

Exploring the Documentation and Preservation of African Indigenous Knowledge in a Digital Lexical Database*

Sonja E. Bosch, *Department of African Languages,
University of South Africa, Pretoria, South Africa (boschse@unisa.ac.za)*
and

Marissa Griesel, *Department of African Languages,
University of South Africa, Pretoria, South Africa (griesm@unisa.ac.za)*

Abstract: Transcending the boundaries of printed lexicographic resources is becoming easier in the digital age, with e-resources facilitating restrictions on the size and type of information that can be included. In this article we explore innovative ways of documenting and preserving African indigenous knowledge, often underrepresented in traditional dictionaries, in an existing digital lexical database. Our approach is based on the extension of the African Wordnet, a lexical database under construction for nine African languages, in this case applied to isiZulu. This article addresses the challenge of consolidating dispersed indigenous knowledge collected from a variety of sources such as conventional dictionaries, interdisciplinary publications and a flat-structured online database, in a digitised hierarchical wordnet structure. A representative sample of traditional domestic utensils in Zulu culture is used to demonstrate the conversion into a set of typical semantic relations in a wordnet structure. By focusing on filling lexical gaps between isiZulu and English as found in the Princeton WordNet, with culturally relevant synsets, the African Wordnet also becomes a useful resource for natural language processing. Finally, it is shown how the hierarchical classification of selected domestic utensils is visually presented in wordnet graphs in the Wordnet-Loom interface.

Keywords: AFRICAN WORDNET, DIGITAL LEXICAL DATABASE, INDIGENOUS KNOWLEDGE, ISIZULU, TRADITIONAL DOMESTIC UTENSILS, SEMANTIC RELATIONS, LEXICAL GAPS, LEXICALISATION

Opsomming: Ondersoek na die dokumentasie en bewaring van inheemse kennis uit Afrika in 'n digitale leksikale databasis. Die digitale era maak dit toenemend moontlik om die grense wat gedrukte leksikografiese hulpbronne stel te oorskry deurdat

* An introduction to this theme was read as a paper at the 24th International Conference of the African Association for Lexicography which was held in Windhoek, Namibia, 26–29 June 2019. A first attempt towards the inclusion of pottery beer vessels in the AfWN was presented in Bosch and Griesel (2019).

e-hulpbronne die beperkings op grootte en tipe inligting wat ingesluit kan word, fasiliteer. In hierdie artikel ondersoek ons innoverende maniere om inheemse kennis uit Afrika, wat dikwels onderverteengewordig is in tradisionele woordeboeke, in 'n bestaande digitale leksikale databasis te dokumenteer en te preserveer. Ons aanslag word op die uitbreiding van die African Wordnet, 'n leksikale databasis wat tans vir nege Afrika-tale ontwikkel word, gebaseer en in hierdie geval op Zoeloe van toepassing gemaak. Die artikel spreek die uitdagings wat die konsolidasie van inheemse kennis uit 'n verskeidenheid bronne soos konvensionele woordeboeke, interdisiplinêre publikasies en 'n aanlyn databasis met 'n plat struktuur in 'n digitale, hiërargiese woordnetstruktuur inhou. 'n Verteengewordigende uittreksel van tradisionele huishoudelike gereedskap in die Zoeloe kultuur word as voorbeeld gebruik om te demonstreer hoe konsepte tot 'n stel semantiese verhoudings in 'n woordnetstruktuur omgeskakel kan word. Deur daarop te fokus om die leksikale gapings tussen Zoeloe en Engels soos dit in die Princeton WordNet vervat is, met relevante kulturele konsepte te vul, word die African Wordnet ook 'n nuttige hulpbron vir natuurliketaalprosessering. Laastens word aangedui hoe die hiërargiese klassifikasie van geselekteerde huishoudelike gereedskap visueel in woordnetgrafieke in die WordnetLoom koppelvlak voorgestel kan word.

Sleutelwoorde: AFRICAN WORDNET, DIGITALE LEKSIKALE DATABASIS, INHEEMSE KENNIS, ZOELOE, TRADISIONELE HUISHOUELIKE GEREEDSKAP, SEMANTIESE VERHOUDINGS, LEKSIKALE GAPINGS, LEKSIKALISERING

1. Introduction

In an article on large-scale lexicography in the digital age, Fellbaum (2014: 378) states:

The Digital Revolution can be fairly said to have shifted the paradigm in lexicology and lexicography. It has opened up new ways of exploring and representing the structure of the lexicon, testing diverse theories of word semantics, and compiling both manually and automatically ever larger and richer resources that reflect multiple dimensions of meaning and lexical organization based on solid empirical data ...

Fellbaum (2014) continues by describing the impact that the digital revolution has had on the construction of lexical resources containing extensive information on aspects of word meaning that are not easily covered in traditional print dictionaries. Gouws et al. (2014: 12) support this statement: "Online accessible corpora and data banks, digitized text editions and electronic editions of older and new dictionaries today offer the lexicographer who depends on sound and comprehensive documentation an ideal working basis".

The most significant limitations that have negatively influenced traditional paper dictionaries are identified as size of databases, access to databases, and methods according to which resources are compiled. When compiling traditional paper dictionaries, lexicographers are often forced to exclude infrequently used concepts and focus on modern language usage, commonly

determined by examining frequency lists extracted from large corpora. This is mostly due to restrictions placed on them by publishers mindful of printing costs and practicality. Since electronic or digital databases are generally not adversely affected by space or size constraints as is the case with printed matter, they lend themselves ideally to the inclusion of additional data, such as indigenous knowledge (IK) that is often underrepresented in conventional dictionaries. Initiatives by the South African Government and the Department of Science and Innovation to bolster the development of digital language resources under less restrictive licenses such as the establishment of the South African Centre for Digital Language Resources (SADiLaR)¹, further promote the accessibility of data. The limitations can therefore be side-lined or overridden in the digital age.

The question now is whether any existing frameworks or digital database structures can be utilised effectively in this move to a more digitally accessible lexicographic working base. Such a framework would need to offer the capacity to include (digitised) data from printed dictionaries and in addition, be easily expandable with less frequently used and underrepresented concepts, bringing the traditional into the digital age. It also needs to be easily accessible for both qualitative and quantitative research or development while being an easy to use reference for language learners or students. Finding such a solution will also address the problem of archiving various forms of missing or dispersed IK in a more sustainable database. Souza et al. (2020: 946) stress that the collection of IK in paper archives and more recently in digital databases is imperative in preserving not only the language, but also traditional customs for posterity.

In this exploratory article we describe a novel application and subsequent expansion of an existing lexical resource for isiZulu which enhances and supplements lexical knowledge from conventional dictionaries and other interdisciplinary sources. We introduce the African Wordnet (AfWN), a prototypical lexical database consisting of words that are grouped into sets of synonyms called synsets, as framework for the digital documentation and preservation of indigenous or cultural knowledge. The opportunities that such a digital lexical database not constrained by size and access can offer with regard to the digital documentation and preservation of indigenous or cultural knowledge, are addressed.

In order to contextualise our study, we start by explaining in general what a wordnet is, including the macrostructure and microstructure of a wordnet such as the English Princeton WordNet (PWN). This is followed by an overview of the African Wordnet (AfWN), currently under active development, and a description of the wordnet editor implemented to provide visualisation features of wordnets. Next, we describe our approach to the design of the AfWN, with a focus on challenges such as lexicalisation and lexical gaps, particularly within the context of the African languages. This is followed by a demonstration of how a representative sample of traditional Zulu domestic utensils, gathered from an interdisciplinary variety of sources, can be transformed into a set

of relations in an electronic lexical resource. We then conclude and provide pointers for future work.

2. Aspects of wordnets

2.1 What is a wordnet?

McCrae et al. (2020: 37) maintain that wordnets have turned out to be one of the most popular types of dictionaries used in natural language processing (NLP) and other areas of language technologies. This can be ascribed mainly to their "structure as a graph of words, that is much easier for computers to understand than the traditional form of a dictionary". Wordnets are primarily built for use by machines in tasks such as automatic text analysis and for artificial intelligence applications. Word sense disambiguation (WSD) and information retrieval (IR), for instance, are performed much more effectively when the semantic relations in a wordnet are used to distinguish between the different meanings of a word in context. It is therefore not surprising that according to Calzolari (2018), wordnets for various languages were among the most cited resources during the Language Resource and Evaluation Conference (LREC) in 2018. Wordnets offer a wealth of information, contained in a machine-readable lexical database where words are grouped into synsets and linked by conceptual-semantic and lexical relations (Miller 1995). An online lexical database for a specific language is, furthermore, an invaluable reference resource for many research and application projects in the linguistics and lexicography domains. Kotzé (2008: 20) states: "[a wordnet is] extremely useful for its accessibility, quick reference and potential for serving as a base or support for other language technological and lexicographical applications".

Each synset in a wordnet is enhanced with lexical information such as the part of speech of the lemmas, a definition and usage example(s) of the concept. Furthermore, in addition to the synonymy relation linking different senses in a synset, other semantic relations between synsets are indicated as well. These relations include the super-subordinate or the hyperonymy, hyponymy relation, the part-whole or meronymy relation, as well as antonymy. Fellbaum (1998: 7) explains: "WordNet is a semantic dictionary that was designed as a network, partly because representing words and concepts as an interrelated system seems to be consistent with evidence for the way speakers organise their mental lexicons".

The first large-scale project to develop a monolingual wordnet was started in the 1990s with the Princeton University WordNet (PWN) for (American) English (Fellbaum 1998) and contains roughly 250 000 synsets for nouns, verbs, adjectives and adverbs. An example of a synset from the PWN can be seen in Figure 1. Note the hierarchical organisation of the synsets using SUMO/MILO classification (cf. SUMO 2002; as well as Niles and Pease 2001) and how this is

easily visualised in the WordnetLoom (Naskret et al. 2018) interface. This makes the PWN and any wordnet developed according to the same ontological structure easy to navigate, both manually and automatically in different digital humanities (DH) applications.

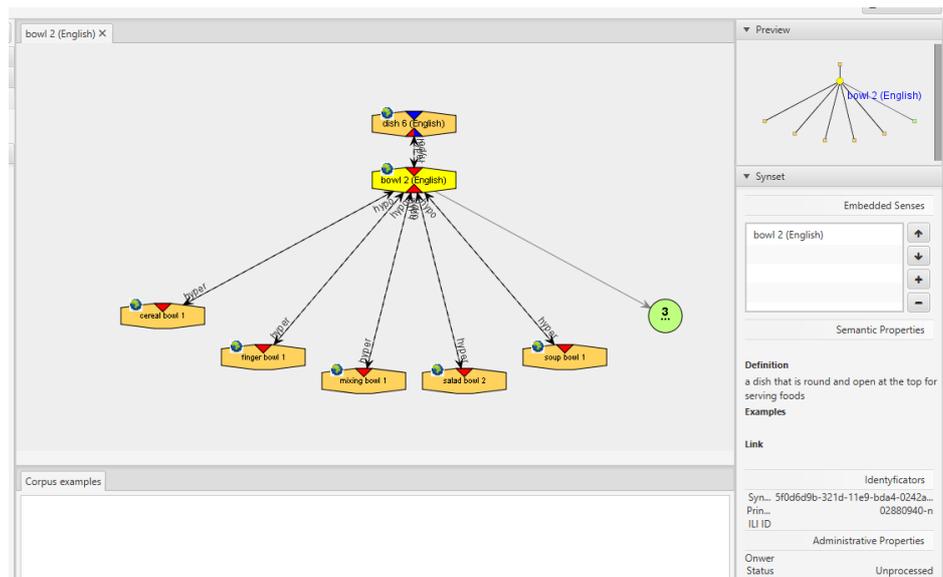


Figure 1: The noun synset 'bowl' in PWN² as visualised in WordnetLoom

The PWN serves as template for many development projects, such as the Hungarian wordnet (Vincze and Almási 2014), the Japanese wordnet (Bond et al. 2008) and the BalkaNet wordnets (Tufis et al. 2004). The hierarchical structure and semantic relations are kept largely unchanged and only the content of each synset, i.e. the lemmas, usage example and definition, is translated into the target language. This, of course, assumes that the target language shares an underlying semantic structure with English as it is captured in the PWN. Ordan and Wintner (2007) and Vossen et al. (2016) refer to this method as the *expand approach* and recommend it for projects with limited lexical resources. If these resources do exist, the so-called *merge approach* is often followed. The PolNet wordnet (Vetulani et al. 2010) was for instance derived from a high quality, monolingual Polish lexicon as base and subsequently aligned with the PWN. Similarly, wordnets are also aligned with the Multilingual Open Wordnet (Bond and Paik 2012) or Global WordNet Grid (Vossen et al. 2016).

2.2 Base concepts

The method by which to develop a wordnet for a new language does not only entail choosing between the expand and the merge approach as described above. Another very important consideration is also which concepts to include, especially in the initial stages of development, to ensure good coverage in the wordnet. To this end, the EuroWordNet and BalkaNet projects created the so called "core base concepts" (CBC) list — a list of seed terms extracted from corpora for various European languages involved in the two projects with which to kickstart wordnet development³. The CBC aims at including those high-level concepts that have many semantic relations with other synsets first, thereby guiding manual inclusion of further synsets in a top down approach. The AfWN also initially followed this internationally accepted method by including most of the CBC in the first version of the wordnets.

However, since the CBC incorporates many concepts that are not lexicalised in the African languages, Anderson et al. (2010: 3763) point out the main disadvantage of the particular approach, namely that "the fundamental WordNet base will be biased to those concepts that are not necessarily core in the new target language". They propose a hybrid approach to building wordnets for African languages, in keeping with the global focus that has "always been on concept hyponymy based on mother tongue speaker understanding" (Anderson et al. 2010: 3763).

Accordingly, the AfWN resorted to incorporating synsets from a more localised seed list — the SIL Comparative African Wordlist (SILCAWL), which was compiled in 2006 by Keith Snider (SIL International and Canada Institute of Linguistics) and James Roberts (SIL Chad and Université de N'Djaména). This list, covering 12 semantic categories, is an English–French bilingual list of lexical data consisting of 1 700 words with glosses, resulting from linguistic research in Africa. Inclusion of seed terms from this list has already resulted in the inclusion of numerous lexicalised concepts such as the elaborate kinship terms in isiZulu and Sesotho sa Leboa in the AfWN in an organised manner (cf. Griesel et al. 2019).

2.3 African Wordnet

The African Wordnet (AfWN) which is still under active development (see Bosch and Griesel 2017; and Griesel and Bosch 2020) aims at steadily growing the number of synsets in wordnets for nine indigenous South African languages. Over the past 10 years of development, the AfWN has grown slowly but consistently with a focus on manually verified, culturally appropriate data that meets international standards. Currently, the AfWN includes roughly 63 000 synsets across the nine languages with 27 000 definitions and 37 000 usage examples added to these synsets. The languages currently included in the project

are isiZulu (ZUL), isiXhosa (XHO), Setswana (TSN), Sesotho sa Leboa (NSO), Tshivenda (VEN), Siswati (SSW), Sesotho (SOT), isiNdebele (NDE) and Xitsonga (TSO). The multilingual wordnet project is a first for these South African languages.

Although there has been an increase in the number and quality of text resources freely available for the South African languages, a recent audit still showed large areas where little to no development has been done and the languages involved in the AfWN can still all be considered resource-scarce. As Moors et al. (2018: 2) conclude in their report on the most recent audit of available resources: "While significant progress has been made since 2009 to develop additional resources across more languages, to develop cutting edge resources, and to develop language independent resources, the more marginalised indigenous languages (particularly Xitsonga, Tshivenda, Sesotho, siSwati, and isiNdebele), remain severely under resourced". This resource scarceness has been one of the main challenges in the development of wordnets and the reason why the *expand approach* was followed, as described in the previous section. Neale (2018) outlines the expand approach as starting point for wordnet construction when few additional resources exist and describes various methods by which automatic extraction of the basic information needed to form a synset can speed up development. However, the basic linguistic and lexicographic resources he describes such as an electronic bilingual dictionary, or even a monolingual lexicon in the target language are not yet freely available for most of the South African languages in the AfWN. This not only makes automatic extraction of synsets impossible, but necessitates a more labor intensive manual process (see Griesel and Bosch (2020) for a discussion of the various methods employed to make the most of the limited resources that do exist for some languages).

The AfWN development team will continue to add more synsets to this base, but we are now ready to delve deeper into language specific questions relating to smaller groups of synsets in order to broaden the coverage and make the AfWN useful not only for language processing applications, but also for human interpretation in the lexicographic process, for instance. An initial experiment into employing the AfWN data in further application saw data from the AfWN used to start populating Kamusi GOLD⁴, a multilingual online dictionary. Furthermore, a free and open mobile dictionary app with an underlying data structure that is not only open for dictionaries, but supports the connection to external resources like the AfWN as well, is also under development (Eckart et al. 2020).

2.4 Editing tool

A wordnet by nature depends on the interconnected relations between concepts as formalised in synsets. Visualising the relations and effortlessly adding

to the network of connections is a crucial aspect of manual development, even more so when working on a multilingual project. At the onset of the current development phase the AfWN was ported to WordnetLoom (Naskret et al. 2018), a freely available, customisable wordnet editor with advanced wordnet visualisation features. The editor has the capability of organising large networks of semantic relations and serves as a browser and development interface. The visual nature of the interface is easy to work in and facilitates visualisation of the connections between concepts. It is envisioned that this tool will also form the basis for a browsing platform to make the AfWN openly accessible and easily searchable as a web service, thereby eliminating the need for specialised software or installation by an expert.

3. Design strategy

Our approach to the design of the AfWN takes place against the background of the wordnet being a source of reference that takes the traditional dictionary to a whole new level, as described by Abubakar et al. (2019). While a dictionary organises words in alphabetical order and can offer information such as meaning, synonyms, parts of speech (POS) and so forth, a wordnet offers the additional feature of synsets, a set of synonyms (synsets) for open word classes, that is nouns, verbs, adjectives and adverbs. Linkage to synsets is provided by various semantic relations such as hypernyms, hyponyms, meronyms, troponyms and antonyms. In the description of our design in this article, we focus on nouns in particular.

In the PWN, which is the point of referral in the expand method described above, the super-subordinate relation (also known as hyperonymy, hyponymy) is the most frequently encoded relation among synsets. This relation links universal synsets like {utensil} to gradually more specific ones like {ceramic ware; pottery; clayware; funnel; server, kitchen utensil etc.}, as can be gleaned from Figure 2. This is an advantage of the expand method since upper levels provide general guidelines, so that the set of relations (hyponyms) can be increased to fill the lexical gaps with indigenous or cultural knowledge concepts of the target language. The vocabulary of a language can be divided into two categories according to Batibo (2016: 135), namely basic vocabulary and cultural vocabulary. He describes basic vocabulary as "the lexical stock which is basic in all human languages. It denotes objects and phenomena that are found universally" while cultural vocabulary is "the lexical stock that a linguistic community develops or adopts through its many cultural experiences after interacting with its physical environment, social milieu and the supernatural world". An aspect of the AfWN that needs to be explored further is the addition of such "cultural" word senses that cannot be linked directly to the PWN. African languages and cultures include many unique word senses that are not easily matched to the core set of meanings in the PWN or for that matter in other wordnets. Such

concepts are not only those that appear in paper dictionaries but also those that may be documented elsewhere and that would be lost to future generations if not preserved digitally and made accessible, ideally in an organised manner.

The mere translation of English concepts contained in the PWN into isiZulu, is not a complete reflection or representation of Zulu cultural knowledge and concepts, especially not on the lower levels of the hierarchy. Therefore, in this study, we take the first steps in creating an isiZulu lexical database that addresses lexicalisation differences. "Lexicalisation differences" are here defined as a) those instances where the source and target languages, here English and isiZulu respectively, lexicalise the same concept with a different kind of lexical unit, be it a word, compound or collocation; and b) those instances where one of the two languages has no lexicalisation for a concept at all and results in a lexical unit in either the source or target language being translated with a description of the concept as a phrase. In the latter case, we therefore have a so-called lexical gap, which Bentivogli and Pianta (2000: 664) define as follows:

A lexical gap occurs whenever a language expresses a concept with a lexical unit whereas the other language expresses the same concept with a free combination of words.

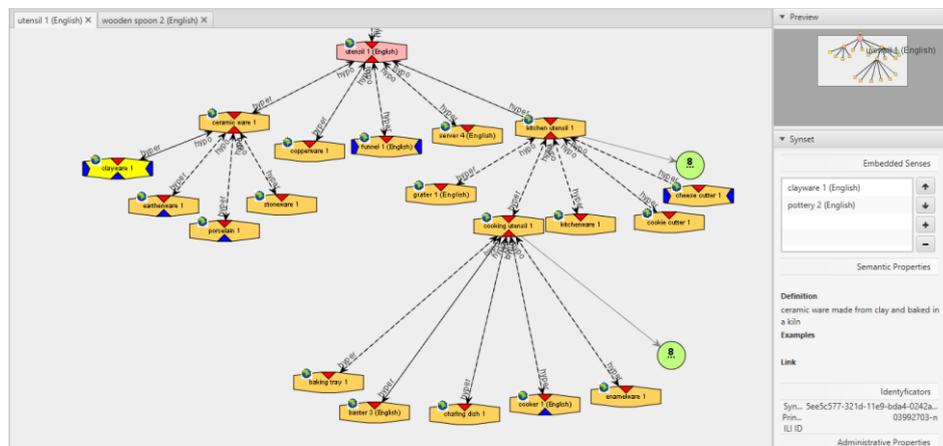


Figure 2: The noun synset 'utensil' in PWN as visualised in WordnetLoom

To fill such gaps, the SIL Comparative African Wordlist (SILCAWL) (Snider and Roberts 2006) is used as benchmark. The items are organised semantically on a continuum, from items relating to human domains at the one extreme, via animate domains, to items relating to non-human domains at the other extreme, and then from concrete items to more abstract items. The twelve main headings are listed in Table 1.

1.	Man's physical being
2.	Man's non-physical being
3.	Persons
4.	Personal interaction
5.	Human civilisation
6.	Animals
7.	Plants
8.	Environment
9.	Events and actions
10.	Quality
11.	Quantity
12.	Grammatical items

Table 1: Headings in the SILCAWL

The headings in Table 1 are subsequently sub-divided into second and third level headings. For example, in the case of *Human civilisation*, the following first level headings are distinguished: SETTLEMENT, CLOTHING AND ADORNMENT OF BODY, FOOD AND DRINK, FOOD PREPARATION, DOMESTIC UTENSILS AND CONTAINMENT, HABITATION, PROFESSIONS AND WORK, and so on. A third level, for instance, in the case of DOMESTIC UTENSILS AND CONTAINMENT includes divisions such as: kitchen utensils, eating utensils and containers, and containment. The parts of speech covered in the SILCAWL are nouns, verbs, adjectives, adverbs, pronouns, interrogatives and conjunctions. Although Snider and Roberts (2006: 4), the compilers of the SILCAWL, concede that they still notice "imperfections and room for improvement (e.g. words that could be deleted, words that could be added, words that could be moved to different semantic domains etc.)", the SILCAWL has proven to be an opportune progression from the CBC list used in the past in the development of the AfWN. The most significant improvement is observed against the background of localisation where the content (of the entries) would be lexicalised within an African environment.

Although the so-called SILCAWL is used as a guide to fill lexical gaps in the African Wordnet, it becomes clear when looking at a fragment of DOMESTIC UTENSILS AND CONTAINMENT in Table 2, that language specific detail needs to be addressed. For instance, in Zulu culture a lexical and conceptual distinction is made between a three-legged cooking pot and a small flat-bottomed cooking pot. Similarly, a lexical and conceptual distinction is made between a wooden spoon for eating or one for serving and stirring, or between

a calabash milk vessel and a calabash beer vessel. These concepts are all lexicalised in isiZulu and will be dealt with in more detail in the next section.

5.5 DOMESTIC UTENSILS AND CONTAINMENT

5.5.1 Kitchen utensils

0610 cooking pot (earthenware)

...

0613 ladle

...

0618 pestle, pounding stick pilon

0619 mortar, pounding pot

5.5.2 Eating utensils

0620 plate

0621 bowl

0622 cup (French: gobelet, coupe, tasse > goblet, cup)

0623 spoon (traditional)

5.5.3 Containers and containment

...

0626 basket

0627 bucket, pail

0628 calabash

Table 2: Fragment of domestic utensils and containment in the SILCAWL

A typical lexical gap that emerges in the wordnet of the source language (PWN), and which can be filled by indigenous knowledge information, is illustrated in Figure 3. Apart from the six hyponyms for "plate", viz. dessert plate, dinner plate, paper plate, salad plate/bowl, soup plate and steel plate, the isiZulu wordnet requires an additional hyponym, namely "earthenware children's plate", translated as *isikhangezo*.

The image shows a screenshot of a WordNet search interface. At the top, it says "WordNet Search - 3.1" with links to "WordNet home page", "Glossary", and "Help". Below this is a search bar with the word "plate" entered and a "Search WordNet" button. The results are listed under the heading "Noun". The first entry is "S: (n) home plate, home base, home, slab where the batter stands; it must b". The second entry is "S: (n) plate (a sheet of metal or wood". The third entry is "S: (n) plate (a full page illustration (usually on slick paper))". The fourth entry is "S: (n) plate (dish on which food is served or from which food is eaten)", which is highlighted with a green box. This entry has sub-entries: "direct hyponym / full hyponym" including "S: (n) dessert plate (a small plate on which dessert can be served)", "S: (n) dinner plate (a plate from which a diner eats during the main course of a meal)", "S: (n) paper plate (a disposable plate made of cardboard)", "S: (n) salad plate, salad bowl (a plate or bowl for individual servings of salad)", "S: (n) soup plate (a deep plate with a wide rim)", and "S: (n) steel plate (a plate of steel)". It also has "direct hypernym / inherited hypernym / sister term" including "S: (n) plate, plateful (the quantity contained in a plate)". The fifth entry is "S: (n) plate, crustal plate (a rigid layer of the Earth's crust that is believed to drift slowly)". A green callout box points to the second entry, stating "Lexical gap: No earthenware plate or children's plate in the Princeton Wordnet".

Figure 3: Example of a lexical gap in the PWN⁵

4. Data and presentation

Cosijn et al. (2002) describe IK as local knowledge that is unique to every culture or society, and that is rooted in community traditions, relationships and rituals. The African continent is rich with sources of IK and Ossai (2010: 2) summarises it as "an embodiment of different modes of thought and epistemology". He also notes the importance of preservation of African IK and points out that it forms part of the decision-making process at the local level for rural communities, involving all aspects of human life — agriculture, health care, food preparation, education, natural resource management, and others. It is therefore important that IK be preserved in such a way as to make it easily accessible to all, including taking advantage of new developments in the digital age, and easily expandable so as to ensure clusters of IK practitioners can add valuable details regarding the intricacies that make African IK so unique and expansive.

According to Cosijn et al. (2002: 94):

Using databases for the representation of IK may offer several advantages. Most importantly, access from a retrieval point of view is much easier in electronic database format than in paper or linear electronic text formats. Secondly, IK can be stored and delivered in multiple copies for those that need it. Furthermore, in

database format, it is possible to annotate IK in various ways from multiple viewpoints to facilitate its analysis. However, in order to realise these advantages, IK in databases must be made accessible.

In this section it will be shown by means of isiZulu examples, how IK concepts gathered from an interdisciplinary variety of sources can be transformed, in some cases from alphabetically ordered entries and in other cases from categorized lists, into a set of relations within the context of a hierarchical wordnet structure. Such a digital knowledge database has almost no physical restrictions and may incorporate, in addition to the conventional wordnet relations, namely synonyms, hypernyms, meronyms and so forth, typical wordnet features such as definitions, usage examples and even dialect information. The data was collected from a variety of sources. Given the fact that dictionaries are the most common resource used for building wordnets, we first of all consulted three authoritative monolingual and bilingual dictionaries (ISZ 2006; ISN 1992 and ZED 1964). To complement the data collected from the dictionaries, two anthropology publications (Krige 1965; and Fowler 2015), diverse cultural publications (Nyembezi and Nxumalo 1966; and Grossert 1985) and an online database, Comparative Bantu Pottery Vocabulary (CBPV) were studied as well.

For the purposes of this study, we will focus on the following aspects of the taxonomy of traditional Zulu domestic utensils, namely function, material (from which the utensil is crafted), size and shape. Krige (1965), Nyembezi and Nxumalo (1966) and Grossert (1985) itemise domestic utensils according to the material from which the utensils are crafted, while Fowler (2015) classifies earthenware or pottery utensils according to their function. Size and shape of the utensils are described arbitrarily in the various sources. In the following discussion we compare descriptions of selected domestic utensils in dictionaries and other resources, point out missing entries in dictionaries — also entries that are not main entries (and would therefore be difficult for users to find), and identify synonyms and meronyms where applicable.

4.1 Kitchen utensils

- **cooking pot** (earthenware)

ikhanzi (plural: *amakhanzi*)

Function	-	cooking
Material	-	pottery
Size	-	large
Shape	-	three-legged, spherical pot with a wide mouth, about 23 centimetres in diameter

In the dictionaries ZED (1964: 381) and ISN (1992: 224), as well as in Nyembezi and Nxumalo (1966: 18) it is stated that *ikhanzi* is a cooking pot. Only ZED (1964: 381) indicates the material as earthenware. ISN (1992: 224) and ISZ (2006: 554) add the additional information that *ikhanzi* is large and three-legged. The shape is

gleaned from other resources, viz. Fowler (2015: 97) who describes it as oval-shaped (accompanied by an illustration), and CBPV as a pot with a wide mouth. CBPV also provides detailed information about size — about 9 inches (23 centimetres) in diameter. This utensil does not feature in Grossert (1985).

isoco (plural: *izoco*)

Function	-	cooking (cereals or vegetables)
Material	-	pottery
Size	-	small
Shape	-	bowl

The dictionary sources ISZ (2006: 936), ISN (1992: 388) and ZED (1964: 634) clearly state that *isoco* functions as an earthenware cooking vessel, while ISZ (2006: 936) and ZED (1964: 634) also comment on the size, characterising it as "small". This is confirmed by Fowler (2015: 98) who adds that it is used specifically for cooking cereals or vegetables, and provides an illustration. This utensil does not feature in CBPV.

isiyoco (plural: *iziyoco*)

Function	-	cooking (cereals or vegetables), frying
Material	-	pottery
Size	-	small
Shape	-	saucepan

The kitchen utensil *isiyoco* does not feature in any of the dictionaries consulted (i.e. ZED, ISN and ISZ), however, it appears in Nyembezi and Nxumalo (1966: 20) where it is described as a frying pan; while Krige (1957: 397) adds the details of size "little saucepan used for cooking, smaller than the ikhanzi". The illustration in Fowler (2015: 98) shows the comparative sizes of *isiyoco* and *isoco*, the former being the smaller of the two cooking pots. This utensil features neither in Grossert (1985) nor in CBPV.

- **ladle**

isikhetho (plural: *izikhetho*)

Function	-	beer-skimmer
Material	-	ilala palm leaves
Size	-	small, flat, curved handle
Shape	-	spoon-like
Meronym:	<i>isibambo</i>	flat and curved handle

Both ZED (1964: 391) and ISN (1992: 527) describe *isikhetho* as a plaited spoon-like utensil used for skimming beer. The former dictionary adds that the utensil

is plaited from palm leaves. It is noteworthy that *isikhetho* does not feature in ISZ (2006).

Krige (1965: 396) provides information on the size of the spoon, namely small, and also specifies the function as that of skimming the beer before it is drunk. Grossert (1985: 23) confirms the function as "When beer is served the *isikhetho* is taken to skim the scum from surface", and she also identifies the palm leaves as those of the ilala palm. The shape of the handle is described by Grossert (1985: 41) as flat and curved.

A meronym or part of the whole is observed in this kitchen utensil, namely a handle *isibambo*. The handle is also made of ilala palm leaves.

imvokoqo (plural: *izimvokoqo*)

Function	-	serving
Material	-	wood
Size	-	large
Shape	-	hollowed out, deep bowl

This utensil is unanimously described by ZED (1964: 837), ISZ (2006: 726) and ISN (1992: 527) as a ladle with a deep bowl. Nyembezi and Nxumalo (1966: 17) concur and add that it is made of wood and used for serving. The serving function is confirmed by ISN (1992: 527). Although the description by Krige (1965: 398) seems to be similar to the ones above, namely a deep-ladled wooden spoon, the term used – *umvokoqa* – is not the same and will therefore not be taken into consideration. This utensil does not appear in Grossert (1985).

4.2 Eating utensils

- bowl

umcakulo (plural: *imicakulo*)

Function	-	eating
Material	-	pottery
Size	-	small (15-20 cm in diameter)
Shape	-	small pot with a wide mouth, shaped like a pudding bowl/basin

Synonym: *umshengele*

All dictionaries consulted, as well as the additional resources, describe *umcakulo* as a small earthenware bowl used for eating. ZED (1964: 100) adds details of the shape, namely "shaped like pudding basin" and "wide mouthed". Grossert (1985: 36) provides details on the size, namely a diameter of 15 to 20 cm. Fowler (2015: 101) adds the specific dishes that are served in these bowls, which include *uphuthu* (mealie meal porridge), *amahewu* (fermented maize porridge drink) and *umdokwe* (porridge).

Nyembezi and Nxumalo (1966: 19); ZED (1964: 737); ISZ (2006: 1102) and ISN (1992: 463) identify *umshengele* as synonym for *umcakulo*. This utensil does not feature in CBPV.

imbenge (plural: *izimbenge*)

- | | | |
|----------|---|------------------------------|
| Function | - | eating |
| Material | - | grass |
| Size | - | small (15-20 cm in diameter) |
| Shape | - | saucer-shaped bowl |

Synonyms: *umhelo*
unyazi

According to Krige (1965: 395) *imbenge* is a basket of woven grass used for serving food, specified by Grossert (1985: 17) as cooked maize, millet or pulses, and by ISN (1992: 295) as boiled maize and sorghum. In ZED (1964: 73) only the size (small) and the material of *imbenge* are described, while Nyembezi and Nxumalo (1966: 21) also describe the shape of the bowl, namely broad and shallow with a wide mouth. Grossert (1985: 17) pays more attention to exact shape and size, namely saucer-shaped and approximately 15 to 20 centimetres in diameter.

It is interesting to note that ISZ (2006: 686) does not include the function of using *imbenge* to serve food, but rather the alternative function of covering food or beer — "*isitsha esakhiwe ngotshani sokwemboza ukhamba noma ukudla*" (a grass bowl for covering food or beer).

Nyembezi and Nxumalo (1966: 21) and Krige (1965: 395) indicate *ingcazi* as synonym, although three of the dictionaries describe *ingcazi* as a different type of utensil. ZED (1964: 550) defines it as a narrow-necked water-pot and not as an eating bowl, while ISZ (2006: 1102) and ISN (1992: 330) define it as a large earthenware pot with a long small mouth, also known as *uphiso*. Until further clarity has been gained, *ingcazi* will not be included as a synonym in the isiZulu wordnet. However, ZED (1964: 73) offers *umhelo* as synonym, even if not as a main entry. ISZ (2006: 919) and ISN (1992: 382) indicate *unyazi* as synonym of *imbenge*.

- **plate**

isikhangezo (plural: *izikhangezo*)

- | | | |
|----------|---|--------------------------------------|
| Function | - | feeding children soft foods, liquids |
| Material | - | pottery |
| Size | - | small (size of a saucer) |
| Shape | - | shallow, shape of a saucer |

Synonym: *umgenqele*

Most sources agree that *isikhangezo* is a small earthenware plate used for feeding children soft foods such as sour milk (ISZ 2006: 552; ZED 1964: 380; ISN 1992: 223; Nyembezi and Nxumalo 1966: 20; and Krige 1965: 397). The shape of the plate is specified by Krige (1965: 397) and ZED (1964:380) as being that of a saucer.

Fowler (2015: 100) is the only source offering a synonym *umgenqele*. This utensil does not feature in CBPV.

- **spoon (traditional)**

ukhezo (plural: *izinkezo*)

- Function - serving
- Material - wood
- Size - average length 30 cm
- Shape - spoon

Meronym: *isibambo*
long handle

It is stated by all sources, except ISZ (2006), that *ukhezo* is a wooden spoon (Krige 1965: 398; Nyembezi and Nxumalo 1966: 17; ZED 1964: 392; and ISN 1992: 231). Grossert (1985: 40-41) gives an average length of 30 cm, and emphasizes the balance between the length and thickness of the handle compared to the size and shape of the bowl of the spoon. Nyembezi and Nxumalo (1966: 17) add additional information to the function, namely a wooden spoon for eating sour milk or other liquid food such as pumpkin porridge.

A meronym or part of the whole is noted in this utensil, namely a handle *isibambo*. It is a long wooden handle as described by Grossert (1985: 40-41).

isixembe (plural: *izixembe*)

- Function - serving and stirring
- Material - wood
- Size - large
- Shape - ladle (broad and flat)

Synonyms: *ingxwembe*
isixwembe
ingxembe
isixembu

The dictionaries ZED (1964: 860) and ISN (1992: 231) indicate the material from which *isixembe* is crafted as wood, while ISN (1992: 231) and ISZ (2006: 1285) indicate the size as large. ISN (1992: 231) adds the function as that of serving. Nyembezi and Nxumalo (1966: 18) supply a more complete description, namely that of a wooden stirring spoon with a broad flat shape in front.

A number of synonyms are put forward by the sources consulted. *ingxwembe* appears in Nyembezi and Nxumalo (1966: 18) and ZED (1964: 860) although not as a main entry in the latter dictionary; *isixwembe* appears in Krige (1965: 398), ISZ (2006: 1285) and ISN (1992: 545); *ingxembe* features in ISN (1992: 545) and *isixembu* appears in ISZ (2006: 1285). Even though Krige (1965: 398) records only *isixwembe*, she describes it as a "large, broad, wooden spoon for serving food or eating *ububendi* (blood and mince)".

4.3 Containers and containment

- calabash

igula (plural: *amagula*)

Function	-	container for sour milk
Material	-	calabash
Size	-	not specified
Shape	-	short with wide body and narrow mouth

All sources consulted (Nyembezi and Nxumalo 1966: 21; Krige 1965: 397; ZED 1964: 274; ISZ 2006: 404; and ISN 1992: 156) agree that *igula* is a calabash vessel used for sour or curdled milk. The dictionaries ZED (1964: 274); ISZ (2006: 404) and ISN (1992: 156) provide additional information on function, namely a container for fermenting milk into sour milk; and on shape, that is, short with a wide body and a narrow mouth.

iphaphasi (plural: *amaphaphasi*)

Function	-	container for food
Material	-	calabash
Size	-	not specified
Shape	-	flat sides, wide, open-mouthed

Two of the dictionaries consulted (ISZ 2006: 957 and ISN 1992: 398), as well as Nyembezi and Nxumalo (1966: 21), all describe *iphaphasi* as a calabash container for food. Concerning the shape, ZED (1964: 648) describes it as a wide, open-mouthed calabash, ISN (1992: 398) describes it as a calabash with dented sides while Nyembezi and Nxumalo (1966: 21) point out that it is dented on the sides to form a flat object. Krige (1965: 397) confirms that *iphaphasi* is a flat calabash with flat sides and top is cut off.

This utensil does not feature in Grossert (1985).

isigubhu (plural: *izigubhu*)

Function	-	container for water, beer or fermented maize porridge
Material	-	calabash
Size	-	not specified
Shape	-	not specified

Although ZED (1964: 274) only describes *isigubhu* as a calabash or gourd, the other two dictionaries (ISZ 2006: 398 and ISN 1992: 60) as well as Nyembezi and Nxumalo (1966: 21), add the function, namely that of a container for water, beer or *amahewu* (fermented maize porridge). The latter specifically point out that it is not used as container for sour milk or food. In Krige (1965: 397) the orthography differs slightly (*isigubu*) but refers to the same type of vessel, namely a gourd or calabash used as water or beer container. Grossert (1985: 37) provides two illustrations of calabash containers which she incorrectly terms *isigubo*. There is no accompanying description.

It is interesting to note that the size and shape of the various calabashes are not specified, possibly because calabashes are natural objects that have different sizes.

- **basket**

isichumo (plural: *izichumo*)

- | | | |
|----------|---|---|
| Function | - | container for carrying beer |
| Material | - | ilala palm leaves |
| Size | - | 25-40 cm in diameter with cylindrical neck, about 7 cm across and 5 cm high |
| Shape | - | globular |

Meronym: *isimbozo*

lid or covering for beer container (*isichumo*) made of ilala palm leaves, fits over the neck like a cap.

The dictionary sources consulted all concur that *isichumo* depicts a container for carrying beer, and that it is crafted from ilala palm leaves (ZED 1964: 116; ISZ 2006: 151 and ISN 1992: 60). ZED (1964: 116) describes it as being bottle-shaped, while ISZ (2006: 151) and ISN (1992: 60) describe *isichumo* as having a long narrow mouth ("esinomlomo omude omncane"). The shape is expressed more accurately by Krige (1965: 395), namely globular. Grossert (1985: 19-20) adds the specific details that the spherical part of *isichumo* is 25 to 40 cm in diameter while the cylindrical neck is about 7 cm across and 5 cm high. ZED (1964: 116) as well as ISZ (2006: 151) add that the material of which this container is made, could also be calabash.

A part of the whole or a meronym is observed in this utensil, namely a lid *isimbozo* for covering the mouth of the container (Krige 1965: 395) and fitting over the neck like a cap Grossert (1985: 21). The lid is also made of ilala palm leaves.

isilulu (plural: *izilulu*)

- | | | |
|----------|---|--|
| Function | - | container for grain storage |
| Material | - | (plaited <i>imbubu</i>) grass |
| Size | - | large, 1 to 1,5 metres in diameter |
| Shape | - | pear shaped, broad base with small opening |

The three dictionaries consulted (ZED 1964: 467; ISN 1992: 398 and ISZ 2006: 670) agree that *isilulu* is a container for storing grain. ZED (1964: 467) and ISN (1992: 398) add that it is made of plaited grass and is large, while (ZED 1964: 467) also gives information about the shape of the container namely "the opening of which is small". Nyembezi and Nxumalo (1966: 22) concur that *isilulu* is a large container for storage of the harvest, made of grass. Krige (1965: 395) provides more detail regarding size, namely very large, with a diameter that can be up to 3 to 4 feet. The most detailed description is given by Grossert (1985: 21), who depicts *isilulu* as the largest of the Zulu baskets with a diameter of 1 to 1,5 metres. She describes it as "pear shaped rather than spherical, with a broad base ... woven with bunches of soft *imbubu* grass".

4.4 Proposed synsets

Proposed synsets to be included in the isiZulu wordnet, based on information for selected domestic utensils as extracted from various sources and analysed in detail above, are summarised in Table 3.

Type of domestic utensil	Definition	Usage example
<i>ikhanzi</i>	a large earthenware three-legged, spherical cooking pot with a wide mouth, about 23 cm in diameter <i>imbiza enkulu yokuphekela yebumba enemilenze emithathu, eyindilinga enomlomo obanzi, engamasentimitha angama-23 ububanzi</i>	<i>sebenzisa ikhanzi ukupheka isibindi nosu eziko</i>
<i>isoco</i>	a small earthenware cooking vessel used specifically for cooking cereals or vegetables <i>imbiza encane yokuphekela esetshenziselwa ukupheka ukudla okuyizinhlamvu noma imifino</i>	<i>letha isoco bakuphekele iklabishi</i>
<i>isiyoco</i>	a small earthenware saucepan, smaller than <i>isoco</i> , used for cooking or frying cereals or vegetables <i>isosipani lebumba elincane kunesoco, elisetshenziselwa ukupheka noma ukuthosa ukudla okuyizinhlamvu noma imifino</i>	<i>ugogo ucela isiyoco azopheka ngaso iklabishi nezindlubu</i>
<i>umcakulo</i> Synonym: <i>umshengele</i>	a small earthenware bowl with a diameter of 15-20 cm, used for serving <i>isitsha sebumba esincane esinedayametha eyi-15 kuye ku-20 cm, esisetshenziselwa ukuphakela</i>	<i>ugogo ucela umcakulo azobanikeza ngawo amahewu, umdokwe noma uphuthu</i>

<i>imbenge</i>	a small saucer-shaped bowl of woven grass, used for serving food	
Synonyms: <i>ingcazi</i> <i>umhelo</i> <i>unyazi</i> <i>isikhangezo</i>	<i>isitsha esincane esakheke okwesosi esakhiwe</i> <i>ngotshani obelukiwe esisetshenziselwa ukuphakela ukudla</i> a small, shallow earthenware plate used for feeding children soft food <i>ipuleti elincane, elibunjiwe elisetshenziselwa ukufunza izingane ukudla okuthambile</i>	<i>letha imbenge sizobekela obaba ukudla</i> <i>unina kaSipho ugeza isikhangezo sosana lwakhe</i>
<i>ukhezo</i>	a wooden serving spoon for eating sour milk or other liquid food <i>isipuni esibaziwe sokuphaka amasi noma okunye ukudla okuwuketshezi</i>	<i>ngiphakele amasi ngokhezo olusha</i>
Meronym: <i>isibambo</i>	the part of an object that is designed to be held in order to use it <i>ingxenye yento enzelwe ukubanjwa ukuze isetshenziswe</i>	<i>isibambo sokhezo sephukile</i>
<i>isixembe</i>	a large wooden serving and stirring spoon shaped broad and flat in front <i>isipuni esikhulu esibaziwe sokuphaka noma sokugovuza esivulekile futhi saba yisibaca phambili</i>	<i>udadewethu ugovuza isijingi ngesixembe</i>
Meronym: <i>isibambo</i>	the part of an object that is designed to be held in order to use it <i>ingxenye yento enzelwe ukubanjwa ukuze isetshenziswe</i>	<i>isibambo sesixembe side</i>
<i>igula</i>	a calabash container for sour milk <i>isitsha soselwa sokufaka amasi</i>	<i>igula likagogo elidala liphukile</i>
<i>isigubhu</i>	a calabash container for water, beer or amahewu (fermented maize porridge) <i>isiqukathi soselwa sokufaka amanzi, utshwala noma amahewu</i>	<i>uthenge isigubhu emakethe yezitsha zesintu</i>
<i>iphaphasi</i>	a flat, wide, open-mouthed calabash container for food <i>isiqukathi soselwa esiyisicaba, esibanzi nesinomlomo ovulekile sokufaka ukudla</i>	<i>bamlethele iphaphasi eligcwele izinkobe</i>

<p><i>isichumo</i></p>	<p>a spherical beer container with a narrow mouth made of lala palm leaves, and provided with a lid of the same material</p> <p><i>isitsha sotshwala esisandingilizi esinomlomo omncane esakhiwe ngelala futhi esinesivalo esakhiwe ngelala</i></p>	<p><i>thathela umkhulu isichumo sakhe ekhishini</i></p>
<p>Meronym: <i>isibozo</i></p>	<p>a lid or covering for a beer container made of the same material and which fits over the neck like a cap</p> <p><i>isivalo noma isembozo sesitsha sotshwala esakhiwe ngelala futhi esingena khaxa esichumweni njengesivalo</i></p>	<p><i>vala isichumo ngesimbozo kungaze kungene impukane</i></p>
<p><i>isilulu</i></p>	<p>a large container with a broad base and small opening at the top, plaited from grass for grain storage</p> <p><i>isitsha esikhulu esivulekile ngaphansi bese sivuleka kancane phezulu, esilukwe ngotshani, sokugcina ukudla okuyizinhlamvu</i></p>	<p><i>isilulu esidala siyavuza</i></p>
<p>Meronym: <i>imvokoqo/ imvongoqo</i></p>	<p>a large wooden serving spoon with a deep or hollow bowl</p> <p><i>isipuni sokuphaka esikhulu esibaziwe esinesitsha esinomgodi oshonayo noma esinomgodi ovulekile</i></p>	<p><i>ubaba ubaza imvokoqo</i></p>
<p>Meronym: <i>isibambo</i></p>	<p>the part of an object that is designed to be held in order to use it</p> <p><i>ingxenye yento enzelwe ukubanjwa ukuze isetshenziswe</i></p>	<p><i>isibambo semvokoqo side</i></p>
<p><i>isikhetho</i></p>	<p>a spoon made of plaited palm leaves to skim beer before it is drunk</p> <p><i>isipuni esakhiwe ngosundu sokuhluza utshwala ngaphambi kokuba buphuzwe</i></p>	<p><i>ukhumbule ukuhluzisa utshwala ngesikhetho ngaphambi kokuba uthethele abantu abadala</i></p>
<p>Meronym: <i>isibambo</i></p>	<p>the part of an object that is designed to be held in order to use it</p> <p><i>ingxenye yento enzelwe ukubanjwa ukuze isetshenziswe</i></p>	<p><i>isibambo sesikhetho siyisicaba futhi sigobile</i></p>

Table 3: Proposed isiZulu wordnet entries for a representative sample of traditional domestic utensils

When added to WordnetLoom, the hierarchical classification of kitchen utensils such as traditional (cooking) pots and ladles is visually presented in the wordnet graphs as shown in Figure 4. Attributes of the particular synset are displayed in the sub-panels on the right of the graph.

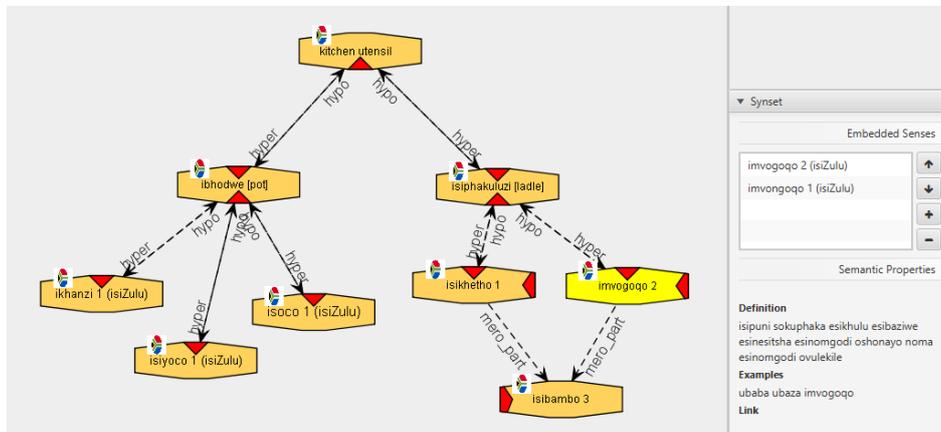


Figure 4: Kitchen utensils represented in WordnetLoom

The arrows in Figure 4 indicate the relevant semantic relations, namely that *ibhodwe* (pot) is in a hypernym relation to *ikhanzi*, *isiyoco* and *isoco* which are all types of cooking pots. *Imvongo* and *isikhetho* are types of ladles and are therefore hyponyms of *isiphakuluzi* (ladle). In the latter examples, a part-whole or meronymy relation exists with regard to *isibambo* (handle). A synonym of *imvongo* is indicated as *imvongo* in the sub-panel of attributes.

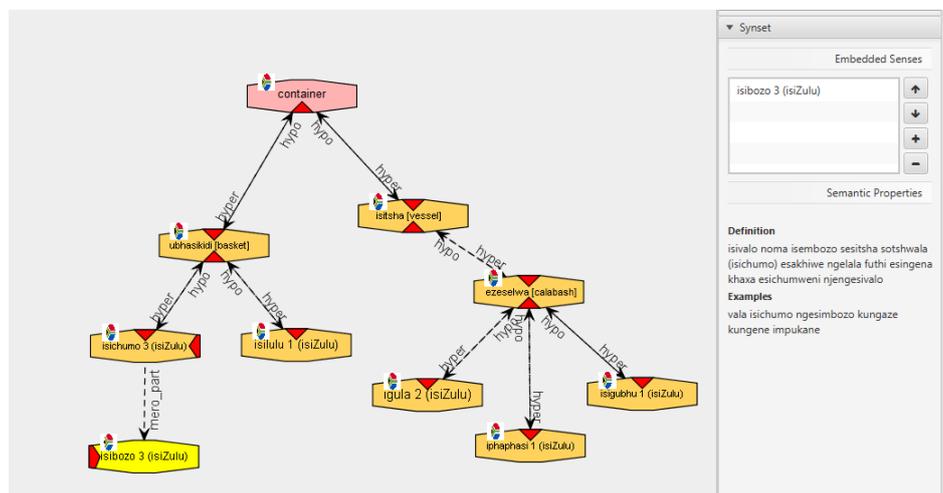


Figure 5: Containers represented in WordnetLoom

The semantic relations in the synsets in Figure 5 indicate for example that *ibhasikidi* (basket) is in a hypernym relation to *isichumo* and *isilulu*. In the case of *isichumo*, a part-whole or meronymy relation exists with *isibozo* (lid for beer container). With regard to the containers resorting under *isitsha* (vessel), it is clear that *igula*, *iphaphasi* and *isigubhu* are types of calabash vessels and are therefore hyponyms of *ezeselwa* (calabash). It is noteworthy that the lexicalized term of a calabash container is determined by the function of the vessel, that is, *igula* is a container for sour milk, *iphaphasi* is a container for water, beer or *amahewu* (fermented maize porridge drink) and *isigubhu* is a container for food.

5. Conclusion and future work

Our core contribution is twofold and ties in with envisioned future work:

Firstly, we describe a new lexical resource for isiZulu which enhances and supplements lexical knowledge from conventional dictionaries and other interdisciplinary sources, thereby addressing the problem of missing or dispersed indigenous knowledge that might be lost for posterity. The nature of the chosen lexical resource, that is the wordnet, furthermore lends itself perfectly to relating the IK to senses used more frequently in modern society, thereby bringing the traditional into the digital age. By focusing on filling lexical gaps between the African language isiZulu, and the English template as found in the PWN with culturally relevant synsets, the AfWN, in turn, becomes an even more useful resource for natural language processing. Digitisation of older manuscripts containing IK is gaining momentum and by including traditional senses in a modern resource, these texts can be automatically unlocked via information retrieval, machine translation and document categorisation. Regarding future work, we envisage expanding the AfWN into a multimodal database including images and sound clips with synsets. Declerck et al. (2020) describe a method by which to add speech data to the Open English Wordnet so that pronunciation information can be added to different senses in each synset. By doing so, they show improvements in disambiguation between different word meanings. This same method could, for instance, be used to add tonal information to the AfWN or, as an additional level, to preserve and document spoken IK such as stories or songs along with the related lexical entries.

Secondly, we lay the foundation for further discussion and development of a common scheme for storing lexical data not only for the South African Bantu languages, but for the Bantu language family as a whole. Schofield (1943), for instance gives a detailed description of the pottery types (with the relevant terminology) of the Nguni and Sotho language groups as well as Tshivenḁa, which could serve as a starting point. Furthermore, the style guides and protocols created for the development of the AfWN according to international best practice, can serve as example for other African languages.

Endnotes

1. See <https://www.sadilar.org>
2. Also available at <http://wordnetweb.princeton.edu/perl/webwn>
3. See a description of the core base concepts list at <http://globalwordnet.org/gwa-base-concepts/>
4. See <https://kamusi.org>
5. Original search results available at:
<http://wordnetweb.princeton.edu/perl/webwn?s=search>

Acknowledgements

The African Wordnet project (AfWN) was made possible with support from the South African Centre for Digital Language Resources (SADiLaR). SADiLaR (<https://www.sadilar.org>) is a research infrastructure established by the Department of Science and Innovation of the South African Government as part of the South African Research Infrastructure Roadmap (SARIR).

Contributions by Ms Lindelwa Mahonga and Ms Celimpilo Dladla, members of the isiZulu AfWN development team, are thankfully acknowledged, while the first author acknowledges the University of South Africa for granting her leave for research purposes.

References

A. Dictionaries

- ISN = Nyembezi, S. 1992. *Isichazamazwi Sanamuhla Nangomuso*. Pietermaritzburg: Reach Out.
- ISZ = Mbatha, M.O. (Ed.). 2006. *Isichazamazwi SesiZulu*. Pietermaritzburg: New Dawn.
- ZED = Doke, C.M. and B.W. Vilakazi. 1964. *Zulu-English Dictionary*. Second Edition. Johannesburg: Witwatersrand University Press.

B. On-line databases

- CBPV = Africa Museum. 2007. Comparative Bantu Pottery Vocabulary. Available:
https://www.africamuseum.be/en/research/discover/human_sciences/culture_society/cb-pottery/browse_zone
- PWN = Princeton University. 2020. WordNet — A Lexical Database for English. Available:
<https://wordnet.princeton.edu/>

C. Other literature

- Abubakar, A.I., A. Roko, A.B. Muhammad and I. Saidu. 2019. Hausa WordNet: An Electronic Lexical Resource. *Saudi Journal of Engineering and Technology* 4(8): 279-285.
- Anderson, W., L. Pretorius and A. Kotzé. 2010. Base Concepts in the African Languages Compared to Upper Ontologies and the WordNet Top Ontology. Calzolari, N. et al. (Eds.). 2010. *Pro-*

- ceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), 17–23 May 2010, Valletta, Malta: 3757-3764. Paris: European Language Resources Association (ELRA).*
- Batibo, H.M.** 2016. The Origin and Evolution of Setswana Culture: A Linguistic Account. *Botswana Notes and Records* 48: 134-149.
- Bentivogli, L. and E. Pianta.** 2000. Looking for Lexical Gaps. Heid, U. et al. (Eds.). 2000. *Proceedings of the Ninth EURALEX International Congress, EURALEX 2000, Stuttgart, Germany, 8–12 August 2000: 8-12. Stuttgart: Institut für Maschinelle Sprachverarbeitung, University of Stuttgart.*
- Bond, F., H. Isahara, K. Kanzaki and K. Uchimoto.** 2008. Boot-strapping a WordNet Using Multiple Existing WordNets. Calzolari, N. et al. (Eds.). 2008. *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), 28–30 May 2008, Marrakech, Morocco: 1619-1624. Paris: European Language Resources Association (ELRA).*
- Bond, F. and K. Paik.** 2012. A Survey of Wordnets and their Licenses. Fellbaum, C. and P.T.J.M. Vossen (Eds.). 2012. *Proceedings of the Sixth Global WordNet Conference 2012 (GWC2012), 9–13 January 2012, Matsue, Japan: 64-71. Brno: Tribun EU.*
- Bosch, S.E. and M. Griesel.** 2017. Strategies for Building Wordnets for Under-resourced Languages: The Case of African Languages. *Literator* 38(1): 1351-1376. Available: <https://doi.org/10.4102/lit.v38i1.1351>
- Bosch, S.E. and M. Griesel.** 2019. Towards the Digital Documentation and Preservation of Indigenous Knowledge in the African Wordnet. Horak, A., K. Osolsobe, A. Rambousek and P. Rychly (Eds.). 2019. *Slavonic Natural Language Processing in the 21st Century: 17-27. Brno, Czech Republic: Tribun EU.*
- Calzolari, N.** 2018. Introduction to LREC 2018. Calzolari, N. et al. (Eds.). 2018. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 7–12 May 2018, Miyazaki, Japan: 1-6. Paris: European Language Resources Association (ELRA). Available: <https://www.aclweb.org/anthology/L18-1088.pdf>*
- Cosijn, E., A. Pirkola, T.J.D. Bothma and K. Järvelin.** 2002. Information Access in Indigenous Languages: A Case Study in Zulu. *South African Journal of Libraries and Information Science* 68(2): 94-103.
- Declerck, T., L. Bajcetic and M. Siegel.** 2020. Adding Pronunciation Information to Wordnets. Declerck, T. et al. (Eds.). 2020. *Proceedings of the Workshop on Multimodal Wordnets (MMWN-2020), 11–16 May 2020, Marseille, France: 39-44. Paris: European Language Resources Association (ELRA).*
- Eckart, T., S. Bosch, U. Quasthoff, E. Körner, D. Goldhahn and S. Kaleschke.** 2020. Usability and Accessibility of Bantu Language Dictionaries in the Digital Age: Mobile Access in an Open Environment. Mabuya, R. et al. (Eds.). 2020. *Proceedings of the First Workshop on Resources for African Indigenous Languages 2020 (RAIL2020), 16 May 2020, Marseilles, France: 9-14. Marseilles, France: European Language Resources Association (ELRA). Available: <https://www.aclweb.org/anthology/2020.rail-1.8/>*
- Fellbaum, C. (Ed.).** 1998. *Wordnet: An Electronic Lexical Database.* Cambridge, MA: MIT Press.
- Fellbaum, C.** 2014. Large-scale Lexicography in the Digital Age. *International Journal of Lexicography* 27(4): 378-395.
- Fowler, K.D.** 2015. Zulu Pottery Technology and Group Identity in the Phongolo Basin, South Africa. *Southern African Humanities* 27:81-111.

- Gouws, R., U. Heid, W. Schweickard and H. Wiegand (Eds.).** 2014. *Dictionaries. An International Encyclopedia of Lexicography. Supplementary Volume: Recent Developments with Focus on Electronic and Computational Lexicography*. Berlin, Boston: De Gruyter Mouton. Available: <https://doi.org/10.1515/9783110238136>
- Griesel, M. and S.E. Bosch.** 2020. Navigating Challenges of Multilingual Resource Development for Under-Resourced Languages: The Case of the African Wordnet Project. Mabuya, R. et al. (Eds.). 2020. *Proceedings of the First Workshop on Resources for African Indigenous Languages 2020 (RAIL2020), 16 May 2020, Marseilles, France*: 45-50. Marseilles, France: European Language Resources Association (ELRA). Available: <https://www.aclweb.org/anthology/2020.rail-1.8/>
- Griesel, M., S. Bosch and M.L. Mojapelo.** 2019. Thinking Globally, Acting Locally — Progress in the African Wordnet Project. Fellbaum, C. et al. (Eds.). 2019. *Proceedings of the Tenth Global WordNet Conference 2019 (GWC2019), 23–27 July 2019, Wrocław University of Science and Technology, Wrocław, Poland*: 191-196. Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej.
- Grossert, J.W.** 1985. *Zulu Crafts*. Pietermaritzburg: Shuter & Shooter.
- Kotzé, G.** 2008. Ontwikkeling van 'n Afrikaanse woordnet: Metodologie en integrasie [Development of an Afrikaans wordnet: Methodology and Integration]. *Literator* 29(1): 163-184. Available: <http://dx.doi.org/10.4102/lit.v29i1.105>
- Krige, E.J.** 1965. *The Social System of the Zulus*. Pietermaritzburg: Shuter & Shooter.
- McCrae, J.P., E. Rudnicka and F. Bond.** 2020. English WordNet: A New Open-source Wordnet for English. *K Lexical News* 28: 37-44. Available: <https://kln.lexicala.com/kln28/mccrae-rudnicka-bond-english-wordnet/>
- Moors, C., I. Wilken, K. Calteaux and T. Gumede.** 2018. Human Language Technology Audit 2018: Analysing the Development Trends in Resource Availability in all South African Languages. *Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists SAICSIT '18: Technology for Change, 26–28 September 2018, Port Elizabeth, South Africa*: 296-304. New York: The Association for Computing Machinery (ACM). Available: <https://dl.acm.org/doi/10.1145/3278681.3278716>
- Miller, G.A.** 1995. WordNet: A Lexical Database for English. *Communications of the ACM* 38(11): 39-41.
- Naskręć, T., A. Dziob, M. Piasecki, C. Saedi and A. Branco.** 2018. WordnetLoom — A Multilingual Wordnet Editing System Focused on Graph-based Presentation. Bond, F. et al. (Eds.). 2018. *Proceedings of Ninth Global WordNet Conference 2018 (GWC2018), 8–12 January 2018, Nanyang Technological University (NTU), Singapore*: 191-200. Singapore: Nanyang Technological University (NTU). Available: <http://compling.hss.ntu.edu.sg/events/2018-gwc/pdfs/gwc-2018-proceedings.pdf>
- Neale, S.** 2018. A Survey on Automatically-Constructed WordNets and their Evaluation: Lexical and Word Embedding-based Approaches. Calzolari, N. et al. (Eds.). 2018. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), 7–12 May 2018, Miyazaki, Japan*: 1705-1710. Paris: European Language Resources Association (ELRA). Available: <https://www.aclweb.org/anthology/L18-1270.pdf>
- Niles, I., and A. Pease.** 2001. Toward a Standard Upper Ontology. Guarino, N. et al. (Eds.). 2001. *Proceedings of the 2nd International Conference on Formal Ontology in Information Systems (FOIS-2001), 17–19 October 2001, Ogunquit, Maine, USA*: 2-9. New York: Association for Computing Machinery.
- Nyembezi, S. and O.E.H. Nxumalo.** 1989. *Inqolobane Yesizwe (The Storehouse of the Nation)*. Pietermaritzburg: Shuter & Shooter.

- Ordan, N. and S. Wintner.** 2007. Hebrew Wordnet: A Test Case of Aligning Lexical Databases across Languages. *International Journal of Translation* 19(1): 39-58.
- Ossai, N.G.** 2010. African Indigenous Knowledge Systems (AIKS). *Symbiosis* 7(2): 1-13.
- Schofield, J.F.** 1943. A Preliminary Study of the Pottery of the Bantu Tribes of the Union of South Africa. *South African Journal of Science* 39: 256-281.
- Snider, K. and J. Roberts.** 2006. *SIL Comparative African Wordlist (SILCAWL)*. Available: https://www.eva.mpg.de/lingua/tools-at-lingboard/pdf/Snider_silewp2006-005.pdf
- SUMO.** 2002. *Suggested Upper Merged Ontology*. Available: <http://www.ontologyportal.org>
- Souza, R.R., A. Dorn, B. Piringer and E. Wandl-Vogt.** 2020. Identification of Indigenous Knowledge Concepts through Semantic Networks, Spelling Tools and Word Embeddings. Calzolari, N. et al. (Eds.). 2020. *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020), 11–16 May 2020, Marseille, France: 943-947*. Paris: European Language Resources Association (ELRA).
- Tufis, D., D. Cristea and S. Stamou.** 2004. BalkaNet: Aims, Methods, Results and Perspectives. A General Overview. *Romanian Journal of Information Science and Technology* 7(1–2): 9-43.
- Vetulani, Z., M. Kubis and T. Obrębski.** 2010. PolNet — Polish WordNet: Data and Tools. Calzolari, N. et al. (Eds.). 2010. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), 17–23 May 2010, Valletta, Malta: 3793-3797*. Paris: European Language Resources Association (ELRA).
- Vincze, V. and A. Almási.** 2014. Non-lexicalized Concepts in Wordnets: A Case Study of English and Hungarian. Orav, H., C. Fellbaum and P. Vossen (Eds.). 2014. *Proceedings of the 7th Global WordNet Conference 2014 (GWC2014), Demonstration Session, Tartu, Estonia, 25–29 January 2014: 118-126*. Tartu: Global WordNet Association.
- Vossen, P., F. Bond and J. McCrae.** 2016. Toward a Truly Multilingual GlobalWordnet Grid. Mititelu, V., C. Forăscu, C. Fellbaum and P. Vossen (Eds.). 2016. *Proceedings of the Eighth Global WordNet Conference 2016 (GWC2016), Bucharest, Romania, 27–30 January 2016: 419-426*. Bucharest: University of Iasi.