



**HAL**  
open science

## How to deal with so-called "vowel verbs with variant non-vowel forms" in a bilingual Zulu lexicon? The case for a pragmatic approach

Michel Lafon, Bolofo Mongezi

### ► To cite this version:

Michel Lafon, Bolofo Mongezi. How to deal with so-called "vowel verbs with variant non-vowel forms" in a bilingual Zulu lexicon? The case for a pragmatic approach. Journal of the Digital Humanities Association of Southern Africa (DHASA), 2022, Proceedings of the 3rd Workshop on Resources for African Indigenous Languages 2022, 4 (1), 7p. halshs-04162019

**HAL Id: halshs-04162019**

**<https://shs.hal.science/halshs-04162019v1>**

Submitted on 13 Jul 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## How to deal with so-called "vowel verbs with variant non-vowel forms" in a bilingual Zulu lexicon ? The case for a pragmatic approach.

Michel Lafon<sup>1</sup>  
research fellow (retired),  
CNRS Llacan Paris, IFAS - Johannesburg  
& CenterPoL, University of Pretoria,  
[maikoro12@gmail.com](mailto:maikoro12@gmail.com)  
& Bolofo Mongezi  
PhD Candidate, Wits School of Education

### Abstract

Whilst lemmatisation in Zulu (and cognate languages) remains to this day a partially unsettled issue, with dictionaries, both bilingual and monolingual, adopting differing strategies in respect of nouns, we wish to focus here on a somewhat minute aspect of the matter: how to lemmatize verbs with optional initial vowels. As the variation occurs at the initial, it is crucial that it receives proper dictionary treatment, lest the user be misdirected or misinformed.

Keywords: isiZulu - lexicography - vowel verbs

We shall start by presenting verbs with optional initial vowels in Zulu, contrasting the lemmatization strategies offered by main available dictionaries, some of which in our view lack consistency. We shall then introduce to the pragmatic solution we adopted in our bi-directional French to Zulu and Zulu to French lexicon, which we claim gives better justice to the language.

### 1 The corpus

Zulu (or isiZulu)<sup>2</sup> is in terms of demography the main African (Bantu) language spoken in South-Africa,<sup>3</sup> being in its different sociolects the first or home language of close to 13 millions people, viz., 22% of a population of at least 59 millions.<sup>4</sup> Although there exists a number of regional or sub-regional dialects, its standard form, taught in schools and propagated through formal situations, including literature, the written and spoken press, as well as official ceremonies, is widely accepted. It also boasts a well-accepted orthography. This paper is based on the standard language and our data is taken mainly from secondary sources, *id est*, dictionaries as quoted in the paper, complemented by random observations taken from formal and informal written documents as well as informal speech, and the native-speaker fluency of one co-author.

Zulu has a rather large inventory of what (Clement Martyn Doke 1992, 131) refers to as "vowel verbs with variant non-vowel forms", such as

**-edlula** or **-dlula**,<sup>5</sup> *pass*<sup>6</sup>

**-esutha** or **-sutha**, *be satiated with food*

The initial vowel is mostly /e/, but /a/ and /o/ also occur, more often as further variants (examples from dictionaries):

**-ehlula**, **-ahlula** or **-hlula**, *conquer, defeat*

**-ejwayela**, **-ojwayela** or **-jwayela**, *be used to*

**-ephula, -aphula, -ophula, or  
—phula, break**

Whereas the presence in the standard of variant forms may proceed from the incorporation of regional features as, although based mainly on «*the Central Zululand Dialect*» (Kubeka 1979, 83), it includes arguably features of other varieties, it seems that the variation has now become evolutionary, as some of Doke's vowel-initial verbs do not seem valid today (but see further down):

**?-efunda for -funda study**  
**?-egcwala for -gcwala, be full**  
**?-emithi for -mithi be with small (animals); become pregnant**  
**?-ojwayela for -jwayela, be used to**

This would point to a phenomenon of vowel erosion. However the fact that some of those verbs, like **-gcwala**, seem to have their source in ideophones (**gcwa**, *of being full to the brim*) which are consonant-commencing suggests this explanation is not sufficient. The vowel would then have been added (or the ideophone would have been obtained by deletion of the vowel). We shall not investigate this issue further however, as it is not relevant to the problem at hand.

Nouns derived from the same root show the same alternative :

**abalusi or abelusi, herders (cf.  
—elusa or -lusa, herd)**

but all possibilities are not always admitted:

**umelaphi<sup>7</sup> healer : ? umlaphi,**

The same variation may apply to extended verbs and nouns derived from them:

**-ehluleka or -ahluleka, fail and isehluleki or isahluleki failure**  
**-efundisa or -fundisa teach, and abafundisi or abefundisi, priests**  
**-efundisa** seems clearly obsolete but the noun **abefundisi** is attested

It also happens that a variant  $\partial$  does not cover all meanings of variant  $\beta$ , suggesting that the verbs were originally different but became conflated due to the phonetic loss of the initial vowel:

**-phuza** has now two meanings:  
**—phuza, drink and -phuza**  
 (variant of **-ephuza**) *be late*

To this category we add the three (or four) verbs with /i/ as the initial vowel, that Doke in the passage quoted above refers to as “*latent-vowel verbs*” due to the fact that /i/ may only appear in conjugated forms when preceded by a vowel /a/ with which it then coalesces resulting in /e/:<sup>8</sup>

**-(i)ma: stand: ngiyema or ngiyama, I stand**  
**-(i)-za, come: bayeza or bayaza, they are coming**  
**-(i)zwa, understand: ubezwa or ubazwa, listen to them**

Since in Zulu /i/ tends to disappear when following any vowel other than /a/ (or merge if the preceding vowel is /i/), no conclusion can be drawn from instances where initial /i/ does not surface, as to what is the variant of the verb stem in any given instance:

**ngimile, I stood: Ingil-liml-lilel or Ingil-lml-lilel**

**uzwile, you heard: lul-lizwl-lilel or lul-lzwl-lilel**

We therefore posit for those verbs a variant form of the stem with an initial vowel, akin to the situation above :

**-ima or -ma, stand**

**-iza or -za, come**

**-izwa or -zwa, hear, understand**

The choice of variant appears related to the tense form, the presence of an extension after the root which seems to result in a preference for the shorter, consonant-commencing form, as well as to the stem itself. The following examples are drawn from our own observations of unsolicited speech:

**ngiyezwa** rather than **ngiyazwa, I hear**

**ngiyazwisisa** rather than **ngiyezwisisa, I understand very clearly**

**kuyezwakala** or **kuyazwakala, it is understandable**

**ngiyeza** rather than **ngiyaza, I am coming**

**ngiyabazisa endlini** rather than **ngiyabezisa endlini, I make them come home**

**ngiyema** or **ngiyama, I stand**

**ngiyamela abafazi** rather than **ngiyemela abafazi, I am waiting for the women**

## 2 Treatment in Dictionaries

How are these facts dealt with in major Zulu dictionaries?

Keeping to items already mentioned, we compare the main dictionaries presently

available: bilingual - Doke & al. Zulu-English and English-Zulu (Clement M. Doke et al. 1999), Dent and Nyembezi's Scholar's Zulu (Dent and Nyembezi 1995), de Schryver's (de Schryver 2015)- and monolingual - Nyembezi's (Nyembezi 1992) and Mbatha's (Mbatha 2006). We are aware that these dictionaries were not compiled within the same framework. Whereas the four older follow with small deviations between them the strategy set out by Doke for Zulu lexicography, which lemmatizes stems rather than words, de Schryver's, besides being corpus-based, introduces an approach which seeks to lemmatize words rather than stems (de Schryver and Wilkes 2008). However, this revolution in Zulu lemmatization does not really affect verbs. Even though Doke and Nyembezi quote verbs under the imperative while other dictionaries refer to the verb stem, appropriately defined in (Marlo 2013) as "an obligatory root, one or more possibly occurring derivational suffixes [also known as extensions]<sup>10</sup> and an inflectional final suffix commonly called the "Final Vowel", the form is segmentally similar, the singular imperative being none other in Zulu than the said stem with a specific tone structure.

We use the following symbols:

A : main entry (as revealed by length of description);

A': main entry although shorter than A [that implies two main entries in same dictionary]

A + a : main entry + cross-reference to at least one variant;

B : shortened entry, with cross-reference to main;

B' : cross-reference to main ;  
0 : no mention of the item

	<b>-phula</b>	<b>-ephula</b>	<b>-aphula</b>	<b>-esutha</b>	<b>-sutha</b>	<b>-egcwla</b>	<b>-gcwala</b>
	<i>break</i>			<i>be satiated</i>		<i>become full</i>	
Doke	B	B	A + a	A'	A	B	A + a
Dent	A	A	A	A'	A	0	A
de Schryver	A	0	0	0	0 <sup>11</sup>	0	A
Nyembezi	A'	A	A'	A'	A	0	A
Mbatha	A	A'	A'	A	A	0	A

	<b>-dlula</b>	<b>-edlula</b>	<b>-hlula</b>	<b>-ehlula</b>	<b>-ahlula</b>	<b>-emitha<sup>12</sup></b>	<b>-mitha</b>
	<i>pass</i>		<i>conquer</i>			<i>become pregnant</i>	
Doke	A + a	B	B	B	A + a	a	A + a
Dent	A	B'	A	A	A	A	0
de Schryver	B'	A	B'	A	B' <sup>13</sup>	0	A
Nyembezi	A	A'	A	A'	A'	A'	A
Mbatha	A	A'	A	A'	A'	0	A

	<b>-ima</b>	<b>-ma</b>	<b>-iza</b>	<b>-za</b>	<b>-imba</b>	<b>-mba</b>	<b>-izwa</b>	<b>-zwa</b>
	<i>stand</i>		<i>come</i>		<i>dig</i>		<i>hear</i>	
Doke	0	A + a	0	A + a	0	A + a	0	A + a
Dent	0	A	0	A	0	0	0	A
de Schryver	0	A		A	0	0 <sup>12</sup>	A	0
Nyembezi	0	A	0	A	0	A	0	A
Mbatha	0	A	0	A	0	A	0	A

### 3 Observations

It seems, even judging from such a limited and haphazard sample, that inconsistency prevails within dictionaries as well as between them, as a cursory look at the table suggests.

#### - Zulu to English.

Since Doke refers specifically to "*vowel verbs with variant non-vowel forms*" one would expect these verbs to be entered systematically under the vowel with an indication that a consonant-commencing variant exists: not so but most instances include a cross-reference, which mitigates the issue. In Dent on the other hand repetition is frequent: **-ephula**, **-aphula** and **-phula** among other such examples constitute three different entries with almost the same definitions repeated without any cross-referencing. De Schryver is more consistent in cross-referencing the (supposedly) less frequent form to the main one.

Regarding the so-called latent vowel or i-commencing verbs, Doke is the only one to systematically make mention of "*latent i*" in the description of all corresponding entries. Dent and de Schryver make no explicit allusion to it while implying it nevertheless through examples: under **-zwa**, Dent gives **ukungezwa nakutshelwa**, *to want to see by oneself ("not wanting to be told")*, where the negative marker **-nga-** alters to **nge** as its /a/ vowel coalesces with the initial /i/;<sup>145</sup>. In a similar fashion, under **-za** de Schryver gives **abantu beza ngobuningi**, *people came in their numbers*, whereas **beza** cannot be

obtained from the subject prefix **ba-** and the verb given as **-za**. Same situation under **zwakala** (separate entry from **-zwa**) with **kuyezwakala** *it is understandable*.<sup>16</sup>

#### - English to Zulu

In all three bilingual dictionaries, only the form given as main entry in the Zulu to English is indicated, with no mention whatsoever of variants. That applies *inter alia* to Dent where, under *break*, only **-aphula** is listed.

#### - Zulu monolingual

Nyembezi and Mbatha provide almost similar descriptions for each variant, each treated as a main entry (see **-dlula**, **-edlula** *inter alia*), occasionally indicating the existence of a variant assimilated to a synonym like in the **-hlula** and **-phula** series. Even if variants are somehow synonyms, it would make sense in a linguistic work to discriminate. And that does not preclude inconsistencies: one of Mbatha's examples under **dlula** is **ukwedlula ngendlu yakhiwa** *not to offer assistance to people working*, which should appear under **-edlula**. Same for **uwepfulile umoya wami lo mfana**, *this boy broke my heart*, which should illustrate **ephula** rather than **-phula**. As for the "latent vowel", no explicit mention whatsoever. However Mbatha includes in his examples conjugated forms where the "latent i" does appear: under **-za**: (...) **uyeza**, (...) *you are coming*; under **-zwa**: (...) **sengathi akezwa** (...) *as if he does not understand*; **ukungezwa ngokutshelwa** (see above and note 14).

#### 4 A pragmatic approach

Our bidirectional French / Zulu lexicon (Lafon and Mongezi 2022) has no pretention to be extensive but rather a handy support whilst one is engaged in conversation. Space and cost were huge concerns. Still we aimed at covering as much vocabulary as possible whilst attempting to remain consistent throughout.

In the French to Zulu part, we opted to place the possible vowel within brackets, so as to offer a maximum of information in a minimum of space, treating i-commencing verbs in a similar fashion:

passer (to pass): **-(e)dlula ;**

aller (to come): **-(i)za**

Thus all variants of conjugated forms can be deduced from the lemmas as they stand.

Obviously the same strategy could not be followed in the Zulu to French part. In order to avoid repetition as well as omission, we opted to lemmatize the form which appeared to us the more common (based on native speaker's intuition), the other or possibly other form(s) being however systematically listed and cross-referred to the one chosen as lemma :

**-edlula : voir [see] -dlula**

**-iza : voir [see] -za**

This strategy counterbalanced at least partially any wrong estimate on our side of the relative frequency of variants. The strategy obviously is not ignored by all of the dictionaries reviewed;<sup>7</sup> rather it is not applied systematically.

#### 5 Conclusion

It would seem, from the review of lexicographical treatment of a few items in old as well as recent Zulu dictionaries, both bilingual and monolingual, traditional as well as recent and corpus- and frequency-based (de Schryver's), that the twin issues of "*vowel verbs with variant non-vowel forms*" and "*latent-vowel verbs*" have been largely overlooked. Minor as this may seem, it remains a blind spot. Our *ad hoc* solution, to consistently include all variants in the simplest and less cumbersome manner, might then go some way towards achieving better coverage in that particular instance. This reminds us that there is no Holy Grail to achieving consistency and user-friendliness in the lexicography of Zulu and cognate languages. Lexicographical issues have to be considered taking into account their specifics, this being the only manner to provide adequate, complete and easily accessible information.

#### 6 References

- Bryant, Alfred T. 1905. A Zulu-English Dictionary: With Notes on Pronunciation, a Revised Orthography and Derivations and Cognate Words from Many Languages; Including Also a Vocabulary of Hlonipa Words, Tribal-Names. Etc., a Synopsis of Zulu Grammar and a Concise History of the Zulu People from the Most Ancient Times. Maritzburg: Davis & Sons.
- Colenso, John William. 1861. Zulu English Dictionary. Vol. 1. Pietermaritzburg.
- Dent, G. R., and C. L. Sibusiso Nyembezi. 1995. Scholar's Zulu Dictionary; English-Zulu, Zulu-English. [1969]. Pietermaritzburg: Shuter and Shooter.
- Doke, Clement M., D. Mc K. Malcolm, J. M. A. Sikakana, and B. Wallet Vilakazi. 1999.

- English-Zulu, Zulu-English Dictionary. [1958]. Wits University Press.
- Doke, Clement Martyn. 1992. Textbook of Zulu Grammar. [1927]. 1 vols. Cape Town: Longmans Southern Africa.
- Hadebe, Samukele, ed. 2001. Isichazamazwi SesiNdebele. Vol. 1. Harare: College Press Publishers and The African Language Research Institute (ALLEX).
- Kubeka, Isaac Sibusiso. 1979. 'A Preliminary Survey of Zulu Dialects in Natal and Zululand'. Univ of Natal, dep of Zulu language & literature.
- Lafon, Michel, and Bolofo Mongezi. 2022. Lexique Français-Zoulou; Isichayamagama SesiZulu Kuya KwisiFulentshi. Moroni: KomEdit.
- Marlo, Michael R. 2013. 'Verb Tone in Bantu Languages: Micro-Typological Patterns and Research Models'. *Africana Linguistica* XIX: 137–234.
- Mbatha, M. O. 2006. Isichamazwi SesiZulu. Pietermaritzburg, South Africa.
- Nyembezi, C. L. Sibusiso. 1992. AZ: Isichazamazwi Sanamuhla Nangomuso. Pietermaritzburg: Reach Out Publishers.
- Samuelson, R.C.A. 1923. The King Cetywayo Zulu Dictionary. Durban.
- Schryver, Gilles-Maurice de, ed. 2015. Funda IsiNgisi. Learn IsiZulu. IsiZulu-IsiNgisi IsiNgisi-IsiZulu. Isichazamazwi Sesikole. School Dictionary. 2nd ed. Cape Town: Oxford University Press.
- Schryver, Gilles-Maurice de, and Arnett Wilkes. 2008. 'User-Friendly Dictionaries for Zulu: An Exercise in Complexicography'. In Proceedings, edited by Elisenda Bernal and Janet DeCesaris, 827–36. Barcelone, Institut Universitari de Linguística Aplicada.
- 
- <sup>1</sup> Corresponding author
- <sup>2</sup> It is considered appropriate in South Africa to refer to local African (Bantu) languages with their prefixes in the name of respecting self-identification. We have however opted to use consistently English language-naming conventions. Thus we maintain French rather than « français » in the text.
- 
- <sup>3</sup> Bantu is a convenient classificatory term for a family of languages indigenous to the African continent. The term has no implication outside the linguistic field. Languages of this family share a number of features, including a noun class system characterized by prefixes and agreement schemes.
- <sup>4</sup> Extrapolated from the 2011 Census which included a question on ethnicity, and the population figures for 2020 (see <https://www.statssa.gov.za>). This figure probably does not give justice to recent immigration. The results of the 2022 census are not published to date.
- <sup>5</sup> Verbs are quoted under their stem, preceded by a hyphen to indicate that any conjugated form - apart from the imperative- requires prepositioned elements. See also below.
- <sup>6</sup> Translations are given so as to facilitate identification of each item but by no means claim to cover the whole shade of meaning of any given word or stem.
- <sup>7</sup> Heard in one episode of the Zulu-spoken soapie **Uzalo** 207, November 2018, SABC 1
- <sup>8</sup> Doke includes **-iva**, *increase*, but we keep it out due to its obsolescence. **-ipha** and **-ikha** are sometimes given as variants for **-pha** *give* and **-kha** *draw water* respectively, but these not being commonly accepted, we refrain from considering them either.
- <sup>9</sup> We did not have access to previous substantial works, such as (Colenso 1861), (Samuelson 1923) or (Bryant 1905), by his own admission main sources of Doke.
- <sup>10</sup> Our addition.
- <sup>11</sup> Item apparently missed or omitted by de Schryver, which is puzzling as it does seem a frequent enough term with forms such as **ngisuthi**, *I am satiated*, etc
- <sup>12</sup> **-emithi** and **-mithi** in Doke.
- <sup>13</sup> No specific entry **\*-ahlula** but passive **-ahlulwa** referred to **-ehlulwa**
- <sup>14</sup> Cf. for instance **ukungavumi ukutshelwa**, *not accepting to be told* where **-nga-** remains unaffected as it precedes a consonant.
- <sup>15</sup> As Dent's verb examples keep to the verbal stem, one is hard pressed to see if instances of "latent i" would surface.
- <sup>16</sup> The failure to identify the presence of the initial vowel in such cases suggests that the same may have occurred elsewhere: hence, the absence of vowel initial verbs in the dictionary does not necessarily imply they were not encountered in the corpus; it would rather point to a failure of the parser to tear them out.
- <sup>17</sup> Same situation in a Zimbabwean isiNdebele dictionary (Hadebe 2001)