Contrastive analysis of word-formation strategies in the translated South African Constitution

Introduction

Of the 11 official languages of South Africa, Xitsonga, Siswati and isiNdebele are three of the four minority indigenous languages, with less access than the other languages to translated materials and a scarcity of the terminology required for current legal, administrative, educational and other official use. Translations are required to provide open access to all documents in the official languages with clear and unambiguous meaning for all speakers of these languages. Moropa (2004:162) observed that before independence, indigenous languages in South Africa received little attention in terminology development; the minimal terminology that was available was created only for basic education. The South African Constitution (1996) gave government a mandate to uplift the status and improve the usage of indigenous languages by translating official documents that were originally prepared in English and Afrikaans. The importance of providing universal access to the Constitution gave rise to its translation from English into all South Africa’s official languages.

When translating texts into a language that lacks relevant terminology, translation equivalence is difficult to achieve, and the accuracy, accessibility and effectiveness of communication of the translated texts may be compromised. This article investigated various approaches to the translation of terminology when no direct translation equivalents were available in the target languages. The researchers performed a comparative analysis of strategies applied in the translation of the South African Constitution from English into Xitsonga, Siswati and isiNdebele. Firstly, Voyant tools were used to identify terms in the English version. Secondly, a multilingual concordancer (ParaConc) was employed to extract translation equivalence. ParaConc allows for contrastive studies on aligned corpora. The strategies that were found were borrowing, paraphrasing, derivation and compounding. Paraphrasing, derivation and compounding are productive strategies, because equivalents are formed using the already existing words in the languages which native speakers clearly understand. The use of borrowing revealed a critical need for terminology development work in these languages to avoid transliterating terms from foreign languages in which the meaning is not directly clear to the native speakers. This article also demonstrated the usefulness of computational approaches in identifying terminology and translation techniques in the context of South Africa’s official languages.

Contribution: The article demonstrates word-formation strategies that were used in the translated South African constitution from English into Xitsonga, Siswati and isiNdebele when no translation equivalents of terminology are available in these target languages.

Keywords: translation; terminology; word-formation strategies; indigenous languages; Voyant tools; ParaConc.
focus on word-formation strategies in which we refer to the methods used by translators when no translation equivalents of terminology are available in these target languages. The English–Xitsonga–IsiZulu–IsiNdebele Parallel Corpus (EXSNPC) is used as the data set for analysis. Voyant Tools (VT) and ParaConc are computational tools used to identify and extract terms from the source language and their translation equivalents in the target languages.

Related works

Van Huyssteen (1999) discussed problems and solutions regarding the scarcity of adequate terminology in South African indigenous languages, with special reference to isiZulu. This study demonstrated that Eurocentrism, lack of time, foreign sounds, standardisation, trendy words, multilingualism, etymological purity, lack of coordination between the various standardisation sectors and the abundance of synonyms are some of the problems identified regarding the lack of terminology for South African languages. Van Huyssteen’s argument was that the lack of terminology in these languages can only be sorted out by standardising and coordinating terminology development activities in these African languages.

Madiba (2001) proposed a model suitable for modernising terminology from English to indigenous languages. The author argued that a pragmatic approach should be adopted for the creation and modernisation of terminology in indigenous languages, particularly Tshivenda. The approach was recommended as it encompasses and acknowledges borrowing and indigenisation as strategies useful in terminology modernisation and accessibility to new terms in science and technology in the indigenous languages. One of the merits of the borrowing strategy is that it fills the terminology gap timeously, where no equivalent terms exist in the target languages. Using the indigenous term technique enables effective communication in less technical or specialised disciplines.

Sineke (2005) explored strategies used to create terms in coping with scarcity of terminology in the isiXhosa- and isiZulu-translated versions of the Constitution. The study showed that derivation, borrowing, compounding, semantic transfer and paraphrasing were the major strategies employed by translators to make up for terminology gaps. Of these, Sineke indicated that paraphrasing was the main strategy used. Regarding borrowing, it was found that all loaned terms were modified based on the orthographical, morphological and phonological rules of the target languages. Sineke (2005) concluded that term-creation strategies are widely used in African language–translated materials with limited resources and that they contribute to the development of these languages.

Both Moropa (2007) and Ndhlouvu (2014) examined parallel corpora using a multilingual concordancer (ParaConc) to analyse the translation strategies used by translators. The English–Xhosa Parallel Corpus was used as a resource in both Moropa’s and Ndhlouvu’s analysis. The corpus was composed of the 1997 Annual Report of the Department of Arts, Culture, Science and Technology (DACST), A Short Guide to the White Paper on Local Government 1998, the 2001–2002 Pan South African Language Board (PanSALB) Annual Report and three manuals on the Promotion of Access to Information Act (PAIA manuals). Both studies were based on the notion that using ParaConc software makes it easier for researchers to identify the translation strategies used in translated texts. Borrowing was the major strategy being used by translators in both studies. The types of borrowing used were pure loaning of acronyms and abbreviations, pure loan words, indigenisation, pure loan words preceded by an explanation and abbreviations preceded by an explanation. Other strategies used were paraphrasing, acronyms, coinage, derivation, compounding and semantic shift using borrowed synonyms.

Mabunda (2013) explored methods of term translation and their effects in circumstances where there is zero equivalence from Xitsonga to English or English to Xitsonga. The study determined that sameness of meaning during translation is influenced by culturally bound contexts, such as dialects, figurative language and referential and geographic contexts, as opposed to the standard language. Hence, to overcome the challenge of zero equivalence during translation, translation by subordinate words, cultural substitution, loan words, illustration, paraphrasing, omission and neologism were found to be useful.

Mojapelo (2018) investigated various strategies utilised in Northern Sotho term creation. The study was prompted by obstacles discovered in the formation of Northern Sotho terminology. The usage of informal language, unavailability of adequate terminology and multiple inconsistent terminology lists obtained from various translators were some of the challenges ascertained. It was also established that inconsistent terms found in orthography and spelling resources that impede the search for translated term equivalents, nonavailability of specialised dictionaries and lack of experience on how terminology development must be conducted contributed to ineffective terminology development in the language. Mojapelo (2018) then noted that borrowing, transliteration, paraphrasing, semantic transfer and compounding were common term creation techniques in Northern Sotho. However, transliteration and paraphrasing were the most employed strategies in the language. These strategies were deemed effective in developing sound terminologies in the target language.

Letsoalo, Mabaso and Gouws (2022) investigated strategies employed in the translation of English robotics terms into indigenous languages, namely Sepedi, Xitsonga and isiXhosa, and determined the appropriateness of the translation equivalents for easier access to robotics
information. The study was undertaken with the intention to show how translation can transform the teaching and learning of robotics programmes in indigenous languages. It was indicated that speakers of indigenous languages could benefit from developing open educational resources, as it gives them a choice to access educational information in their language of interest. It was recommended that strategies such as compounding, semantic transfer and derivation should be considered first in terminology development. Borrowing must be the method least adopted, because it allows a source language to encroach on a target language, and this could lead to indigenous languages sounding foreign. From these strategies that were used for robotics terminology development, it was evident that there is no specific strategy that is used uniformly in all the South African indigenous languages.

As far as researchers know, no other study has embarked on exploring and comparing word-formation strategies used in the texts translated into Xitsonga, Siswati and isiNdebele. Scholars who studied translation strategies in South African languages focused on translation from a source language into one target language. There is still a gap in investigating the strategies used in the translated texts when no direct translation equivalents of a source language exist in the multiple target languages. Therefore, this article fills the gap in translation studies by examining and comparing the strategies used in the translation of the Constitution from English as a source language into Xitsonga, Siswati and isiNdebele as target languages with limited terminology developed.

Methodology
This section describes the methodology employed in the investigation, beginning with a description of the data set used. Next, the different procedures used to identify translation equivalents are explained, and the different word-formation strategies are described.

Data sets
The versions of the Constitution in English, Xitsonga, Siswati and isiNdebele were downloaded in PDF format from the Department of Justice website. Next, these versions were converted into plain text format to be conveniently analysed using computational tools. This resulted in the creation of an EXSNPC. The EXSNPC is a multilingual corpus which was manually aligned into paragraphs, sentences, phrases and words. The corpus consists of a total of 173,351 tokens (including numbers but excluding punctuation) with 47,351 tokens in English, 57,258 tokens in Xitsonga, 34,054 tokens in Siswati and 34,488 tokens in isiNdebele.

Procedures
Voyant Tools, accessed online from https://voyant-tools.org, were used to identify terms in the source language. Welsh (2014:96) advised that VT can provide term occurrences and trends that can be retrieved by users to investigate statistical information of a given corpus. Voyant Tools was used owing to its ability to display lexical terms (which this article is concerned with) separately from function words. The English Constitution in a text format

![Voyant Tools term list.](image)
was uploaded in VT, and the feature terms were selected to display all lexical terms found in the corpus as a list, as shown in Figure 1.

Purposive sampling was then used to select 30 terms from the VT term list. The selected terms from the English corpus were then searched in ParaConc to extract translation equivalents for the target languages. ParaConc was utilised because of its ability to semi-automatically retrieve the translated equivalents of the lexical terms in the target languages Xitsonga, Siswati and isiNdebele, as presented in Figure 2.

After the equivalent terms in the target languages were identified through the assistance of ParaConc, a manual analysis was then conducted in which the terms were compared to establish word-formation strategies employed by translators.

**Word-formation strategies**

Word formation is a research area in linguistics that examines processes in which users of a language create new lexical units (Fandrych 2004:6). Nasser (2008:71) defined word formation as a morphological and grammatical process that aims at generating new words in a language. The purpose is to promote effective language usage through the establishment of words when terminology is lacking. In the next subsection, the strategies of borrowing, paraphrasing, derivation and compounding, which are mainly used when no translation equivalents of terminology are available in a target language, are discussed in more detail.

**Borrowing**

Taljard (2016:34) described borrowing as a process in which ‘linguistic elements are taken over from one language to another’. Taljard distinguished two types of borrowing in South African languages: loan words and transliteration. Loan words are lexical items that are taken as a unit from a certain language with their meaning being maintained, while transliteration involves words being entirely adapted to the various linguistic levels of the target language. Indigenous languages do not only borrow from foreign languages but also make use of words taken from other African languages.

**Paraphrasing**

Paraphrasing expresses phrases or a single word to clarify, describe and define concepts when part of the meaning is contentious; it is significant for exploring natural language semantics (Danielsson 2007:33). Meanwhile, Baker (2011) differentiated between two categories of paraphrasing: translation by paraphrase using a related word and translation by paraphrase using unrelated words. Concerning translation using a related word, Baker notes that this strategy is used when the source element is lexicalised in the target language, and the frequency with which a specific word is used in the source language is more than the frequency of use in the receiving language. When translation takes place using unrelated words, the term in the source text is not lexicalised in the receiving language. When the meaning of the source text is complex, this strategy may be used based on modifying a super-ordinate or simply making explicit the meaning of the source text. These types of paraphrasing vary in the way that they function in the context, but the key purpose of using paraphrasing is to explain, describe, interpret, or clarify an expression of the source text in the target text in the case of

![FIGURE 2: English term and its equivalents in target languages.](http://www.literator.org.za)
nonequivalence in the target language, and as such, it helps the users to understand the meanings of a given text.

**Derivation**

Yule (2006:53) described derivation as a strategy used to create new words by adding morphemes to the word's stem. Derivation is a word-formation strategy in which words are created from already existing words by attaching affixes. The two types of affixes involved in African languages are prefixes and suffixes. Prefixes are attached at the beginning of a word stem, while suffixes are added at the end of a word to form a derivative word (Alberts 2017; Moropa 2007; Sineke 2005). This strategy is highly productive, as it creates several word categories with different meanings from existing words.

**Compounding**

Compounding is a morphological process in which new words are formed by combining different words or lexical elements (Ten Hacken 2017). Moreover, Mashamba (2011:49) explained compounding as a strategy in which two or more separate words or free morphemes are ideally combined to form a new lexical term. Compound words consist of different parts of speech. Languages differ by the ways in which existing words are merged to form compound words. Compounds may be compositional, in which two or more words that used to create new items have related meaning with the compound as their conventional meanings. In noncompositional compounds, the meaning of the new word in the compound differs from its traditional meaning (Salehi, Cook & Baldwin 2014:472). Consequently, the meanings of compositional compounds can be derived, while the meanings of the component words cannot be derived in the case of noncompositional compounds.

**Analysis of results and discussion**

**Results**

This subsection shows the results of strategies used to translate the 30 sampled terms that were extracted from the data sets. Only examples of each strategy that are applicable in the respective languages are given in the tables.

**Borrowing**

Table 1 shows all examples of borrowing in the target languages within the 30 selected terms for analysis. Indigenous languages mainly borrow from English.

<table>
<thead>
<tr>
<th>English</th>
<th>Xitsonga</th>
<th>Siswati</th>
<th>isiNdebele</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commission</td>
<td>khamixini</td>
<td>ikhamishani</td>
<td>ikomithini</td>
</tr>
<tr>
<td>Office</td>
<td>Hofisi</td>
<td>Lithovisi</td>
<td>i-ofisi</td>
</tr>
<tr>
<td>Republic</td>
<td>Riphabiliki</td>
<td>iriphabilihuli</td>
<td>iriphabiliki</td>
</tr>
<tr>
<td>Parliament</td>
<td>Palamendle</td>
<td>ipalamendle</td>
<td>ipalendle</td>
</tr>
<tr>
<td>Court</td>
<td>Khato</td>
<td>iKantolo</td>
<td>ikhatho</td>
</tr>
<tr>
<td>Referendum</td>
<td>iRerendamunum</td>
<td>-</td>
<td>iRerendamunu</td>
</tr>
<tr>
<td>Summons</td>
<td>Masamanisi</td>
<td>-</td>
<td>Amasamoni</td>
</tr>
<tr>
<td>Budget</td>
<td>-</td>
<td>iBhajethi</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 1 shows all examples of borrowing in the target languages within the 30 selected terms for analysis. Indigenous languages mainly borrow from English.

However, there are cases in which they borrow from Afrikaans and other languages. The borrowed terms are usually adapted to phonological and morphological rules of the target languages. In all borrowed nouns, prefixes are attached. In Table 1, the Siswati and isiNdebele borrowed terms begin with singular prefix /si-/ except the term lihovisi that has singular prefix /li-/ and amasamoni that has plural prefix /ama-/ . Xitsonga borrowed terms have no overt prefix as a result of the required phonological and morphological rules, except for the term masamanisi with the plural prefix /ma-/ . All the target languages have borrowed terms through transliteration from English, while the term ‘court’, which is inkantolo in Siswati, was transliterated from the term kantoor in Afrikaans. The whole process of borrowing through transliteration is not productive, apart from the terms being modified to the phonological and morphological rules of the target languages. This notion is also supported by Mphahlele (2004:341), who observed that ‘transliteration often fails target users because the foreign concept is not easily transmitted into the target language’. He further explained that in most circumstances, the use of transliteration is detrimental to the target users, as they cannot understand the semantic concept unless the term used in the source language has been accepted and is well-known in the target language.

**Paraphrasing**

Table 2 shows instances where paraphrasing was used to close lexical gaps in the target languages. In the table, the term ‘suffrage’ was paraphrased as mfanelo yo vhota ‘the right to vote’ in Xitsonga, kuwumela kutsi uvote ‘to be allowed to vote’ in Siswati and ilungelo lokuvowuda ‘a right to vote’ in isiNdebele. Translation by paraphrasing using related words was employed. For example, the word ‘detained’ was translated as ku pfaleriwa ‘to be locked up’ in Xitsonga, kuwaleluwa ejele ‘to be locked up in prison’ in Siswati and kuwaleluwa ejele ‘to be locked up in prison’ in isiNdebele. The term ‘senate’ was translated using unrelated words like Huvo ya rixaka, ‘house of the nation’ in Xitsonga, Indlu yetiMphungu ‘house of elders’ in Siswati and Indlu yeSimethi ‘house of senate’ in isiNdebele.

**TABLE 2: Examples of paraphrasing.**

<table>
<thead>
<tr>
<th>English</th>
<th>Xitsonga</th>
<th>Siswati</th>
<th>isiNdebele</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suffrage</td>
<td>Mfanelo yo vhota</td>
<td>Kuwumela kutsi uvote</td>
<td>Ilungelo lokuvowuda</td>
</tr>
<tr>
<td>Swear</td>
<td>Ku tvisisa vutshembeki</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Detained</td>
<td>Ku pfaleriwa</td>
<td>Kuwaleluwa ejele</td>
<td>Kuwaleluwa ejele</td>
</tr>
<tr>
<td>Privilege</td>
<td>-</td>
<td>Tiphatsimandla tetemtsefto</td>
<td>Limpaphatsiwa zomthetho</td>
</tr>
<tr>
<td>Judiciary</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Legislation</td>
<td>Huvo ya endla milawu</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Servitude</td>
<td>Xyimo xa vuhlonga</td>
<td>Kusebenta matima njengesigcila</td>
<td>Ukukatelelwa ukukhuluphiha amanye</td>
</tr>
<tr>
<td>Misconduct</td>
<td>Matikhomelo yo ka ya nga ri kahle</td>
<td>Kutiphasa kobi kakhulu</td>
<td>Ukukatelelwa kumbi okudansako</td>
</tr>
<tr>
<td>Senate</td>
<td>Huvo ya Rixaka</td>
<td>Indlu yetiMphungu</td>
<td>Indlu yeSimethi</td>
</tr>
<tr>
<td>Sovereign</td>
<td>Mfumo lowu thiyimeleke</td>
<td>Live lelitibusako</td>
<td>Mbuso aziyameleko</td>
</tr>
</tbody>
</table>

Derivation

Table 3 shows examples of derivation in the target languages. Words can be derived from different parts of speech by attaching prefixes and suffixes. In Table 3, all the nouns are derived from the verbs. In Xitsonga, the term xihlambanyo was derived from the verb ilambanya ‘to declare’. The prefix /xi-/ was added to the verb with the nominal suffix /-o/ (used when a noun is derived from the verb) to form the equivalent. Meanwhile, Siswati kufunga and isiNdebele ukufunga were derived from the verb funga ‘swear’ by prefixing /ku-/ and /uki-/ to form the nouns, respectively.

Compounding

Table 4 shows all examples of compounding used in the target languages. The terms in the table have been translated into the target languages by combining two or more separate words to form a single larger word. Currently, Xitsonga, Siswati and isiNdebele compounds are formed as a single term without using hyphenation (-) between the combined lexical categories. However, in the translation of the Constitution, some compound terms were hyphenated in the target languages because of the spelling and orthography rules used when the Constitution was translated.

All the terms from Table 4 in the target languages were translated using compositional compounding. For example, the term ‘Auditor-General’ was translated using compounding, in which three words were used to form a single word. It was translated as Mukambelatinkotankulu in Xitsonga. The term was formed by combining the verb kambela ‘examine’, the noun tinkota ‘audit’ and the adjective kulu ‘big’. Since Auditor-General refers to a human being, a prefix /mu-/ was attached to comply with the system of noun classes in Xitsonga. In Siswati, it was translated as UmCwaningimabhubuku-Jikelele with the noun UmCwaningi ‘examiner’, the noun madhubuku ‘books’ and the adjective jikelele ‘general’ being merged to form the compound. Meanwhile, in isiNdebele, the term was translated as umHloliincwadizombele. The term was formed through the combination of the noun Umhloli ‘inspector’, the noun incwadi ‘books’ and the adjective zombele ‘general’. In Xitsonga, the term for ‘Executive’, which is Huvonkulu, was formed by combining two words huvo ‘house’ and nkulu ‘big’. The term ‘Legislature’, which is Sishayamtsetfo, was coined by joining sishaya ‘maker’ and umsetfo ‘law’ in Siswati. In isiNdebele, the term Isibethamthetho was also created by combining the words isibetha ‘maker’ and umthetho ‘law’. Therefore, compound nouns can be formed by two or more words.

Discussion

Various strategies are used by translators of the Constitution for the target languages. This article has found that all word-formation strategies that were used when no equivalent translation was available are typically recognised in other indigenous languages. Some of the strategies are more productive than others in closing the lexical gaps between a source and target language in the absence of relevant terminology. From the study, we observed that the strategies used in the translations of the Constitution seem to be a common pattern used by translators of other South African indigenous languages, as already found in the studies by Sineke (2005), Moropa (2007) and Ndhllovu (2014). Therefore, this extends the work done in South African indigenous languages which did not include Xitsonga, Siswati and isiNdebele. The strategies of borrowing, paraphrasing, derivation and compounding were employed in these target languages.

Paraphrasing, derivation and compounding were found to be productive strategies used by translators of the target languages. These strategies were applied using already existing words in the target languages. Paraphrasing was used by translators to clarify and describe concepts using related and unrelated words. Regarding derivation, equivalent terms in the target languages were formed from existing words by attaching affixes. The terms were
mainly derived from verbs to form nouns within the target language concerned. Compositional compounds were formed by combining two free morphemes, and others were formed by joining three lexical terms. In the context of the African languages, these strategies are productive in creating relevant terminology when no direct translation equivalents are available in the target languages. Therefore, they fill the existing gap in equivalent terms between the source and target languages.

Borrowing was used as a strategy in the target languages, in which terms were borrowed mainly from English and Afrikaans through transliteration. The borrowed terms were modified to the phonological and morphological rules of the target language. This strategy does not provide an ideal contribution to close lexical gaps that exist with the target languages, as observed by Mphahlele (2004).

Conclusion

This article described the strategies used in translating the South African Constitution from English as a source language into Xitsonga, Siswati and isiNdebele when no translation equivalents of terminology are available in the target languages. We showed that borrowing, paraphrasing, derivation and compounding were the different strategies employed. In instances where borrowing was used, the target languages borrowed terms mainly from English, and in a few instances from Afrikaans, through transliteration. The borrowed terms were adapted following the spelling and orthographical rules of the receiving languages. The use of borrowing has limitations, because the meaning is not directly clear to the native speakers who are not familiar with the loaned term which has not yet been lexicalised in the target language. The study recommends the use of productive strategies such as paraphrasing, derivation and compounding to narrow the lexical gaps that exist within the indigenous South African minority languages.

In the future, we will investigate how translators working with the other indigenous languages dealt with the issue of equivalence in translated texts, as these languages are perceived to be reasonably well-resourced compared to the minority languages. We are also interested in investigating the preferred strategies used by translators and the reasons that inform their preferences in contemporary translated texts in specialised fields.

Acknowledgements

The authors would like to thank the following people for their helpful comments and feedback during the writing of this article: Menno van Zaanan and Juan Steyn.

Competing interests

The authors have declared that no competing interest exists.

Author’s contributions

All authors contributed equally in the writing of this paper.

Ethical considerations

This article followed all ethical standards for research without direct contact with human or animal subjects.

Funding information

This publication was made possible with the support from the South African Centre for Digital Language Resources (SADiLaR). South African Centre for Digital Language Resources 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).

Data availability

The data that support the findings of this study are available in the Constitution of the Republic of South Africa at: https://www.justice.gov.za/legislation/constitution/

Disclaimer

The views shared in this paper are those of the authors and do not represent the opinions of others or the associated authors’ organisation.

References


Taljard, E., 2016, Terminology, University of Pretoria, Pretoria.


