

## File S1: South Bantu Data for PI-Effects

The material in this supplementary file, which is organized by language group, is both wide-ranging and yet also selective and abbreviated. Many of the data sets are of two extremes, both of which could take more space — either languages have more examples than a reader needs or too few (and then each possible example needs discussion). Because PI-effects in class 5 are still morphophonemic in Nyanjic and Shonic languages (and dialectally in Gitonga) a complete list of nouns for those languages would not be possible. But for most other groups, an effort has been made to provide a fairly full list of the secure examples. Please note that proto-forms are often provided, but the absence of a proto-form should not be taken as indicating that it is unknown. The launch of the online South Bantu Etymological Dictionary (SBED) in 2025 will give an opportunity to include more material than is provided here, especially to give reconstructions for more words than was possible here.

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

The languages are arranged in order of Guthrie codes:

Nyanjic: N.15 Malawian Tonga, N20 Tumbuka, N31b Chewa, N31x Ncheu, N42 Chikunda, N43 Nyungwe, N44 Sena

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

Shonic (S10): S11 Korekore, S12 Zezuru, S13 Manyika, S14 Karanga, S15 Ndaou, S16 Tjikalanga, Ikalanga, Nambya.

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): S51 Tswa, S53 Changana-Tsonga, S54 Rhonga

Copi (S61)

Gitonga (S62)

**RECONSTRUCTIONS:** Unless otherwise noted, all reconstructions are mine for Proto-Eastern-Bantu (PEB), based on those in *Bantu Lexical Reconstructions 3* (BLR) with appropriate zones. BLR has occasionally different reconstructions and often lists other classes or meanings, so it should be consulted directly. Note that reconstructed roots are written here without a preceding or following hyphen and without \*j, following Wills (2022), and with the use of \*s, since “Instead of /c/, one might just as well use the symbol /s/” (Meeussen 1967: 83). Thus, I reconstruct \*ísò ‘eye’ rather than the form \*jícò 5/6 ‘eye’ at BLR 3405 — and my reconstruction is for PEB without any claims for PB. Certain reconstructions can only be made for a more limited group of languages and are labelled as such, e.g. Proto-South-Bantu (PSB) in cases where reflexes of a reconstruction are limited to languages in zones NPS. Further explanation of them will be provided in SBED.

**REFLEXES:** For each language, forms are listed by proto-phoneme, usually in the order p-t-k-b-d-Z (the last being reflexes of PEB \*y, \*g and other). For comparison, the unconditioned reflex is given before the PI-reflex. In languages where there are only a few examples or they seem uncertain, the etymologies are provided consistently; otherwise, only occasionally since the pattern of reflexes is usually clear from class 5/6 stem variations.

### Nyanjic Group (N21, N30, N42-44)

The languages in zones N30 and N42-44 (Chewa-Nyanja-Chikunda-Nyungwe-Sena) are very closely related and show similar PI-effects: aspiration of unvoiced stops and strengthening of voiced consonants. Also grouped here are N21 Tumbuka, which has been heavily influenced by Chewa but historically was probably not too different, and N15 Malawian Tonga, heavily influenced by Tumbuka.

## N.15 Malawian Tonga

Malawian Tonga (N15) also has many interesting forms with PI-effects that deserve a separate study. Briefly, the old PI-effects looks like Shonic: unvoiced PEB stops show voicing, e.g. *bika*, 'cook', *dana* 'call', *dete/matete* 'reed', *ga/maka* 'ember, charcoal'; and voiced ones show strengthening (Turner 1952). But Malawian Tonga also has many aspirated forms, e.g. *khofi*, *makofi* (5/6) 'slap', probably influenced by neighboring Chewa and Tumbuka. More research is needed to identify which parts of the language are inherited and which are due to contact, and only then can its historical grouping be more clearly stated.

## N.21 Tumbuka

CLASS 5: PI-effects produce aspiration on unvoiced stops \*p, \*t, \*k and fortition of \*b in the usual Nyanjic pattern. There are about a dozen surviving examples of class 5/6 stem difference (e.g. *phili/mapili* 5/6 'mountain') but also evidence for levelling between those two classes due to generalizing a pattern of *ø/ma-* with extension of the sound change into the plural, e.g. *phapu/maphapu* 5/6 'lung', *jambuko/majambuko* 5/6 'fording place' < *ambuka* 'cross a river'. But in a few monosyllabic stems, the old augment *li-* is retained as a stabilizer as well: *libwe/mawe* 5/6 'stone', which was sufficiently irregular to generate a new plural *malibwe* (based on the old pattern *ø/ma-*). In a parallel way, or under the influence of Tumbuka, Malawian Tonga (N15) has a strong pattern of nouns with initial *li-* in class 5 and *mali-* in class 6.

SOURCES and ORTHOGRAPHY: There are competing orthographies for Tumbuka. Most singular/plural noun examples here are taken from Vail (1971) using his orthography (with  $\beta$  and c). Other words come from Kishindo and Lipenga (2005) and Turner (1952).

### REFLEXES:

\*p > p

\*I-p > ph

phili/mapili 5/6 'mountain' < \*pidi 'hill; heap' (BLR 2507)

phasu/mapasu 5/6 'rafter'

phumba/mapumba 5/6 'boil'

phika 'cook' < \*ip-ik 'cook, boil' (BLR 3496)

\*t > t

\*I-t > th

thete/matete 5/6 'reed'

thama/matama 5/6 'cheek'

thoci/matoci 5/6 'banana'

\*k > k

\*I-k > kh

khenya/makenya 5/6 'thread'

khongono/makongono 5/6 'knee'

khutu/makutu 5/6 'ear'

likho/malikho 5/6 'dirt' < \*ikó 'dirt' (cf. Guthrie C.S. 1093 \*kó)

khala 'live, sit, stay' < \*ikad (BLR 3441)

-kha 'alone'

\*b >  $\beta$  [written *w* in some orthographies]

\*I-b > b

bele/ma $\beta$ ele 5/6 'breast' < \*béèdè 5/6 (BLR 125)

bingu/ma $\beta$ ingu 5/6 'cloud'

banga/ma $\beta$ anga 5/6 'spot, blemish'

libwe/mawe or malibwe 5/6 'stone'

\*d > l/r

\*I-d > d

dango/marango 5/6 'law' [Turner]  
dindi/madindi 5/6 'tomb' [Turner] < \*dindí 5/6 'pit' (BLR 993)  
\*g > ø (or glide)  
mayiro 6 'yesterday' < \*gidò (BLR 1379)

Z-reflexes: both *j* and *z*. The second appears to show Chewa influence but there are probably other factors too.

jiso/maso 5/6 'eye'  
jino/mino 5/6 'tooth'  
jembe/mayembe 5/6 'hoe'  
jawo/mayawo 5/6 'cassava'  
jani/mani 5/6 'leaf'  
jiβala/mabala 5/6 'spot'  
jisumbi & sumbi/masumbi 5/6 'egg'  
josi 5 'smoke'  
maji 6 'water'  
jendayekha 5 'solitary animal' (< enda 'go' + -ekha 'alone')  
juzi 'the day before yesterday'  
jula 'open'  
—  
zina 5 'name'  
zeru 5 'wisdom'  
zuŵa 5 'sun, day'  
dazi/ma-dazi 5/6 'day, sun'  
iza 'come'

### N.31b—CHEWA

PI-effects are aspiration on unvoiced stops \*p, \*t, \*k and fortition of \*b in the usual Nyanjic pattern. There are numerous examples of class 5/6 stem difference (e.g. *phili/mapili* 5/6 'mountain') so the list below is not complete, just enough to show some of the issues.

SOURCES and ORTHOGRAPHY: Unless otherwise stated, Chewa forms are from Paas (2024, marked SP), but tones and variations are from Mtenje (2001, marked Mt) or Watkins (1937), which is useful because it is based on the speech of a single speaker. The use of *r* or *l* next to front vowels varies by local tradition and author. Paas (2024) does not indicate dialectal variation, but historically Nyanja (N31a) has been distinguished from Chewa (N31b) by more affricates (*bv*, *mbv*, *pf*, *ps*).

#### REFLEXES:

\*p > p

\*I-p > ph

phàngà/màphàngà 5/6 'cave' < \*pàngà (BLR 2400)  
pházì/mapázì 5/6 'foot, footprint'  
phéwà/màpéwà 5/6 'shoulder' < \*pègà (BLR 2433)  
phìrì/màpìrì 5/6 'mountain, hill'  
phìkà 'cook, boil' < \*ìp-ik (BLR 3496)  
phà 'kill' < \*ìp-ag 'kill' (BLR 3494)

\*t > t

\*I-t > th

thákò/màtákò 5/6 'buttocks' < \*tákò (BLR 2741)  
thùmbà/màtùmbà 5/6 'pocket, burlap sack'  
thèwèrà/màtèwèrà 5/6 'narrow loincloth worn by women'  
thupi/matupi 5/6 'pocket'

\*k > k

\*I-k > kh

khàlà 'sit, remain, be' < \*ikad (BLR 3441)  
khutira 'appreciate, be satisfied'  
khasu/makasu 5/6 'hoe'  
khúmi/màkúmi 5/6 'ten'  
khàlà 5 'ember, coal' (often màkàlà pl. only)  
khólò/màkólò 5/6 'parent'  
-ekha/-okha 'only, alone' (Watkins 1937: 122-123)

\*b > w, ø

\*I-b > b [6]

ba 'steal' < \*ib 'steal' (BLR 3387)  
bánjà/ màbánjà & maanja 5/6 'house, household, family'  
bám̀bà/mà̀bám̀bà 5/6 [Watkins] 'scar', bàm̀bà/mamba 'scales of a fish'  
bàngà/ màwàngà 5/6 [Mt: pl. manga] 'spot'  
bángò/ màwángò 5/6 'reed'  
bére [Mt: béle] 5, pl. màwère [SP, Mt], màère [Mt], mabere [SP] 'breast, mother's milk'  
bóndò/màóndò 5/6 'knee'  
bìlá 'dive' < \*ib-id 'sink' (BLR 3397)  
bùkà 'come up in the water' \*ib-uk 'come out of the water' (BLR 3399)

\*d > l

\*I-d > d [d]

dùwà/màlùwà 'flower' 5/6 < \*dòbà (BLR M1158)  
dòndò/màdòndò 5/6 'thick bush, forest' [Mt; but SP: 'Chishona']  
dòngó/màdòngó 5/6 'white clay, used for making pots' [SP: 'Chitumbuka']  
dám̀bò/madám̀bò & malambo 5/6 'small stream; marsh; meadow'  
lipiro/malipiro & dipo/malipo 5/6 [Watkins dipirò/ màlipirò; Mt: màlipò] 'wages'  
dá 'be black or dark' < \*id 'get dark' (BLR 6142)

\*s > s -sèkà 'laugh' < \*sèk (BLR 522)

\*I-s > ts ? (see "Fricativization" below)

tsiku/masiku 5/6 'day' (only Chewa has 5/6; all 3 languages have cl. 9)  
tsamba/masamba 5/6 'leaf, letter' (Chewa, Nyungwe, Sena)  
tsoka/matsoka 5/6 'misfortune' (Chewa; Nyungwe & Sena ts/ts)  
tsimbe/masimbe 5/6 'ember, coal' (Chewa also cl. 9, Nyungwe, not Sena)

\*g > ø dzànjà/mànjà 5/6 'hand' < \*gànzà 'hand, palm of hand' (BLR 1329)

Z-reflexes (dz) and other:

dzínà/màínà 5/6 'name'  
dzínò/màínò & meno 5/6 'tooth'  
dzíkò/màdzíkò & maiko 5/6 'country, land'  
dzàni/màyàni 5/6 'leaf'  
dzàna & zàna/màdzàna 5/6 'hundred' < \*gàna 5 (BLR 1314)  
(homonym with dzàna 5 'day before yesterday' < \*àná 5 (BLR 1566))  
dzàlà 'be full'  
dzìwà 'know'  
—  
diso/maso (various tones) 5/6 'eye' (not \*Z, but rather PI-effect on augment \*di)  
—  
pfùkò & fùkò/mafùkò 5/6 'tribe, mother's clan' (Nyungwe pf/pf)  
tchèndè/màchèndè 5/6 'testicle' [SP], chende/machende [Mt]  
tsìka 'descent' < ?? \*ik (BLR 3436)  
tsèkà 'shut' < ?? \*ig-ad 'shut' (BLR 3415)

tsègùlà 'open' < ?? \*ìg-ud 'open' (BLR 3416)

### N31x Ncheu

Ntcheu is a district located in the Central Region of Malawi, midway between Blantyre and Lilongwe. All material here comes from the thesis by Crosby (1966). This variety is almost identical with N31b Chewa, but that has become a mixed language, so it is useful to list this well-defined variety separately. All nouns are class 5/6 unless otherwise stated.

REFLEXES:

\*p > p

\*I-p > ph

phìlì/màpìlì 'mountain'  
phéwà/màpéwà 'shoulder'  
pházi/màpázi 'sole of foot'  
phési/màpési 'stalk of dry maize'  
phàngà/màpàngà 'cave'  
phàpò/màpàpò 'lung'

\*t > t

\*I-t > th

thàkàdzò/màtàkàdzò 'trail'  
thèwèlà/màtèwèlà 'loin cloth'  
thékà/màtéka 'portion'  
thùkùtà/màtùkùtà 'sweat'

\*k > k

\*I-k > kh

khòsì/màkòsì 'neck'  
khásù/màkásù 'hoe'  
khómà/màkómà 'wall'  
khwángwàlà/màkwángwàlà 'crow'  
khòswè/màkòswè 'house rat'  
khúmì/màkúmì 'ten'  
khádàbò/màkádàbò 'fingernail'  
khútù/màkútù 'ear' < \*kó (cl. 15) - \*tó (BLR 3030)

\*b > w, ø — lùwà/màlùwà 'flower' < \*dòbà (BLR M1158)

kù-ónà 'see' < \*bón (BLR 266)

\*I-b > b [b]

bàlà/màbàlà 'wound' < \*bàdà (BLR 28)

bélè/màbélè 'breast'

bóndò/màbóndò 'knee'

\*d > l

\*I-d > d [d]

dípò/màlípò 'payment'  
dìmbà/màlìmbà 'garden'  
dòvù/màlòvù 'spittle'  
dúmbò/màdúmbò & (less common) màlúmbò 'quarrel'  
dúngú/madúngú & & (less common) màlúngú 'surliness'  
*But also levelling (or other origins):*  
dámbo/màdámbo 'plain'  
dòkò/màdòdò 'sand bank'  
dòsà/màdòsà 'long feather'  
dòmbòlò/màdòmbòlò 'ransom'  
làmulò/màlàmulò 'law'

làmbà/màlàmbà 'leather belt'  
lùwà/màlùwà 'flower'

\*s > s

\*I-s > tsh (or secondary?)

tsháyà/màsáyà 'cheek'  
tshàmbà/màsàmbà 'leaf'  
tshíkù/màsíkù 'day'  
tshòkà/màsòkà 'misfortune'

Z-reflexes (dz) and other:

dzínà/màínà 'name'  
dzínò/mánó & mánú 'tooth'  
dzíkò/màíkò 'country, land'  
dzénjé/màénjé & màdzénjé 'pit, hole'  
dzulò 'yesterday'

—

zàrà/màzàrà 'hundred' < \*gàrà 5 (BLR 1314)  
dìsò/másó 'eye' (not \*Z, but rather PI-effect on augment \*dì)  
líú/máú 'word' (not \*Z but maintenance of augment \*dì)

—

chèndè/màcèndè 'testicle' < \*kèndè 'testicle' 5/6 (BLR 1775)  
chèngà/màcèngà 'species of insect'  
chíndù/màcíndù 'roof'  
dzilà/màzilà 'egg'

#### N.42 — Chikunda

Chikunda (centered at the confluence of the Luangwa and Zambezi Rivers) is at the western end of the Nyungwe continuum and shares all the same sound changes with Nyungwe. A publication on its historical phonology (including PI-effects) is expected from the OriKunda project.

#### N.43 — Nyungwe

The PI-effects are aspiration of unvoiced stops and strengthening of voiced consonants.

CLASS 5: There has been significant levelling with only a minority of class 5/6 nouns now showing stem variation. For example, Nyungwe *khosi/makhosi* 'neck' < \*kòti (BLR 1963) shows a *kh/k* difference in other Nyangic varieties. Nyungwe has no surviving augment but the class 5 concord remains *li-*. Class 11 has merged with class 5 and so generated a number of class 5 words beginning with *li-*, e.g. *lirime* 'tongue'.

VERBS: In each Nyanjic language, there are number of verb stems with aspirated onsets: in Nyungwe some are denominatives, and some have unaspirated equivalents that suggest dialectal differences, e.g. *khudzula* = *kudzula* 'pluck', *khonjedza* = *konceza* 'unite' (transferring the aspiration from the stem in *konkha* 'unite?').

SOURCES and ORTHOGRAPHY: All the Nyungwe data comes from Martins (1991). Nyungwe has no tones. All nouns are class 5/6 unless otherwise stated.

REFLEXES:

\*p > p

\*I-p > ph — In Nyungwe, there is no inherited source for initial *ph* except from PI-effects, but Martins has almost fifty class 5 nouns beginning with *ph*, of which the following 8 have class 6 plurals in *map*-:

phanda/mapanda 'flattening of the head on the sides'  
phaza/mapaza 'hoe'  
phere/maphere 'leprosy'  
phiri/mapiri 'hill'  
pholo/mapolo 'testicle'  
phoza/mapoza 'custard apple'  
phundu/mapundu 'lumps in food; resentment'  
phutu/maputu 'cheekbone, face'  
Also *phà* 'kill' & *phaya* < \**ip*-ag 'kill' (BLR 3494)

\*t > t

\*I-t > th — Martins' dictionary has over sixty class 5 nouns beginning with *th*-, with 15 having plurals in *mat*-.

thaka/mataka 'land'  
thako/matako 'buttock'  
thalala/matalala 'hail'  
thambiko/matambiko 'steps, ladder'  
thanga/matanga 'pumpkin' (vs. *thanga/mathanga* 'corral')  
thenga/matenga 'feather'  
thende-bunde/mathende-bunje & matende-bunje 'fruit of a climbing plant sp.'  
Cf. also *thende/matende* 1/6 'a rich man, lord'  
there/matere 'old dwelling site, ruins' (vs. *there/mathere* 'comparison')  
thinji/matinji 'inner lips of the vulva'  
thomati/matomati 'tomato'  
thope/matope 'mud'  
those/matosi 'bird droppings'  
thowa/matowa 'edible fruit of the *mtowa*'  
thowo/matowo 'edible fruit of the *mtowo*'  
thunduziro/matunduziro 'instigation of enmity' (cf. *thunduzira* 'instigate a fight')

\*k > k

\*I-k > kh — Almost fifty class 5 nouns in Martins' dictionary begin with *kh*-. But only eight have plurals in *mak*-:

khaka/makaka 'cucumber' (and *khakanamba/makakanamba* 'scales (of snake)')  
khamba/makamba 'cassava'  
khanda/makhanda & makanda 'skin'  
khomba/makomba 'grasshopper, cicada (sp.)'  
khuka/makuka 'pile gathered in the field for burning'  
khuli/makuli 'tuber; great member of family'  
khumbu/makumbu 'cereal casing on cob'  
khumi/makumi 'ten'  
-ekha/-okha 'only, alone'

\*b > ø (or glide) — *ona* & *wona* < \**bón* (BLR 266)

\*I-b > b [6] — with levelling

boko/maboko 'arm' < \**bókò* (BLR 260)  
bondo/mabondo 'knee'  
ba 'steal'

\*d > l

\*I-d > d [d] — almost no examples of stem alternation:

dambe/malambe 'fruit sp.'  
diwa/maliwa 'stone mousetrap'  
dindi/madindi 'pit, ditch' < \*dindí (BLR 993)

Spirants:

pfigu/mafigu 'banana'  
pfuwa /mafuwa 'hearth stones' < \*pígà (BLR 2566);  
pfuta/mafuta 'piece of fat, oil' < \*kútà (BLR 2138)  
pfupa/mafupa 'bone'  
pfuzi/mafuzi 'fart'

Z-reflexes: dz

dzina/madzina 'name'  
dzino/mano 'tooth'  
dziko/madziko 'country, region'  
dzulo 'yesterday'  
dzala 'be full'  
dziwa 'know'  
diso/maso 'eye' (not Z, but rather PI-effects of augment \*dɪ)

**N.44—Sena**

Different Sena sources from different regions and dates show differing levels of class 5/6 levelling, but there has clearly been a fair amount, e.g. *kholongo/makholongo* 'throat' (vs. Chewa *kholingo/mak-*) < \*kòdò (BLR 1886). Several generations ago, Alves (1939) reported stem variation in *diwa/maliwa* 'stone mousetrap' but that has been lost, so there is now either *diwa/madiwa* [PA] or *liwa/maliwa* [DS, PA]. The language has a number of verbs which begin with *ph*, *th*, or *kh*, which might be investigated for possible vowel-initial or reflexive origins.

CLASS 5: Unique for this group of languages, the class 5 concord is *i-* not *li-*, so Sena *diso ibodzi* 'one eye' rather than Nyungwe *diso libodzi*. This is possibly a preservation of the class 5 nominal prefix in some way or more likely an analogical adaptation of the class 9 concord *i-*. Class 11 words in *li-* have merged with class 5, so there is both *lirime likuru* 'big tongue' (PA) and *lirimi ibodzi* 'one tongue' (DS), which probably helped reassign the concord *li-* to nouns beginning with *li-*.

SOURCES and ORTHOGRAPHY: Unless otherwise noted, the basic Sena data comes from Simbe (2004), but most of it was checked with Pita Alfandega, a native speaker who provides class 5/6 stem alteration more often than Simbe (2004). Also noted are forms from Moreira (1924) with adapted spelling. Simbe's spelling of *ch-* and *tch-* (in the style of modern Chewa) has been changed to *c-* and *ch-* to accord with the standard orthography of Ngunga and Faquir (2012). Sena has no tones.

AA = Alves (1939)

DS = Simbe (2004)

Mor = Moreira (1924)

PA = Pita Alfandega (p.c.)

All nouns are class 5/6 unless otherwise stated.

REFLEXES:

\*p > p

\*I-p > ph

phika 'cook' (vs. ku-pika 'take the oath')  
phata 'take hold' (vs. ku-pata 'to be thick (as of forest)')  
phewa/mapewa 'shoulder' [PA, not in DS]  
phiri/mapiri 'mountain'



phadza/mapadza 'hoe' [PA map-; DS ma-]  
phala/mapala 'porridge'  
phali/mapali 'piece of pottery, tile'  
phanga/mapanga 'robber'

\*t > t

\*I-t > th

thako/matako 'buttocks' [PA mat-; DS ma-] < \*tákò 5  
thambo/matambo 'sky' [PA; not in DS]  
thanga/matanga 'pumpkin' [PA mat-; DS ma-] < \*tàngà (BLR 2793)  
thangwi/matangwi 'cause, reason' < thanga 'begin'  
thesi/matesi 'species of frog' [PA, Mor mat-; DS ma-]  
thope/matope 'mud' [PA mat-; DS ma-]  
thika/matika 'hyena' [Mor mat-; DS ma-]  
-athu 'our' 1pl. possessive

\*k > k

\*I-k > kh

khala 'live, sit, be' (vs. ku-kala 'be missing')  
khakha/makaka 'cucumber' (note the double aspiration)  
khanda/makanda 'skin, hide'  
khando/makando 'bamboo basket'  
khoko/makoko 'bark of tree, skin of fruit, scales of fish'  
kholo/makolo 'family of belonging' (vs. kholo/makholo 'mother hen')  
khosi/makosi 'neck' [PA; not in DS]  
khudzi/makudzi 'wrinkle on the forehead'  
khumbi/makumbi 'kennel, little house'  
khumi/makumi 'ten'  
khuni/makuni 'turn'  
khureya/makureya 'gull; stork'  
khutu/ makutu 'ear'  
-ekha/-okha 'alone, only'

\*b > w, ø

\*I-b > b [b] (with levelling)

ba 'steal'  
bere/mabere 'breast' < \*béèdè 5/6 (BLR 125)  
bondo/mabondo 'knee'  
banda/mabanda 'mortar' (cf. Nyungwe *banda* 'crush (stone, food)')

\*d > l

\*I-d > d [d] (with levelling)

dezi/malezi 'bird of prey' [AA]  
dongo/madongo 'clay' < \*dòngó (BLR 1138)  
dotha/madotha 'ash' < \*dòtà 3 (BLR 7330) — irreg. aspiration  
dembe/madembe 'abandoned former home' [PA]  
dalo/madalo 'place to spend the night' [DS]  
dala/madala 'house built above the ground (for protection)' [PA]  
dezi/madezi 'ointment to heal a wound' [PA]  
dimba/madimba 'small garden for vegetables'  
dungo/madungu 'sound'

Spirants:

pfufu/mafufu 'wall of a house' [PA; DS ma-] ? < \*tùtù (BLR 9515)  
pfukwe/mafukwe 'pit, animal space; residence' [DS]  
pfigu/mafigu 'banana' [PA]  
pfuri/mafuri 'base (of a table, tree)'

pfuta/mafuta 'piece of onion or solid fat' < \*kútà (BLR 2138)  
pfuwa /mafuwa 'hearth-stones' < \*pígà (BLR 2566)

Z-effects (dz) and other:

dzino/mano 'tooth'

dzana/madzana 'hundred'

dzana 'day before yesterday'

dzanja/madzanja 'hand'

dziko/madziko 'country'

dzuwa 'sun'

dziwa 'know'

—

diso/maso & madiso 'eye' (not Z, but rather PI-effects on augment \*dɪ)

### P.30 Makhuwic group

The Makhuwic group was divided by Guthrie into P.31 Makhuwa, P.32 Lomwe (Mozambique), P.33 Llomwe/Ngulu (mostly Malawi), P.34 Cuwabo, although the first three seem to be a continuum. Takwane has usually been labelled a Makhuwa variety but the data below aligns more closely to Cuwabo in several ways. Makhuwa itself is a broad cluster of varieties (Kisseberth 2003), with significant phonological differences, perhaps as much or more than is seen in some entire groups like N30-40, e.g. 'pleasant': *osiva* (Wamphula), *ociva* (Esaaka), *oziva* (Enahara), *odhiva* (Esankaci) (Ngunga & Faquir 2012). Unfortunately, most individual varieties are not well described and older scholarship on Makhuwa often does not identify the variety described, or variations. This makes it very difficult to generalize about PI-effects since dialects apparently had different reflexes, as discussed about gemination in Ikorovere (in Tanzania) at the end of this group.

CLASS 5: Presumably the class 5 prefix *ni-* < \*dɪ-i-. A certain number of class 5 forms show PI-effects, but there has been an almost complete leveling of class 5/6 differences.

INITIAL-VOWELS and REFLEXIVES: Initial vowels in verb roots are often preserved. Reflexive examples are rare, but Schadeberg (1999: 386-87) discusses the model with the verb pair *o-vitha* 'hide (tr.)', *w-i-pitha* 'hide oneself' < \*pís 'hide, cover' (BLR 2563).

SOURCES and ORTHOGRAPHY: The current orthography is that of Ngunga & Faquir (2012) based on the variety Wamphula spoken in Nampula Makhuwa group; earlier versions of the *Padronização da Ortografia de Línguas Moçambicanas*, based on a series of seminars, appeared in 1989 and 2000. Forms are cited for these varieties:

*Makhuwa*:

Mc = Makhuwa Central: from SIL Moçambique (2010a), based on the variety in and around Nampula, the capital city of the province of Nampula, with no tones marked.

Vm = dictionary of Matos (1974), with a different orthography.

Na = Enahara: from van der Wal (2009) based on the Enahara variety, following the NELIMO orthography of 2000, with tones marked.

Kot = Ekoti (coastal variety in Angoche with significant Swahili influence): from Schadeberg & Mucanheia (2000)

Kor = Ikorovere (Tunduru district in Tanzania): from Kisseberth (2003) and Kisseberth & Casimjee's unpublished wordlists with tones.

Mi = Imithupi (Masasi district in Tanzania): from Kisseberth (2003) and Kisseberth & Casimjee's unpublished wordlists with tones.

*Lomwe*:

Lo = Lomwe: from Marques et al. (2017), no tones marked.

Ll = Ellomwe: from Sapitwa Digital Library (2023), no tones marked.

*Cuwabo*:

Cs = Cuwabo: from Shrum (2018), no tones marked.

Cg = Cuwabo (Guerois 2015), tones marked.

Tk = Takwane: from Shrum (2017), no tones marked.

#### REFLEXES:

Basically, the PI-effects are aspiration for voiceless stops and strengthening of the other voiced consonants, with devoicing in Makhuwa.

\*p > v — ‘give’: *o-váha* (McNaLoLlCgTk), *o-vaya* (Cs) < \*pá (BLR 2344)

\*I-p > ph (Makhuwa-Lomwe), p (Cuwabo), but some mixture

‘cook, boil’: *w-aápéya* (McLoLl), *o-piya* (CsTk) < \*ip-ik (BLR 3496)

‘kill’: *w-iipha* (LoLl), *wiva* & *wipha* (Vm), *o-pa* (CgCsTk), *w-iíva* (MiMc), *w-ivva* (Na), *w-ivva* & *w-inva* & *w-iiva* (Kor: Schadeberg 1999) < \*ip-ag ‘kill’ (BLR 3494)

‘be long, tall’: *o-lápa* (CgCsTk), *o-lepa* (Kot) (vs. *rakama* & *rekama* in Makhuwa) < \*dàì-p ‘be long’ (BLR 784)

‘foam’: *nipheru/mapheru* 5/6 (McVm), *niphweru* (Lo), *niphoru* (Ll), *nipweru* (Tk), *nípòrú/mápòrú* 5/6 (CgCs), < \*pódù (BLR 2595)

\*t > r — ‘send’: *o-rúma* (McNaKorMi), *ó-rumá* (Cg) < \*tóm (BLR 3055)

\*I-t > tth [tʰ] (Makhuwa-Lomwe), tt/dd (Cuwabo) ?

‘call, name’: *w-iitthana* (Mc), *w-iithana* (TkLl), *w-iittànà* (KorMi), *wiichana* (Lo), *w-iittana* (Cs) < \*it-an ‘call’ (BLR 3379, 2708, 2710)

‘pour (out)’: *w-ittha* (Kor), *w-iittha* (LlKot), *w-iitthela* (NaKot) < \*it-id (BLR 3504)

‘drying platform, granary’: *ntatha* 5/6 (McLo), *ntthala/matthala* (Kot), *nddara* (Cs), *ddara* (Tk) < \*tádà (BLR 2720)

\*k > ø (+ glide insertion) — ‘squeeze’: *o(w)-ama* (McLoTk) < \*kám (BLR 1689)

‘gather’: *w-ukela* (McVm), *w-uugela* (Tk) < \*kóng (BLR 2055)

(\*ng > Makhuwa k, Cuwabo g)

‘tail’ *mw-iila* (McLoTk), *mw-iilla* (Ll), *mw-iilra* (Cs) < \*kídà 3/4 (BLR 1793)

\*I-k > kh (Makhuwa), k (Cuwabo)

‘charcoal’: *nì-khàlà/mà-* 5/6 (KorLo), *makhála* (McNa), *nikhalla* & *mankhalla* (Ll), *makala* (Tk), *nìkàlràl/má-* (Cg), *makalra* (Cs) < \*kádà (BLR 1662)

‘be, live’: *o-khala* (McNaKorMiLo), *o-khalla* (Ll), *o-kála* (CgCsTk) < \*ikad (BLR 3441)

‘ten’: *khumi* (LoLl), *kúmi/ma-* (CgCsTk) < \*i-kómì (BLR 2027), replaced in C. Makhuwa by *loko*.

‘ladle, spoon’: *niikho* 5 (Mc), *likho* & *liikho* /*makho* & *meekho* 5/6 (Kot), *liiko* 5 (Tk), *múkò* 3 (CgCs) < \*ikò (BLR 3443)

‘fireplace, cooking fire’: *vaikho* (Mc), *v-iikho*, *veikho* & *vaikho* ‘on the fire’ (KorMi), *viikho* & *vayiko* (Lo), *veeko* (Tk) < \*pa 16 + \*ikò 5/6 (BLR 3444)

‘only, alone’: Makhuwa: *-eekh-* ‘alone’, a stem of *meekh-* or *veekh-* combines with possessives to form an adverb meaning ‘alone’ or ‘by oneself’, e.g. *meekháawé* ‘she by herself’ (3s -*áwe*), *veekháú* ‘you alone’ (2s -*áú*). (van der Wal 2009: 46). It appears that the cl. 16 and 18 locative markers became fixed as the first components (cf. Swahili *peke* ‘alone’, Ikalanga *pógà* adj ‘different’) and then a new set of personal markers was suffixed. — Lomwe: *-ekha-* ‘alone, only’, with a stem of *mekha-* combining with pronominal bases, e.g. 1sg *mekhaka* ‘I alone’, 2sg *mekha*, 3sg *mekheiye*, 1pl *mekhihu*, 2pl *mekheinyu*, 3pl *mekhaya* (Bawman 1949: 203); *yeekha* (adj.) ‘only’ (Kalinde 2018).

— Cuwabo: *-eká/-oká* ‘alone’, combining with pronominal bases, e.g. 1sg *meeká* ‘I alone’, 2sg *weeká* ‘you alone’, 1pl *nooká* ‘we alone’ etc. (Guérois 2015:233).

\*b > ø (+ glide insertion) — boil': *w-iíla* (CgCs) < \*bid (BLR 176)  
 'be bitter': *u-wááwa* (McKoMi), *yooŋvaawa* adj (Cs) < \*báb (BLR 5)  
 \*I-b > p (Makhuwa), b (Cuwabo)  
 'breast, udder': *ni-pele* (McLo), *ni-pelle/ma-* (LI), *ní-belé* (CgCsTk) < \*béèdè 5/6 (BLR 125)  
 'bear child': *ó-báála* (CgCs) < \*i-bad (BLR 3391), no effect in *o-yara* (McNa)

\*d > l — 'cultivate, hoe': *o-líma* (McLoCgCsTk), *o-llima* (LI) < \*dim (BLR 968)  
 \*I-d > ? tt [t] (Makhuwa), dd [d] (Cuwabo)  
 'ash(es)': *n-tuurwa* (LI), *n-ddoora* (Tk) *mu-nddúura* 3 (CgCs), < \*dòtá 3 (BLR 7330) — but  
 possible class 9 influence *ettúural/ittúura* 9/10 (McNa), *ecuurwa* (Lo)  
 'darkness': *é-ddíima* (CgCs) < \*id-imà (BLR 3411)

\*s > th (Makhuwa), t (Cuwabo) —  
 'laugh': *o-theya* (Mc), *o-theya* (NaLoLI), *o-téya* (CgCsTk) < \*sèk (BLR 522)  
 'back; backbone': *mu-thana* (Mc), *nthána* (Na) < \*cánà 3/4 (BLR 443)

\*Z-reflexes > a wide range depending on variety and stem, perhaps reflecting different language influences: often s (Makhuwa), dd (Llomwe), z (Cuwabo), but sometimes combinations. Cuwabo *dh* looks old, e.g. *múdhúulu* 'grandchild', and reflects secondary \*Z elsewhere, e.g. *odhólíha* 'anoint' (cf. C. Shonic *zòra*, Nyungwe *dzola*, Zulu *súla* 'wipe') supposedly from \*tí-ud (BLR 2907). As discussed in the text, in vowel initial verbs, Z-reflexes are not consistent, e.g. reflexes of \*bi-ik 'put away; bury' (BLR 200) show a contrast between the ø reflex in Cuwabo (Cs) *w-eeka* and the Z-reflex in *o-sikha* (Mc). Cuwabo has both long and short and versions of the PB stem for 'know', with and without Z-reflexes: *o-ziwa* and *w-iidí*.  
 'name': *n-sina* (McLo), *ntsíná* (Na), *nddina* (LI), *ízíná* (CgCs), *ndina* & *mudhina* (Tk) < \*ínò (BLR 3464)  
 'know': *o-suwela* (McLo), *o-tsuwela* (Na), *o-dduwella* (LI), *vuwa* (Tk), *o-ziwa* (CgCs), *w-iidí* (Cg) < \*íyib (BLR 6209)  
 'water': *maasi* 6 (McKorMi), *maátsi* (Na), *mahi* (Lo), *maaddi* (LI), *maánje* (Cg), *manje* (Cs), *mahinje* (Tk) < \*íyi (BLR 3433)  
 'be full': *o-sareya* (McLo), *u-cára* 'fill up' (MiKor), *ó-dháála* (CgCsTk), *o-dhaalla* (LI) < \*íyad (BLR 3429)  
 'trash heap': *n-sala* (Mc) < \*àdà 5 (BLR 1557)  
 'evening, yesterday': *o-silo* (Lo), *oddillo* (LI), *cilo* (KorMi), *n-dilo* (Tk), *n-zilo* (Cs) < \*gìdò (BLR 1379)  
 'day before yesterday': *nsuri* (Mc), *ntsúri* (Na), *n-cúri* (ma-) Kor < \*gòdò 5 (BLR 1485)  
 'come': *o-rwa* (McLo), *o-wa* (NaLI), *ú-wa* (KorMi), *o-dha* (CgCsTk) < \*íy (BLR 3425)  
 Other Cuwabo Z-reflexes: *ízánà* 'day before yesterday', *ízùwá* 'sun', *ízù* 'voice'.

GEMINATION: In several Makhuwa varieties, an interesting phenomenon occurred that demonstrates how PI-effects could arise independently at any stage of a language. One of well-known sound changes of Makhuwa is the change of voiced \*NC to unvoiced C, e.g. \*imb 'sing' (BLR 3361) > *ipa* or *iipa* in most dialects. But, as discussed by Schadeberg (1999), in Ikorovere in Tanzania we find *ippa*. After a vowel-initial *i* (especially a bi-moraic *i*), Ikorovere has geminated the following consonant in other words too, e.g. \*bín 'dance' > *iina* in other dialects, but *inna* in Ikorovere; the regular loss of \*b in Makhuwa allows the verb to become vowel initial and available for a kind of PI-effect. Likewise, \*íba 'steal' > *iya* in other dialects but *iyya* in Ikorovere (and *iyya* in passives instead of *iya*). Since only one dialect has this gemination and it affects structures that arose inside the history of Makhuwa, it is clearly late and independent of early South Bantu PI-effects. One example makes that distinction particularly clear: \*ipag 'kill' has various dialectal reflexes in Makhuwa: *iiva* (with regular \*p > v), *ipha* (the expected PI-effect), and *inva* (nasalization and the influences of high closed vowels seem somehow related in Makhuwa). Predictably, Ikorovere has a geminated form of the regular *v*,

i.e. the common *iiva* > *ivva*, which is a completely different result than the old PI-effect seen in dialects with *ipha*. Van der Wal (2009: 16) adds similar data from Enahara in Mozambique where the first of two consonants in a sequence is moraic: *wiwwa* 'hear', *miwwa* 'thorns', *ninno* 'tooth' — all instances where the stem begins with *i* — as well as *wumma* 'be dry' < \*óm (BLR 3616). Another possible PI-effect or nasal effect in Makhuwa is the class 10 gemination noted by Schadeberg (1999: 386), "in the Eeratti dialect as observed by Kisseberth (pers. comm.), all stem initial consonants are geminated in class 10. For example, the word for 'drum' in Eeratti is *ekoma* 9/*ekkoma* 10 (< \*ji-ngoma / ji-ngoma), where other dialects have *ekomalekoma* or *ekoma/ikoma*."

## S.10 – Shonic group

The group is composed of the following members (with abbreviations in parentheses):

Eastern Shonic (E. Shonic)

S.15 Ndau (Nd)

Central Shonic (C. Shonic)

S.11 Korekore (Ko), S.12 Zezuru (Z), S.13 Manyika (M), S.14 Karanga (Kar)

Western Shonic (W. Shonic)

S.16A Tjikalanga (Tj) — mostly in Zimbabwe

S.16C Ikalanga (Ik) — mostly in Botswana

S.16B Nambya (Na)

The Shonic group underwent, or at least has preserved, a wide range of PI-effects in both verbs and nouns and even grammaticalized features of them in classes 5/6. For the unvoiced stops the final result was voicing or implosion (\*p > b, ɓ; \*t > d, ɗ; \*k > g), and for the voiced consonants the final effect was strengthening or implosion (\*b > b, ɓ; \*d > d, ɗ). Implosives only occur in E.-C. Shonic, so the initial stage of general development was probably voicing/strengthening with the advent of implosion as a final change in that Eastern-Central group.

CLASS 5: PI-effects impacted countless class 5 nouns of all Shonic varieties to the extent that it became productive, so the forms below are given only as examples. The historic class 5 nominal prefix \*i was usually lost to synzesis and then PI-effects, but its trace is apparently seen in the variant *ne-* of the comitative (coalescence of \*na-i), some other sandhi cases, and the copula prefix. For stems with initial unvoiced stops, the pattern of voicing in the singular class 5 form became the rule and was extended to unvoiced affricates generated by so-called Bantu Spirantization (see below). But in stems with the inherited voiced stops (or continuants) \*b and \*d, we see systematic levelling so that the class 5/6 forms now show no difference in their onsets. Perhaps those difference were seen as too irregular after the re-analysis of the rule as "class 5 voicing" based on the voiceless stops.

REFLEXIVE: The current reflexive pronoun in Shonic is segmentally identical with the class 8 object pronoun in all varieties (E.-C. Shonic *zvi*, W. Shonic *zwi*) so it must have replaced the inherited PSB reflexive \*i at an early stage. But in addition to PI-effects in familiar PSB verbs (e.g. *gúta* 'become satisfied' < \*i-kúta), there are a number of verbs with PI-effects that could come from lemmatized reflexives from the South Bantu stage, e.g. *dèva* 'follow' KZ in contrast to the more common form *tèvera* 'follow' KKOMZ with the applicative suffix.

SOURCES: For the varieties cited in the text and below, the sources are:

E. Shonic (Ndau): Wilder et al. (1915).

C. Shonic (Korekore, Zezuru, Manyika, Karanga): Hannan (1984).

W. Shonic (Ikalanga): Mathangwane (1999).

W. Shonic (Tjikalanga): Ngulani (2017).

W. Shonic (Nambya): Moreno (1988).

Due to the differences in orthography among these languages and the fact that numerous examples are cited in the texts, I have limited the list here to Central Shonic. Further examples on class 5 in

Central Shonic: Fortune (1985, 1984), Doke (1931: 125-131). For Western Shonic, see the forthcoming SBED.

REFLEXES:

**Central Shonic**

\*p > p

pá 'give' < \*pá 'give' < (BLR 2344)  
póra 'heal (of a wound), cool' < \*pód (BLR2589)

\*I-p > b [b]

bángá/mà-pángá 5/6 'knife' < \*pángà (BLR 2399)  
bàdzá/mà-pàdzá 5/6 'hoe'  
báni/mà-páni 5/6 'plain, open grassland'  
bànzú/mà-pànzú 5/6 'sharp wooden peg'  
bázi/mà-pázi 5/6 'branch'  
bàpìrò/mà-pàpìrò 5/6 'wing'  
bèrè/mà-pèrè 5/6 'hyena'  
bòfù/mà-pòfù 5/6 'blind man'  
bòndé/mà-pòndé 5/6 'reed, mat'  
bìka 'cook'  
bàya 'kill'  
báta 'hold, grasp'

\*t > t

támba 'play, dance'  
téma 'chop, decide'

\*I-t > d [d]

dákó/mà-tákó 5/6 'buttock' < \*tákò (BLR 2741)  
dàmbá/mà-dàmbá 5/6 'klapper apple'  
dàngá/mà-tàngá 5/6 'corral for animals'  
démó/mà-témó 5/6 'axe' < téma 'chop'  
dòmbó/mà-tòmbó 5/6 'rock'  
dúra/mà-túra 5/6 'granary'  
dùrí/mà-tùrí 5/6 'wooden mortar'  
dána 'call, love one another'  
dáidza 'call'  
dìra 'pour'  
daba (Na), dá(v)ira (KKoMZ), dàbílà (Ik) answer' < \*ít-ab-(íd) (BLR 3381, 6031)  
-édú 'our' 1pl. possessive < \*-áítú

\*k > k

káma 'milk, comb (hair)'  
kúra 'grow'

\*I-k > g

gàrá/mà-kàrá 5/6 'ember' < \*kádà (BLR 1662)  
gárwe/mà-kárwe 5/6 'crocodile'  
gòmbá/mà-kòmbá 5/6 'pit'  
gòmò/mà-kòmò 5/6 'hill, mountain'  
gònye/mà-kònye 5/6 'worm'  
gòré/mà-kòré 5/6 'cloud, year'  
gùmbò/mà-kùmbò 5/6 'leg'  
gàra 'sit, live, stay' < \*ikad (BLR 3441)  
gúta 'become satisfied (with food)' < \*í-kut (BLR 3445)  
-ega/-oga 'alone, only', combining with pronominal bases, e.g. ndógà 'I alone' (all parts of Shonic (for Ikalanga, see Chebanne and Schmidt 2010: 84-85)

\*b > v [β, v], ø

váva 'be bitter, itch' < \*báb (BLR 5)

vìrí, -ìrí 'two' < \*bidí (BLR 190)

vénga 'hate'

vèza 'carve'

\*I-b > b [β]

bere 5 'udder' (M; other varieties replace with *zamu*) < \*béèdè (BLR 125)

bùwè/màbùwè & bùvè/màbùvè 5 'spider' < \*bòbì (BLR 290)

bá 'steal'

bereka 'give birth to' < ì-bed-ik (BLR 3394)

\*d > r

rára 'sleep, lie down'

rúma 'bite'

\*I-d > d [d'] is rare, and examples of *d* in Manyika and Ndaue are perhaps through Nyanjic influence. Some Venda examples in *d* suggest a Shonic borrowing. Otherwise, where there is a clear etymology, usually *r* in both singular and plural of class 5/6 nouns, which could be levelling but must be early since it holds for all varieties. Possibly \*I-d > d > r occurred on a separate path from \*I-t > th (?) > d.

\*dòtà 'ashes' (BLR 7330) > dòtà/mà-dòtà 5/6 'ash' (KMZNd) & ròtà/mà-ròtà 5 'ash' (MZNd).

Cf. Nyungwe & Sena *dotha/madotha* 5/6, but also the old Shonic compound *zàràdòtà* 5 'ash heap or pit' (KZ).

\*dómb 'ask for' (BLR 1112) > dòmbè/mà-dòmbè 5/6 'beggar' (M), ròmbè 5 'strolling beggar' (KMZ), mùròmbè 1 'poor person; beggar' (KZ)

\*dindí 'hole in ground, pit' (BLR 993) > Venda *dindí* 5 (cf. Venda *mulindí* 3 'hole in ground') where Manyika also has *dindí* 5 (mad-) 'hole in the ground' (cf. *gomba* elsewhere in C. Shonic) — however, Nyungwe also has *dindi/madindi* and Chewa *dzènjé/maenje* & *madzenge*.

\*dòngó 'clay; clod; soil' (BLR 1138) > dòngó/màdòngó 5/6 'clay (for pottery); clay pit' (M)

\*dòbà 'flower' (BLR M1158) > rùvà/mà-rùvà 5/6 (KMZ), rùvà 5, pl. nduva 9 & maruva 6 (Nd), i-luba (Na), lùbà 5 'cotton wool; flower' (Ik), dùvà 5 'florescence of leguminous plant' (M),

\*díbà 'trap (falling, flat stone)' (BLR 955) > rívà 5/6 (KMZ)

\*dòpà 'blood' (BLR 1141) > ròpà 5 'blood' (KKoMZ), lòpà 5 (Ik), ì-lòpà (Na)

\*i-dimà 'darkness' (BLR 3411) > rímà 5 (KMZ), límà (Ik); but díma 5 (M), chidima 7 (Nd), chadima (Z) — cf. *cidima* 7 (Nyungwe, Sena), *mdima* 3 (Chewa, Nyungwe)

dòrò/màdòrò 'beer' (KKoMZNd) — no clear etymology.

\*s > s but sh- common in Ndaue, Karanga, and W. Shonic, sometimes through influence of class 9 sèka 'laugh' (KKoMZTj); shèkà & sèsà 'cause to laugh' (Nd)

sòna 'sew' (KKoMZNd)

sàmba 'wash self' (M), with various semantics: tsàmba (M), shàmba (KNdNa), shàmbira (KKoZ)

\*I-s > s – sh variation apparently by variety but there could be other factors

sángó/ma-sángó 5/6 'bush country' (KKoMZ), shàngò 9 (Tj), shàngó (Nd)

simbá/ma-simbá 5/6 'strength, power' (KMZTj)

For \*isò 'eye', Shonic varieties typically have medial *s*, but note *zhishó* (Ikal), *hhíhó* (Tj)

## VOICING OF AFFRICATES AND FRICATIVES:

The power of the current class-5-voicing rule is seen in its partial extension to unvoiced affricates. Many Shonic fricatives and affricates were generated by so-called Bantu Spirantization and palatalization, and as a rule those fricatives and affricates are unaffected by PI-effects, for example:

\*tu, ku > pfu (fu Karanga, thu Ikalanga)

pfúndó/mà-pfúndó 5/6 'knot' (KMZ) < \*kúnd 'tie knot' (BLR 2128)

pfúdzì (ZNd), fúdzì 5 (KM), thúdzì (Ik) 'shoulder' < \*túdì (BLR 3987)

\*bu > vhu [vu]

vhùdzì/mà-vhùdzì 5/6 'hair' (KMZ) < \*bùdì 5/6 (BLR 369)

vhú/mà-vhú 5/6 'soil' (K), ì-vhú (KoMZ) < \*bú 3,5,6 (BLR 355)

\*di > dzi

dzìvàmà-dzìvàmà 5/6 'pool, lake' (KKoMZ) < \*dìbà 5/6 (BLR 1025)

\*du > bvù (du Ikalanga)

bví/mà-bví 5/6 'knee' (KNd), ì-bví (KoMZ) < \*dúí (BLR 1239)

bvùrè/mà-bvùrè 5/6 'shade' (KoM), bvùrí (Nd), mùbvùrí 3 (MZ) < \*dúdì 3/4

As can be seen, no PI-effects are visible in the typical class 5 forms above, only Bantu spirantization. But at some stage, the PI-effects on voiceless stops were re-analyzed as a rule for class 5 voicing which led to some examples of additional voicing of affricates. Doke (1931: 129) notes that *pfeni/mapfeni* 5/6 'baboon' is "very rarely used; the augmentative *bveni* is more common, with plural *mabveni*". Similarly, the augmentative *bvupa/mabvupa* instead of *pfupa/mapfupa* 'bone' < \*kúpà (BLR 2132) and *dzvatsvátsva/matsvatsvátsva* 5/6 'spider sp.'. The voiced affricates, of course, are unchanged.

Shonic languages also developed an affricate from palatalizing the unvoiced velar (\*k<sub>l</sub> > tʃ), and that affricate could then be subject to class 5 voicing: *ch* > *j* [dʒ]

jékó/mà-chékó 5/6 'sickle' < chéka 'cut'

jírá/mà-chírá 5/6 'cloth'

jèmbèrè/mà-chèmbèrè 5/6 'large old woman'

Before stems commencing with *ts*, *f*, *s*, *sv*, and *sh*, the change to a voiced affricate depressor is only realized as a secondary augmentative prefix (Fortune 1955:50):

dzvíngá 5 'large bundle' < svíngá 5 'bundle of wood'

dzìrù 5 'big, fat heifer' < tsìrù 5 'heifer'

dzíkàná 5 'large girl' < mù-síkàná 1 'girl'

jìrì 5,9 'large bird' < shìrì 9 'bird'

The usefulness of the voicing rule is obvious in the first two examples where even an inherited class 5 noun can be further marked, with voicing marking the augmentative (seen also in 'baboon' above).

Z-reflexes > z

zísó 5 'eye' (KKoMZ), dzísó (Nd), zhíshò/ mèshó (Ik), ìzhìshó (Na); mèsó 6 (KKoZ), màsó (M), màzísó (KMZ)

zínó 5 'tooth' (KKoMZNd), zhínò (Ik), ìzhínò & ìzhènó (Na), ìnyó 5 (manyó) & zínyó 5 (M); mènó 6 (KKoIk), màzínó (KMZ)

rw-ìzì 11 'river' (KMZ), rùzì (M), lwizi (Na) Nambya; njìzì 10 (pl) (Ik)

zàná 5 'hundred' (KKoMZNd)

zíkó 5 'hearth' (M), but the usual word is *choto* in all dialects

zòngóròrò 5 'millipede' (KMZ) zhòngólòlò (Ik) < \*gongododo (BLR 1454)

zà 'come' (MZNdK(á)), zhà (Ik), zhá (Na)

zára 'become full' (KKoMZNd), zhálà (Ik), zhàlá (Na), dzara (Nd)

zíva 'know' (KMZ), zíbà (Kb)

## S.21 – Venda

It is well known that Venda experienced massive contact from Shonic (S10) varieties due to the long political domination of the kingdoms of the Zimbabwean plateau. The numerous doublets ("New Venda") that resulted in Venda from contact with Karanga-Kalanga are discussed at the end of this section on Venda. Accordingly, the first part below is limited to the direct Venda reflexes of Proto-South-Bantu ("Old Venda"), which are also quite numerous. Venda's unconditioned and PI-effect reflexes are similar to the Sotho-Tswana group, and they clearly underwent several changes together, although Venda's later borrowings from Sotho complicate the story a bit. Nevertheless, the Venda data gives a sense of how vigorous the PI-effects must have once been in the Sotho-Tswana group, too.



CLASS 5: Until recently, Venda had no remaining segmental augment or nominal prefix in class 5 (just PI-effects), although the nominal prefix \*i- is perhaps preserved in *ìpfí* 'voice'. But lately the prefix *lì-* has been spreading, especially to mark adoptives, and it must come from the concord series. It features the uncommon phoneme *l*, which is also seen in *lú* 'eat' (\*dì-) and its derivatives (*mu-lí* 1 'eater', *lisa* 'feed'), but it is not clear if it is from \*dì-i- or from \*dì-V in pronominal environments. There is occasional leveling of the singular and plural stems, but in general Venda speakers maintain many allomorphs in both the verbal and nominal systems. The adoption of many Shonic forms with class 5/6 stem difference would have reinforced the inherited difference. Despite some levelling, Venda is particularly useful for roots with initial \*b, since several other languages have levelled them more strongly.

SOURCES and ORTHOGRAPHY: Venda has a distinctive standard orthography used here and in the major lexical sources: van Warmelo (1989), Wentzel and Muloiwa (1982), and Westphal (1946); cf. also Wentzel (1983). All nouns are class 5/6 unless otherwise stated.

#### OLD VENDA REFLEXES:

\*p > fh [ɸ]

\*I-p > f

fúrí & fhúrí/màfhúrí 'pumpkin' < \*pòdí (BLR 2631)  
 fáló 5 'strong purgative; enema medicine' (cf. fhala 'scrape')  
 fàfhú [rare, dialectal] & fhàfhú/màfhàfhú 'lung' < \*pàpó (BLR 9444)  
 fára 'seize, grasp, hold' < \*i-pat (cf. BLR 2414)  
 fiha 'work self into a rage; fume' ? < \*ip-ik 'cook, boil' (BLR 3496)

\*t > r

\*I-t > sh [=ʃ]

shámá/màráamá 'cheek' < \*támà (BLR 2744)  
 shàmbá/màràmbá 'fruit of the muramba' < \*támbá (BLR 8419)  
 shàngà/màràngà 'spherical calabash' < \*tàngà (BLR 2793)  
 shámbó/màràmbó 'bone' < \*támbó (BLR 8426) — (clearly Old Venda, because this root is only in Southeast Bantu, not in Nyanjic or Shonic)  
 sháhó/màráhó 'buttock' < \*tákò (BLR 2741)  
 shèvhò/màrèvhò 'dressed hide, kaross' (Cf. Karanga dèbwè, Ikalanga dèbgè)  
 shúndá/màrúndá 'nape of the neck'  
 -ashu '1s pl possessive' < \*aitu  
 shèla 'pour' < \*it-id (BLR 3504)  
 shùwa 'be startled' < \*í-toka (BLR 3511)

\*k > h

\*I-k > s, f (before back vowels)

sádzíngò 5 'large pot with small mouth' (cf. -hadzinga 'roast dry' < \*kád-ing BLR1680)  
 sálá/màhálá 'ember' < \*kádà (BLR 1662)  
 sáḍá [rare] & sháḍá/màháḍá 'shoulder'  
 —  
 fòbvù/mahòbvù 'habitual thief'  
 fòndò/màhòndò 'common rat' < \*kònjò 5/6 (BLR 6759)  
 fúlú 'big' cl. 5 adj., obsolete form now replaced by *lìhulu*  
 fúmí, màhúmí 'ten' < \*i-kómì (BLR 2027)  
 fúyù/màhúyù 'wild fig (fruit)'  
 lù-fò/mpfò 11/10 'wooden spoon' < \*ikò 'ladle' (BLR 3443)  
 fúra 'be satisfied' < \*í-kot (BLR 3445)

\*b > vh [β]

\*I-b > v [v]

vánḑá 5 'enclosure at entrance of village' < \*bánzá (BLR 97)  
 vèlé/màvhlèlé 'maize' < \*bèdé 'millet, sorghum' (BLR 122)  
 véngó 5 'hatred' < -vhénga 'hate' < \*beng (BLR 154)  
 vóhó/màv'hóhó 'arm, foreleg of animal' < \*bókò (BLR 260)  
 vumbá 5 'clay' < \*bómbà (BLR 335)  
 vúnú 5 'rectum', màv'húnú 'buttocks' < bónò 'waist' (BLR 337)

\*d > l

\*I-d > dz

dzáló 5 'sleeping place, usually in bush' < \*lála 'lie down' < \*dáad (BLR 795)  
 dzimà 5 'hoeing time' < \*dim- 'cultivate' (BLR 968)  
 dzòfhà/màlofhà & màdzòfhà 'large clot of blood'; pl. 'blood, various pains' < \*dòpà 3 (BLR M1141)  
 dzùvhà/màlùvhà 'flower, blossom' < \*dòbà (BLR M1158)  
 dzábí/malabi & madzabi 'old dirty rag', derog of labi 5 'rag' (Afr. Lappie) — this borrowing shows the productive pattern.

\*s > ʈ — ʈàlula 'single out' < \*càd-ud- 'to choose' (BLR 402)

ʈàmba 'wash' < \*càmb (BLR 434)

lù-ʈàhá/ʈàhàhá 11/10 'jawbone' < \*càyá (BLR 497)

\*I-s > s

sámbá/mà-ʈámbá 'word of abuse, insult'

sanzwo 5 'ritual washing or bath by relatives of a person dead of consumption' (cf. -ʈanzwa, old causative of ʈamba 'wash')

sambi 5 (pl. mit- 4, maʈ- 6, mas- 6) 'herd of cattle (esp. mit-), flock of goats, birds'

\*g > ø

àna 'take an oath' < \*gàn (BLR 8264)

Z-Reflexes (ḑ) and other

ḑòrà 'thirst' < \*yòòtà 5, 9 'thirst' (BLR 7055)

ḑúvhá 'sun' < \*óbà 5 'sun' (BLR 1614)

lí-ʈà-ḑúlú 5 'heaven, sky' < \*gòdò (BLR 1486)

má-ḑí 6 'water'

ḑà 'come'

ḑála 'full'

ḑívha 'know'

—

ító & lító 5 'eye' — no Z reflex

línó 5 'tooth' — no Z reflex

dzíná 5 'name' — Shonic influence or reflex before front vowel?

#### INFLUENCE FROM SHONIC:

Venda is well-known for having a series of doublets, e.g. "the co-existence of differently-shifted descendants of the same root, such as *-hadzinga, sadzingo / -kadzinga, gadzingo; -tshea / dzheko, dzhesi; -hoha, khokho / -kokodza, goko; khoro / mukoto; muraga/danga*", etc. (van Warmelo 1989: pref. 19). Venda doublets are both so widespread and so systematic that they seem to arise from two different grammars. The source is not hard to identify — for centuries, the Karanga and Kalanga kingdoms of the Zimbabwean plateau dominated and became the chieftain class of much of the Venda populations living in the Limpopo River valley. Accordingly, van Warmelo states "the really basic fact about Venda: that it is a fusion or amalgam of Sotho-type and Shona-type components" (pref. 6). But that fusion is so deep that it affects the grammar as well as vocabulary. Attention to the multiple traditions involved (including the impact of Tsonga refugees) has mostly been made by ethnographers rather than linguists but important exceptions are Jones-Phillipson (1972) and Wentzel (1983), who emphasizes the closer connection of Venda to W. Shonic (Kalanga) than C. Shonic (Karanga).

The reason that that Shonic influence on Venda has been difficult to specify is that the impact on the language was so thorough and so significant that it has been internalized into the core vocabulary and productive grammar. Even the basic words ‘big’ and ‘all’ have doublets, *-húlú /-kúlú* and *-óthé /-ósé*, although the latter Shonic variants have become obsolete. In general, Venda’s regional dialects are not so different, and most attention (both socially and scholarly) is given to the difference between the speech of commoners (*vhalanda*) and the specialized vocabulary of the court (*musanda*). While one might assume that the court language is filled with the Shonic lexicon of the conquerors, it is in fact a list of words with varied origins (van Warmelo 1971), reflecting a deliberate social effort to find alternative idioms to the standard speech, in a way that recalls Nguni *hlonipha* politeness avoidance. Whatever their specific origins, court language has helpfully maintained some lexical items or variants that might otherwise have been lost.

Accordingly, the lexical and phonological doublets are usually seen as distantly related variation inside Venda, i.e. van Warmelo’s “co-existence of differently-shifted descendants of the same root”. For example, he gives these reflexes of Proto-Bantu \*k: “ka, ko, ku, kw > h, -, hw, k, kw” and for the class 5 changes “li+k > s, sh, f, g” (van Warmelo 1989: pref. 20). Sotho and Shonic forms are often mentioned but treated as parallels to one variety or another rather than in a historical paradigm (Wentzel 1974). All this must be sorted out into what I would call “Old Venda” and “New Venda”, acknowledging that the Shonic influence is a more recent layer (Janson 1991-92: 91). From the sociolinguistics of conquest, one might consider a parallel with modern English after the Norman conquest and cultural Latinization, where the borrowed or influenced vocabulary is now about half the word-stock of English and Latin derivational processes are often more productive than inherited ones. Nevertheless, from the current lexicon of English, one can partially sort out Anglo-Saxon and Latin roots by Grimm’s Law. For Venda, we have both unconditioned sound changes and PI-effects as a guide to distinguishing Old Venda and New Venda.

The inherited consonant series of Old Venda has many similarities to Sotho (and historically Venda was probably in a continuum with that group) and is quite distinct from that of New Venda, (which is basically the same as Shonic). Using the traditional orthography, I have set out the unconditioned sound changes in classes 5 and 6 from PEB to Old Venda and New Venda below.

#### Unconditioned and PI-effect reflexes of consonants in Old and New Venda

PEB >	Old Venda		Old Venda examples (cl. 6 — cl. 5)
*p >	fh [ɸ]	*I-p >	f      mà-fhúrí — fúrí ‘pumpkin’
*t >	r	*I-t >	sh [ʃ]      mà-rámbó — shámbó ‘bone’
*k >	h	*I-k >	s, f      mà-hálá — sálá ‘ember’
*b >	vh [β]	*I-b >	v      mà-vhádzí — vádzí ‘scar’
*d >	l	*I-d >	dz      mà-lùvhà — dzùvhà ‘flower’
*g >	ø	*I-g >	ɖ      tsh-àndà 7 ‘hand’ — ɖàndà (21) 5 ‘big hand’

PEB >	New Venda		New Venda examples (cl. 6 — cl. 5)
*p >	p	*I-p >	b      mà-pàkò — bàkò ‘cave’ = Karanga
*t >	t	*I-t >	d      mà-tèmbà — dèmbà ‘type of calabash’ = Karanga
*k >	k	*I-k >	g      mà-kándá — gándá ‘peel’ = Karanga
*b >	vh [β]	*I-b >	b      mà-vhádzí — bádzí ‘scar’
*d >	l	*I-d >	d      mù-lìlò 3 ‘fire’ — didò 5 ‘huge fire’

I have chosen a number of New Venda terms which are identical in Karanga to show the similarity. While these two patterns exist, they are complicated by Venda’s further development of this material (and we have not even mentioned the many nasal and spirantization combinations). The usual class 5/6 singular and plural leveling can apply, so inside Old Venda there can be extension of the singular stem to the plural. Besides the regular *shámbó/màrámbó* 5/6 ‘bone’, also reported are *rámbó* and *lì-rámbó*. Likewise, the original class 5 form *fúlú* of the adjective ‘big’ (\*I-k > f before back vowels) has been levelled to *lì-húlú* with the rest of the stem. ‘Wild fig’ can be *fúyù* or *húyù* or *lì-húyù*, all with pl. *mà-húyù*. But this does not happen too much, perhaps because Venda speakers are used to having

great diversity in their systems. However, the two systems do overlap with some sounds (e.g. \*b, \*d and certain affricates), so one can expect analogical influence between the systems or mergers. An example of the combined systems is provided by \*pángà ‘bush-knife’ (BLR 2399) which produces the *lùfhángá/phángá* 11/10 ‘knife’ by regular Old Venda correspondences. To this was added a diminutive *tshìpángá* 7 ‘small knife’ and an augmentative *bángá/màpángá* 5/6 ‘big knife; slash-knife’ (all borrowed or based on Shonic).

The strongest tendency is for Shonic voicing to be applied to other stems which have voiceless stops or affricates. This is particularly true for creating class 5 augmentatives via voicing, for example, the Old Venda *lùhúni/khúni* 11/10, ‘a piece of firewood’ has an augmentative with New Venda voicing *gúni* 5/6 ‘large piece of firewood’ (cf. Ikalanga *gúni* ‘log’) and levelling in the plural, both the expected *màkúni* and also *màgúni*. For some of the Venda augmentatives, there might be (or have been) the same form in Shonic but I have not found them, e.g. *gùmbú/màkùmbú* 5/6 ‘big calabash’ (cf. *tshikùmbú* 7). Venda *dzhílá* 5 ‘big tail’ was created by voicing the stem of *mù-tshílá* 3 ‘tail’ (\*kída BLR 1793), but Shonic languages have only the class 3 form of this root, and nothing in class 5.

The scale of Shonic impact on Venda can be appreciated from a single data point: in van Warmelo’s (1989) dictionary, there are over 150 class 5 nouns with initial g- and most have plurals in *mak-*. Some have obvious Karanga or Kalanga equivalents, some are dialectal, some are flora, some are foreign borrowings (e.g. *gùgá* 5/6 ‘cook boy’), but some are words that would typically seem old (e.g. *gàil/màgàil* & *màkàil* ‘nit, egg of lice’) without this phonological trace. An interesting form is the ordinary Venda word for ‘egg’ (which should be a stable lexical item): *gùmbá/màkùmbá* 5/6, which clearly has Shonic voicing. Although it has no obvious Shonic parallel, it could possibly be derived from the ancestor of Nambya *ihùmbá* 9 ‘snail’ and Tjikalinga words for ‘snail shell’ (*kumbu*, *n’umba*, *whomba*) which then evolved inside Venda. In general, one cannot accept a form as inherited (Old Venda) if there is any possibility of borrowing a Shonic form. Some of these excluded forms may be inherited or old and adapted but, since the Shonic impact on Venda was a long time ago (over 500 years ago), even today’s Venda forms with varying semantics can sometimes be borrowings.

Helpful cases for assessing the difficulty in sorting out the layers of Venda come from vowel-initial stems. Although the data is limited, Venda seems to align with the S30 group in having fewer Z-reflexes for vowel-initial nouns. In addition to the usual body parts (*itó*, *litó* ‘eye’, *linó* ‘tooth’), Venda has no hiatus-filler in *lì-ìvhà* 5 ‘pigeon, dove’ (cf. N. Sotho *le-eba* vs. Swati *lí-tùbà* & *lí-jùbà* ‘dove’). Shonic only has class 9 forms like Ikalanga *njibà*. Vowel-initial verbs follow a pattern (either inherited from a common Venda-Sotho ancestor or influenced by Sotho) where class 5 deverbative nouns begin with a velar. This has been generalized from verbs which once had initial \*g that was lost. For example, the Venda verb *ámusa* ‘suckle’ comes from an extension of \**ámu* ‘suck’ (BLR 3178), and it has an augmentative deverbative *gámúsò* 5 ‘excessive secretion of milk’. However, there is also *ḁámú* 5 ‘breast’, where the *ḁ* looks like Old Venda, but the word is probably a nativization of Karanga *zámú* 5 ‘breast, udder’. That stem for ‘breast’ is restricted to Shonic, where in all varieties it has replaced \**bèèdè* 5/6 ‘breast; udder’ — otherwise it is one of the most stable words in the history of Bantu.

Nativization or camouflaged borrowing among related languages can be particularly deceptive. In fact, I even wonder whether the Venda word *ḁáná* ‘hundred’ was such an adaptation of Karanga *zànà*. As we saw above, the irregular (*d*)z- forms for that word in the languages in the lower Limpopo (Tsonga *dzana*, Copi *dì-dzana* and GiTonga *lí-zana* — where we expect initial *t*) can most easily be explained by the trading power of the Zimbabwean kingdoms. It would be surprising for Venda not to have been influenced too. Could *ḁáná* be a replacement for the standard regional word seen in Sotho-Tswana *lekgoro* and Nguni *ikhulu*? Every Venda speaker must have known the equivalence of Venda *ḁ* and Karanga *z* through the equivalence of the so-called class 21 augmentative prefixes Venda *ḁi-* (*ḁithu* ‘big monster’, *ḁandá* ‘big clumsy hand’) with Shonic *zi-* (*zìngóndi* ‘big baboon’ KM, *zànzà* ‘big handful’ Z). In another core word, Venda *dzíná* ‘name’ (instead of *ḁ-*) also looks like a rather direct adaptation of Kalanga *zíná*. Finally, Venda has two words meaning ‘pool in river’, with two different reflexes of \**dìbà* 5/6 ‘pool, deep water’ (BLR 1025): *dzìvhà/màdzìvhà* (like Shonic) and *tìvhà/màtìvhà* (like Tsongic) and both are used in traditional idioms about a chief’s death — so it is not easy to separate

the layers at first glance. In short, in a language where such deep contact has reached a “fusion”, one must be very careful, and a quantitative approach to the lexicon, even core vocabulary, has its limits.

### S.30 – Sotho-Tswana group

The Sotho-Tswana group is famous for its numerous consonant changes, but they are mostly due to preceding nasals and subsequent glides. The other two possible sets of consonant changes (changes from PI-effects or subsequent vowels) are less common. In general, the Sotho-Tswana frequently preserves initial vowels in verb stems, which indicates the lesser frequency by which those same vowels underwent synizesis in this group.

The S30 languages at present have no augment and only a standardized single class 5 prefix *li-*. Thus, the forms with PI-effects must have arisen either from an earlier augment + prefix combination (\**di-i-*) or from an unaugmented prefix \**i* which was later replaced by the plain concord (\**di*). Although the current Tswana class marker *li-* is never contracted (cf. [lǃnó / mǃnǃ] 5/6 ‘tooth’), it may be deleted in a small number of nouns: (*le*)*lapa* ‘yard’, (*le*)*gae* ‘home’, (*le*)*tsatsi* ‘day, sun’, (*le*)*fatshe* ‘earth, country’, (*le*)*nong* ‘vulture’. Krüger (2006: 60) notes that there are “no specific morphological or phonological reasons for this elision”. Likewise in Southern Sotho, Guma (1971:51) notes a limited number of class 5 nouns that can be used without a prefix or suffix as “adverbs of place”: *hae* ‘home’, *ifo* ‘hearth’, and *khotla* ‘court’. Perhaps similarly in Zulu (S42) some of the few non-negative instances when an augment is not used are class 6 adverbs of time and class 17 nouns of place and time (von Staden 1973: 173). A possible scenario for the Sotho-Tswana languages is that the original class 5 prefix \**i-* was generally lost in the transformative process of consonant mutation, just as happened in closely related Venda and the neighboring Shonic and Tsongic languages. The resulting gap in Proto-Sotho-Tswana (with only classes 5 and 1a lacking a prefix) was then healed by the adoption of the concord *li-* as the standard nominal class marker, with only a few relics surviving from an earlier period.

Although the PI-reflexes are very similar throughout the group, the data for the major subgroups is summarized separately: Tswana (S31), Kgalagadi (S311), Northern Sotho (S32), Southern Sotho (S33), each of which has several varieties. Reconstruction is further complicated by the fact that certain varieties show Nguni, Tsonga and Venda influence.

### S.31 – Tswana

The alternations of stems in classes 5/6 and reflexive verb forms of Tswana have long been noted (Cole 1955; Malepe 1966) and the data was particularly studied by Creissels (1999). For Tswana consonant changes in general, see Creissels (2007) and Dickens (1984).

CLASS 5: The class 5/6 forms have been subject to significant levelling, as can be seen in the lists below and especially in Matumo (1993) and Krüger (2006). Cole (1955: 83) notes “this process of phonological change is obsolescent, and many people now use the same form of the stem for both singular and plural. There is thus a tendency towards standardization of the stem, but in selecting the standard, the original singular form of the stem is preferred in some dialects, while the original plural form is preferred in others.” For example, for *lesama/marama* 5/6 ‘cheek’, Cole also notes the existence of *letshama*, *leşama*, *letšhama*, *lerama*; *masama*, etc. with some alternatives due to secondary sound changes. To demonstrate the PI-effect sound correspondences, I have selected contrasting singular and plural forms, which in Cole’s time were usually also the most common and apparently continue to be according to Krüger (2006: 64) who note that both shifted and unshifted forms can take the plural, but the tendency is that “the palatalised (or alveolarised) form is preferred in the singular while on the other hand the depalatalized (or dealveolarised) form is preferred in the plural.” The fact that the contrasts are not productive is helpful for reconstruction: any nouns showing them must have been in the language and in classes 5/6 at an early age. Three deverbatives (with final a) suggest that was an old pattern: *letshoroma* ‘fever’ < *roroma* ‘shiver’; *letsapa* ‘clay’ < *bopa* ‘mould’; *letsapa* ‘tiredness’ < *lapa* ‘be tired’.

REFLEXIVE: The few examples of Tswana verbs with PEB initial \*-i- (not of reflexive origin) show reflexes similar to those of class 5, so we should expect that same pattern for reflexive verbs. But it is easy to see that the changes caused synchronically by the reflexive prefix *i-* to the initial consonant of a verb stem are the same as those caused by a preceding nasal inside a syllable, e.g. *bôna* 'see' > *i-pôna* 'see oneself', *loma* 'bite' > *i-toma* 'bite oneself', *rata* 'love' > *i-thata* 'be conceited'. Thus Cole (1955: 41, 232-4) posited a reflexive prefix *i(N)-* and that combination is confirmed by Kgalagadi where the prefix is *i-* before consonant-initial verb stems but *in-* before vowel-initial stems, e.g. *ab/inab* 'divide/divide oneself', *arġaba/inarġaba* 'answer/answer oneself' (Dickens 1986: 171; 1984: 107; cf. the Tswana cognates *araba/karaba*). When and how exactly that nasal treatment arose is unclear, but at some stage of S30 there must have been a process (presumably analogical) which led speakers to insert a nasal or copy its treatment in the prefixal element to create the current Tswana reflexive pattern.

*agêla* 'build for' > *ikagêla* 'build for oneself'

*ana* 'venerate, honor' > *ikana* 'take an oath'

*êtla* & *kêtla* > *ikêtla* 'go slowly, carefully'

*utlwa* 'hear, feel' > *ikutlwa* 'feel well, fit'

This apparent intrusion of nasals into reflexives is paralleled in occasional noun forms, e.g. \*bùè 5/6 'stone' > *li-ntšwè* / *mà-džè*, where the nasal is only in the class 5 singular which shows the influence of both nasal and PI-effect (Creissels 1999: 332).

#### ORTHOGRAPHY AND SOURCES:

Forms are given in the standard Tswana orthography (with š). Phonetic citations are from Creissels (n.d.). Where sources disagree in an interesting way, I have noted them with these abbreviations:

Cole = Cole (1955)

Cr = Creissels (n.d.) — unpublished Setswana-French dictionary, which is the source for all the phonetic transcriptions

H = Hurutshe dialect as recorded by Malepe (1966).

Kr = Krüger (2006)

M = Matumo (1993), the major source for variants

#### REFLEXES:

\*p > h, f (dialectally)

\*I-p > tsh(w); s, š, f before u

tshwara 'seize' [tsh<sup>hw</sup>ará] < \*í-pat (cf. BLR 2414)

letshwafo/ma- 5/6 'lung' [M: with variants lekghwafo/ma-, lekhwafo; liq<sup>hw</sup>àfó]

tshola 'serve (food)' [tsh<sup>h</sup>òlà] < \*ip-ud (BLR 3500)

ntsha 'take out' [Cr ntshà] with irregular nasal < \*ip 'clear away' (BLR 3492)

—

lešulô/mahulô 5/6 'foam, froth' [M: lesulô & lefulô 'foam'; losulô 11/6 'froth' but no forms with *h* which others give; Kr: lesulô & le-fulô & lehulô/mahulô & mafulô; Cr lõfúlo]

lešudu/mahuđu 5/6 'a place for threshing grain' [M: & lefudu; Cole: stem -hudu;

cf. lohudu 'a hole or niche in a ground squirrel's hole' [M]

\*t > r

\*I-t > s; tsh before back vowels (cf. Bennett 2019)

lesaka/maraka 5/6 'cattle kraal', pl. 'cattle post' [Krüger: \*leraka does not occur; Cr: lisàká / màràká] < \*tàngá 5 'cattle-post' (BLR 2795)

lesama/marama 5/6 'cheek' [Cole: letshama, lešama, letšhama, lerama; masama etc.; Cr: lisámá; H: lèráma/lèsámá & màsámá/màráma] < \*támá 5 'cheek'

lesagô/maragô 5/6 'buttock' [M: & lerago, masago; Cole: & letshago, letšago, lešago, leragô; Cr: lisáχó & lisáχó] < \*tákò (BLR 2741)

lesapô/marapô 5/6 'bone' [M: & lerapô; Cr: lisápô/màr] < \*támbo (BLR 8426)

ǃàjà & ǃàà [H] 'give a name' < \*ít-ag ?

—

leshoo & letšhoo/maroo 5/6 ‘paw’ [M: & leroo; H: lèróò/mà- & lètshòò/ma-; Cr: lìshóú & lìróú]  
— etymology?

letšhoroma ‘fever’ < roroma ‘tremble, shiver’ [Cole; Cr lìtshòróma]

lešôlê/marôlê 5/6 ‘large calf’ [Cole; M: lesôlê/masôlê & marôlê; Cr lìsôlê]

lešôphi/marôphi 5/6 ‘blister’ [M: lesophi/ma- & lerôphi/ma-; Cole and Cr note lerôphi as more common; Cr lùrôphí but Malepe tones LHL]

lešope/marope 5/6 ‘ruin, old dwelling place’ [Cole: & lerope; H: lèsóp’è/ma- & lèróp’è/ma-]

— cf. Venda *shúbí/marúbí* ‘ruin, deserted village site’, Shonic (KKoMZ) *dùmbà* 5 (*mat-*) ‘temporary shelter’,

tshoga ‘be afraid’ [tshòχà] < \*í-tuka (BLR 3511)

tshela ‘pour, flow’ [tshèlà] < \*ít-id (BLR 3504) — nasal influence?

\*k > g [χ]

\*I-k > s, š

mosadi/basadi 1/2 ‘woman’, [mòsádí] < \*kádí 1. The PEB stem is seen in *bogadi* [bòχádí] ‘cattle given as a dowry’, so Creissels (1999: 333) proposes \*i-kadi here, but folk etymology connects to *sálá* ‘remain’.

leiso/maiso 5/6 ‘hearth’ [lùsò] < \*íkò (BLR 3444)

loso/dintsho 11/10 ‘ladle’ & leso 5 [lùsò]; lesô/mašo ‘spoon’ [M] < \*íkò (BLR 3443)

—

lešomê/magome 5/6 ‘ten’ [lìsómé; Cole: mašome more common than magome]

lešolobitla/magolobitla ‘type of lizard’ [Cole]

—

Cf. kgora ‘become satisfied’ [qʰórá] — this is regular nasal effect of reflexives, so this was still perceived as reflexive, cf. Kgal. tjhóra ‘be satisfied’.

\*b > b

\*I-b > ts, tš before back vowels

letsêlê/mabêlê 5/6 & lebêlê ‘breast’ [-bélê]

letsêlê/mabêlê 5 & lebêlê ‘grain of sorghum’ [-bèlê]

letsôpa ‘clay for pots’ [Cole tš; M, Cr: ts] < \*bómbà (BLR 335)

letswele/mabole 5/6 ‘fist’ [M; Cole tšw; Cr lìtswíli & libóli]

lentšwê/majê 5/6 ‘stone’ [Cole]. Creissels (1999: 333) notes that the plural stem is the regular reflex of \*bùè “whereas the singular stem implies the influence both of a nasal and of a vowel sequence beginning with \*i”

letšôgô/mabôgô 5/6 & lebôgô ‘arm’ [M & Cole tš; Cr: lìtsoχó & libóχó] < \*bókò (BLR 260)

\*d > l

\*I-d > ts

letsema 5 ‘work party in the fields’ [lìtsìmà] < \*dìm ‘cultivate’

letsapa 5 ‘tiredness’ [lìtsàpà] < làpà ‘be tired’ < \*dàmb ‘be hungry’

Z-reflexes: tl, ts (before front vowels) — There are almost no Z-reflexes in class 5 nouns. More common is hiatus (often with unwritten glide), e.g. *lè-ítlho* ‘eye’, *lèíná* ‘name’.

tla ‘come’

tlala ‘be full’ [tlálá]

utlwa ‘hear’ [útlwá] < \*ígù

mutlwa/mitlwa 3/4 ‘thorn’ [mútlwá] < \*ígùà

mo-tlogolo [mòtlóχóló] ‘nephew, niece’ < \*íyòkòdò ‘grandchild’ (BLR 3435)

—

ítse ‘know’ [ítsí] < \*íyib (BLR 6209)

metsi [mètsí] ‘water’

motse ‘village’

letsatsi/malatsi 5/6 ‘sun, day’ [Cole, M; lìtsàtsí]

letseba/matseba 5/6 'rock pigeon' [M] —if cognate with leeba 5/6 'pigeon' < \*ibà (BLR 1591), this would be a rare Z-reflex for a vowel-initial noun in Tswana.

### S.311—Kgalagari

CLASS 5: The traces are fewer than in Tswana, presumably through leveling in favor of class 6 forms, so there is no remaining sign of PI-effects in class 5 words like Kgalagari *lerama* 'cheek', *lerapo* 'bone' or *lebele* 'breast'. But in a few cases the singular/plural alternations have been levelled in favor of class 5 as Dickens (1986: 206) notes, "the Cl 6 plural of *lešoru* 'thief' is not \**maχoru* (without sibilantization) but *mašoru*".

REFLEXIVE: As in Tswana, the reflexive marker is *i-* before consonant-initial stems with the same influence on the following consonant as a nasal, but Kgalagari has *in-* before vowel-initial stems, e.g. *ab/inab* 'divide /divide oneself', *ařāba/inarřāba* 'answer/answer oneself' (Dickens 1986: 171; 1984: 107; cf. the Tswana cognates *araba/karaba*

SOURCES and ORTHOGRAPHY: unless otherwise noted, citations are from Lukusa & Monaka (2008). Also important is Dickens (1986). L&M has a similar orthography to Tswana but note sh (Dickens sj) = [ʃ], tj = [tʃ], tsh = [tʃʰ], zh (Dickens zj) = [ʒ].

#### REFLEXES:

\*p > h — pa- cl. 16 prefix < \*pa

\*I-p > tjh

le-tjhwaho/ma- 5/6 'lung'

le-tjhwagola/ma- 5/6 'young castrated goat'

tjhwára 'hold'

\*t > r [r]

\*I-t > sh (perhaps tsh before back vowels)

le-shelwa/ma- 5/6 'watermelon'

cf. bétjho 'ours' (2pl) [Dickens] (< \*bá-ito) [Shetjhauba -etʃu]

\*k > g [χ]

\*I-k > sh (but verbs apparently tsh)

le-shada/ma- 5/6 'shoulder' [Dickens lesjádá 5 'shoulder-blade']

le-shome/ma- 5/6 'ten'

le-isho/ma- 5/6 'hearth' [van der Merwe & Schapera (1943)], preserved in locatives *mo-isho* 'on the fire', *ha-isho* 'by the fire' (L&M)

lo-sho/rintjho [lôʃo] 11/10 'spoon' [Dickens: leʃo 5]

lesjorzu/mašorzu 'thief' with levelling (see Dickens 1986: 204-5), cf. Tswana *legodu* [lìχòdù] and PI-reflex in Venda *fòbvù/mahòbvù* 'habitual thief'.

—

tjhóra 'be satisfied' [Dickens, not in L&M]

tjhóxa 'be frightened' > tjhosja 'frighten' [Dickens, not in L&M]

\*I-b > zhw [žw]

lezhwe/ma- 5/6 'stone'

iba 'steal' — no effect

Z-reflexes: d, z (before front vowels and glides?) — cf. Venda d

da 'come' [Dickens]

dála 'become full' [Dickens]

—

zé 'know' [Dickens; Shetjhauba ízí]

maze 6 'water'



uzwa 'hear' < \*ígù  
 midzwa 4 'thorns' < \*ígùà  
 —  
 leiro/maro 5/6 'eye'  
 leino/meno 5/6 'tooth'

### S.32—Northern Sotho

Northern Sotho, i.e. Sesotho sa Leboa, is a group of related varieties including Sepedi, Khelobedu, Setlokwa, and others. Many Sepedi examples have been listed since Eiselen (1923/24) and Meinhof and van Warmelo (1932). The PI-effects are basically the same as for Tswana but seem to indicate a mixture of dialects.

CLASS 5: The original variation between class 5 (with PI-effects) and unaffected class 6 forms has undergone significant levelling in both directions. Mokgokong (1966: 69) reported, "In Pedi as spoken today the original and palatalized plural forms are generally used interchangeably, e.g.

[matsw'elɛ] and [maβelɛ] (breasts)  
 [maʃoβa] and [maroβa] (holes)  
 [maʃáyo] and [maráyo] (buttocks)

This tendency suggests that the modern plural is derived by analogy of the palatalized singular form." But more recent dictionaries I have consulted regularly show the generalization of the class 6 stems into the singular, e.g. *lerago* 'buttocks', *lerapo* 'bone', *lerope* 'ruins'. It would be useful to have a more detailed study of the dialectal varieties of Northern Sotho, which is sometimes a catch-all for whatever is not S31 or S33. To demonstrate the PI-effect sound correspondences, I have selected contrasting singular and plural forms, but in most cases there are analogical alternatives.

REFLEXIVE: As in other Sotho-Tswana languages (discussed above), the current output of the reflexive *i-* is identical to nasal strengthening rather than to PI-effects, e.g. *i-pôna* 'see yourself' < \*bóna 'see'. Likewise for the 1sg object, e.g. *m-pona* 'see (for) me'.

SOURCES and ORTHOGRAPHY: The major sources used for Northern Sotho are Mokgokong (1966), Poulos and Louwrens (1994), Kriel (1988). The standard orthography uses *g* = [x], *kg* = [kx<sup>h</sup>], *š* [ʃ]. Unfortunately, some sources do not mark vowel distinctions so I have removed all vowel markings to avoid inconsistency.

REFLEXES:

\*p > φ, f (dialectally)

\*I-p > tshw/sw

tshwara & swara 'hold'  
 leswafo 5 'lung'

\*t > r

\*I-t > š [ʃ]

lešapo/marapo 5/6 'bone' (also lerapo)  
 lešope/marope 5/6 'ruins' (also lerope)  
 lešago/marago 5/6 'buttock' (also lerago)  
 lešoba/maroba 5/6 'hole' (also maʃoba)  
 lešete/marete 5/6 'testicle' (also lerete)  
 lešoo/maroo 5/6 'claw, paw' (also leroo)  
 lešama/marama 5/6 'cheek' (also lerama or lethama)  
 -ešo 1pl poss. 'our' < \*aito

\*k > g [= χ, γ, h dialectally]

\*I-k > s, tš

lesome/ masome 5/6 'ten'

- letšogo '5 shoulder (of meat)  
cf. khora 'have plenty, be satisfied' < í-kot (BLR 3445)
- \*b > b
- \*I-b > ts
- letsopa 5/6 'clay'  
letsogo 5/6 'arm'  
letswele/mabele 5/6 'breast' (also matswele)
- \*d > l
- \*I-d > tš
- letšema 5 'working party'  
letšoba/maloba 5/6 'flower' (and leloba & matšoba)

### S.33—Southern Sotho

In addition to the varieties of Southern Sotho in South Africa and Lesotho, there is the closely related language Lozi (K21) which migrated to southwestern Zambia in the 19<sup>th</sup> century. Lozi will not be discussed here, but Gowlett (1967: 107-112) gives a well-sorted catalog of about 80 reflexive verbs in Lozi (some showing initial-consonant changes and some not) so there is no need to repeat them here. Lozi is a 5-vowel language and the reflexive prefix /i/ is overt. He notes (108), "There is a great deal of irregularity about the occurrence of these changes, however, and the process appears to be obsolescent. In some instances their occurrence is morphologically determined, while in other cases they occur as free, idiolectal or dialectal alternants. They rarely, if ever, occur with adoptives."

CLASS 5: Southern Sotho has strongly levelled the PI-effect variations between classes 5 and 6, but the levelling has not been unidirectional. So, the surviving stem can reflect the PI-effect expected for class 5 (e.g. *letsoho* 'arm, hand') or the unaffected stem originally in class 6 (e.g. *lerama* 'cheek'). One factor that helped preserve PI-reflexes was a number of class 5 nouns *singularia tantum* which had no competition from class 6: e.g. *letsopa* 'clay', *letšolo* 'hunting party', *letsoku* 'ochre' (Guma 1971: 51).

SOURCES and ORTHOGRAPHY: I have used the orthography common in South Africa, for which the Lesotho equivalents are: tjh [SA] = ch [Les], tsh [SA] = tš [Les], y [SA] = e [Les], w [SA] = o [Les]. The major sources for citations are the online dictionary of Olivier (2009) and Guma (1971).

#### REFLEXES:

- \*p > f
- \*I-p > tsh
- letshwafo /matshwafo 5/6 'lung'  
tshwara 'seize, hold'  
tshola 'dish up food' < \*ìp-ud (BLR 3500)
- \*t > r
- \*I-t > s
- lesaka/masaka 5/6 'enclosure, kraal' < \*tàngá 5 'cattle-post'  
lesupi/masupi 5/6 'ruin'  
lesapo/masapo 5/6 'bone'  
-eso 1&2pl collective pronouns
- k > h
- \*I-k > sh [ʃ]
- leshala/mashala 5/6 'coal' < \*kádà (BLR 1662)  
leshodu/mashodu 5/6 'thief'  
leshome/mashome 5/6 'ten'
- 
- leifo/maifo 5/6 'hearth' (& -ifo with locatives) < \*íkò 5/6. Cf. Lozi liso/maiso 'hearth'.  
kgora 'become satisfied' < \*í-kot (BLR 3445)

\*b > b

\*I-b > ts

letsoho/matsoho 5/6 'arm, forearm'

letsopa 5/6 'clay'

letswele/matswele 5/6 'breast, udder'

Z-reflexes and other:

tla 'come'

tlatsa 'fill'

metsi 6 'water'

motsana/metsana 3/4 'little village'

—

leihlo/mahlo 5/6 'eye'

leino/meno 5/6 'tooth'

#### S.40 Nguni

In general, the Nguni languages show few PI-effects, perhaps because they have maintained or generalized a vocalic version of \*i, both in the class 5 nominal prefix \*i and in vowel-initial verbs (see §5). The general Nguni pattern of aspirating unvoiced stops may also cover any aspiration due to a mutation of \*I-p and \*I-t. Despite the rarity of the surviving examples, there seems no simple way to explain them except through PI-effects, which were perhaps once more prevalent.

SOURCES AND ORTHOGRAPHY: The languages (abbreviations in parentheses) used here have standard orthographies:

S41 Xhosa (Xh): Tshabe (2006)

S42 Zulu (Zu): Doke et al. (2014) and Mbatha (2006)

S43 Swati (Sw): Rycroft (1981)

There are a number of other Nguni varieties (Phuthi, Bhaca, Lala, Northern Ndebele, Southern Ndebele, Zimbabwean Ndebele, Nhlanguwini) but they do not seem to provide alternative reflexes.

REFLEXES:

\*p > ph — *phá* 'give' (XhZu) < \*pá (BLR 2344)

\*I-p > ph (like Tsonga) which merged with \*p > ph?

No distinctive evidence in seen in likely candidates, e.g. *íphikò* 5 (Zu), *líphiko* (Sw) 'wing' < \*pikò 5/6; *phèka* 'cook' (XhZu), *íphàphú* 5 'lung' (Zu).

\*t > th, tsh, tf — Zunda: *thathu* 'three' (Zu); Tekela: *tshátfù* (Sw) < \*tátò (BLR 2811)

\*I-t > s? th?

sàbèlà 'answer' (Zu) < \*ítabid (BLR 6031)

No distinctive evidence is seen in other likely candidates, e.g. *thà* 'give a name to', *thà* 'pour', -*ethu* 'our'.

\*k > kh — *kháma* 'press, squeeze' (XhZu) < \*kám 'squeeze' (BLR 1689)

\*I-k > hl [h]; before *u* > sh, s

It appears that \*I-k > \*s (and \*I-ku > ʃ?) (perhaps common to Nguni and Sotho-Tswana?) and then Proto-Nguni \*s > hl [h]. As a result, Nguni *hl* is the reflex of both PEB \*s and \*I-k.

(i)hlà (Xh), èhla (Zu) 'go down' < \*ik (BLR 3436)

hlàla (XhZu), hlàlà (Sw) 'sit' < \*ikad (BLR 3441)

íhlômbè 5/6 (Zu), líhlômbè (Sw) 'shoulder' < \*kómbè 'shoulder blade' (BLR 1922)

—

í-shùmì/ámáshùmì 5/6 (Zu), líshùmì (Sw) 'ten' < \*i-kómì (BLR 2027)

sútha & esútha (Zu), ésútsa (Sw) ‘be satiated’ < \*í-kot (BLR 3445). Xhosa *hlútha* seems to have permitted the further shift \*s > hl. The tone pattern shows Xhosa preserved the initial vowel for some time (Rycroft 1980).

\*b > b [6] — ḅila ‘boil’ (XhZuSw) < \*ḅid (BLR 176)

\*I-b > b — apparently same as \*b, but often the preceding vowel is preserved in verbs.

èbà ‘steal’ (ZuSw) < \*íb ‘steal’ (BLR 3387)

íbélè/ámàbélè ‘breast; udder’ (Zu), líbélè (Sw) < \*béèdè 5/6 (BLR 125)

\*d > l — ḷima ‘cultivate, hoe’ (XhZuSw) < \*ḍim (BLR 968)

\*I-d > d — ídòngà 5/6 ‘clay (hlonipha term)’ Zu < \*dòngó (BLR 1138)

\*s > hl [l] — hlakùla ‘weed with hoe’ (XhZuSw) < \*sàk-ud (BLR 429)

\*I-s > s — ili-so/amehlo (Xh) ‘eye’, i-so/amehlo (Zu), ḷisó/éṃhlo (Sw) < \*ísò (BLR 3405)

\*g > ø — ís-ándḷà 7 ‘forearm & ám-ándḷà 6 ‘strength’ (Zu) < \*gànzà ‘palm’ (BLR 1329)

Z-reflexes: z (Zunda varieties), t (Tekela varieties)

íẓínyò 5/6 (Zu), líf̣ínyò (Sw) ‘tooth’ < \*ínò (BLR 3472)

íẓùlù 5 (Zu), líṭùlù (Sw) ‘sky, weather’ < \*gòdò 5 ‘sky, top’ (BLR 1486)

bàza (XhZu), bàta (Sw) ‘work wood’

## S.50 Tsongic Group

This group of closely related languages has three main subgroups: S.51 Tswa (Xitshwa, Citshwa), S.53 Tsonga (Xitsonga) or Changana (Xichangana), and S.54 Ronga (Xironga). The language known in English as Tsonga or Changana is called Xitsonga in South Africa (varieties in Baumbach (1974), Xichangana in Mozambique, and Shangani in the constitution of Zimbabwe. Ronga is spoken just south of Maputo in Mozambique but also extends a little into South Africa. There are frequent variations and doublets in the sources, due both to dialectal divergence and the impact of Nguni and Shonic languages which border Tsonga. Furthermore, Pastor Andre Chitlango (p.c.) says there is a problem with the influential “Shangaan” bible translation done in the late 1800s by the Swiss missionaries in Valdezia near the border with the Venda because the translators were Sotho.

**SOURCES AND ORTHOGRAPHY:** Unless otherwise marked, the data here comes from Tsonga in detail, with limited notes on the other two subgroups. The reader should understand that the actual phonetic values can vary widely in the varieties and subvarieties.

*Tsonga:* The Mozambican orthography is used, following Ngunga & Faquir (2012).

BS = Siteo (2011), based on Mozambican varieties; provides tones.

EB = Baumbach (1987), based on South African varieties.

AC = Andre Chitlango (p.c.), pastor based in Maputo, originally from the south side of the Save River (on the border of Tswa and Shangaan).

Eis = Eiselen (1923/24)

*Ronga* (labelled “R”): Citations from Siteo et al. (2008), a Ronga-Portuguese dictionary with tones, following the orthography of Siteo and Ngunga (2000). There is now a handy trilingual dictionary (Ronga-Portuguese-English) online, but using another orthography and without tones: <https://dr.xizronga.org>. The most recent grammar is Bachetti (2006). Note that the retroflex alveolar fricative (from \*t) has had various spellings over the years including *rhi* (used here), *dji*, *ḍyi*, *ri*, and *zr* (online dictionary).

CLASS 5: In Tsonga, the usual class 5 nominal prefix is  $\emptyset$ , but sometimes seen with PI-effects on the initial consonant due to the original class prefix  $*i$ . General levelling of singular and plural forms has taken place so that few nouns now show class 5/6 differences in the stems, and in those surviving cases analogy is continuing to work. For example, we see the expected aspiration in the singular of *khòndlò / màkòndlò* 5/6 ‘rat’ but *k* is also seen in the singular (BS) and *kh* in the plural (EB), and AC reports singular-plural uniformity of consonants in all three subgroups of S50 but with different tones. BS and AC have *khàlà / màkhàlà* ‘coal’ but Eiselen (1923/24) had the plural with *k*. EB gives the historically expected difference in *khàtlà / màkàtlà* ‘shoulder’, but BS has *k* singular and AC reports both *k* and *kh* in the plural while noting that the idiom ‘to have a child on the shoulders’ is always with just *k* in the plural. There are very many nouns with *kh*- stems, so probably this pattern of aspiration was once productive in some varieties.

Modern Tsonga lacks an augment, but traces of it are probably seen in monosyllabic stems where the prefix group developed  $*di-i > ri-$  and became part of the stem, e.g. *rìfù / màrìfù* ‘cloud’  $< *tù$  ‘cloud’ (BLR 3093), *ribye / maribye* ‘stone’, *riwa / mariwa* ‘precipice’, *rito / marito* ‘voice’. But it is also possible that *i-* was initially preserved in monosyllabics and, since it appeared irregular, was then reformed to *ri-* on the basis of the concord. The occasional words 5/6 with independent *ri-* (e.g. *risindza / masindza* ‘bracelet’, *rihlampfu / mahlampfu* ‘fence’) are perhaps adaptations of cl. 11 or *li-* from other languages.

In Ronga, the class 5 prefix is usually  $\emptyset$  with rare stem differences from class 6 (perhaps only *khúme / màkúme* ‘ten’). Class 5 and 11 have been merging so the class 11 prefix *li-* has now spread into class 5 and appears in nouns like *lihleko / mahleko* 5/6 ‘laugh’ vs. the older pattern seen in Tsonga *hlèkò / màhlèkò*. Many contemporary speakers now consider *li-* as the class 5 marker (Dimande-Chimbutane 2022). The class 5 concord is *dri-* (associative *dra-*).

REFLEXIVE: The reflexive marker in all these languages is now *ti-*, segmentally equivalent to the class 10 object marker, with no ongoing PI-effects.

REFLEXES: The unvoiced stops  $*p$  and  $*k$  became aspirated (*ph*, *kh*) but the effect on the voiced stops is less clear (see below), and there are a variety of Z reflexes. In addition, there are several instances of voicing in Tsonga which must be due to Shonic influence or a dialect with that trait. The voicing is seen in augmentative pairs (Baumbach 1987: 41), e.g. *gòxá* 5 ‘extremely old person’, cf. *xikòxà* 7 ‘old woman’; *gúgùlúmè* 5 ‘extremely big rooster’, cf. *nkúkú* 3 ‘rooster’, *gúngúvá* 5 ‘pied crow’, cf. *xikúngúvá* 7; *gónádzó* ‘important way, street’, cf. *nkónádzó* 3 ‘footstep’. But sometimes there is no difference in meaning, e.g. *bòfù* 5 & *xiphofu* 7 ‘blind person’ or *gòmbè* 5 ‘shoulder blade’ ( $*kómbè$  5/6 BLR D1922). As in Shonic, this voicing also affected aspirates, so there are numerous cl. 5 nouns beginning with *dz-*, some of which have unvoiced equivalents, e.g. *dzándzá / màdzándzá* 5/6 & *ntsándzá* 3/4 ‘log, beam’, *dzúví* 5 & *ntsúví* 3 ‘water in which stamped mealies have been soaked’, *dzòvá* 5 & *ntsòvá* 3 ‘firewood not yet dry’; *dzándzàwà* 5 & *cánjàwà* 5 ‘arid land’. As in Venda, there are even Tsonga class 5 doublets with both the expected aspirated form and voiced forms reported, e.g. *khùdzù* & *khùjù* & *gùtsù* & *rigùtsù* ‘big toe, thumb’; *khútlà* & *gutla* ‘bull-frog’; *gón’wà* & *khón’wà* ‘lily’.

$*p > h$  -ha ‘give’  $< *pá$  (BLR 2344)  
 hólá ‘cool down’  $< *pód$  (BLR2589)

$*I-p > ph$   
 phàphú / màphàphú 5/6 ‘lung’, but also *hàhu*, and an older pattern *phapu / maphapu* (Eis)]  $< *pápú$  (BLR 9444)  
 phàngá / màphàngá 5/6 ‘a hollow; empty space’  $< *pángá$  ‘cave’ (BLR 2400)

—but also *sv* [retroflex fricative]?

*svéká* ‘cook’ (also R), which apparently derives from  $*ip-ik$  ‘cook, boil’ (BLR 3496), but how?  
 Tsonga cl. 8 prefix *svi-*  $< *bi-$ , so is *svéká* a development from something like Shonic *bika*? Cf. Tsonga *svárhá* ‘be burnt’ ( $< *pí-ad$  ?) from  $*pí-$  (BLR 2491).

\*t > rh — e.g. *rhárhù* ‘three’ < \*tátò (BLR 2811)

\*I-t > tsh, ch

tshàngá/màtshàngá 5/6 ‘cattle kraal’ < \*tàngá ‘cattle-pen; cattle-post’ (BLR 2795)

tshálá/màtshálá 5/6 ‘platform on which granary stands’ < \*tádà (BLR 2720)

tshùrí/màtshùrí 5/6 ‘mortar (for pounding)’ < \*tódí (BLR 3042)

chá & thá ‘tell, play’ < \*ít ‘call’ (BLR 3379, 2708, 2710)

chèlà ‘pour’ < \*ít-id (BLR 3504)

— but also:

dzùkà ‘be startled’ < \*í-toka (BLR 3511)

— No effects seen in several words where they might be expected, for example:

-erhu 1pl. possessive ‘our’ < \*áitú

rhámá/màrhámá ‘cheek’ < \*támà (BLR 2744)

\*k > k -khámá ‘squeeze’ < \*kám (BLR 1689)

\*I-k > kh

khúmè/màkúmé 5/6 ‘ten’ (also Eis, R) < \*i-kómì ‘ten’ (BLR 2027)

khálá/màkhálá 5/6 ‘coal’ (also R; Eis: makala) < \*kádà (BLR 1662)

khàlávàtlà/màkhàlávàtlà 5/6 ‘watermelon’ (but Eis had *kh-/k-*)

khàlè adv ‘in the past’ (also R) < \*dài 12 ‘long ago’ (BLR 9158), cf. Zulu *ikade* 5 ‘olden time’

khélè/màkhélè 5/6 ‘frog’ (also chélé) < \*kédé (BLR 1759)

khèlè/màkhèlè 5/6 ‘pit, ditch’ (also chélé)

khálá ‘protest’, perhaps < \*íkad ‘stay, dwell’ (with a semantic shift as the sense ‘sit, stay’ was taken over by *tshámá*)

— but also:

xúrhá ‘be satisfied’ (also R) < \*í-kut (BLR 3445) — Perhaps the rule is \*I-ku > xu, but then why not also for *khúmè/màkúmé* ‘ten’? Is the word for ‘ten’ remodeled from other languages through trade? Or do verbs have a different pattern?

\*b > v [v] — e.g. *vilà* ‘boil’ < \*bíð (BLR 176)

\*I-b > b — but there are few solid examples

bòwù/màbòwù 5 ‘baobab fruit’ (EB: *bowu* & *vowu*) < \*bùyó 3/4 ‘baobab tree’ (BLR 354)

bàwà 5 ‘gadfly’; cf. \*bùgò (BLR 4961), Karanga *vùvò*, Ikalanga *bùwè* ‘gadfly’

— No such effect is seen in several words where it might be expected, for example:

b’èlèkà [b — probably from Nguni] & *vèlèkà* ‘give birth; carry a child’

yívá ‘to steal’ < \*íb ‘steal’ (BLR 3387)

vélé/mavélé 5/6 ‘breast’ < \*béèdè 5/6 (BLR 125)

\*d > l, r (\*dɪ > Tsonga *r*, Ronga (*d*)*z*) — e.g. *rimà* ‘cultivate’ < \*dìm (BLR 968)

\*I-d > ts

tsìmà 5/6 ‘work party’ < \*dìm- ‘cultivate’

tsòlò/matsòlò 5/6 ‘knee’ (cf. Zulu *ídòlò*)

\*g > ø — màndlà 6 ‘hands’ (BS, EB) < \*ma-andla, cf. *xìyàndlà* 7 ‘hand, palm’ < \*gànzà ‘hand, palm of hand’ (BLR 1329)

Z-reflexes: t (also dz, which may be borrowed)

tíhló/máhló & mà-tíhló ‘eye’ 5/6 < \*ísò (BLR 3405)

tíkó/màtíkó 5/6 ‘country’

tínyó/màtínyó & mènýó & màtínyó 5/6 (BS; EB: tino/matino & meno)

má-tí 6 ‘water’ < \*iyi (BLR 3433)

mù-tì/mì-tì 3/4 ‘home, village’ (R: mùntì & mùtì/mì-) < \*gìì ‘village’ 3/4 (BLR 6466)

rì-tó/mà-tó 5/6 ‘voice’ (R: rhító & rhítú/mà-)

tàndrá/mà-tàndrá & màndrá 5/6 ‘egg’ (R)

tòlò ‘yesterday’ (also R) < \*gòdò (BLR 1421)

tà ‘come’ (also R)

—  
dzàrà/màdzàrà 5/6 'hundred' < \*gàrà 5 (BLR 1314)

dzànù/mànù 5/6 'sense, common sense' (R: dzàrà/màdzàrà) < \*gàrà 6 'wisdom' (BLR 1317)

### S.61—Copi

Copi has considerable regional and idiolectal variation (Bailey 1995), the most noted of which is the now generally extinct kiLenge (cl. 7 *ki-* = central ciCopi *ci-* = Gitonga *gi-*) which may even be better described as a separate language. But the greatest challenge for reconstruction is the significant relexification from the Tsongic group, which were themselves also impacted by Nguni and Shonic trade routes. As a result, it is not obvious whether the two types of PI-reflexes we see in Copi (aspiration and voicing) are due to older differences in Copi dialects or to contact. But aspiration of unvoiced stops seems to be the pattern of both Lenge and central Copi, with strengthening for voiced consonants (with devoicing in Lenge). The few instances of voicing (e.g. *dìgúmí* 'ten') would then be explained by contact, even though that is surprising for a basic word like *bhikà* 'cook'.

CLASS 5: There are only a few words still with differences in class 5/6 stems (noted under the Z-reflexes), so I have limited the list to those words with secure etymologies, although there are many more likely candidates. Smyth and Matthews (1902) state that in Lenge the class 5 prefix *dì* is often omitted. But historically we might say that the concord prefix was actually in the process of being added then — since the incorporation or effect of the original nominal prefix *i-* into the initial consonant had left the nouns appearing prefixless.

SOURCES and ORTHOGRAPHY: The major published sources for Copi are dos Santos (1941, 1949) [DS], who worked on southern Copi, and Smyth and Matthews (1902) [S&M] on Lenge — both of which require some care to use. In particular, the latter rarely marks aspiration. Fortunately, Richard Bailey [B], who has conducted Copi field work, was kind enough to share some lexical notes with tones, which are the basis for the entries below. The orthography here follows Ngunga and Faquir (2012): bh = [b], b = [β], dh = [d], d = [d̥], g = [g].

#### REFLEXES:

\*p > h -há 'give'

\*I-p > ph, but also bh [b]

làpha 'become long' < \*dàip- (BLR 784)

dì-phèlè 5 'cockroach' < \*pèdè (M2426); DS also records *hele*, which could be generalized from cl. 6 or influence of Tsonga & Ronga *hèlè/mà-* 5/6

dì-phùhó 5 'wind' < \*pòòpò (BLR 2661), cf. *hùhà* 'fly, blow' < \*pòòp (BLR 2660)

dì-phùtú 5 'foam' < \*pùdú (cf. Gitonga *liphùdzú*) < \*púdò 5 (BLR 2677)

dì-phiyélò 5 'broom' < *hiyéla* 'sweep' < \*piág-id (BLR 2499)

phùla 'take off the fire' < \*ip-od (BLR 3500)

—  
dì-bhòfù 5 'blind person' < \*pòkù (BLR 2600) — cf. Common Shonic *bòfù* — likewise Tsonga has both *xiphofu* 7 'blind person' and *bòfù* 5

*bhika* 'cook' < \*ip-ik (BLR 3496) — cf. Common Shonic *bika*

\*t > r

\*I-t > th, but also d [d̥]

dì-thùdí 5 'mortar' < \*tódí (BLR 3042)

dì-thàlá 5 'attic' < \*tádà (BLR 2720)

dì-thùmbá 5 'boil, abscess' < \*tòmb 'swell' (BLR 3057)

dì-thamá 5 'cheek' (also *ràamá*) < \*ámà (BLR 2744)

thèla 'pour' < \*it-id (BLR 3504)

-athu 'our (1pl poss.)' < \*-áitú

dì-thavi 5 & dì-dháví & lì-dháví 11 'branch' < \*tábì (BLR 2713), cf. Karanga *dávì/mat-* 5/6

—  
dhána 'call' < \*ít-an 'call' (BLR 3379, 2708, 2710), cf. Shonic (KZIkTj) *dána*  
dhùka 'get a fright' < \*í-toka (BLR 3511), cf. Zezuru *dùka*

\*k > k

\*I-k > kh, but also g

m'-khò/mì-khò 3/4 'ladle' < \*ikò (BLR 3443)  
c-ìkhó 7 'hearth' < \*ikò (BLR 3444)  
khàla 'stay, be' < \*ikad (BLR 3441)  
khúra 'become full, satisfied' < \*í-kot (BLR 3445)  
dì-khèlè 5 'frog' (& chèlè) < \*kédé (BLR 1759) but cf. Tsonga *khèlè* & *chèlè* 'frog'  
-ekha/-okha & ekhelela/-okhelela 'alone, only' (dos Santos 1941: 117-19)

—  
dì-gúmí (& e)/màkúmè 'ten' < \*i-kómì (BLR 2027), cf. C. Shonic *gúmí/màk-*  
dì-gómbé 5 'large wooden mug', cf. Shonic (KZ) *gómbè/màk-* 5 'big ladle'  
dì-gwàná 5 'wooden spoon', cf. Ikalanga *lù-gwàná*  
cf. also unexpected voicing in *hà-gàrí* 'in the middle', cf. C. Shonic *pàkàtí* 'in the middle' < \*pà  
+ \*kàtí 'time' (BLR 1735)

\*b > v [v], w (near back vowels)

\*I-b > p, but also b

pá 'steal' (& *ba* in some dialects) < \*íb 'steal' (BLR 3387)  
(dì-)pòkò 5 'the whole arm' (B: Lenge — otherwise *dì-bòkò* 5 'arm' & *dì-wòkò* and *dì-bùmbá* 5  
'clay' & *dì-wùmbá*) < \*bókò (BLR 260)

—  
béleka 'carry on the back' (& *vèleka* 'give birth') < ì-bed-ík (BLR 3394)

\*d > l, but \*dì > d (-dì-mà 'cultivate', lùdì-mí 11 'tongue', class 5 prefix di-)

\*I-d > t, but possibly also d

(dì-)tìmà 5 'garden work party' (B: Lenge dialect vs usual *dì-dì-mà*)  
dì-tembe & dilémbè 5 'year' (B: younger *di-tembe*) < \*dèmbè 'epoch' (BLR 7549)  
dì-tàmbò (& -ù) 5 'sun' (cf. Ronga *dàmbù/màdàmbù* 5/6)  
taho 5 'sleeping mat' (S&M: Lenge) < \*dàgò (BLR 816)  
-ite, perfect suffix allomorph < \*ide

—  
cì-démá 7 'darkness' < \*id-imà (BLR 3411) — or is this \*d > d before front vowel?  
dì-démá 'bat' < \*dí-mà (BLR 916) — or is this \*d > d before front vowel?

\*g > y, ø

lòya "be a witch" < \*dòg (BLR 1100)

\*s > s -sèkà 'laugh'

\*I-s > tsh ?

dì-tshìkú/mà-sìkú 5/6 'day'  
dì-tshàn-zú 5 'branch' < \*sán-zú 'branch; leaf' (BLR 461, 491)  
m'-tshì 3 'pestle' < \*isì or \*insì (BLR 6054, 3475)

Z-reflexes > t, but also z due to contact

dì-tàlà 5 'rubbish heap' < \*àdà 5 (BLR 1557)  
dì-tikò 5 'country' < \*ikò 'fireplace, country' (BLR 3444)  
dì-tòrá 5 'thirst' < \*yóòtà 5, 9 'thirst' (BLR 7055)  
dì-tò/mà-dìtò 5/6 'voice, word'  
tínà 5 'name' (Lenge and general southern Copi) & *dìnà/mà-dìnà* (B: northern)  
nya-tulu T 'day before yesterday' < \*gùdò (BLR 1485)  
mì-tì 4 'villages'  
mà-tí 6 'water'



- tà ‘come’  
 tálá ‘become full’  
 tívá ‘know’ (S&M on Lenge) & (d)zívá  
 vátà ‘work wood’ < \*bàiy ‘work wood’ (BLR 8930)
- Vowel-initial stems with class 5 di-  
 dìsò 5 ‘eye’ — no Z-reflex  
 dìnó/mànó 5/6 (& di-dìnó dial.) ‘tooth’  
 dìná/mà-dìná (B: northern); tíná 5 ‘name’ (Lenge and general southern Copi)  
 dànzà/mànzà 5/6 ‘palm, hand’ (DS: dandza/mandza, dialectal di-andza/mandza) & c-ànzà 7 <  
     \*gànzà ‘hand, palm of hand’ (BLR 1329)  
 danda/manda 5/6 ‘egg’ (DS: also dianda/manda; B: northern dialect now dídàndá/màdàndá)
  - Forms with z, perhaps due to contact.  
 dì-zànà 5 ‘hundred’ (DS: didzana/madzana; S&M on Lenge: zana, dzana, tsana). Note  
     irregular *dzànà* in Tsonga as well. For the source, cf. Common Shonic *zànà*.  
 dì-zòngònónò 5 ‘millipede’ (B: dizongololo in Lambwe dialect). As a source for contact, cf.  
     Shonic (KMZ) *zòngóròrò* or Sena *dzongololo* rather than Zulu *ishòngòlòlò*.

## S.62 — Gitonga

Gitonga, spoken around the Inhambane Bay, provides substantial evidence for PI-reflexes especially through the distinction of class 5 and 6 stems. Gitonga forms are numerous and often cited in the main text, so the items below are just a selection.

CLASS 5: Some levelling of class 5 and 6 stems has taken place and continues to, since Lanham (1955) has more examples of difference than later sources. Richard Bailey (p.c.) says that the 20-year-olds now have no class 5/6 stem alternations in their urban Gitonga, although their parents do in their Khumbana dialect. In fact, instead of *lìthàgó/màràgó* 5/6 ‘buttock’ (Lanham) the younger speakers use *lìràgó/màràgó* (based on the plural stem) in their vernacular and for politeness they generalize the singular stem *lìthàgó/màthàgó*. In Gitonga, preceding nasals had the same aspirating effects as PI-reflexes for unvoiced stops, e.g. \*N-p, \*I-p > ph. Accordingly, when class 11 was recently incorporated into class 5, some new aspirated stems entered class 5, but usually these words have no alternations (Lanham 1955: 71).

SOURCES and ORTHOGRAPHY: Unless otherwise stated the entries here are from Amaral (2007) and Lanham (1955), with tones kindly provided to me by Richard Bailey (p.c.). The orthography follows Ngunga and Faquir (2012): *b* [b], *bh* [b], *d* [d], *dh* [d], *g* [ɣ], *gh* [g], *v* [see below], *vh* [v], *vb* [β], and aspiration in *ph*, *th*, *kh*, *pʰh*, *tsh*. A detailed phonology of Gitonga can be found in Lanham (1955).

A = Amaral (2007)

B = Bailey (p.c.)

C = Cabrá (2012)

L = Lanham (1955)

### REFLEXES:

\*p > vb [β]

\*I-p > ph

lìphàdzá/màvbadzá 5/6 ‘hoe’ [B: people older than 40 years of age use mavbadza, but younger speakers use maphadza]

lìphàlà/màvbàlà 5/6 ‘pit, cave’ [B, C; but A: liphala/ma-]

lìphàndzé/màvbàndzé 5/6 ‘prawn, shrimp’ [C; Bailey reports lìphàndzí, livb-, maph-, mavb- but plural preferred]

lìphùdzú 5 ‘foam’ [B], cf. Copi diphútú

phàrà ‘lay hold of, take’ < \*í-pat (BLR 2414 \*pát ‘hold’)

phula ‘cook’ < \*ip-ud ‘take (food or pot) off the fire’ (BLR 3500)

làphà 'be tall, high' < \*dài-p 'be long' (BLR 784)

—

pàya 'prick, kill' < \*ip-ag 'kill' (BLR 3494)

\*t > r

\*I-t > th

lithàmbó/màràmbó 5/6 'bone' [L; but A: lirambo/mar-; B: variants in 3/4 and 7/8] < \*támbo (BLR 8426)

lithàgó/màràgó 5/6 'buttock' < \*tákò (BLR 2741)

lithòngò/màròngò 5/6 'tear' [L; A has lirongo; B has both] < \*tòngò (BLR 3001)

athu '1pl. possessive' < \*-áító

tháná 'respond to a call' < \*ít-(an) 'call' (BLR 3379, 3508)

[B: tháná 'respond when called' vs. ráná 'call (someone)']

thèla 'pour' < \*ít-id (BLR 3504)

thùga 'get lost, get a fright' <

\*k > g [ɣ]

\*I-k > kh

likhùmì/màgùmì 5/6 'ten' [C: likhume & likhumi] < \*i-kómì (BLR 2027)

likhàtá/màgàtá 5/6 'shoulder' [L: ligata/magata]

likhendre/magendre 5/6 'testicle' [A] < \*kèndè (BLR 1775)

likhòndzò/màgòndzò 5/6 'rat' [B] < \*kònjò 5/6 (BLR 6759)

likhùyú/màgùyù 5/6 'fig' [B] < \*kúyó (BLR 2086)

khàlà 'sit, dwell' < \*ikad (BLR 3441)

khúrá 'be satisfied' < \*í-kot (BLR 3445)

-ekha/-okha & -enga/-onga, 'alone, only' (Lanham 1955: 122) [B: younger informants use *eni nyéénga*, but know others use *eni nyóókha*]

Cf. lì-kànà /màgànà 5/6 'mouth' < \*kà-nùà 12 'mouth, lip' (BLR 4709)

\*b > v — According to Bailey (p.c.), this is a labiodental approximant alternating with a labiodental flap /v/ [v ~ ʋ] e.g. [va:tʰu] 'people', rather than "implosiva bilabial vozeada" (Ngunga and Faquir 2012). The reflex of \*b is /w/ before \*u, \*o, e.g. *guwoná* 'see'..

\*I-b > b [β]

libèlè/màvèlè 5/6 'breast' [L: libele/mab-] < \*bèèdè 5/6 (BLR 125)

libògó/màwògó 5/6 'arm'

libùmbá/màwùmbá 5/6 'clay'

bá 'steal'

\*d > l

\*I-d > d [d]

lidàhó/màlàhó 5/6 'sleeping mats' [L: without h; A: lidahu & lidawu/malawu] < \*dàgò (BLR 816)

lidìmbà/màlìmbà 5/6 'door' [A; B: lidìmbà/màdìmbà]

lidìmò/màlìmò 5/6 'work of a digging party' [A: -a]

lidówó/màlówó 5/6 'skin, hide'

gìdémá 'darkness' 7 < \*id-ìmà (BLR 3411)

\*s > h -hèga 'laugh' < \*sèk (BLR 522)

\*I-s > s

lìsò/màhó 'eye' < \*ísò (BLR 3405)

Z-reflexes: t

lì-tìgó/màìgó 5/6 'country' [B: never \*matigo] < \*íkò (BLR 3444)

lì-tàndrá/màndrá 5/6 'egg' [A; B: also màtàndrá, and ddandá, madandá]

lì-tàlà 5/6 'garbage heap'

lì-tó/mà-lító 5/6 'voice, word'

mà-tí 6 'water'

tà 'come'

tálá 'be full'

tùlà 'open'

tí 'know'

— no Z-reflex:

lisó/màhó 'eye'

liná 'name'