
Olivier, J. A. K. (2009). *Online Sesotho—English dictionary—Bukantswe ya sesotho le senyesemane sesotho*. Available online: http://bukantswe.sesotho.org/search.php (accessed on 31 January 2009).
Overton, R. M. (2018). *Vowel hiatus resolution in simbiti and the orthographic implications* [Master's dissertation, Redcliffe College].
Paas, S. (2024). *Chichewa-english/english-chichewa dictionary*. Available online: https://www.chichewadictionary.org (accessed on 31 July 2024).
Polak, L. (1983). Le réfléchi en bantou. *Africana Linguistica*, *9*, 271–304. [CrossRef]
Polak, L. (1986). Les infixes ("préfixes objets") du bantou et leur reconstruction. *Africana Linguistica*, *10*, 365–421. [CrossRef]
Ponelis, F. A. (1973). *Studies in northern sotho phonological dynamics*. University of South Africa (UNISA).
Poulos, G., & Louwrens, L. J. (1994). *A linguistic analysis of northern sotho* (1st ed.). Via Afrika Limited.
Prinsloo, D. J. (1992). Lemmatization of reflexives in Northern Sotho. *Lexikos*, *2*, 178–191. [CrossRef]
Rycroft, D. K. (1980). Nguni tonal typology and Common Bantu. *African language studies*, *17*, 33–76.
Rycroft, D. K. (1981). *Concise SiSwati dictionary: SiSwati-english, english-SiSwati*. J.L. van Schaik.
Sapitwa Digital Library. (2023). *Ellomwe—English dictionary*. [Webonary.org. SIL International]. Available online: https://www.webonary.org/ellomwe/ (accessed on 31 July 2024).
Schadeberg, T. C. (1999). Katupha's Law in Makhuwa. In J.-M. Hombert, & L. M. Hyman (Eds.), *Bantu historical linguistics: Theoretical and empirical perspectives* (pp. 79–394). CSLI (Center for the Study of Language and Information) Publications.
Schadeberg, T. C., & Mucanheia, F. U. (2000). *Ekoti: The maka or Swahili language of angoche*. (East African Languages and Dialects, 11). Köppe.
Shrum, J. (Ed.). (2017). *Takwane dictionary*. [Webonary.org. SIL International]. Available online: https://www.webonary.org/takwane/ (accessed on 31 July 2024).
Shrum, J. (Ed.). (2018). *Chuwabo—Portuguese dictionary*. [Webonary.org. SIL International]. Available online: https://www.webonary.org/chuwabo/ (accessed on 31 July 2024).
SIL Moçambique. (2010a). *Vocabulário de Echuwabo*. Sociedade Internacional de Linguística.
SIL Moçambique. (2010b). *Vocabulário de emakhuwa central*. Sociedade Internacional de Linguística.
Simbe, D. (2004). *Dicionário chisena-português*. Imprensa Universitária.
Sitoe, B. (2011). *Dicionário changana-português* (2nd ed.). Texto Editores.
Sitoe, B., Mahumana, N., & Langa, P. (2008). *Dicionário ronga-português*. CIPROMETRA.
Sitoe, B., & Ngunga, A. (Eds.). (2000). *Relatório do II seminário sobre a padronização da ortografia de línguas moçambicanas*. Livraria Moderna.
Smyth, W. E., & Matthews, J. (1902). *A vocabulary with a short grammar of Xilenge: The language of the people commonly called Chopi, spoken on the east coast of Africa between the Limpopo River and Inhambane*. Society for Promoting Christian Knowledge (SPCK).
Torrend, J. (1891). *A comparative grammar of the South-African Bantu languages: Comprising those of Zanzibar, Mozambique, the Zambesi, Kafirland, Benguela, Angola, the Congo, the Ogowe, the Cameroons, the lake region, etc.*. Paul, Trench, Trübner.
Tshabe, S. L. (2006). *The greater dictionary of IsiXhosa* (Vols. 1–3). University of Fort Hare.
Turner, W. Y. (1952). *Tumbuka-Tonga-English dictionary*. Hetherwick press.
Vail, L. (1971). The noun classes of Tumbuka. *African Studies*, *30*, 34–59. [CrossRef]
van der Merwe, D. F., & Schapera, I. (1943). *A comparative study of Kgalagadi, Kwena and other Sotho dialects*. (Comm. from the School of African Studies, new series, 9). University of Cape Town (UCT).
van der Spuy, A. (1990). Phonological relationships between the southern Bantu languages. *African Studies*, *49*(1), 119–147. [CrossRef]
van der Vlugt, N. (2023). *The development of lateral obstruents in Nguni languages (S40, Southern Bantu): A diachronic study* [Master's thesis, Leiden University].
van der Wal, J. (2009). *Word order and information structure in Makhuwa-Enahara*. LOT.
van Warmelo, N. J. (1971). Courts and court speech in Venda. *African Studies*, *30*(3/4), 355–370. [CrossRef]
van Warmelo, N. J. (1989). *Venda dictionary: Tshivenda-english* (2nd ed.). J.L. van Schaik.
von Staden, P. M. S. (1973). The initial vowel of the noun in Zulu. *African Studies*, *32*, 163–181. [CrossRef]
Wandl, F. (2020). On the relative chronology of the II regressive and the progressive palatalizations of Common Slavic. *Russian Linguistics*, *44*, 79–108. [CrossRef]
Watkins, M. H. (1937). A grammar of chichewa: A Bantu language of british central africa. *Language*, *13*(2), 5–158. [CrossRef]

Wentzel, P. J. (1974). A new practical method of describing sound changes in Venda. *Limi*, *2*(1), 45–65.

Wentzel, P. J. (1983). *Nau dzaba Kalanga/A history of the Kalanga* (3 vols.). University of South Africa (UNISA).

Wentzel, P. J., & Muloiwa, T. W. (1982). *Venda-Afrikaans-English improved trilingual dictionary*. (Comm. from the Dept. of African Languages, 25). University of South Africa (UNISA).

Werner, A. (1919). *Introductory sketch of the Bantu languages*. Kegan Paul, Trench, Trubner & Co.

Westphal, E. (1946). *A scientific analysis of the phonetics, morphology and syntax of the Venda language* [Master's thesis, University of the Witwatersrand].

Wilder, G. A., Dysart, J. P., Fuller, C. C., & Orner, A. J. (1915). *Chindau-English and English-Chindau vocabulary, with grammatical notes*. (American Board of Commissioners for Foreign Missions (ABCFM)). Kennedy-Morris Co.

Wills, J. (2022). Sorting out Proto-Bantu **j*. In K. Bostoen, G.-M. de Schryver, R. Guérois, & S. Pacchiarotti (Eds.), *Reconstructing proto-Bantu grammar* (pp. 59–101). Language Science Press.

Yukawa, Y. (1989). A tonological study of Nilamba verbs. In *Studies in Tanzanian languages* (pp. 405–450). Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

Zungu, E. M. (1999). *A Comparative phonological and morphological analysis of the north and south lala dialects of Tekela Nguni* [Ph.D. dissertation, University of South Africa].