Abstract
In promoting a multilingual South Africa, the government is encouraging people to speak more than one language. In order to comply with this initiative, people choose to learn the languages which they do not speak as home language. The African languages are mostly chosen because they are spoken by the majority of the country’s population. Most words in these languages have many possible senses. This phenomenon tends to pose problems to people who want to learn these languages. This article argues that the African WordNet may the best tool to address the problem of sense discrimination. The focus of the argument will be on the primary sense of the word ‘hand’, which is part of the body, as lexicalized in three indigenous languages spoken in South Africa, namely, Tshivenda, Sesotho sa Leboa and isiZulu. A brief historical background of the African WordNet will be provided, followed by the definition of the word ‘hand’ in the three languages and the analysis of the word in context. Lastly, the primary sense of the word ‘hand’ across the three languages will be discussed.

1 Introduction
Thoughtful lexicography work for indigenous African languages of South Africa commenced just after the introduction of democratic elections in 1994. With the establishment of the Pan South African Language Board, national lexicography units were constituted in all the official languages of South Africa. The lexicography units were tasked with the duty of establishing dictionaries in the different official languages of South Africa. Although many of the dictionaries are bilingual, they give very little information regarding sense discrimination, especially for non-mother tongue speakers who are interested in learning indigenous African languages. The South African government encourages people to learn one indigenous African language in addition to their first language. Lexicography work in African languages produced so far does not address the needs of indigenous African language learners because the equivalents provided do not address the problem of sense discrimination. Similarly, indigenous African language learners take it for granted that a lexical item has the same sense across these languages, whereas sometimes the sense of a word is different in these languages even if languages are related.

This paper argues that African WordNet could be a viable tool to address problems such as those mentioned above. The equivalents of ‘hand’ in Tshivenda (Venda), Sesotho sa Leboa (Northern Sotho) and isiZulu (Zulu) are tshanda, seatla (letsogo) and isandla, respectively. Indigenous official languages of South Africa belong to the same family of languages; they are Bantu languages belonging to the Niger-Congo family. They are further divided into groups that are, to a certain extent, mutually intelligible. The Nguni language group and the Sotho language group, for example, are not mutually intelligible whereas languages within any of the two groups are. A
majority of the people in the country is multilingual but they may nevertheless not be competent in all the languages. Being a rainbow nation with a myriad of people and languages, everyday life dictates that one has some understanding or awareness, however limited, of other languages. The fact that official African languages in the country belong to the same family often tempts people, knowingly or unknowingly, to clamp them together with the saying ‘if you know one you know them all’ – and this is far from the truth. The lexicons and the senses reflect some similarities, overlaps and unrelatedness to an extent that they may result in miscommunication unless sense discrimination is taken care of.

We have used the English word ‘hand’ to demonstrate lexicalisation and sense discrimination in the languages, Sesotho sa Leboa (Northern Sotho), Tshivenda (Venda) and isiZulu (Zulu). Whilst there are other examples that could be used in the African WordNet to indicate sense discrimination across the indigenous African languages of South Africa, the choice of the word ‘hand’ stems from its cultural significance in the African value system. The word ‘hand’ has as its underpinning in the ‘Ubuntu’ (a value system that promotes humanity to others) element which regards humanity as a fundamental part of the eco-systems that lead to a communal responsibility to sustain life.

2 African WordNet defined

African WordNet is based on the Princeton WordNet. It is a multilingual WordNet of official indigenous languages of South Africa. WordNets for African languages were introduced with a training workshop for linguists, lexicographers and computer scientists facilitated by international experts in 2007. The development of WordNet prototypes for four official African languages started in 2008 as the African WordNet Project. This project was based on collaboration between the Department of African Languages at the University of South Africa (UNISA) and the Centre for Text Technology (CTeXT) at the North-West University (NWU), as well as support from the developers of the DEBVisDic tools at the Masaryk University. The initiative resulted in first versions of WordNets for isiZulu [zul], isiXhosa [xho], Setswana [tsn] and Sesotho sa Leboa [nso], all members of the Bantu language family (Griesel and Bosch, 2014). Currently Tshivenda is the fifth of the nine official African languages of the country that are part of the project.

3 Word sense

Sense is defined as one of a set of meanings a word or phrase may bear especially as segregated in a dictionary entry (Miriam Webster Online). Frege (1892) argues that sense is the mode of presentation of the referent. There are multiple ways of describing and conveying information about one and the same referent; and to each of these ways correspond a distinct sense. Every word is associated with a sense, and the sense specifies the condition for being the word’s referent.

According to Fellbaum (1998) in WordNet, each occurrence of a word form indicates a different sense of the word, which provides for its unique identification. A word in a synset is represented by its orthographic word form, syntactic category, semantic field and identification number. Together these items make a “sense key” that uniquely identifies each word/sense pair in the database. The sense of a word can be derived from the semantic relations that it has with other words. The manner in which word sense is viewed has a great appeal for the discussion of the word ‘hand’ in this article.

The underlying hypothesis of this paper relies on previous studies that used multiplicative models of composition by exploring methods to extend the models to exploit richer contexts. Studies by Gale et al., (1993) and Dagan et al., (1991) have used parallel texts for sense discrimination to identify semantic properties of and relations among lexemes (Dyvik, 1998). Whilst there are different approaches to sense discrimination, this paper adopts an approach by Akkaya, Wiebe and Mihalcea (2012) which is to cluster target word instances, so that the induced clusters contain instances used with the same sense.
4 The primary sense of ‘hand’ in the three African languages

The primary meaning of a word is its literal meaning. This section looks into the dictionary equivalents of the primary meaning of the English word ‘hand’ in the three languages Tshivenda, Sesotho sa Leboa and isiZulu. The concept under discussion in this paper is defined in WordNet as “the (prehensile) extremity of the superior limb”. It is sense 1 of the domain Anatomy and SUMO Bodypart [POS: n ID: ENG 20-05246212-n BCS: 3].

4.1 Tshivenda

The equivalent of hand in Tshivenda is tshanđa. Whereas hand in English refers to the part at the end of a person’s arm, including the fingers and thumb (Longman Dictionary of Contemporary English, 1995), tshanđa in Tshivenda refers to both arm and hand taken as one. Tshivenda does not separate between arm and hand as languages such as English do, both are taken as one.

There is a slight difference among the Tshivenda lexicographers in defining the lexical entry tshanđa. Wentzel and Muloiwa (1982:65 and 173) define tshanđa and ‘hand’ differently. They define tshanđa (pl. zwanđa) as arm, hand; whereas hand is defined as tshanđa (pl. zwanđa). According to these lexicographers, tshanđa has got two senses, that of the whole arm, and the part at the end of a person’s arm.

The same applies to Tshivenda – English Thalusamaipфи Dictionary (2006); the equivalent of hand is tshanđa and the equivalents of tshanđa are hand and arm. It would seem Tshivenda – English Thalusamaipфи Dictionary (2006) adopted the definitions of the two lexical entries direct from Wentzel and Muloiwa (1982). To them both hand and arm are called tshanđa. Van Warmelo (1989:388) on the other hand provides the equivalent of tshanđa as hand. He does not differentiate between arm and hand; according to him the whole limb is tshanđa. However, he also refers to the upper arm as tshishasha. Tshikota (2012a) and Tshikota (2012b) in his two monolingual dictionaries, Thalusamaidomа ya luamboluthihi ya Tshivenda (Tshivenda monolingual dictionary of idioms) and Thalusamaipфи ya luamboluthihi ya Tshivenda (Tshivenda monolingual dictionary) define tshanđa as follows:

- tshanđa dzin tshipiđa tsha muvhili tshi re na minwe mišana tshine tsha shumiswa u fara ngatsho (Tshikota, 2012a:57)
  
  ‘part of the body with five fingers, which is used to hold’

- tshanđa (zwanđa) dzin 1 tshipiđa tsha muvhili tshi re na minwe mišana tshine tsha shumiswa u fara ngatsho (Tshikota, 2012b:258)
  
  ‘part of the body with five fingers, which is used to hold’

The definitions of the lexical entry tshanđa in the two dictionaries are similar, and they refer to the English word hand. Lexicographers in these dictionaries were influenced by the English definition of hand. They do not reflect what the word tshanđa refers in the spoken language. The word tshanđa in spoken Tshivenda refers to English arm plus hand. This is attested by Wentzel and Muloiwa (1982), Van Warmelo (1989) and Tshivenda – English Thalusamaipфи Dictionary (2006). The word tshanđa also refers to the palm.

4.2 Sesotho sa Leboa

The word for ‘hand’ in Northern Sotho is seatla (plural: diatla). Ziervogel and Mokgokong’s (1975) trilingual dictionary gives entries in Northern Sotho and equivalents in Afrikaans and English. The English equivalents of the word seatla in the dictionary are ‘hand’, ‘palm of hand’, ‘handwriting’. The dictionary then continues to use the word in various linguistic contexts in order to lay bare different senses. Of the three English equivalents mentioned above, only ‘handwriting’ seems to be non-literal, not representing the sense under the domain - Anatomy. The first two equivalents refer to the physical part of the body. Only the first equivalent has a conceptual one-to-one with the concept defined in WordNet as “the (prehensile) extremity of the superior limb”. The other equivalent ‘palm of hand’ is part of the whole concept defined above. Another trilingual
dictionary (Northern Sotho Language Board, 1988) gives entries in English and equivalents in Afrikaans and Northern Sotho. The latter is not only a dictionary, but a terminology and orthography standardizing document as well. The entry ‘hand’ has a Northern Sotho equivalent seatla. Following this entry is a number of English compound nouns and two-word entries which include ‘hand’. Of these entries seven are clearly built on the primary meaning of ‘hand’. The seven entries reflect that ‘hand’ is also referred to as letsogo in Northern Sotho. For example, the Northern Sotho ‘hand’ is also referred to as seatla, meaning of ‘hand’. The seven entries reflect that word entries which include ‘hand’. Of these entries seven are clearly built on the primary meaning of ‘hand’. The seven entries reflect that ‘hand’ is also referred to as letsogo in Northern Sotho. For example, the Northern Sotho equivalent of ‘handwork’ is modiro wa diatla, ‘hand muscle’ is mošifa wa seatla, ‘hand movement’ is tshepedišo ya letsogo, ‘hand drill’ is borotsogo, and ‘handbag’ is sekhwamatso.

4.3 IsiZulu

Mbattha (2006: 9) in his isiZulu monolingual dictionary defines ‘hand’ as isitho somuntu okuyisona abamba ngaso ‘a body part which a human uses to hold’. Mbattha’s definition dears to the lexicographic feature in providing the quality of definition required to give clarity. However, Doke and Vilakazi (1972: 9) in their Zulu-English dictionary define ‘hand’ as forearm (including the hand). From the definitions of these lexicographers, it is apparent that they define the concept not exactly the same. Mbattha seems to be focusing mostly on the functional aspect of the word ‘hand’ than striving to describe its meaning. Mbattha’s definition dears to the lexicographic feature in providing the quality of definition required to give clarity. The definition by Doke and Vilakazi on the other hand, is not detailed enough. When considering Doke and Vilakazi’s definition, it lacks the defining criteria and the characteristics that are necessary to understand what the word means. What makes Doke and Vilakazi’s definition incomplete is that it does not give enough information about the word. In Collins English Dictionary (1991:704) the word hand is defined as ‘the prehensile part of the body at the end of the arm, consisting of a thumb, four fingers and a palm’. Considering the definitions given by Mbattha, and Doke & Vilakazi, it becomes clear the information that they have provided has a tentative validity.

5 Discussion

Across the three languages, the primary sense of ‘hand’ is a physical part of the human body. Lexicographers have to constantly strive to enhance the quality of definitions in monolingual dictionaries to best suit the needs and level of their target users (Gouws 2001:143). Landau (2001:162) also maintains that the definition must define and not just talk about the word or its usage. It is clear from the argument given above that they do not provide the answer to the question ‘what it is’ that is being defined as Gouws (Ibid) suggests. Lombard (1991:166) pinpoints defining criteria that would result in good definitions namely completeness, clarity, accuracy, consistency, independency, objectivity and neutrality. Although words for ‘hand’ in the three languages may refer to the different parts of the limb, starting at the end of the shoulder and ending at the fingers, the parts constitute the same limb. Whereas in Tshivena and isiZulu, ‘hand’ is referred to as tshanga and isandla respectively, in Sesotho sa Leboa it is referred to as seatla or letsogo. In Tshivena, tshanga is that part of the human body starting from the shoulder to the fingers. This means that the whole limb is referred to as tshanga. The sense in isiZulu is slightly different from that in Tshivena because isandla refers to the forearm including the wrist, fingers. Whereas Tshivena tshanga refers to the whole limb, isiZulu isandla refers part of the limb, i.e. forearm. Sesotho sa Leboa refers to the whole limb as letsogo ‘arm’, to the ‘hand’ as seatla; additionally ‘hand’ is referred to as letsogo. Seatla is part of the whole limb, a synonym of letsogo ‘arm’, but also used synonymously with letsogo. Unlike Tshivena and Sesotho sa Leboa, isiZulu recognises the forearm as part of the hand, which is referred to as isandla. In Tshivena and Sesotho sa Leboa, the palm of the hand is referred to as tshanga and seatla, respectively. The diagram in appendix 1 illustrates the situation sketches above.

It emerges from the Northern Sotho dictionary definitions and equivalents that the concept is lexicalized as seatla and/or letsogo. The English dictionary equivalent of Northern Sotho letsogo is ‘arm’. Letsogo refers to the whole superior limb, which includes seatla ‘hand’. The two are understood to be in a
holonym-meronym relationship, while being used as synonyms as well.

6 Conclusion

The empirical conclusion in this paper provides a new understanding of words with different senses which pose a challenge to the different speakers of the indigenous South African languages, particularly the three languages mentioned. Considering the hypothesis posed at the beginning of this paper, it can be concluded that the primary sense of hand in the three languages, although related, is different. People learning these languages should not conclude that because they are grouped as African languages the senses of their lexicons are similar throughout. It is also noted that the sense of hand in English is different from that in the African languages. WordNet is a good tool to investigate the sense of African languages’ lexicons, in that the word ‘arm’ has a comparable sense and an ID, namely, arm: 1 [POS: n ID: ENG 20-05245410-n BCS: 3] and belongs to a specific domain: Anatomy.

The discussion in this paper has gone some way towards enhancing our understanding of the degree to which African WordNet can be a tool that can be used to differentiate word sense. This research has thrown up many questions in need of further investigation regarding the other sense such as the metaphoric use and the idiomatic expression of the word in discussion. It became evident from the discussion that the same word can have different senses in the different.

References


Tshikota, S.L. 2012. Ṭhalusamaidioma ya Luamboluthihi ya Tshivenda. Ṭhohoyand ou: Tshivenda National Lexicography Unit.
Tshikota, S.L. 2012. Ṭḥalusamaipfì ya Luamboluthihi ya Tshivënà. Ṭhohoyanqou: Tshivënà National Lexicography Unit.


