

Automation and the Necessity of Expert Intervention in Lexicography: The Case of Kasahorow's French-Zulu Pocket Dictionary

Michel Lafon^{1*} 

¹Retired French Researcher, CNRS Paris, France

IFAS Johannesburg and CenteRPol, University of Pretoria, South Africa

maikoro12@gmail.com

Mongezi Bolofo² 

²University of South Africa, South Africa

moeketsiebolis@gmail.com

Received: 19/02/2026

Accepted: 09/06/2026

Published: 30/06/2026

***Corresponding author**

Citation:

Lafon, M., & Bolofo, M.
(2026). Automation and the
Necessity of Expert
Intervention in
Lexicography: The Case of
Kasahorow's French-Zulu
Pocket Dictionary. *Studies
in Linguistics and Language
Education*, 2(1), 38–59.

***Studies in Linguistics and
Language Education (SLLE)***

© 2026 The Author(s).

Published by the University of

Oran 2 Mohamed Ben

Ahmed, Algeria.

This is an open access article

under the [CC BY](https://creativecommons.org/licenses/by/4.0/) license



ABSTRACT: *This paper offers a critical assessment of a Zulu to French glossary which was most probably compiled with the help of machine translation. It evolved from a simple review to a full paper, given the wealth of the issues identified. We start by providing information on the book's editors cum publishers which we deemed relevant to the case in point; we pursue by offering a brief sketch of Zulu lexicographic conventions and attending debates as background to our critics, so that readers unfamiliar with Zulu lexicography, may construe their own opinion in respect of the critics advanced. We then come to the core of the paper. We substantiate our remarks by numerous examples extracted from the glossary. We conclude by stressing that relying solely on machine translation without human expert intervention in such endeavours as bilingual glossaries, is a recipe for substandard products.*

KEYWORDS: Bilingual Dictionary, French, isiZulu, Kasahorow, Lexicography, Machine Translation.

¹ Michel Lafon has conducted research on shiNgazidja or Great Comorian and ciShona, both Bantu languages on which he (co-)authored lexicographical work (Lafon, Saadi, and Zaina 1991), (Lafon 1995). He also contributed, on behalf of IFRA, to the 1996 Shona monolingual dictionary (<https://www.dokpro.uio.no/allex/gsd/fm/2-ndiani.htm>) (Lafon 1994). Since 2000 he has turned his attention to isiZulu. ML. & BM have co-authored a French to Zulu lexicon (Lafon and Mongezi 2022), the 2nd edition of which they are presently editing. The comments on the work at stake are based on our research on Zulu and our own respective knowledge of both languages.

Context of the paper

In the course of the compilation of a French-Zulu lexicon in 2021 (Lafon and Mongezi 2022) we took notice of a French-Zulu dictionary published under the aegis of the Kasahorow foundation, a concern we had not heard about before.² So as to avoid the risk of being influenced even unwittingly in our own work, we refrained from looking at it. We eventually ordered it in late 2025 whilst preparing the 2nd edition of the lexicon, as our own methodology, corpus and so forth, were set and, at that stage, were no longer susceptible to be influenced.

The Kasahorow foundation.³

The Kasahorow foundation presents itself as a Ghana-based NGO, with branches in Germany (which provides funding) and the US (which commercialises products), and is concerned with the noble goal of, in their own words, “promot[ing] inclusion through languages”; it focuses on communication between people from different linguistic and cultural backgrounds, initially within Africa itself or between Africa and the rest of the world; since 2011 it has expanded across all continents wherever there is a demand, that is, or so we are made to understand, an affiliated individual requesting for the inclusion of a certain language, so to “reflect the diversity of our world”. To quote from the website again: “We can further customize our products between any 2 languages you desire. For example: "Amharic & Akan Animals Activity book" or "Swahili & Wolof Learner's Dictionary", etc.” The Kasahorow shop webpage (Bakuwa) reinforces this aim: “With thousands of language combinations available, our dictionaries, activity books, and learning resources make it easier to bring your languages into everyday life—at home, at school, and beyond.” We understand this came out of a concern for transmitting first languages in the home in spite of the pressure of the multilingual world many live in in Africa, which by all means is an objective one can only concur with. Browsing through the webpage, we found ca100 languages, many from Africa, accessible with French as a member of the pair. Zulu, arguably the main African language in the country, is an obvious candidate, and the pairing with French makes particular sense considering the many francophone immigrants in the country.⁴

It is indeed a notable and praiseworthy feature of the Kasahorow initiative to offer linkage between any pair of languages as per request, even if the bulk of publications listed gives, as could be expected, as the main languages in any pair, English followed by French and German. Their publications, including children learners’ dictionaries and dictionaries proper, are commercialised since 2023 in PDF format text. This paper however is strictly concerned with the French-Zulu dictionary; no inference should be made to other Kasahorow products.

² Kasahorow, ed. 2025 [2014], Dictionnaire Zoulou de Poche Zoulou-Français, Français-Zoulou. <https://bakuwa.kasahorow.com/products/dictionnaire-multilingue-de-poche?variant=50780536635679>

³ Mostly taken from the Kasahorow foundation webpage (<https://www.kasahorow.org/about>, visited during late December 2025 and January 2026).

⁴ The 1996 South-African Constitution recognizes 11 official languages, among which 9 are African and belong to the Bantu family (see below for the use of this word). Zulu dominates largely in terms of numbers and use: self-declared home or first language by nearly 25% of the country’s population of 62 millions (https://en.wikipedia.org/wiki/2022_South_African_census), it is understood at various degrees by another 20% whose Nguni language varieties display a close proximity with it, while it is practised as a second language by many Africans of non-Nguni other linguistic background, including recent immigrants from neighbouring countries. It has de facto become the African lingua franca in many urban settings.

A Brief Overview of Zulu grammatical structure and its implications for lexicography⁵

A preliminary word on Zulu language structure, its spelling system and attending consequences on lexicography, is required. In Zulu, as in Bantu languages in general,⁶ the major word categories, except ideophones,⁷ viz., verbs, nouns, adjectives, pronouns, etc., consist of a lexical item (i. e., roots, stems, themes or bases)⁸ preceded by at least one varying element (prefixes for nouns, concords for adjectives, pronouns and the like, indexes for verbal forms, etc.). Such terms may be called “variable”. The noun determines the concords that apply to any term governed by it. As the concords are either identical or bear similarity to the noun prefix, Bantu languages were initially deemed alliterative.

In Zulu, the full noun prefix consist of an initial vowel (aka preprefix or augment),⁹ plus the prefix proper. To take a simple example:

umuntu, person, human being: |u-mu-ntu|, augment – prefix - root

Zulu Nouns are distributed among 15 so-called “classes”.¹⁰

If the prefix can be void,¹¹ no class is deprived of the augment. Borrowed nouns are customarily inserted into a class, through the addition of an augment.¹² Many lexical items also have suffixes. There is often coalescence or merging at the boundary between the affixe(s) and the lexical item. Hence Zulu and Bantu languages in general are considered agglutinative. Their spelling systems in most cases reflect this feature, with written words drawn-out as they incorporate prefixed (and suffixed) elements to the root in what is dubbed conjunctive writing. This is very much the case for Zulu after Colenso’s suggestion influenced individuals and institutions involved in the production of texts; it is reflected in the settlement reached in 1934 by the Central Orthography Committee under the auspices of the Inter-Universities Committee for African Studies (Clement Martyn Doke 1954, 16);¹³ the prescriptions arrived at remain the basis of the present orthography, as presented in successive orthographical sketches, culminating to date with (Maphumulo 2021). The Zulu word structure however has remained a challenge in Zulu lexicography. If a given noun verifies only a limited number of prefixes, it is not so for verbs in which members of various paradigms, expressing status and tense markers,

⁵ The following intends merely to shed light on aspects relevant to the points discussed herein. Zulu has inspired a significant number of grammatical publications, Doke’s Textbook, initially published in 1920, remaining a must (Clement Martyn Doke 1992). De Schryver’s dictionary (de Schryver 2015) includes a good measure of grammatical information along roughly the same line. (Cope 1984), unfortunately not easily available, offers a different approach as does our own conversation manual in French (Lafon and Mongezi 2021).

⁶ We use the term Bantu simply to refer to a family of languages spread over a large part of Africa south of the Sahara and sharing a significant number of grammatical features and lexical cognates. The scientific literature covering Bantu languages is abundant. For compendiums, see (Guthrie 1967), (Alexandre 1981) for a summary in French, and more recently (Nurse and Philippson 2003).

⁷ Ideophones (also known as mimetics or expressives) flourish in languages from Asia, Africa and the America, but not so, seemingly, from Europe. They invite iconic associations (see Wikipedia). In Zulu, they constitute a category numerically significant; bar a few, their usage is now fading away, but their importance lies in that they constitute the root of many verbs and nouns.

⁸ These terms may serve to indicate slightly different lexical units, according to the grammarian own academic language and school of thought; these distinctions are irrelevant here. On verbal bases, see below.

⁹ This initial vowel anticipates the quality of the prefix vowel.

¹⁰ The overall number varies slightly according to sources, as the existence of some classes in Zulu is disputed (see for instance cl. 17 below).

¹¹ This is the case for subclasses 1a, 3a, 9a (see (Canonici 1990)).

¹² This is not to say that the augment is always expressed in real utterances.

¹³ Colenso, first Anglican bishop of Natal, serving from 1853 to his death in 1883, became an influent promoter of the use of the language. On unifying the orthography, see (Natal Missionary Conference 1911).

subject and object infixes, in particular, appear before the root; combined with different suffixes, this feature makes listing all occurring forms of a given verb well-nigh impracticable. Various strategies are resorted to, with no overall consensus so far and dictionaries still differ, even if until recently Doke's prescriptions set the tune.¹⁴ Let us examine how the main dictionaries have dealt with the problem.¹⁵

In Bryant's massive dictionary (Bryant 1905), lexical items commence with a capital which determine their place in the dictionary. Nouns are listed however with their "full" prefix, separated by a hyphen from the lexical item; the full prefix consist of the augment plus the prefix proper. Adjectives are listed without any concordial agreement and verbs are given under their base, consisting of the root terminated by the neutral desinence -a. The part of speech being indicated in italics.

Bryant: **umu-Ntu** *n.*, human being, person; **Hle** *adj.*, good; **Hamba** *v.*, go, travel¹⁶

Breaking away partially from Bryant's method -or furthering it, depending on appreciation- and although Bryant's was by his own admission a major influence on his own dictionary, Doke advocated for a dictionary of stems, rather than words, a strategy which he fully applied to nouns. Hence in his Zulu-English dictionary whose 2nd edition appeared in 1953 (and which, now combined with English-Zulu, has become the lexicographical reference for Zulu) (Clement M. Doke et al. 2014). Nouns herein appear stripped of any prefix and introduced by an hyphen indicating that the lemma is not a complete utterance, while the entry lists the noun(s) with its prefix(es) in full.¹⁷ Doke's sheer wealth in coverage and meanings has made it unsurpassable so as to "stall(ing) all future lexicographic efforts for Zulu" as de Schryver observed somewhat ironically (de Schryver 2008a, 64), imposing along the way its lexicographic strategy. Dent & Nyembezi's later dictionary (Dent and Nyembezi 1962), as well as Nyembezi monolingual (Nyembezi 1992), adopt Doke's lemmatization strategy for nouns, while keeping to the prefixes rather than the full word in the entry, followed by an hyphen:¹⁸

Doke: **-ntu (umuntu, abantu)** *n.*; Dent & Nyembezi: **-ntu (umu-, aba-)**, (*n.*);
Nyembezi 1992 **-ntu (úmú -, ábá-)**; person

Doke: **-phathi (umphathi, abaphathi)** *n.* ; Dent & Nyembezi: **-phathi (um-, aba-)** *n.*;
Nyembezi 1992: **-pháthi (úm-, ábá-)**; one in charge

The same applies to adjectives, pronouns and other categories taking concordial agreement, only that the prefixes are not listed exhaustively, if at all, merely subsumed under the initial hyphen.¹⁹

¹⁴ The Kasahorow dictionary that we comment on is dated 2025 (date of purchase), but the webpage, visited mid-January 2026, refers to a 2023 revised edition (although we suspect this refers rather to a general change in the layout of this type of Kasahorow dictionaries). The first page of the text indicates a 2014 copyright (©2014), but shows no ISBN. Dictionaries and papers published prior to 2014 are therefore especially relevant, but the edition dated 2025 *could* have been amended in view of recent publications, as late as 2023 at least.

¹⁵ Since the late years of the 19th century, Zulu language has received much attention, resulting in a flurry of publications, from individuals of European extraction and native speakers alike. Since the early years, there has been a marked interest in lexicography.

¹⁶ Quotes from published dictionaries attempt to be identical reproduction of the original, including bold or italics but save the font. We limit ourselves to a common enough translation.

¹⁷ The lemmas here do not equate strictly with lexical items as the former may include a suffix; in some instances, even a nominal prefix (ex. below).

¹⁸ Translation and word category are taken from the bilingual dictionaries. We do not repeat the translation for each source; Nyembezi monolingual marks tone on each syllable.

¹⁹ See de Schryver below however on adjectival forms.

Doke: **-hle** adj. (**isihle, ubuhle, uMuhle, kahle**);²⁰ Dent & Nyembezi, Nyembezi 1992: **-hle** (adj.) good, pretty

For verbs, Doke departs only in theory from Bryant. He claims to quote verbs under the singular imperative, so he can indicate the tonal scheme of each, as he screened their utterance;²¹ it so happens that this form equates graphically the verbal base. Although Doke appears not to have carried the point, being followed by (Nyembezi 1992) only among our sample of dictionaries, the end result is very much the same, save the addition of the hyphen as in Dent & Nyembezi and the recent (de Schryver 2015).²²

Doke: **hamba** v.; Nyembezi: **hámba** sz;²³

Dent & Nyembezi: **-hamba** (v); de Schryver: **-hamba**, verb, travel

It is important to note that the hyphen here is an artificial construct, used out of convenience, to convey the critical information that the lemma as quoted is incomplete (we have seen though that this may be controversial for verbs), and not by any means a part of a word. Hence it is neutralised in the alphabetical order in all dictionaries that resort to it, for whichever category of terms.²⁴ Such use of the hyphen has now become standard in Bantu lexicography at large for terms which take a variety of prefixes, particularly verbs (see for instance recent online dictionaries, such as The General Shona Dictionary, <https://www.dokpro.uio.no/allex/gsd.html>, or the Swahili-English Dictionary <https://africanlanguages.com/swahili/>).

Interestingly Mbatha monolingual dictionary avoids completely the use of the hyphen (Mbatha 2006).²⁵ Verbs are quoted under their infinitive, a nominal form characterized by the cl. 15 prefix *uku-*. As for the terms which agree with nouns and hence take concords (adjectives, pronouns, etc.) they are (conspicuously) absent from the dictionary.²⁶ Mbatha's is therefore essentially a dictionary of nouns including verbal infinitives, and ideophones. Nouns are not lemmatized by their prefix but, *à la Bryant*, by the stem or root initial, using italics for the prefix and bold for the lemma: thus, *umuntu*, person, *isikole*, school, *ukuvala*, close, *ukuvusa*, wake up, etc. De Schryver aligns with Mbatha for nouns *stricto sensu*, only that they are listed according to the prefix initial letter: **umuntu**, **isikole**, etc. For adjectives and the like, the

²⁰ The bracketed terms correspond to words derived from the root, and appearing as such in the volume; **uMuhle** comes under **-Muhle**, where **-mu-** is integrated into the lemma as the class is deemed to be 1a rather than 1, on account of its plural formation **oMuhle**. **UMuhle** is glossed by Doke as the name given to a district Native Commissioner. It would be a secondary formation.

²¹ This is partly theoretical as the plausibility of imperatives such as *na!* (rain!), is questionable.

²² De Schryver first edition is dated 2010.

²³ Nyembezi uses Zulu abbreviations. sz stands for **isenzo**, action.

²⁴ The hyphen in isiZulu also serves in borrowed nouns commencing in a vowel, to avoid an orthographical hiatus by isolating the augment: **i-apula**, *apple*. The hyphen here is part and parcel of the word, and belongs to the spelling system. We are not concerned with this usage.

²⁵ Although not explicitly indicated in the foreword to the dictionary itself, this dictionary is the product of the government funded isiZulu Lexicography Unit, as referenced in their own webpage, a testimony to the government's commitment to promoting African languages. (<http://www.isichazamazwi.co.za/isichazamazwi-3/>).

²⁶ The official reason, comments (de Schryver 2008a, 65), is that the dictionary is limited to "content words", and adjectives and the like are ascribed to the grammar and are not deemed meaning carriers. We suspect that the perceived difficulty to lemmatize them might also have played a part. Indeed, the lemmatization principle explicitly rests on full words: "**ukuthi igama negama libhalwe liphelele**" (Mbatha 2006, ii) ("*so that each and every word be written in full*", translation ours). Neither excuse however can justify such exclusion in a general dictionary.

strategy is to quote them with a concordial agreement,²⁷ limiting however to forms reaching at least 50 occurrences in his corpus (de Schryver 2008a, 69, 70): thus agreement in cl. 14 and occasionally cl. 11 may be excluded.²⁸ Verbs are quoted through the use of the hyphen preceding the verbal base.

Coming to the point: The Kasahorow French-Zulu pocket dictionary.

General

The digitalised copy of the dictionary, purchased in October 2025 from Kasahorow Baquwa shop,²⁹ comes, once the PDF is printed, as a medium sized volume of 189 pages, 92 pages for each section (Zulu to French, French to Zulu). Pages' format is magazine -21x29.6- rather than book. Entries appear in two columns in each page, all introduced by a square (see the reproduction of page 1 annexed). Each lemma in bold is followed by an indication of the speech part it is ascribed to through an abbreviation, then follows the translation in French, by a single word, justifying its qualification as a word-for-word dictionary, if such a concept is acceptable (see conclusion). If a word is deemed to deserve more than one translation, it is repeated. It includes an estimate of 3000 entries, including reduplicated ones. The volume has neither any indication of individual or institutional authorship, nor introduction, or any prefatory matter as lexicographers world over are wont to provide, on, say, works consulted, the corpus and data collection method, classification and lemmatization strategies, language registers, and the like. Still as we have seen Zulu lexicographical conventions remain an academic battlefield, and different schools justify different outputs. Clarity of intent here is of the essence. In a language such as Zulu the lexicographical practices of which still fluctuate, as illustrated above, one should either follow what has become dominant, not to say standard, or spell out the template followed, explaining the reasons supporting the deviation, and, above all in our view, remain consistent throughout.

What is the intended public of the piece? This is difficult to say. The price is reasonable. As it is not searchable on line, it seems the intention is that users will print it, in which case it is not a pocket, portable dictionary but rather a desktop one, which suggests in turn it aims more an academic public than, say, travelers or tourists. This public, if we are correct, would imply a more thoroughly researched piece.

We consider only the Zulu to French section of Kasahorow's dictionary (p3 to 95), as our interest lies in the treatment of Zulu.

Zulu lemmatization strategy in Kasahorow French-Zulu dictionary

Zulu entries are divided in two unequal chunks: the first 10 pages consist of terms commencing with an hyphen. They include some verbs and verbal forms, adjectives, pronouns (including possessives and demonstratives). This is seemingly in accordance with the conventional use of the hyphen in Zulu lexicography for variable terms (see above). However, if the second chunk, which makes up the bulk of the dictionary, of about 84 pages, does include, as expected, nouns, it also includes verbs, some being repeated from the first part, with

²⁷ The concord here is the so-called relative agreement, which is itself a construction as de Schryver acknowledges (de Schryver 2008a, 78) (for a critical view, see Lafon 2026).

²⁸ Also, somewhat surprisingly, cl. 2 for **-de**, tall as in (**abantu**) **abade**, tall people.

²⁹ The purchase is submitted to a commitment to respect copyrights by not circulating copies.

meanings similar: **-vala** and **vala**, ‘fermer’ (close), or different: **-vusa**, ‘relancer’ (start again), and **vusa**, ‘remuer’ (move, stir), among a plethora of similar examples.³⁰ We come back to this issue when we address verbs as a category.

A closer look at the entries does little to save the day. We were interested mainly in establishing the lemmatization strategy. The dictionary volunteering no indication whatsoever of any presiding principle, it can only be extracted from the data, viz., examining how lemmas are quoted and labelled. Examples were taken by cherry picking among pages of both sections (hyphenated and not). Obviously we retained entries we claim are problematic and did not dwell on accurate ones. What emerged are mis-categorizations aplenty, leading to inconsistencies, as will be shown hereunder. We also considered briefly some translation issues as well as words which have no place in such a dictionary, since they are not Zulu. A handful of instances taken from the main categories will suffice, lest we reproduce a large portion of the dictionary. By guess estimate, close to half of the entries are not fully accurate one way or the other, be it the lemma labelling or the meaning. To avoid repetition, we mention spelling mistakes and similar problems as we proceed with the examples, even though the issue here is obviously of less weight. We however refrained from picking words exhibiting mere *tipos*, the above being already overwhelming.

Abbreviations for word categories are reasonably transparent, a clear advantage in the absence of any abbreviations’ chart:

- act: (? action) presumably for verbs
- adj: clearly for adjectives
- adv: clearly for adverbs
- cjn: clearly for conjunction
- det.: clearly for determiners
- exc.: clearly for exclamatives
- nom: clearly for nouns
- pos: clearly for possessives
- pre: clearly for prepositions
- pro.: clearly for pronouns

We now look at a handful among the items that pose problems, within various categories. We focused on items representative of the shortcomings of the categorization.

i. Act., i. e., verbs

Act. (verbs) includes various verbal constructions, some legitimate, some not, and even some terms belonging to other, non-verbal, categories. Now, the lemmatization practices for verbs which have been used in reference dictionaries (see above), include

- ✓ the base preceded by the hyphen,
- ✓ the base without hyphen (allegedly the imperative),
- ✓ the infinitive under its nominal trappings.

Either one can claim legitimacy, providing it is used throughout. It is not the case here, as all three strategies are applied, in different proportions. We provisionally consider all of them accurate. We return to the issue in the final table (table A) where their respective frequency is assessed.

³⁰ We quote the terms as they appear in the list. Translation of French or Zulu terms into English, given between round brackets, is ours. As French terms may also have a plurality of English equivalents and no indication appears, we used our discretion, taking obviously into account the Zulu term, to decide which was intended.

- Accurate lemmas:
 - verbs under their bases:
 - with hyphen: **-azisa**, ‘informer’ (to inform);³¹ **-vusa**, ‘relancer’ (start again);³² **-zwa** ‘se sentir’ (feel like);³³
 - without hyphen: **vula**, ‘ouvrir’ (open); **vusa**, ‘remuer’ (move, stir)
 - verbs under infinitive: **ukopha**, “saigner” (to bleed);³⁴ **ukuzola**, “se calmer” (to get calm);
- Inaccurate lemmas:
 - Conjugated verbal forms or constructions:
 - √ **ubilise**, ‘bouillir’ (boil); **ubilise**, as in **ubilise uphuthu**, can be a recent past “you (or she, he) boiled the pap (maize meal)”;³⁵ it can also be a subjunctive, “boil the pap”³⁶
 - √ **-ziyakubonisa**, ‘signifier’ (signify); **ziyakubonisa**, litt. “they (cl. 8 or 10, eg. **Izintombi** the girls or **izindoda**, the men) show you”, is a present affirmative. It is a complete phrase that needs no hyphen.
 - √ **uchama** ‘uriner’ (urinate); **uchama** may be a short present, but then needs a complement: **uchama kakhulu**, *you piss a lot (!)*; **uchama** as a lemma is incorrect
 - √ **sihloniphe** ‘honorer’ (respect, honor); **sihloniphe** is a subjunctive, with imperative value; -si- indexes the 1st pers. pl., in object or subject function: ‘honor or respect us’; ‘let us honor or respect’, ‘let us be respectful’
 - Mis-categorizations: terms neither verbs nor verbal:
 - √ **ukhula**, ‘désheber’ (to weed): **ukhula** is a cl. 11 noun, meaning ‘weed’; there is a verb **-khula**, duly listed as such, meaning ‘to grow’.
 - √ **udokotela** ‘falsifier’ (falsify); **udokotela** is a noun, “medical doctor” and does appear as such; the meaning ‘falsify’ probably comes from an extrapolation of Eng. “to doctor”, but this extension of meaning (and transfer of verbal category) is absent in Zulu.

ii. Adj., i. e., adjectives

Adjectives are arguably a controversial category in Zulu, to the point that there occurred a debate as to whether it should be maintained (discussed in (Lafon and Mongezi 2025)). Doke confined them to terms which take “nominal” concords, of which they are only a few, while those which take “verbal” concords he labelled relatives.³⁷ No such discussion here. Kasahorow’s adjectives include, besides some of the acknowledged adjectives and Doke’s relatives, verbal forms, nouns and nominal constructions, most of which cannot by any standard be construed as adjectives even though they often do qualify nouns.

³¹ Same entry also erroneously translated by a noun, ‘valeur’ (value).

³² **-vusa** is usually associated with ‘wake [someone] up’; it also appears without hyphen with a different meaning (below).

³³ On **-zwa** see also under ‘verbal forms that cannot pass for adjectives’ further below

³⁴ **Ukopha** also appears correctly as a noun, translated by “haemorrhage”

³⁵ **ubilise** is the short or conjunct form of the recent past (as opposed to full form **ubilisile**), constrained by the presence of an external object, itself part of the same phrase.

³⁶ Intonation would discriminate in speech.

³⁷ (Lafon and Mongezi 2021, 49–50) group both (sub)categories under an all-inclusive adjectival category, further divided into nominal and verbal adjectives

- Verbal forms
- verbal forms that may pass for adjectives

√ **-ahlukene**, ‘different’. **abantu abahlukene**, different people

√ **thembekile** ‘fiable’ (reliable); **abantu abathembekile**, reliable people

Both **-ahlukene** and **thembekile** constitute the long forms of the recent past of **-ahlukana**, ‘differ’ and **-thembaka**, ‘be reliable’, respectively; the recent past has often an inchoative meaning, sometimes being dubbed resultative and such forms may indeed pass for adjectives as long as they are hyphenated.³⁵

- verbal forms that cannot pass for adjectives

√ **-gulayo**, ‘être malade, malade’ (to be sick; sick); **-gulayo** is the relative present of **-gula**, “be sick”; a literal translation would read “[who/which] is sick”; the final **-yo** is the relative suffix, hence it cannot be accepted as an adjective.

√ **Ukwethusa**, ‘effrayant’ (scary); **Ukwethusa** is the infinitive of **-ethusa**; its inclusion under adjectives is all the more surprising as the neutral or stative (not transitive) **ukwethuka** is correctly labelled verb.

√ **abangasebenzi**, ‘chômeur’ (jobless, unemployed); **abangasebenzi** ‘those who do not work’ is a negative relative present, presuming a cl. 2 noun such as **abantu**, ‘people’; it could also be analysed as nominalized.³⁸

√ **engathandeki**, ‘indésirable’³⁹ (d*); **engathandeki** is a negative relative present, with a cl. 1 (or 9) concordial agreement, “[person, man] who is not likeable”;

√ **akuthandeki**, ‘laid’ (ugly); **akuthandeki** is a negative present, with a cl. 15 (infinitive) or 17 (impersonal) subject, ‘it is not likeable’;⁴⁰ ugly is context translation, from which the category has been deducted.

√ **vula**, ‘ouvrir’ (open); **vula** is repeated and labelled a verb a line above; possibly a mere typo.

√ **sondela** ‘avance’ (go forward);

vula and **sondela** are either imperative (as in the French translation for **sondela**) or stand for verbal bases (as in the French translation for **vula**). In no way are they adjectives.

√ **-zwa kahle**, ‘bien’ (good); if **kahle** does mean ‘good’ as an adverb (built on adjective **-hle**), **-zwa** is a full verb, with a variety of meaning ‘feel, hear, understand’: **ngizwa kahle**, “I hear well”; **angizwa kahle**, “I don’t understand”; **ngiyazizwa kahle**, “I am feeling well”; one fails to see how the clause was reduced to an adjective.

- nouns
- nouns in connective constructions

Zulu, like many Bantu languages, uses a connector [a], labelled by (Cope 1984, 256) “A of relationship which fuses with the class concords”, to link two nouns. This connector merges with the noun initial vowel (i. e., the augment or preprefix, see above and below). The construction is qualificative but it is a far stretch to label it ‘adjective’.

³⁸ At the very least, the plural (it is a cl. 2 agreement) should be indicated in the French translation ‘chômeurs’.

³⁹ Sic for ‘indésirable’.

⁴⁰ Cl. 17, limited to concordial agreement markers, is not formally distinguishable in Zulu from cl. 15, to the point that some grammarians argue for their merging, as (Creissels 2011, 45) did for Tswana.

√ **-yeshumi nambili**, ‘douzième’ (twelfth, “ten + two”); **-wamashumi amathathu**, erroneously translated as ‘treizième’ (thirteenth): **amashumi amathathu** means in fact 30 (tens three); the two entries show an unexplained class difference: 12th suggests an agreement in cl. 9 as in **[imoto] yeshumi nembili**, the 12th [car]; 30th an agreement in cl. 1 (or 3):⁴¹ **[umuntu] wamashumi amathathu**, the 30th [person]. It is not clear also why only those two ordinal numbers appear.

○ Self-standing nouns

√ **ishumi**, ‘dix’ (ten); **ishumi** is a cl. 5 noun making its plural in cl. 6: **amashumi**; it appears in higher numerals: **ishumi nambili** ‘douze’ (twelve) (see above), **ishumi nantathu**, ‘treize’ (thirteen); **amashumi amathathu**, 30, etc. By all means **ishumi** is not an adjective.

√ **-yingozi**, ‘dangereux’ (dangerous); cl. 9 noun **ingozi** means ‘danger’; prefixed with **yi-** it becomes a « copulative » in Doke’s term (Clement Martyn Doke 1992, 215)⁴²: **yingozi**, “it is a danger”, hence, “it is dangerous” from where, presumably, the labelling as adjective is derived.

√ **-phezulu**, ‘maximum’; **phezulu** is probably best treated as an adverb – and it does appear as one, glossed ‘en haut’ (above). **phezulu** is a frozen locative noun (cl. 16), related historically to cl. 5 **izulu** ‘sky’. To label it adjective is incorrect. The meaning ‘maximum’ may occur in context, but ‘above’ is the overall meaning.

iii. *Adv., i. e., adverbs*

Adverbs do exist in Zulu, even if there is no overall agreement on their definition and scope. One undisputed feature though is that adverbs are stable (do not take concordial agreement). We find in the dictionary a medley of terms, few (if any) of which fitting the feature. Some are verbal imperatives.

√ **-zinsuku zonke**, ‘quotidiennement’ (daily occurrence); **izinsuku**, days, is a cl. 10 noun. **izinsuku zonke** means all days, every day. In a connective construction: **umsebenzi wezinsuku zonke**, ‘work of every day’, hence the meaning ‘daily’ and the (far-fetched) labeling as adverb.

√ **siza**, ‘s’il vous plait’ (please); **-siza** is a verb, ‘help’; in the imperative, ‘help!’ may be translated by ‘please!’ although **ngiyacela**, ‘I am asking [you]’, would be closer. In no way can it be considered an adverb.

√ **sheshisa**, ‘rapidement’; **sheshisa** is the causative of **-shesha**, “do fast”; **sheshisa** as an imperative means “hurry!”.

iv. *Nom., i. e. nouns*

We have seen that Zulu dictionaries lemmatize nouns either with their full prefix, or, after Doke, under their stem or root. One could also consider removing the initial vowel, although we have yet to encounter such strategy. We do find it here though, alongside the more conventional strategy of lemmatizing the whole noun, that is, with its full prefix. We consider as semi-accurate nouns without the initial vowel. This inconsistency nullifies any trappings of a chosen lemmatization process. It applies to both indigenous or borrowed nouns. Some nouns

⁴¹ cl. 3: **[umuthi] wamashumi amathathu**, the 30th [tree]

⁴² We prefer to call this element “presentative”.

are lemmatized with the initial locative marker |e-|, with no indication whatsoever of being locativized. We however consider them as accurate.

- nouns lemmatized without augment
- indigenous:

Khanda, ‘tête’ (head) instead of **ikhanda**.

Sifungo, ‘serment’ (promise) instead of **isifungo**

Phutha, ‘erreur’ (mistake), instead of **iphutha**

khasi, ‘page’ (d°), instead of **ikhasi**.

- borrowed

phayinaphuli, ‘ananas’ (pineapple), instead of **uphayinaphu**;

phepha, ‘papier’ (paper) instead of **iphepha**

sulumane, ‘musulman’ (muslim) : instead of **iSulumane** ; **Sulumane**, ‘Islam’ instead of **inkolo yamaSulumani**

Somalia, ‘Somalie’ (Somalia): instead of **uSomalia**

Eritrea, ‘Erythrée’: instead of **u-Erithrea**⁴³

- nouns lemmatized with their augment
- indigenous

izinyo, ‘dent’ (tooth); **isiqu**, ‘diplôme’ (degree);

uhlazo, ‘disgrace’ (shame)

- borrowed

imoto, ‘véhicule’ (vehicle)

ukhumini, ‘cumin’ (caraway seed)

In many such nouns, the augment is wrongly separated from the lexical item: **i-dinosaure**, ‘dinosaure’ (dinosaur), **i-ukraine**, ‘Ukraine’; etc.⁴⁴

We have mentioned that all nouns in Zulu must have an augment. And indeed it is a strategy to enter otherwise foreign terms in the language. The dictionary makes extensive use of it, by listing several pages of terms, many technical, **ijiyometry** [sic for *ijiyometri*], ‘géométrie’ [sic for *géométrie*], **ikhiyubhu**, ‘cube’, **ilebhu**, ‘labo’ [sic for *laboratoire* (laboratory)], etc. In the absence of any frequency count to decide on entries, one has to rely on likelihood and importance: these words have arguably a very feeble likelihood to appear in any general intercourse, as they belong to specialized vocabulary.

- Locatives

Some nouns are lemmatized with the initial locative marker |e-|, which goes with or without the -|ini| suffix,⁴⁵ the use of the locative is not reflected in the translation.

⁴³ Possibly the initial e- has been taken as the initial locative marker that applies to countries’ name (see below and table B).

⁴⁴ For the orthographic use of The hyphen see note 24 above.

⁴⁵ The distribution of the suffix is mostly lexical, some nouns refusing it while most require it; no clear distribution pattern has emerged.

emhlabathini, ‘terre’ (earth) ; **umhlabathi** means ‘sand, earth’; **emhlabathini** means “in, at, from, the sand”; it is also a given place name.

empumalanga, ‘est’ (East) ; **impumalanga** means East ; **eMpumalanga** is often heard as the name of a province to the East of Gauteng in phrases such as **ngihlala/ ngiya eMpumalanga**, “I live in/ go to Mpumalanga”

We find an inordinate number of countries’ and cities’ names, a feature somewhat unexpected in such a small list of Zulu terms:

eAccra; **e-Eshiya**,⁴⁶ **Asie** (Asia); **eCameroon**, **eCongo**, **eCote d’Ivoire**, **eGhana**, **eKenya**, **eLandani**, **eNigeria**, **Etiopia**, **eTogo**

- terms dubbed nouns which are not properly nouns
- nominalized constructions

omnyama ‘ebene’ (sic for *ébène*) (ebony):⁴⁷ **omnyama** is the cl. 1 “relative” agreement of adj. **-mnyama**, “black”; as such it qualifies a noun: “[the person] who is black, the black person”; it can stand on its own, “the black [person]”; however a nominalized adjective is not a noun.

owesifazane, femelle (female); **owesifazane** is a complex construction, which involve the cl. 1 relative marker (o-) followed by the connective a-, the lexical item being cl.7 **isifazane**, which Doke translates as ‘women-folk’. It is a nominalized form rather than a noun.

- Verbs

A not insignificant number of verbal bases (or imperatives) are labelled nouns, disregarding their possible entries as verbs as well: **Khetha**, ‘choix’ (choice) ; **thimula** ‘éternuement’ (sneeze); **shesha**, ‘vitesse’ (speed); **-hambisana nomculo**, ‘musical’ (musical); **-hambisana [nomculo]** is a reciprocal built on a causative verb, that can translate as ‘go along [the music]’; ‘musical’ may be reached in context but the phrase is no noun. We also found conjugated verbs: **siyakuhalalisela**, ‘félicitations’ (congratulation); **siyakuhalalisela** means “we congratulate you”. The inconsistency is clear enough.

v. *Det., viz., determiners*

We found at least two entries labelled det., which we refer for their oddity.

Ø det, ‘la/le; Ø det., un/une

Zulu has no explicit article, be it definite or indefinite;⁴⁸ Doke’s sophisticated English to Zulu dictionary specifies under entry **a** (English indefinite article) “no equivalent” whereas it spells out a full clause with an object infix for the definite, under **the**, a strategy repeated in Dent. Doke indeed argues that the cooccurrence of an object infix and the noun, is the closest to definite article in Zulu (Clement Martyn Doke 1992, 299).

The Kasahorow entry suggests a mechanical mapping of French grammar onto Zulu, regardless of the grammatical structure of the later.

⁴⁶ Rather questionable in Zulu

⁴⁷ For a discussion on the meaning, see below.

⁴⁸ The language naturally possesses ways to express definiteness but they fall well outside the scope of this paper.

vi. *Pro., viz., pronouns*

Personal (autonomous) pronouns are correctly lemmatized; some translations are however problematic:

mina, 'je', 'moi', 'moi-même' (I, me, myself); **wena**, 'tu', 'toi' (you, *you*); **thina**, 'nous', 'nous' (we, us); **nina**, 'vous', 'vous', 'toi' (you pl., you pl., you sg.): no major reproach here. Each lemma is repeated at least two times, for the simple and emphatic forms. In the case of **nina**, you pl., it is also translated by the sing.: we suspect a (misplaced) retroaction from French which uses a polite plural for a single person.

yena, 'sien', 'elle' (his/hers, she); **bona**, 'leur', 'elles/ils' (their, they). For reasons we cannot fathom, the 3rd person is translated *also* as a possessive. Still the possessives **akhe** 'sa/son' (his/her) as well as **abo** 'leur' (their) both appear.

Some so-called pronouns are in fact phrases:

akekho, 'aucun' (no one): **akekho** is a cl. 1 negative predicative construction meaning "he/she is not here".

owesilisa, 'il', 'lui' (he, him); **owesifazane**, 'elle' (she): those are complex constructions, see above under **owesifazane**, and refer to male and female, rather than pronouns.

owethu, 'les notres' (ours): this is a nominalized possessive construction, that cannot pass for a pronoun.

vii. *Pre., viz., prepositions*

Few words fall under prepositions. We spotted it for three frozen locative constructions, two of which implying a complement, as can be deduced from the final hyphen:

phakhathi kuka-, 'en' (in); **phakhathi kwa-** 'parmi' (among)

phambi kwa-, 'avant' (before)

phansi kwa, 'sous' (under)

The use of "pre" is not consistent: similar structures appear labelled adv.:

phansi kwa- 'en dessous de' (below sthg)

phezulu, 'en haut' (above)

"Pre" appears also for the possessive marker specific to cl. 1a nouns **ka-**:

ka, 'de' (of); this refers to such phrase as **indlu kababa**, "my father's house", where **ka-** is used in lieu of the connective **-a-** due to **ubaba** belonging to cl. 1a.

However the two other examples of **ka-** construction that we came across are ascribed to different speech parts as **ka-** is made part of the lexical item, dubbed adjective for one and pronoun for the other:

adj.: **kagesi**, 'électrique' (electrical): **ugesi** in a cl.1a noun meaning electricity, altogether absent from the Zulu list

pro.: **kabani**, 'dont' (whose): **ubani** is cl. 1a interrogative noun who?; **ubani** is given as pronoun meaning who (not interrogative).

viii. *Exc., viz., exclamatives*

Not many such labels.

-lungile, ‘d’accord’ (agreed, ok); **-lungile** is the recent past (resultative) of **-lunga**, ‘be correct’;⁴⁹ it normally needs a subject index, a frequent occurrence being with cl. 17 which carries an indefinite value: **kulungile**, ‘it is ok’, a construction sometimes abridged in speech to **lungile**, hence the meaning. However it can also be used with a person index: **sebelungile**, ‘they are now well’. It seems the labeling as exclamative is derived from the rather approximate translation.

ah, no translation offered. ‘ah’, as an interjection, is common in French; implying a sigh, it would rather correspond to non standard Zulu **eish**

ix. Questionable lemmas

The presence of some entries is to say the least puzzling, their relevance in an altogether short Zulu dictionary defeating any rationale

- West African entities

Some lemmas are presumably proper names, from outside Zulu or even South African culture, in all likelihood from West Africa, where Kasahorow originates.

Kente; Takoradi; Yaw, Waache: no translation or gloss offered; **shekere** ‘Shekere’; **oware kwaba**, ‘oware’

Jolofu, ‘jollof’: refers to a West African rice recipe

- unexpected realities:

i-dinosaure, ‘dinosaure’ (dinosaur), **i-ukraine**, ‘ukraine’; **i-tarantula**, ‘tarantule’ (tarantula); **i-vodka**, ‘vodka’; **i-hiplife**, ‘hiplife’⁵⁰; **i-harmattan** ‘Harmattan’ (harmattan)⁵¹; **i-bhotela le-shea**, ‘beurre de karité’ (shea-butter)⁵²;

Thai, ‘thai’ (thai): **thai** belies Zulu morphology and spelling rules: **umthayi** could possibly refer to a Thai individual. The same applies to **Haiti**, ‘Haiti’ (Haiti);

Spin-phezulu, ‘toupie’ (top, spinning top); why should spinning top be included? In any case the makeshift coinage is not genuine in Zulu.

We have already mentioned the unexpected presence of names of countries not so often alluded to in South Africa (see above, nouns with locatives). Paradoxically, South Africa itself, Great-Britain or the US, arguably frequent occurrences in general parlance not to mention France in a Zulu/French dictionary, some of which do have a Zulu coinage, are omitted.

⁴⁹ On past tense, see above under **ubilise** and attending note

⁵⁰ Hiplife is by all means not a French term. In South Africa, it refers to hip-hop musical genre that bears connections with Ghanaian hiplife. Clearly an unverified assumption of shared terminology (from https://www.google.com/search?q=hiplife+music+south+africa&client=safari&hs=nwMp&sca_esv=17b8a313ff376b22&rls=en&ei=coeUaZ-jfLU3hbIPiZy1iAI&biw=1137&bih=755&oq=hiplife+music+south%C2%A0&gs_lp=Egxnd3Mtd2l6LXNlcjAIFWhecGxpZmUgbXVzaWMgc291dGjCoCoCCAAyBRAhGKABMgUQIRigAUjMSFCuBFiINnAAeAOQAQCYAbYKoAGAK6oBCTQtMy4xLjEuMrgBACgBAPgBAZgCCaACqy7CAgQQABhHwgIHEAAyGAQYE8ICBAAGBMYFhgewgIEECEYFZgDAIgGAZAGCJIHDTEuMS40LTMuMS4xLjKgB7oYsgcJNC0zLjEuMS4y uAfdLMIHCzItMi40LjEuMS4xyAeQAoAIAA&scient=gws-wiz-serp).

⁵¹ Harmattan, the name of a strong wind in West Africa, has entered French as spoken in West Africa; its relevance here is questionable.

⁵² Shea-butter is widely used in West Africa.

x. *Non-Zulu words*

Some lemmas are just non-Zulu words, nor even 'possible' words; besides French and English, we found Finnish and Portuguese words disguised as Zulu entries, without, for the English words, any attempt at adaptation. How did such words end up in a French-Zulu dictionary? one can imagine a contamination with another list being worked on simultaneously.

French: **imbécil**, 'imbecile' (idiot): French imbécile constitutes the Zulu entry, being translated by itself.

English: **accounting**, 'comptabilité' (accounting).

Finnish words: ⁵³ **Sivistynyt**, 'civilisé' (civilized), Fin. educated person; **vakooja**, 'espionner' (to spy), Fin. to spy, or a secret agent; **Selkä**, 'colonne vertébrale' (column), Fin. the back (of the body)

Portuguese words: ⁵⁴ **adivinar**, 'deviner' (to guess); **continuar**, proceder [sic for proceder] (to proceed); **cooperativa**, coopérative, (cooperative); **cordero** agneau 'lamb'; discípulo^{disciple (disciple)}; **descender**, 'descendre' (go down); **espátula**, 'spatule' (spatula); **hacha** 'hâche' (saw); **hacha** is not a Portuguese word; probably an attempt to coin a Portuguese word from the French ... ⁵⁵

xi. *Translations*

Translations are not always accurate; some are pure nonsense.

Osebenza ebhange, 'banquier' (banker) : **osebenza ebhange** is a relative construction meaning someone who works at a bank, regardless of the position, whereas 'banquier' (banker) implies someone in a measure of control of the bank. Minor offence here.

umuntu 'être' (? being, to be); 'être' here stands for 'être humain' (human being), but 'être' on its own, especially in a dictionary context, would refer to the verb to be. Minor offence here again.

Emuva-kwekhanda (sic), ⁵⁶ 'nuque' (nape of the neck); the Zulu word for this part of the body is **isiphundu**. No reason to opt for a description.

omnyama 'ebene' (see above): Fr. ébène refers to an African hard wood, and colloquially, bois d'ébène (ebony wood) to slaves, as Africans taken for enslavement were infamously called during the times of the Atlantic slave trade. No such tradition is attested in South Africa.

Sithandwa, 'voyant' (seer): **isithandwa** is a loved one; we see not how one can arrive at 'seer'.

Omese, 'silex' (silex): **omese** is the cl. 2a pl. of **umese**, knife; knives are indeed contusive instruments and so are silexes, but that does not suffice to blend the two, unless we validate syllogisms.

bheka-off [sic], 'se défaire' (get freed from, get loose); **-bheka** means 'look, observe' and is absent herein; **-beka** (listed as **beka**) means 'put, place'; clearly the confusion between **beka** and **bheka** has been aggravated by the addition of the English suffix -off to

⁵³ Finnish was identified thanks to a Wikipedia search.

⁵⁴ Own data

⁵⁵ The three c-commencing entries appear in sequence, which suggests there was a mix-up with a Portuguese list...

⁵⁶ The hyphen is out of context

express the distancing, probably inspired from break off, ... This construction is unheard of in Zulu.

Empumalanga-yesikhathi,⁵⁷ 'Timor Oriental' (East Timor); **Empumalanga-yesikhathi** would literally mean 'East of the time'; we suspect Timor was unduly segmented and mapped onto English word time, which led to the translation as East Timor.

Imoto, 'gari': **imoto**, a legitimate Zulu noun meaning 'car', is correctly translated as 'voiture' (car), 'véhicule' (vehicle) but also as 'gari', which is Swahili for car.

iklabishi 'Caraïbes' (Carib or Caribbean); **iklabishi** refers to cabbage and this meaning is given; probably a confusion between the two English words, mapped onto the French translation...

Umpe kwakungu, 'singe' (monkey); **umpe** is translated by 'juice' or 'nectar'; the phrase is at the very least incomplete and the relation with 'monkey' uncertain. The common term for monkey being **inkawu**.

What to Make of the Kasahorow Dictionary?

The errors fall into three main categories: miscategorisations, incorrect citation forms, and wrong meanings. Note that those types are not mutually exclusive.

Mis-categorizations, that is, words allocated to a part of speech they don't belong to, are in abundance and this review focused on this issue. This is no minor drawback in what depicts itself as a dictionary, when such a work is supposed to inform generally on the targeted language, beyond meanings, to give indication on how to use a word. It seems that many isiZulu words were categorized according to their French "equivalents". This is obviously not acceptable in this day and age, when bilingual dictionaries imply "contrastive lexicography" (Benzehra 2007, 31, 41), where lexicographers decide, from their knowledge of the languages at stake, which word-categories may be equalized or otherwise.⁵⁸ This implies familiarity with the language's lexicographical practices as well as its morphology. Mis-categorizations go largely hand in hand with incorrect lemmatizations, that is, a wrong quotation form for an item. We summarized these drawbacks in tables B & C below. We also have items not belonging to isiZulu, about which there is no need to comment further. More crucially in a dictionary, we spotted a few wrong translations.

The overall impression suggests that the dictionary was completed with minimal expert human intervention (if at all).⁵⁹ Indeed, in a private correspondence, the publishers did mention that experts for all pairs of languages required by their affiliates to be linked, were not always at hand, hence one had to make do with what resources they could muster. Recourse to machine translation is therefore most likely, as strongly suggested by a few examples commented upon such as **ukhula**, 'désheber' (to weed) construed as a verb, or **udokotela** 'falsifier' (falsify), where the whole English scope of the word has been falsely attributed to Zulu, or **Empumalanga-yesikhathi** which ends up as East Timor. We checked the entries analysed in

⁵⁷ **empumalanga** : see above under noun.

⁵⁸ Google translate, that was probably used to search for translations (below), is no recourse here as it does not offer information on parts of speech.

⁵⁹ The presence of a computer-generated square in front of each entry in the final product supports the point.

google translate as the most common system available,⁶⁰ taken either from Zulu or from French: some correspond literally, some do so with slight adaptation. Hereunder what we obtained⁶¹

Entries identical

Kabani > dont (whose)
Ukhumini > cumin
déssherber > **ukhula**
chômeur > **abangasebenzi**
indésirable⁶² > **engathandeki**
Ukraine > **i-Ukraine**⁶³
félicitation > **siyakuhalalisela**
phezulu > en haut
dinosaur > **i-dinsaure**
tarantula > **i-tarantula**
vodka > **i-vodka**
harmattan > **i-harmattan**
beurre de karité > **i-bhotela le-shea**

Entries adapted

Kagesi > d'électricité : Kas. électrique
Dangereux > **kuyingozi** : Kas. -yingozi
Maximum > **esiphezulu** : Kas. -phezulu
Owesifazane > une femme (a woman) : Kas. 'femelle' (female)
Quotidiennement > **nsuku zonke** : Kas. -zinsukuzonke
fiable > **okuthembekile** : Kas. -thembekile
malade > **abagulayo** : Kas. -gulayo
douzième > **kweshumi nambili** : Kas. yeshumi nambili
owethu > la notre; Kas. 'les notres'

While the reviewers are in no position to evaluate whether more specific or sophisticated programs could have offered a better-informed outcome, we would contend that a dictionary requires, and definitely required at the time of its completion, expert human intervention throughout its various stages.⁶⁴ Corpus extraction and data compilation can easily, even

⁶⁰ Visited January 2026. The version of google translate that would have been resorted to, predates 2014 when the dictionary was first published. It is certainly prior to Google Neural Machine Translation (GNMT), introduced in 2016

(<https://search.brave.com/search?q=history+of+google+translate+inaccurate+or+bizarre+translations&source=web&summary=1&conversation=08bd44388dff49cfc330ae496075bca86461>).

⁶¹ We have to admit that google translate was often more to the point than what is found here (but see note above).

⁶² Google does not skip the accent (é)

⁶³ Kasahorow has no capital.

⁶⁴ No-one can fully anticipate what AI in the future may come out with in terms of bilingual dictionaries and we acknowledged that AI comments helped us improve this paper. It seems however that, in spite of considerable progress, even GNMT is marred by "inaccurate or bizarre translations" (<https://search.brave.com/search?q=history+of+google+translate+inaccurate+or+bizarre+translations>)

preferably, be carried out electronically [for Zulu see (de Schryver and Prinsloo 2000)] but this automatism needs to be contained within parameters and these parameters can only come from expert decision based on the desired outcome and applicable lexicographic principles, some of which language-specific. In other words, the dictionary needs to be carefully planned in its various stages, and its outcome carefully spelt out, all the more so when it is based on machine compilation of data. De Schryver's Zulu dictionary is an example in point where, even though the data was compiled through search engines and automated frequency counts, a large team was constituted which worked several years to attain the result, under guidance from field experts, as exposed in (de Schryver 2010). Words were retained if they reached a certain frequency threshold, decided upon by the lexicographers.⁶⁵ Specific papers deal with the lemmatization strategy for problematic categories: adjectives, ideophones, pronouns, etc. (de Schryver 2008a) (de Schryver 2008b), (de Schryver 2009). Only when machine's role is naught or kept to a bare minimum, can prior explicit planning be possibly overlooked. That was naturally the case for yesteryear dictionaries; even if arguably not ideal it still applies to some of today's artisanal products, as the outcome is *de facto* constantly monitored by the lexicographers while progressing in their work.

The argument of the lack of individual experts familiar with both languages, which obviously would be a recurring situation when any pair of languages is considered, can easily be overcome by a team of at least two individuals, each familiar with one of the languages at stake, in the very least in respect of its structure and lexicographic conventions and issues. Since team work does not imply any more temporal and spatial simultaneity, this is easily attainable and no excuse in our view can be accepted for not resorting to it when completing such dictionaries.

More generally, the very term of dictionary, even word-for-word, for the piece under review is questionable, if by dictionary one means more than a list of words. In any language the vast majority of words possess several meanings, depending *inter alia* in their context of use, especially collocations, etc. A bilingual dictionary cannot be content to provide a word-to-word translation (or attempt thereof). As (Harouni 2004, 77) remarks "bilingual lexicography should pay more attention to the polysemic nature of lexical units"; he goes on to suggest that entries should contain several meanings duly ordered, a far cry from what is found here. On this line this piece would qualify at most as a tentative bilingual vocabulary, rather than word-for-word dictionary.

As a conclusion to this rather critical review, while acknowledging Kasahorow's initiative we would insist that without adequate expert involvement, such type of lexicographical work risks producing inaccurate and misleading, even sub-standard products. This would ultimately end up, through inappropriate handling of words and the inclusion of non-existing ones, as a disservice to users and to the language itself.⁶⁶

[ns&source=web&summary=1&conversation=08bd44388dff49cfc330ae496075bca86461](#)), visited January 2026. This suggests that human final checking remains of paramount importance.

⁶⁵ We make the case in our introduction to the 2nd edition of the French-Zulu lexicon, forthcoming, that, for African languages at least, frequency should not be the ultimate criteria. For the dictionary to depict the language fully, culturally relevant terms should appear, regardless of their corpus frequency. Also (Lafon and Mongezi 2025)

⁶⁶ After all, one could contend, if no human being monitors the final outcome, why expect that humans will eventually resort to use it at all?

Annex: Tables of Entries' Distribution in Kasahorow's Word Categories

Taking entries that fall under the main labels (act., aka verbs, adjectives and nouns), we examine their adequacy with accepted lexicographical conventions and practices for Zulu, as delineated in the paper, what we call accuracy. Accuracy is understood as full adequacy between a quotation form and its allocated category. We also consider consistency, viz., labelling items similarly throughout one single piece of work. The first table (table A) considers the whole Zulu to French section for all terms labelled act., aka verbs, which are therefore ignored by the following tables B & C. Those two tables are restricted to the hyphenated section in its entirety plus letters **a** through to **f**, which is slightly wider in size than the first section; for ease of comparison, those last two tables part entries according to the presence of an hyphen.

There are obviously degrees between plain accuracy and plain inaccuracy, and this is illustrated by some of the few forms discussed in the body of the paper, hence their qualification herein as semi-accurate. It is not feasible in a two-columns table to discriminate precisely according to a specific discrepancy. The borderline column contains those entries which fall in-between, leaning on either side (accuracy or inaccuracy). We remain conscious that there is a measure of subjectivity involved here; still arguably this emerges from the very lack of accuracy and consistency.

Some clarifications on the building of the tables are in order.

If the same item appears in both hyphenated and not-hyphenated sections (like **-vala** and **vala** both as act.), it is counted under each column (table A). Same applies to items differing only by the hyphen and the label (**-abantu**, adj. 'human', and **abantu**, noun, 'humanity'). However, if, within the same section, one entry is repeated due to several meanings, it appears only once (like **-sebenzisa**, 'utiliser' (use) and **-sebenzisa**, 'dépenser' (spend)); similarly, extended forms of verbs co-existing with the neutral form are counted only once under the same label (like **-enza**, 'agir' (act), and **-enziwa**, 'constituer' (constitute)); likewise for adjectives and their diminutives.

A few examples to help understand our accuracy judgment, especially for what regards 'borderline' examples:

Inaccurate adjectives: connective or copulative constructions; relative verbal forms with relative suffix; unhyphenated resultative verb forms as adjectives are variable terms; verbs:

-abantu, 'human'. **abantu** is a noun placed in a connective construction as in **izwe labantu**, the country of the people;

-yingozi 'dangerous'; **yingozi** is a copulative construction « it is the/ a danger, it is dangerous »;

thembekile 'fiable' (reliable);

-enzeka, 'possible': this is the verb **-enza**, do, with the stative extension

-gulayo, 'who/which is sick'

Borderline adjectives: resultative verb forms if hyphenated:

-ahlukene, 'different'

Entries non-existent in the language, such as foreign words, and some nonsensical words, have been excluded from further analysis (lists in footnotes).

A. Verbs aka act. (whole Zulu to French dictionary)

It has been seen that Zulu lexicography illustrates the three possible quotation forms for verbs (see above): hyphenated verbal bases, not-hyphenated verbal bases and not-hyphenated infinitives. All are deemed accurate. Inaccurate lemmas include: conjugated verbal forms which can be complete as **-ziyakubonisa**, given as ‘signify’, which means ‘they show you’, or incomplete, hyphenated infinitives, nouns and nominal clauses (connective, associatives, etc). Inaccurate forms represent close to a quarter of total entries labelled act. (399), which is significant, but the main failure in our view consists however in the inconsistency in the quotation forms, not acceptable for one single piece of work.

	Hyphenated	Non-Hyphenated	total	
accurate				
Verbal bases	78	196	274	
Infinitives		47	47	
			321	321
inaccurate				
Conjugated forms	2	30	32	
Infinitives	2		2	
Nouns	4	30	34	
Adjectives	3	1	4	
Other	3	3	6	
			78	78
				399

B. Hyphenated entries

	accurate	inaccurate	Borderline	Excluded	total
Adj.	24	67	16	1 ⁶⁷	108
Nom.		2			2

Nouns included here are inaccurate by nature as they should not start with an hyphen.

C. Non-Hyphenated entries (sample : letters a-, b-, c-, d-, e-, f-)

	accurate	inaccurate	Border line	Excluded	total
Adj.	1	24	1	1 ⁶⁸	27
Nom.	53	36	4	22 ⁶⁹	

⁶⁷ -**akhazngiweyo**, ‘desséché’ (withered, dry)

⁶⁸ **Atlantic**,

⁶⁹ **accounting, alergia, Ama, Apodidae, aprendiz, ayoyo kwaba, azonto, blog, bungalow, cedi, chojambula, cooperative, cordero, Deutsch, dikhagoni, dimi, discipulo, dodi, donut, e-cilantro, eid, espátula, ezimanzi-inkathi, fufu.**

All adjectives included here are inaccurate by nature as they should start with an hyphen; however we accepted adjectival stems with no concord as they are only minimally wrong (**bomvu**, 'rouge' (red)).

Nouns starting with the initial prefix consonant, rather than the initial augment vowel, are classified as borderline ; nouns without augment are inaccurate (**Afrika**, 'Afrique'; **baba**, 'père' (father)); nouns with locative marker e-, are accepted as accurate; nouns not following Zulu spelling are considered borderline (e-Accra, eCote d'Ivoire (Ivory Coast), etc.

Acknowledgment

Successive versions of the paper were run through ChatGPT and Deepseek; both AI provided useful comments and suggestions for amendments, thus increasing its scope and relevance, and we are clearly indebted to them. We also benefitted from comments made by an anonymous reviewer. In view of the aforesaid, it may be worth adding that we refrained from running the piece reviewed here through an AI, lest we found ourself outperformed and redundant And be discouraged to pursue.

References

- Alexandre, P. (1981). Les langues bantu. In J. Perrot & G. Manessy (Eds.), *Les langues dans le monde ancien et moderne* (pp. 353–375). CNRS.
- Benzehra, R. (2007). Some implications of contrastive lexicology for bilingual lexicography and teaching translation (with special reference to English-Arabic). *Revue des Sciences Humaines*, 18(1), 31–43. <https://asjp.cerist.dz/en/article/3664>
- Bryant, A. 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*. Marianhill Mission Press. <https://archive.org/details/zuluenglishdicti00brya/mode/2up>
- Canonici, N. N. (1990). Subclasses of Zulu nouns. *South African Journal of African Languages*, 10(2), 52–58.
- Cope, A. T. (1984). *Zulu: A comprehensive course in the Zulu language* (Rev. ed.). University of Natal.
- Creissels, D. (2011). Tswana locatives and their status in the inversion construction. *Africana Linguistica*, 17, 33–52.
- Dent, G. R., & Nyembezi, C. L. S. (1962). *Compact Zulu dictionary: English-Zulu; Zulu-English* (2nd ed.). Shuter & Shooter.
- Doke, C. M. (1954). *The Southern Bantu languages*. Oxford University Press for the International African Institute.
- Doke, C. M. (1992). *Textbook of Zulu grammar* (6th ed.). Longmans Southern Africa.
- Doke, C. M., Malcolm, D. Mc K., Sikakana, J. M. A., & Vilakazi, B. W. (2014). *English-Zulu, Zulu-English dictionary*. Wits University Press. (Original work published 1958)
- Guthrie, M. (1967). *Comparative Bantu: An introduction to the comparative linguistics and prehistory of the Bantu languages*. Gregg.
- Harouni, Z. (2004). Contrastive lexicology, bilingual lexicography and translation. *Revue des Sciences Humaines*, 15(2), 73–80. <https://asjp.cerist.dz/en/article/90712>

- Lafon, M. (1994). The African Languages Lexical Project of African Languages and Literature of University of Zimbabwe (ALLEX). In Z. J. Duri (Ed.), *Highlights of the Research Programme at the University of Zimbabwe* (Vol. 4, pp. 23–41).
- Lafon, M. (1995). *Le Shona & les Shonas du Zimbabwe: Recueil d'informations sur la langue et la culture*. Mambo Press; L'Harmattan.
- Lafon, M., & Bolofo, M. (2021). *Manuel de conversation français-zoulou et zoulou-français*. Kom-Edit.
- Lafon, M., & Bolofo, M. (2022). *Lexique français-zoulou: Isichayamagama sesiZulu kuya kwisiFulentshi*. KomEdit.
- Lafon, M., & Bolofo, M. (2025). Purism or laxity? A lexicographer's dilemma in a diglossic situation, with isiZulu as an example. In H. E. Wolff & J. Roux (Eds.), *Contextualising African language dynamics of change* (pp. 219–233). African Sun Media.
<https://doi.org/10.52779/9781991260581>
- Lafon, M., Saadi, A., & Said Zaina, D. (1991). *Lexique français-comorien (shingazidja)*. L'Harmattan.
- Maphumulo, A. M. (Ed.). (2021). *Ukuvamisa imithetho yokubhala nobhalomagama lwesiZulu lonyaka wezi-2021*. University of KwaZulu-Natal Press.
- Mbatha, M. O. (2006). *Isichazamazwi sesiZulu*. New Dawn Publishers.
- Natal Missionary Conference. (1911). Report of the proceedings of the Zulu Orthography Conference 1907. In *Proceedings of the Zulu Orthography Conference 1907*. P. David & Sons.
- Nurse, D., & Philippson, G. (Eds.). (2003). *The Bantu languages*. Routledge.
- Nyembezi, C. L. S. (1992). *AZ: Isichazamazwi sanamuhla nangomuso*. Reach Out Publishers.
- de Schryver, G.-M. (2008a). A new way to lemmatize adjectives in a user-friendly Zulu-English dictionary. *Lexikos*, 18, 63–91.
- de Schryver, G.-M. (2008b). The lexicographic treatment of quantitative pronouns in Zulu. *Lexikos*, 18, 92–105.
- de Schryver, G.-M. (2009). The lexicographic treatment of ideophones in Zulu. *Lexikos*, 19, 34–54.
- de Schryver, G.-M. (2010). Revolutionizing Bantu lexicography: A Zulu case study. *Lexikos*, 20, 161–201.
- de Schryver, G.-M. (Ed.). (2015). *Funda isiNgisi. Learn isiZulu. isiZulu-isiNgisi, isiNgisi-isiZulu: Isichazamazwi sesikole = School dictionary* (2nd ed.). Oxford University Press.
- de Schryver, G.-M., & Prinsloo, D. J. (2000). Electronic corpora as a basis for the compilation of African-language dictionaries, Part 1: The macrostructure. *South African Journal of African Languages*, 20(4), 291–309.