

List of concords and affixes in isiZulu

C. Maria Keet

Department of Computer Science, University of Cape Town, South Africa,
mkeet@cs.uct.ac.za

October 7, 2022

Abstract

This document aims to collect in one place all the concords and noun class-relevant affixes for the noun classes of isiZulu. The current version includes: the oral prefix, augment, prefix, quantitative concords, enumerative concord, relative concord, possessive concord, adjectival concord, demonstrative pronouns with this/these and that/those, emphatic pronouns, possessive pronouns, subject concord, negative subject concord, object concord, and relative tense concords.

1 Overview

Knowing which concords and noun class-relevant affixes there are, with which values for each class, is useful for natural language generation, be it data-to-text, information-to-text or knowledge-to-text. It also could infuse machine learning, deep learning, neural MT algorithms with some basic knowledge to improve on their machine translation algorithms. Or for a generative grammar for other purposes, such as the development of a rule-based spellchecker.

This document first describes some information on how to read the two tables, notably the abbreviations used, notational conventions, and then all the sources used for the different concords included in the two tables.

Abbreviations

The following abbreviations are used in the tables below:

- NC: Noun class
- oral: oral prefix
- AU: augment
- PRE: prefix
- QC: quantitative concord (2x, for ‘for all’ and for part of the ‘some’).
- EC: enumerative concord
- RC: relative concord
- PC: possessive concord
- AC: adjectival concord
- DP: demonstrative pronouns, this/these and that/those
- EP: emphatic pronouns / Pronomial
- PP: possessive pronouns
- SC: subject concord

- NEG SC: negative subject concord
- OC: object concord
- RT: relative tense concords (4x)

Additional notes to the tables

A “?” in a table field means that I don’t know the value at the time of writing. Different lists of concords and affixes do not all have the same set of noun classes listed, largely due to datedness of the source. The most common omissions are 3a and 9a, and to some extent 17 as well. Occasionally, a(n old) source still lists 16, but my understanding is that that noun class has been definitely deprecated in the meantime and therefore has not been included in the following tables.

A single, lone, “-” in a table field means that the string is empty.

A “-” appended to characters provides an indication if it’s a proclitic or enclitic or prefix or suffix. e.g. *ye-* means that something must still follow this and *-khe* that it is appended at the end. Sometimes, it’s variable, notably with the OC: it’s typically in the middle of something, so it could be indicated accordingly (e.g., *-m-*), but in some constructions, the OC moves to the start of a word; in that case, position hints are omitted.

Brackets denote variants. For instance, with *um(u)-*, the full prefix is either *um-* or *umu-*. (in this case, selection of either depends on whether the stem is monosyllabic or not).

Sources

PRE, CQs, RC, PC, SC, NEG SC, OC, EP/PRON, EC, oral, AU are from or updated from the tables in [2, 5, 7, 6] (any highly unusual deviant cases of prefixes not included), where oral from [7] was originally from [3] and nc17’s OC from [8]. For QCall, only the contracted form is listed.

The AC’s and this/that DP’s from [9], p197 and p149, respectively, except for their respective NC 17 whose values come from [8].

The yonder DP values were taken from [8].

The possessive pronoun PP values were taken from [1] and [8].

RT relative tenses are further described by [4] and summarised in tables 1-4 in that document. Only the “contracted form”, i.e., what remains after the elision process, is included in the table below, not the “full form”. The RT shortly and long after are indeed the same in their contracted form (it’s not a typing up mistake), but not in their full form.

Table 1: Various concords and affixes; see text for details.

| NC | Full PRE | QC (V) | QC (Ξ) | EC | PC | AC | RC | DP (this/ these) | DP (that/ those) | DP (yonder) | EP | PP | oral | AU | PRE |
|----|------------------|-----------|-----------------|-----|------|--------|------|------------------------|------------------------|----------------|--------|------|------|-------|--------|
| 1 | um(u)- | wonke | ye- | mu- | wa- | om(u)- | o- | lo | lowo | lowaya | yena | -khe | u- | u- | m(u)- |
| 2 | aba- | bonke | bo- | ba- | ba- | aba- | aba- | laba | labo | labaya | bona | -bo | ba- | a- | ba- |
| 1a | u- | wonke | ye- | mu- | wa- | om(u)- | o- | lo | lowo | lowaya | yena | -khe | u- | u- | - |
| 2a | o- | bonke | bo- | ba- | ba- | aba- | aba- | laba | labo | labaya | bona | -bo | ba- | o- | - |
| 3a | u- | wonke | ye- | mu- | wa- | ? | o- | ? | ? | ? | wona | ? | u- | u- | - |
| 2a | o- | bonke | bo- | ba- | ba- | aba- | aba- | laba | labo | labaya | bona | -bo | ba- | o- | - |
| 3 | um(u)- | wonke | wo- | mu- | wa- | om(u)- | o- | lo | lowo | lowaya | wona | -wo | u- | u- | m(u)- |
| 4 | imi- | yonke | yo- | mi- | ya- | emi- | e- | le | leyo | leyaya | yona | yo- | i- | i- | mi- |
| 5 | i(li)- | lonke | lo- | li- | la- | eli- | eli- | leli | lelo | leliya | wona | lo- | li- | i- | (li)- |
| 6 | ama- | onke | wo- | ma- | a- | ama- | a- | la | lawo | lawaya | wona | wo- | a- | a- | ma- |
| 7 | isi- | sonke | so- | si- | sa- | esi- | esi- | lesi | leso | lesiya | sona | so- | si- | i- | si- |
| 8 | izi- | zonke | zo- | zi- | za- | ezin- | ezi- | lezi | lezo | leziya | zona | zo- | zi- | i- | zi- |
| 9a | i- | yonke | yo- | yi- | ya- | ? | e- | ? | ? | ? | yona | ? | i- | i- | - |
| 6 | ama- | onke | wo- | ma- | a- | ama- | a- | la | lawo | lawaya | wona | wo- | a- | a- | ma- |
| 9 | i(n)-, i(m)- | yonke | yo- | yi- | ya- | en- | e- | le | leyo | leyaya | yona | yo- | i- | i(n)- | - |
| 10 | izi(n)-, izi(m)- | zonke | zo- | zi- | za- | ezin- | ezi- | lezi | lezi | leziya | zona | zo- | zi- | i- | zi(n)- |
| 11 | u(lu)- | lonke | lo- | lu- | lwa- | olu- | olu- | lolu | lolo | loluya | lona | lo- | lu- | u- | (lu)- |
| 10 | izi(n)-, izi(m)- | zonke | zo- | zi- | za- | ezin- | ezi- | lezi | lezi | leziya | zona | zo- | zi- | i- | zi(n)- |
| 14 | ubu- | bonke | bo- | bu- | ba- | obu- | obu- | lobu | lobo | lobuya | bona | bo- | ba- | u- | bu- |
| 15 | uku- | konke | ko- | ku- | kwa- | oku- | oku- | lokhu | lokho | lokhuya | khona | -kho | ku- | u- | ku- |
| 17 | ku- | lonke | lo- | | kwa- | oku- | olu- | lokhu- | lokho- | lokhuya | khona- | -kho | | | ku- |

Table 2: Concords for verbs; see text for details.

| NC | SC | NEG SC | OC | RT (shortly before coding time) | RT (long before) | RT (shortly after) | RT (long after) |
|----|-----|--------|----|---------------------------------|------------------|--------------------|-----------------|
| 1 | u- | aka- | m | ube- | waye- | wobe- | wobe- |
| 2 | ba- | aba- | ba | bebe- | babe- | bobe be- | bobe be- |
| 1a | u- | aka- | m | ube- | waye- | wobe- | wobe- |
| 2a | ba- | aba- | ba | bebe- | babe- | bobe be- | bobe be- |
| 3a | u- | aka- | wu | bewu- / ubu- | wawu- | wobu- | wobu- |
| 2a | ba- | aba- | ba | bebe- | babe- | bobe be- | bobe be- |
| 3 | u- | awu- | wu | bewu- / ubu- | wawu- | wobu- | wobu- |
| 4 | i- | ayi- | yi | beyi- / ibi- | yayi- | yobi- | yobi- |
| 5 | li- | ali- | li | beli- | lali- | lobel li- | lobe li- |
| 6 | a- | awa- | wa | abe- / aye- | aye- | obe- | obe- |
| 7 | si- | asi- | si | besi- | sasi- | sobe si- | sobe si- |
| 8 | zi- | azi- | zi | bezi- | zazi- | zobe zi- | zobe zi- |
| 9a | i- | ayi- | yi | ? | ? | ? | ? |
| 6 | a- | awa- | wa | abe- / aye- | aye- | obe- | obe- |
| 9 | i- | ayi- | yi | beyi- / ibi- | yayi- | yobi- | yobi- |
| 10 | zi- | azi- | zi | bezi- | zazi- | zobe zi- | zobe zi- |
| 11 | lu- | alu- | lu | belu- | lwalu- | lobe lu- | lobe lu- |
| 10 | zi- | azi- | zi | bezi- | zazi- | zobe zi- | zobe zi- |
| 14 | bu- | abu- | bu | bebu- | babu- | kobe ku- | kobe ku- |
| 15 | ku- | aku- | ku | beku- | kwaku- | kobe ku- | kobe ku- |
| 17 | lu- | kwa- | ku | beku- | kwaku- | kobe ku- | kobe ku- |

Acknowledgements

My main collaborator in NLG and, indirectly with that, trying to structure the linguistic knowledge, is Langa Khumalo, and I thank him for the collaboration.

This work is based on the research supported in part by the National Research Foundation of South Africa (Grant Number 120852).

References

- [1] Anonymous. noun class table. Online material, 2022. Wiktionary, https://africanlanguages.com/zulu/noun_class_table.html; last accessed: 6-10-2022.
- [2] J. Byamugisha, C. M. Keet, and L. Khumalo. Pluralising nouns in isiZulu and similar languages. In A. Gelbukh, editor, *Proceedings of CICLing’16*, volume 9623 of *LNCS*, pages 271–283. Springer, 2018.
- [3] J. Goldsmith and G. Buthelezi. The Zulu Language – Fall 2005. Online course material, Fall 2005. University of Chicago, <http://hum.uchicago.edu/jagoldsm/ZuluLanguage/>; last accessed: 24-12-2014.
- [4] L. Hall. The -be relative tenses of Zulu. MA thesis, Faculty of Humanities, University of Pretoria, 2005.
- [5] C. M. Keet and L. Khumalo. On the verbalization patterns of part-whole relations in isizulu. In *9th International Natural Language Generation conference (INLG’16)*, pages 174–183. ACL, 2016. 5-8 September, 2016, Edinburgh, UK.

- [6] C. M. Keet and L. Khumalo. Grammar rules for the isizulu complex verb. *Southern African Journal of Language and Linguistics*, 35(2):183–200, 2017.
- [7] C. M. Keet and L. Khumalo. Toward a knowledge-to-text controlled natural language of isiZulu. *Language Resources and Evaluation*, 51(1):131–157, 2017.
- [8] Multiauthor. Appendix:Zulu concords. Online material, 2022. Wiktionary, https://en.wiktionary.org/wiki/Appendix:Zulu_concords; last accessed: 6-10-2022.
- [9] A. Wilkes and N. Nkosi. *Teach yourself Zulu*. Hodder Education, UK, 2003.