UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

UNIVERSITY OF PRETORIA

TERMINOLOGICAL ISSUES IN THE TRANSLATION OF CHEMISTRY TERMS FROM ENGLISH TO NORTHERN SOTHO

M.J NCHABELENG

2011

Mini-dissertation submitted in partial fulfilment of the requirements for the degree MA (African languages) Course work

Supervisor: Prof. Elsabé Taljard
Department of African Languages
University of Pretoria

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

1.1 Background to the study

In the late nineties a project was started by the *Suid-Afrikaanse Akademie vir Wetenskap en Kuns* that had as one of its aims the compilation of a bilingual (English/Afrikaans) explanatory chemistry dictionary. The identified target users of this dictionary were senior secondary school learners and undergraduate chemistry students. Subject field experts, in cooperation with a team of lexicographers, compiled a lemma list of 500 chemistry terms. The English definitions of these terms were written first followed by the equivalent Afrikaans terms and definitions.

The *Akademie* then approached the Department of African Languages at the University of Pretoria to assist with the translation of the terms and their definitions into Northern Sotho (as representative of the Sotho Languages) and Zulu (as representative of the Nguni Languages). This project was subsequently incorporated into the postgraduate program of the Department of African Languages.

Each student participating in the project received 50 terms and their definitions in English (and Afrikaans). The final aim is to have these terms translated into Northern Sotho and Zulu and simultaneously, compile a list of chemistry terms that can eventually be standardized by the relevant language bodies.

The value of this project is that Northern Sotho and Zulu learners will have access to terms and concepts in their mother tongue. In some cases term equivalents already exist in Northern Sotho, but the learner is unaware of this. A case in point is the term 'potassium permanganate'. Learners do not make the connection between this term and the Northern Sotho word *kgakgakga*, something that most of them are familiar with, as it is a substance that is traditionally used to repel reptiles, especially snakes, during the summer.

Aside from playing a major role in the conceptualization of basic chemistry concepts, this project will also make mother-tongue speakers of Northern Sotho aware of the fact that a number of scientific terms already exist in their language. This would help to dispel the notion that Northern Sotho cannot be used as a language of science.

1.2 Aims and objectives

The aim of this mini-dissertation is to give an account of the terminological processing of a technical text before any translation work, with specific focus on the terminological problems experienced by the translators when translating technical texts from English as a source language, to a language of limited diffusion such as Northern Sotho. It needs to be stated clearly that technical translation must be preceded by proper terminological groundwork in order to ensure a final product of high quality.

The description of the different strategies and procedures that have been devised are investigated to ensure that the final translation has a sound terminological basis. The study therefore focuses on the terminological groundwork prior to translation, and not on the translation itself.

It further aims to contribute to the terminological development of Northern Sotho, as a term list of Chemistry terms will be produced as a by-product of the translation. This list will be submitted to the National Language Body (NLB) for Northern Sotho for standardization purposes.

1.3 Procedures and strategies

As a first step all terms appearing in the definitions, which are to be translated, need to be isolated or extracted, before Northern Sotho term equivalents can be sought. Term extraction is done both semi-automatically and manually.

Secondly, different sources of Northern Sotho term equivalents must be consulted to find translation equivalents for the extracted terms. Possible sources of term equivalents include Language for General Purposes (LGP) dictionaries, Language for Special Purposes (LSP) dictionaries, as well as informal term lists compiled by individuals, of which a surprisingly large number exists.

In view of the previous terminological marginalization of Northern Sotho, it is to be expected that the existing terminological sources will be inadequate in providing term equivalents for all the isolated terms. Consultation with subject field experts who are mother-tongue speakers of Northern Sotho is therefore necessary. The input of these field experts is three-fold: First, they may be able to provide term equivalents in cases were no translation equivalents can be found in existing terminological sources; secondly, they will be able to give guidance to the preferred term in cases where more than one term equivalent exists; and thirdly, the syllable structure which is best preferred. In cases where a term needs to be coined, they can assist in explaining the concept to which the term refers and in coining an appropriate term.

Only after having compiled a detailed term list can the translation process commence. Having a term list ensures terminological consistency in the translation – something that is of utmost importance in technical translation.

1.4 Discussion of relevant concepts and procedures as explained in the literature

1.4.1 Technical translation

Technical translation involves translation of texts belonging to technical fields or specialized subject fields. Technical texts are characterized by the use of LSP, i.e. a language that is used to discuss specialized fields of knowledge. Bowker and Pearson (2002:27) point out that any LSP characteristically contains specialized vocabulary, collocations and highly specific stylistic features, all of which need to be taken into consideration when translating a technical text.

It is common knowledge that the translation of technical texts into Northern Sotho (or any other African language) is problematic due to a lack of properly standardized terminology. One of the aims of this study is to indicate that the lack of terminology can be compensated for by the proper and intensive mining of all possible sources of terminology. It needs to be mentioned that language proficiency and cultural competence in both languages are of utmost importance when one attempts to do technical translation. Computer literacy is also a necessity as research needs to be done and records need to be saved.

1.4.2 Term extraction

As was pointed out previously, the translation of a technical text must be preceded by proper terminological groundwork. This implies that before translation can commence a list of terms, which appear in the source text, must be compiled. The compilation of such a list implies that the relevant terms must be extracted from the source text. This can be done semi-automatically or manually. These terms must then be provided with term equivalents in the target language. The end product of both processes will be a glossary, containing the terms extracted from the source text with their equivalents in the target language.

a. Semi-automatic term extraction

Semi-automatic term extraction implies access to electronic corpora. Dedicated computer software is then used to extract terms from the corpus. The fact that the method is called 'semi-automatic' does not mean that the computer can fully automatically extract terms on its own. What the computer software initially extracts are only term candidates. Candidates may not all be terms. Therefore, a person, preferably a subject field expert, needs to go through the list to isolate the real terms. Also, some terms from the corpus may have been overlooked by the programme and will not appear

on the list of candidates. Thus a person must verify all work done by the computer. For these reasons it is called 'semi-automatic' and not fully automatic.

b. Manual term excerption

As in the case of semi-automatic term extraction, manual term excerption also presupposes access to a corpus; however, for manual excerption, the corpus does not have to be in electronic format. It can consist of manuals, textbooks, magazines, reports, etc. Terms are manually excerpted by going through the corpus and physically reading through it in order to identify terms which are relevant to that specific subject field, in this case Chemistry.

To sum up, manual term excerption simply means reading through a text and then identifying the source terms. The obvious disadvantage of this method is that it is time-consuming and the margin for human error is rather big.

Semi-automatic term extraction and manual term excerption are not mutually exclusive. They can both be used to produce an accurate term list. As mentioned earlier, after extracting candidate terms semi-automatically, one needs to physically go through the list in order to extract real terms and leave out non-terms. On the other hand, the computer may have left out some real terms from the corpus, and one needs to go through the corpus manually in search of 'left-out' terms. Thus both manual and semi-automatic term extraction methods will be used in this project.

1.4.3 Sources of terminology

Sources of terminology are divided into two categories: i.e. treated terminology and untreated terminology.

Treated terminology has terms that belong to the LSP and have been lemmatized in a proper manner. Some of the treated terminologies have definitions and translation

equivalents. Even though treated terminology is found in LSP dictionaries, some terms can also be found in LGP dictionaries.

Untreated terminology refers to those terms that form part of any spoken or written material in a specific subject field. These terms have not yet been isolated according to certain criteria, and are thus regarded as new terminological material. Untreated terminology can be found in textbooks, information brochures, question papers, newspapers, etc. Both treated and untreated terminology need to be mined in the search for terms.

1.4.4 Term formation strategies

A distinction is made between two strategies for the formation of new terms: i.e. primary term formation and secondary term formation.

Primary term formation means that the terminologist identifies a new concept and then searches for a term suitable for that concept. This approach is termed the onomasiological (naming) approach. Thus primary term formation accompanies concept formation and is therefore monolingual.

Secondary term formation occurs when a new term is created for a known concept. This can happen as a result of knowledge transfer from one linguistic community to another or when there is a revision of terminology. Due to the nature of the project, secondary term formation is the strategy which is relevant for this study, since it concerns the transfer of knowledge from the English speaking community to that of the speakers of Northern Sotho.

1.4.5 Term formation strategies in the African languages

Term formation strategies are especially relevant with regard to secondary term formation because the concepts are already in existence; what is needed is a term in a specific language in question. A thorough knowledge of these processes is needed for the coining of proper terms.

1.4.5.1 Language internal word-formation strategy

This includes the following:

a. Semantic transfer

Semantic transfer is the process of attaching new meanings to existing words by modifying their semantic contents. This can turn an LGP word into an LSP term. Semantic transfer is divided into two categories: i.e. semantic specialization and semantic generalization.

Semantic specialization takes place when a word from the LGP attains the status of an LSP term. An example in this regard is the LGP term *kgokong ye ntsho* 'black hartebeest'. When used in the LSP of traditional medicine, it becomes a term referring to *muti* made from human body parts.

Semantic generalization is the opposite of semantic specialization. This often takes place in the case of brand names which are generalized to refer to a whole class of objects. An example is the brand name *Colgate*, which is used to refer to any toothpaste, regardless of the specific brand.

b. Paraphrase

Paraphrase is a way of explaining or describing a concept by making use of a phrase or even a sentence. For example: *monyanya wa Paseka* for 'Passover'. The main disadvantage of paraphrasing is that it is not economical when a phrase has to be used repeatedly.

c. Compounding

During compounding a new term is formed by joining more than one word e.g.
tšwela + pele = tšwelopele 'progress' or *dikarolo tša polelo = dikarolopolelo* 'parts of language or speech'.

1.4.5.2 Borrowing

Borrowing is the process whereby linguistic elements are taken over from one language to another. Different types of borrowing are:

a. Loan words/foreign words

Loan words are borrowed as wholes, both morphologically and phonologically and their meanings are retained intact, e.g. T.V., remote, dairy.

b. Transliteration/adoptives

Adopted words are completely adapted to the language system of the borrowing language: syntactically, morphologically, phonologically and tonologically. The spelling is adapted to suit the syllable structure of the borrowing language. Examples are:
'machine' – *motšhene*
'stove' – *setofo*

1.5 OUTLINE OF CHAPTERS

Chapter 2 deals with the isolating of terms from the source text. Semi-automatic and manual term extraction procedures are described and the results of these procedures are listed in the form of tables. Both single word and multiword terms are treated.

In chapter 3 the provision of term equivalents for the extracted source terms is discussed. The difference between treated and untreated sources of terminology is described and the procedure for the sourcing of term equivalents from both standardized and non-standardized sources is described. A brief discussion of the current standardization procedure of terminology in South Africa is also included. Chapter 4 is devoted to the different term formation strategies which are available to the Northern Sotho terminologist for the coining of new terms, whereas chapter 5 investigates the utilization of users' preferences as a means of preliminary term standardization. Chapter 6 contains the conclusion.

CHAPTER 2 SEMI-AUTOMATIC AND MANUAL TERM EXTRACTION

2.1 Introduction

In this chapter term extraction of source terms from a technical text is explained. The two major methodologies, i.e. semi-automatic term extraction and manual term excerption are explained and applied to the source text. Thereafter follows the results of the application of these two procedures for terms to be extracted from a technical text on chemistry, prior to the translation of the text.

It should be clearly stated that before any translation of a technical text can be attempted, there should be proper terminological groundwork. Firstly, a term list should be compiled by extracting all terms from the source text in order to provide term equivalents in the target language for these source terms. The value of such a term list, which in this case, would be a bilingual English – Northern Sotho one, is to ensure terminological consistency in the target text. After having gone through the proper standardization and validation procedures, such a list could also contribute to terminological development, and thus be available for future translation work.

2.2 Isolating terms in a text

Before trying to isolate terms from non-terms, one should be sure of which words are terms and which are not. It is usually not an easy thing to do but it is of utmost importance as some words may pose as terms even though they are not. This is because some terms may, in some instance, lose their term status and become words, whereas words may sometimes gain a term status in some particular fields. In ‘sports language’ the word ‘field’ is a term but in general language, it is just a word which refers to an open space of land.

A term can be defined as a linguistic label for a concept belonging to a specific subject field or domain. Furthermore, a concept is regarded as an abstract entity that exists in people's minds. Concepts need to be created or come to exit, and may be material or immaterial. Terms are therefore symbols that represent objects (created concepts). This suggests that there should be a concept before a term can be formed.

There is a difference between a word and a term. Terms are generally associated with LSP (language for special purpose). A word is general and can be said to form part of LGP (language for general purpose). LGP is a language used for everyday communication, by people of diverse social or occupational backgrounds in different situations. This is a language that can be understood by everyone, child and adult, literate and illiterates, etc. It has no boundaries and it can be called simple basic language. On the other hand, LSP cannot be understood by anyone. It needs to be kept in mind that the demarcation is not an absolute one, and that there is considerable overlap between these concepts.

Terms are furthermore associated with special subject fields. Examples of special subject fields or domains are literature, physics, chemistry, medical field, agricultural field, computer science, etc. There is no clear-cut difference between subject fields, they usually overlap. An overlap between different subject fields implies an overlap in terminology. Therefore, a term like 'climate' can be found in agriculture, geography and even in the medical field.

Terms can be single words e.g. atom, iron or multiword terms e.g. electrical circuit, magnetic field etc. Terms may also be acronyms (SADTU, PanSALB), abbreviations (DP, IFP), chemical symbols (H_2O = water) or formulas (Area = L x B).

In order to have a term list, terms are extracted either manually or semi-automatically from a source text. It should be stated that manual and semi-automatic term-extraction supplement each other. During manual term excerption, the terminologist may overlook some terms, but they can be recovered during semi-automatic term extraction. On the

other hand, the computer software may extract non-terms which implies that the terminologist has to go through the term list and manually detect the non-terms which had been incorrectly extracted and delete them from the list of candidate terms. The standard procedure is to do the computer – assisted semi-automatic term extraction first, then follow it up with manual term excerption.

A technical translator should be able to differentiate between an ordinary word and a term. Some words function as terms when used in specific domains and remain ordinary words when used in general domains. According to Pearson, (1998:1), "terms are labels of concepts, which are abstract entities, isolated from the text". Pearson suggests that the communicative setting may be inspected to determine whether a lexical item functions as an ordinary word or as a term. This suggests that the terminologist should always consult with the subject-field experts in order to produce a proper term list.

2.3 Semi-automatic term extraction

In this paragraph, the process of semi-automatic term extraction is explained with particular reference to the computer software used to extract term candidates during semi-automatic term extraction. When using electronic devices (computers) a corpus is used from which terms are extracted. According to Bowker and Pearson (2002:1) a corpus can be described as "a large collection of authentic texts that have been gathered in electronic format according to a specific set of criteria". A text can be called authentic if it represents the real, daily language between people. A text is in an electronic format when it is machine readable and can be processed electronically.

A text in electronic format (a corpus) has two advantages according to Taljard (2004: 11). She points out that it can be queried by dedicated corpus analysis tools. These tools help to manipulate data and allow access to the information contained in the texts. This information can be displayed in a variety of ways, i.e. alphabetically, according to frequency, keyness etc. A second advantage is that electronic texts can be gathered and

consulted more quickly than printed texts. With the help of the Internet, texts can be searched for and downloaded in a matter of seconds.

There are, in particular, two main types of corpora, i.e. general / reference corpora and special purpose corpora. According to Bowker and Pearson (2002:11) there are many different types of corpora that can be utilized in research, e.g. general reference corpora versus special purpose corpora, written versus spoken corpora, monolingual versus multilingual corpora, synchronic versus diachronic corpora, open versus closed corpora, etc. For the purpose of our study, the difference between general reference corpora and special purpose corpora will briefly be discussed. According to Bowker and Pearson (2002:11) a general reference corpus is one that can be seen as being representative of a given language as a whole. It can therefore be used to make general observations about that particular language. This type of corpus typically contains written and spoken material, a broad cross-section of texts types, e.g. newspapers reports, fiction, radio and T.V. broadcasts, and focuses on language for general purposes.

A special purpose corpus is one that focuses on a particular aspect of language. It could be restricted to the LSP of a specific subject field e.g. chemistry, linguistics, biology, etc. It could also be restricted to a specific text type or a particular language variety or a language used by a certain demographic group, e.g. young urban blacks. This kind of corpus cannot be used to make general observations about language, but it can be used in a comparative fashion to identify those features of a specialized text that differ from general language. A special purpose corpus is usually smaller than general reference corpora and is usually designed and compiled by terminologists and translators for specific terminological purposes. In order to semi-automatically extract terms from a source text, a general reference corpus and a special purpose corpus are needed.

For the purpose of this study, the existing English definitions of the 50 chemistry terms which is the source text that is to be translated into Northern Sotho, automatically constitute the special purpose corpus, whereas the University of Pretoria English Internet Corpus (PEIC) is used as the general or reference corpus. The special purpose corpus

consists of 1225 tokens and 158 types, whereas the reference corpus has approximately 12, 5 million tokens and 118 193 types.

The computer program used for this process is WordSmith Tools; available from http//www.oup.co.ut/elt//catalogue/multimedia/WordSmith Tools 3.0. Although the initial purpose of this program was not aimed at term extraction, it has been proven by Taljard and De Schryver (2002:51-54) that the Wordsmith Tools KeyWord function is quite effective in extracting terms from electronic texts.

As a first step, the Wordlist function of Wordsmith Tools is used to compile a single word frequency list of both the general purpose corpus and the special purpose corpus. These wordlists are then used as inputs for the KeyWord function, which is used to semi-automatically extract single-word terms from the special purpose corpus. Scott (1997: 236) defines the term 'KeyWord' as 'a word which occurs with unusual frequency in a given text'. Unusual frequency can be related to outstandingness and implies that a word has an unusually high (or unusually low) frequency in a text (or sub-corpus) in comparison to its occurrence in a reference corpus of some kind.

The application of the keyword function to our electronic data resulted in 81 term candidates being isolated from the special corpus. See **Table 1** which constitutes a list of these term candidates. It needs to be noted that symbols, e.g. H (for hydrogen) and N (for nitrogen) and abbreviations will also be thrown up in such a list.

TABLE 1 **Term candidates and their frequency of appearance**

No.	Term candidate	Frequency	No.	Term candidate	Frequency
1	A	70	41	electron	6
2	acid	3	42	electroplating	2
3	an	22	43	element	5
4	aqueous	2	44	endothermic	2
5	atoms	7	45	enthalpy	3

6	atomic	4	46	formula	4
7	atoms	4	47	G	21
8	bond	4	48	gas	5
9	carbon	3	49	H	5
10	cell	5	50	HCI	2
11	chemical	7	51	homogenous	3
12	CI	3	52	hydrogen	5
13	colloidal	3	53	ions	8
14	compound	7	54	is	32
15	concentration	3	55	liquid	5
16	current	5	56	molecule	4
17	debye	4	57	molecules	3
18	decomposition	3	58	monoxide	2
19	density	4	59	N	46
2	diatomic	2	60	negative	5
21	diffusion	3	61	orbital	2
22	dilute	3	62	oxidation	2
23	dimmer	2	63	particles	3
24	dipole	4	64	periodic	3
25	diprotic	3	65	point	7
26	dispersion	4	66	process	6
27	dissociation	2	67	protons	6
28	dissolved	5	68	reaction	7
29	distillate	2	69	separation	2
30	E	23	70	solute	2
31	electric	6	71	solution	6
32	electrochemical	3	72	solvent	3
33	electrode	2	73	spin	4
34	electrodes	4	74	substance	11
35	electrolysis	3	75	sulphate	2

36	electrolyte	3	76	symbol	7
37	electrolytic	3	77	synonym	3
38	electromotive	2	78	titration	2
39	electron	11	79	unit	4
40	electronegative	3	80	V	11
			81	water	9

The word 'term candidate' is used in this regard, because not all words thrown up by the keyword search are necessarily real terms. The only way to separate the terms from non-terms is through manually scanning of the candidate list and the deletion of the non-terms. Contrary to what one may think, it is quite easy for a non-specialist to separate terms from non-terms. On the other hand, the terminologist always has the option of consulting with subject field experts whenever there is confusion concerning the term status of a particular term candidate. Due to the density of the text, it was found that very few of the term candidates offered by the search were not actually terms. The term status of 71 terms was confirmed by manually scanning the keyword list. See the list of terms as shown in **Table 2**.

TABLE 2 **Semi-automatically extracted terms**

Term	**Term**	**Term**	**Term**
acid	diatomic	electrolytic	molecules
aqueous	diffusion	electromotive	monoxide
atom	dilute	electron	negative
atomic	dimer	electronegative	orbital
bond	dipole	electroplating	oxidation
carbon	diprotic	element	particles
cell	dispersion	endothermic	periodic
chemical	dissociation	enthalpy	point
CI (symbol)	dissolved	formula	symbol

colloidal	distillate	gas	synonym
compound	E (symbol)	H (symbol)	titration
concentration	electric	HCI (symbol)	unit
current	electrochemical	homogeneous	water
debye	electrode	hydrogen	
decomposition	electrolysis	ions	
density	electrolyte	liquid	

As was pointed out above, terms do not only consist of single (orthographic) words; therefore one also has to make provision for multiword terms. The keyword function can also be run to isolate two-word terms from an electronic text. Therefore, the same procedure is followed with regard to two-word terms. In this case, the candidate list offered 72 candidates. The manual scanning of the list resulted in identification of 18 two-word terms, which are listed below:

TABLE 3 **Two-word terms**

Term	**Term**
atomic number	electromotive force
carbon monoxide	electron configuration
chemical compound	electron density
copper sulphate	empirical formula
dipole moment	end point
dissolved substance	enthalpy change
electric current	homogeneous mixture
electrochemical cell	metallic element
electrolytic cell	periodic table

The total number of terms extracted semi-automatically therefore comes to 71 + 18 = 89. These terms will form the basis of the bilingual term list which will be used in the actual translation of the definitions. However, semi-automatic term extraction does not succeed

in extracting all terms from the source text. According to Taljard and De Schryver (2002:56), semi-automatic term extraction has a recall rate of approximately 60% of terms in a running text. Therefore, computational extraction needs to be supplemented by manual term excerption.

2.4 Manual term excerption

Manual term excerption implies close scrutiny of a text in order to identify terms which are relevant to a specific subject field. Here, one has to physically read (scan) through a text (or corpus) and mark a term if identified. This is done, because the purely mechanical processing of the text by means of the relatively unsophisticated software that we use is not able to isolate all terms. Even though the computer software can facilitate the process, human judgment cannot be disregarded.

Before the advent of electronic text corpora and the availability of computer software to manipulate and search these corpora, manual term excerption used to be the standard terminological procedure. As a matter of fact, this method of term excerption is still used by many terminologists and (technical) translators in South Africa. Although a tried and tested method, manual term excerption is not without potential pitfalls, e.g.

- It is time consuming to go through a text and excerpt terms.
- It is labour-intensive and potentially repetitive because of the need to recover different kinds of terminological data.
- Human readers are prone to mistakes - they can miss out on terms or take incorrect decisions as regards to possible term status of items.
- It becomes easier if the translator is also a subject field expert, which is practically impossible. Therefore the translator has to work in close cooperation with the subject field experts, which will consume time and other resources.
- Specifically with regard to the South African situation, there is often a lack of commitment among subject field specialists who are also mother tongue speakers of African languages to develop their own languages. They are often unwilling to assist with the identification of terms on a list of potential terms.

Using manual term excerption in combination with semi-automatic term extraction by means of computer software, however, gives terminologists and translators the benefits of both methods. Doing manual term excerption after the computational term extraction also drastically reduces the time that has to be spent on the identification of terms.

An example of one of the definitions is shown below to illustrate how terms are manually excerpted from definitions.

> **Endothermic** (*a*)
> A chemical reaction is endothermic when energy is absorbed from the environment, e.g. the formation of carbon monoxide and hydrogen from coke and water; the formation of magnetite from iron and water.

Terms that were identified semi-automatically are the following (Compare Tables 2 and 3 in this regard):

1. chemical
2. reaction
3. endothermic
4. carbon monoxide
5. hydrogen
6. water

Manual perusal of the text (in this case, the definition) resulted in the identification of 6 extra terms, i.e.

1. chemical reaction
2. energy
3. environment
4. coke
5. magnetite

6. iron

These 6 terms were overlooked by the computer software and could only be manually excerpted. The terms manually excerpted should therefore be added to the list of terms, resulting in a total of 12 terms being excerpted from the example definition.

For the special corpus in its totality, manual term excerption resulted in the identification of 40 single word terms and 19 two word terms, i.e. 59 terms in total which were not picked up during semi-automatic term extraction process. In the process of manual extraction, both single and multiword terms were extracted. Table 4 contains the terms manually excerpted from the definitions.

TABLE 4 **Terms manually extracted from definitions**

Term	**Terms**	**Term**	**Term**
motion	charged particles	overall	magnetite
electric circuit	emission	conductors	strength
time unit	carbonated water	via	reduce
decomposition reaction	organic compound	electrochemical reaction	ethylene
decompose	trivial name	cations	resultant mixture
coke	ethoxy name	cathode	chemical reaction
hydration	dissociation	anions	radio waves
volume	iron	anode	region
hydrogen chloride	control rods	molten	layer
diamagnetism	nuclear reactors	immersed	stable
property	lantherides	propagated	zero
magnetic field	radioactive	X-ray	electrons
diamagnetic	emulsify	visible light	atoms
cyclotron	donating	gas exchange	electrodes

device	actinides	energy	
accelerate	equivalent point	movement	

After manual term excerption, the total number of raw terms to be translated is 92+59=151. At this point, it was necessary to carry out a basic lemmatization procedure, since the terms isolated thus far are raw terms, i.e. some terms which are nouns, appear in the plural form. In other cases, both the singular and plural forms were thrown up as terms. Nouns were therefore lemmatized under the singular form only. Lemmatization resulted in a final list of 148 source terms (**electrodes, electrons** and **atoms** were excluded as their singular forms are already included in the list that was semi-automatically excerpted), which will form the basis of the bilingual glossary. The next step in the terminological processing of the source text is the provision of term equivalents for the 148 terms in Northern South. This process will be discussed in the next chapter.

2.5 Summary

In this chapter, term extraction as terminological pre-processing of technical texts to be translated was explained. It was indicated that both semi-automatic and manual term extraction need to be done to extract the maximum number of terms from the text prior to its translation. It is clear that these two procedures are interdependent and that semi-automatic term extraction can reduce the amount of manual input, but that it cannot completely eliminate the need for human intervention.

In the following chapter, focus will be on treated versus untreated terminology, standardization of terms and finding translation equivalents for the extracted terms in both standardized and non-standardized sources.

CHAPTER 3 PROVISION OF TERM EQUIVALENTS: CONSULTATION OF TREATED AND UNTREATED SOURCES

3.1 Introduction

In the previous chapter, the first phase in the terminological processing of the source text, i.e. the extraction of terms was explained. The next step in the processing of the source text is the finding of term equivalents in Northern Sotho for the source terms by mining all possible sources of terminology.

Provision of term equivalents in Northern Sotho for the source terms will result in a bilingual (English: Northern Sotho) glossary to be used as a translation aid during the translation process. This is necessitated by the fact that very little standardized chemistry terminology is available for Northern Sotho - what is available is mostly restricted to a rather outdated Terminology and Orthography List published in 1988, non-standardized term lists and glossaries compiled by individuals working in the specific subject field, and a few LGP dictionaries as sources of technical terms.

The bilingual glossary forms the basis for an internally standardized term list to be made available to all translators who participate in the larger project. Having a standardized list ensures terminological consistency within the translations of the different translators. It will furthermore be submitted to the NLB as standardizing body for official standardization, thus contributing to terminology development on a larger scale.

3.2 Treated versus untreated terminology

After all terms have been extracted (both manually and semi-automatically) one has to find their translation equivalents in Northern Sotho. A language is often gauged against other languages with regard to its richness in vocabulary and its ability to coin new terms. There are different sources that can be consulted to finally get the translation equivalents of the terms in question.

According to Taljard (2004:70) there are two main types of sources of terminology, i.e. treated and untreated terminology. Sources such as term lists, dictionaries and data banks in which terms appear, are regarded as treated terminology. Terms appearing in these sources have already been identified as belonging to a certain subject field; they have been lemmatized in a proper manner and in some cases might even be linked to translation equivalents and/or terminological definitions. Language for special purpose (LSP) dictionaries is one of the prime sources of terminology, but terms are also found in Language for General Purpose (LGP) dictionaries.

According to Taljard (2004:7) untreated terminology refers to terms appearing in any kind of spoken or written material on a specific subject field. Terms appearing in these sources have not been isolated or identified in any way and can be regarded as raw terminological material. Examples of raw terminological source are textbooks, pamphlets, information brochures, question papers, study materials, radio/ T.V. broadcasts, internet material, etc. Any material carrying information on a specific subject field falls under untreated terminology. Mining sources of untreated terminology for term equivalents therefore requires a concerted effort from the translator/ terminologist to find these term equivalents, and to evaluate them for appropriateness.

The main difference between treated and untreated terminology centres around the issue of standardization. It is usually assumed that terms found in dictionaries and official term lists have been standardized by the official body responsible for standardization. As will be pointed out below, this is not always true within the South African context. From the information provided in the table 5 below, it is clear that standardization process in South Africa is rather flawed, and that this leads to the proliferation of the terms. On the other hand, should a translator / terminologist opt for a term equivalent sourced from an untreated source, he /she should be aware of the fact that such a term has not been standardized, and that it may – during the later standardization process – not be recognized as a standard term. Since the issue of standardization of terms is central to the

sourcing of term equivalents for the source terms in any given terminological activity, a brief discussion of the standardization process will be given in the paragraph to follow.

3.3 Standardization procedure in South Africa

Before the democratic dispensation in South Africa, only two languages (English and Afrikaans) were given official status. According to section 6 of chapter 1 of the Constitution of the Republic of South Africa; 1996 (Act 108 of 1996), Sesotho sa Leboa (Northern Sotho), Setswana (Tswana), Sesotho (South Sotho), Tshivenda (Venda), XiTsonga (Tsonga), Afrikaans, English, SiSwati (Swati), IsiNdebele (South Ndebele), IsiXhosa (Xhosa) and IsiZulu (Zulu) are regarded as official languages.

In accordance with the Constitution (Act 108 of 1996) and the National Language Framework Policy (NLFP) and Implementation Plan [5], the government aims to promote South Africa's linguistic diversity and to achieve this by means of an approach of functional multilingualism. One of the aspects of the promotion of the functional multilingualism is the development and standardization of terminology for all the official S.A. languages. The National language bodies which fall under the auspices of PanSALB is one of the most important role-players in terminology development. Unfortunately, these bodies do not seem to be functioning at full capacity. Members of these NLBs work on a voluntary basis, most of them having another full time occupation. This implies that members can only do National Language Board (NLB) work in their spare time. The NLBs have until recently, not undergone formal terminological training and thus have had to rely on their mother tongue intuition and any other incidental linguistic training they may have had. As a result, decisions taken with regard to terminology are often idiosyncratic and not linguistically well motivated. This issue has however in the meanwhile been addressed, by providing terminological training to the NLBs. This training is done by experts from PanSALB. Dr Alberts presented all the terminology training at PanSALB.

The actual standardization process is done by the Terminology Coordination Section of the National Language Services, which falls under the auspices of the Department of Arts and Culture. According to Alberts (2008) the Terminology Coordination Section (TCS) manages terminology as follows:

- The terminographers excerpt terminology in the source language (SL) which is usually [but not necessarily] English.
- The terms are then supplied with definitions, example sentences and relevant information in the source language.
- These terms and relevant information are discussed with subject field specialists to confirm the contents.
- After the SL terms and relevant information have been finalized, the information is translated into the target language [TL].
- The information in the TL is also discussed with collaborators and subject field specialists before the final terminological product, which may be a dictionary or a term list.

After finalization of a specific terminology list, the relevant NLBs of PanSALB are requested to verify and authenticate the terminology. The relevant terminologists then change the database according to suggestions.

After the database is finalized, the relevant terminology list can be published and disseminated to target users. The procedure above is also followed for standardization purposes.

Apart from the shortcomings mentioned above, the procedure for term standardization seems sound. However, the practical implementation of this procedure seems inadequate. In the first instance, dissemination of terminological products such as term lists and glossaries does not seem to be very efficient. Schools, especially secondary schools and tertiary institutions would be a logical distribution point for these products. Distribution of these products should furthermore be linked to an awareness campaign making

potential target users aware of the existence of such products. Production costs, which are often cited as a reason for inadequate dissemination could substantially be lowered by making these products available electronically, preferably via a website.

Secondly, language practitioners, specifically translators are in the course of their activities, continuously creating terms. These are often documented in the form of glossaries, but for personal use by the translator only. Language practitioners are generally unaware of the possibility of having these terms standardized. Ideally, language practitioners should be encouraged to submit these glossaries to either the TCS or the relevant NLB for standardization purposes. After having been standardized, these terms should then be added to the terminological database. However, the functionality of such a database is directly linked to its accessibility to target users. Serious considerations should therefore be given to setting up a terminological database, which is freely available to target users.

The functionality and efficiency of the standardization procedure is of particular importance to this study. After having completed the bilingual term list, the aim is to submit this list to the NLB for Northern Sotho, via PanSALB's head office. The NLB will be requested to validate the terms, both content-wise and spelling-wise. Where multiple translation equivalents exist, the idea is to retain all possible TEs, leaving the choice for the standardized one up to the members of the NLB. However, the list will be organized in such a way that the most preferred terms are put first, as indicated by subject field specialists and educators.

3.4 Provision of term equivalents from standardized sources

The first preference when searching for a term equivalent should be perusal of standardized sources. Sources like dictionaries and official Terminology and Orthography lists should be the first to be consulted. A term is standardized when it is linked to a particular concept and its meaning and definition is well documented in an accepted terminological document of that particular language. It is assumed that if a term

does not appear in some official terminological document such as the Terminology and Orthography of that particular language, the term has not yet been standardized, or may be standardized but not yet documented.

Due to the shortage of terminological sources in Northern Sotho for chemistry, both standardized and non-standardized sources (e.g. textbooks, journals, any written document, etc.) had to be consulted. For the purpose of this study, a number of existing dictionaries and the Terminology and Orthography (1988) are regarded as standardized sources. Term lists compiled by individuals are regarded as non-standardized sources.

The following sources are used as examples of standardized sources;
1. South African Multi-language Dictionary and Phrase Book [READERS DIGEST] SAMDPB.
2. Learners English-Northern Sotho Dictionary (LD)
3. Terminology and Orthography (Northern Sotho) 1988 (T & O)

TABLE 5 **Source terms and their equivalents from standardized sources**

	Source term	**SAMDPB**	**LD**	**T & O**
1	acid	esiti	-	esiti, sedilana
2	accelerate	-	-	akgofiša
3	atom	-	atomo	atomo
4	bond	setlamo	kamanyo, tlemagano	pofo, tlamo, kgwerano
5	carbon	-	-	khapone
6	cell	sele	sele, okwana	sele
7	chemical	-	sa chemise, khemikhale	khemikhale
8	colloidal	thulano	-	-
9	compound	-	-	tlhakantšhetšo
10	concentration	-	motswakoti	motswakoti
11	current	moela	moela	moela

12	cyclotron	sesesedi	-	-
13	decomposition	polo	kamologano	popolo, tlharamollo
14	density	kitlagano	kitlagano	kitlano, teteano
15	diffusion	-	-	phatlalatšo, kgašanyo
16	dilute	hlaphola	-	hlaphola
17	dispersion	phatlalatšo	phatlalatso	tšitlano, phatlalatšo
18	dissociation	kgaogano	tlogelano	tšhwalalano, tlogelano
19	dissolved	tologa	nyaoša, nyaoga	nyaoša, nyaoga
20	distillate	hlwekiša	hlwekiša	sekiša
21	electric	mohlagase	mohlagase	mohlagase
22	electron	-	-	elektrone
23	element	element	-	setho, tokollo, elemente
24	energy	-	-	maatla, mooko, mafolofolo, enetši
25	formula	-	fomula, kaelo	fomula, kaelo
26	gas	gase	gase, kese, moya	gase, kese
27	hydrogen	-	-	haetrotšene
28	immersed	inetše	-	-
29	iron	tshipi	-	-
30	liquid	seela	seela	seela
31	magnetic field	-	-	lepatlelokgogedi, karolokgogedi
32	molecule	-	molecule	molecule
33	motion			tšhikinyo, tšhišinyego
34	movement	-	-	tshepelo, tšhišinyego, tšhikinyego
35	negative	-	-	ganetšago, latolago
36	particle	-	karolonyana, seripana, lerathana	seripana, sekgawana,

				setsekana
37	periodic	-	lebaka, paka, nako, sebaka	ka mabaka
38	point	ntlha	ntlha, khutlo	khutlo, ntlha
39	process	-	-	tshepetšo
40	property	-	-	seeng
41	protons	thoto	-	-
42	reaction	-	phetogo	phetogo
43	reduce	-	fokotša	fokotša
44	region	-	-	selete, tikologo
45	separation	karogano	karogano, kgaogano	karogano
46	solution	-	-	setološwa
47	solvent	-	setološi, setološa	setološi, setološa
48	stable	se sa fetogego	-	-
49	strength	maatla	-	-
50	substance	selo	-	selo, nto
51	symbol	sešupo	-	seka, sešupo
52	synonym	-	-	lehlalošetšagotee, sinonime
53	unit	botee	-	motšo, botee
54	volume	bolumo	-	-
55	water	meetse	meetse	meetse
56	x-ray	eksrei	-	-
57	zero	lefela	-	-

From the total of 148 source terms, TEs for only 57 STs could be retrieved from standardized sources. This means that only 39 % of STs can be covered by terms from standardized sources. For 32 terms, i.e. 22 % of the STs, multiple equivalents were found. In some cases [cf. atom, current, liquid, etc] the different sources were in

agreement with regard to term equivalents. In other cases however, cf. dissolved, density, dissociation, etc. different TEs were provided by different sources. This problem persisted and even became more and more problematic as the effort to provide translation equivalents for the source terms progressed.

3.5 Provision of term equivalents from non-standardized sources.

As was indicated in the previous paragraph, translation equivalents for roughly a third of the extracted terms could be gleaned from standardized sources. Therefore, the terminologist / translator has no choice but to also make use of non-standardized sources. For the purpose of this project, untreated sources consisted mainly of informal term lists compiled by individuals working in the field of chemistry, who can also speak Northern Sotho.

From 91 terms of which translation equivalents could not be found from standardized sources, 19 TEs were harvested from non-standardized sources. Table 6 presents a list of those terms and their TEs.

TABLE 6 **Term equivalents from non-standardized sources**

	TERM	**TRANSLATION EQUIVALENTS**
1	anion	ayone
2	anode	anode
3	aqueous	sa meetse
4	carbon monoxide	monokside wa khapone
5	cathode	katode
6	cation	katayone
7	chemical reaction	kgohlagano-khemikhale
8	copper sulphate	salfate ya koporo
9	decompose	bola
10	electrochemical cell	lelahle la mohlagasekhemikhale

11	electrode	electrode
12	endothermic	endothemiki
13	hydration	meetsefatša
14	hydrogen chloride	tloride ya haetrotšene
15	ion	ayone
16	layer	leyara
17	oxidation	oksitšenefatšo
18	solute	setologi
19	sulphate	sulfate

At this stage, the number of terms remaining without TEs has been reduced to 72. Although the terms harvested from non-standardized sources do not have any official status, an analysis of some of these terms reveals that the person(s) responsible for the provision of these terms has a deep-seated knowledge not only of chemistry, but also of the term formation strategies of Northern Sotho. In many cases, the terms succeed in correctly reflecting the concept which is named by the term. This is particularly true for the term equivalents of the multiword terms, where the order of 'base + qualifier' often causes problems when new terms are coined in the African languages. In the case of English source terms, the qualifier precedes the base, e.g. **copper sulphate, carbon monoxide** and **hydrogen chloride,** where **sulphate, monoxide** and **chloride** respectively represent the base forms. When coining new TEs for these multiword units, terminologists often opt for a simple left to right literal translation, as is found for example in the TE for 'carbon dioxide', which is given as *khapone taeoksaete* in the official Terminology and Orthography, 1988. However, according to the grammar rules of Northern Sotho, in this language the base is followed by the qualifier, where the qualifier usually appears in a so-called possessive structure. A term such as *taeoksaete ya khapone* 'dioxide of carbon' which follows this rule, would therefore correctly reflect the concept that carbon dioxide is a specific kind of dioxide, thus aiding the process of conceptualization. This principle has consistently been applied by whoever coined the multiword terms in Table 6. Compare for example the term equivalents for **copper sulphate:** *salfate ya koporo* 'sulphate of copper', **carbon monoxide**: *monoksaete wa*

khapone ‘monoxide of carbon’ and **hydrogen chloride**: *tloride ya haetrotšene* ‘chloride of hydrogen’. However, the suggested spelling of especially some of the transliterated terms seemed not to be in line with the accepted spelling rules of Northern Sotho, and the spelling of some of these terms does not correctly reflect the Northern Sotho pronunciation. It was therefore deemed necessary to adapt the spelling of some of the terms, in order to narrow the gap between pronunciation and orthographical representation. Where possible, the guidelines given in the Terminology and Orthography were followed, even though these rules have serious shortcomings. The spellings of the following terms were consequently adapted:

monokside > *monokosaete*
katode > *khatote*
katayone > *khateayone*
sulfate ya koporo > *salefeiti ya koporo*
electrode > *eleketerote*
endothemiki > *entothemiki*
tloride ya haetrotšene > *tleloraete ya haeterotšene*
molecule > *molekhule*
oksitšenefatšo > *okositšenefatšo*
sulfate > *salefeiti*

It can therefore be concluded that although non-standardized sources of terminology need to be treated with the necessary circumspection, these sources often provide well-formed and appropriate TEs. To standardize these terms will therefore serve in the best interest of terminology development of Northern Sotho.

By utilizing both treated and untreated sources of terminology, TEs for a total of 76 STs could be found. A number of source terms for which no direct translation equivalents could be found represent derived forms of STs for which TEs had already been found. To illustrate: a TE for the term ‘atom’ was found in one of the dictionaries used as sources of terminology. By making use of this information, TEs could be provided for the terms

'atomic' and 'atomic number' as well. The total number of terms coined in this way is 21. Compare Table 7 below for a list of these terms and their equivalents.

TABLE 7 **Coined term equivalents**

	Term	**Term equivalent**
1	atomic	seka-athomo, [-]go ba le athomo
2	atomic number	athomopalo
3	carbonated water	meetse ao a tšhetšwego khapone
4	chemical compound	tlhakantšhetšo ya sekhemise
5	Cl (symbol)	Cl (seka)
6	conductor	sesepediši
7	decomposition reaction	phetogo ya polo
8	dissolved substance	selo se se tološitšwego, setološwa
9	E (symbol)	E (seka)
10	electric current	moela wa mohlagase
11	electrochemical	mohlagasekhemikhale
12	electrochemical reaction	phetogo ya mohlagasekhemikhale
13	gas exchange	neeletšano ya moya
14	H (symbol)	H (seka)
15	HCl (symbol)	HCl (seka)
16	monoxide	monokosaete
17	periodic table	papetla ya pheriotiki
18	radio waves	maphotho a radio
19	spin	dikologa, dikološa
20	time unit	motšonako
21	trivial name	leinatlwaelo

At this stage, the original list of 148 STs which needed TEs had been reduced to 51 STs without TEs.

TABLE 8 **Terms without translation equivalents**

1	actinides	14	dissolution	27	electroplating	40	magnetite
2	charged particle	15	donating	28	emission	41	metallic element
3	coke	16	electric circuit	29	empirical formula	42	nuclear reactors
4	control rods	17	electrolysis	30	emulsify	43	orbital
5	debye	18	electrolyte	31	end point	44	organic compound
6	device	19	electrolytic	32	enthalpy	45	overall
7	diamagnetic	20	electrolytic cell	33	enthalpy change	46	propagated
8	diamagnetism	21	electromotive	34	ethoxy ethane	47	radio active
9	diatomic	22	electromotive cell	35	ethylene	48	resultant mixture
10	dimer	23	electron configuration	36	equivalent point	49	titration
11	dipole	24	electron density	37	homogeneous	50	via
12	dipole moment	25	electron density	38	homogeneous mixture	51	visible light
13	diprotic	26	electronegative	39	lanthanides		

Sourcing TEs for the 148 source terms isolated from the text that needs to be translated, results in 97 term equivalents being found by utilizing a variety of sources. This means that 66% of source terms are covered by mining both standardized and non-standardized sources of terminology. Table 5 reveals one of the challenges facing the terminologist / translator who searches for term equivalents: rather than a lack or absence of TEs, the terminologist is faced with a significant number of cases where multiple TEs exist. This is the direct result of inefficiency in the standardization process of terms in Northern Sotho.

At this stage in the project, there are still 51 terms without TEs. As a third step in sourcing translation equivalents, consultation with subject field specialists is considered.

3.6 Provision of term equivalents by subject field experts

It was obvious from the discussion in paragraph 3.4 that most terms do not have equivalents from formal sources (standardized sources). After having consulted standardized sources, informal sources, (e.g. non-standardized term lists compiled by individuals) were consulted and these terms were added as TEs. After consulting formal and non-formal sources, consultation with subject specialists was regarded as the next option in order to have most, if not all translation equivalents. Consultation with subject field experts who are mother tongue speakers of the target language is an accepted method of finding suitable translation equivalents for source terms.

Initially, people working in medical professions were identified as potential participants due to their knowledge of Chemistry as a science subject. It is assumed that Chemistry formed part of their medical training. Several doctors working around Marble Hall were approached but only one was willing to participate in an interview. Five nurses were also approached and indicated their willingness to participate. One pharmacist was also willing to participate. The above potential participants are all mother tongue speakers of Northern Sotho.

Although the above-mentioned people initially showed their willingness, their participation was unfortunately never realized. Interviews were scheduled but never actually took place. Participants cited time constraints and work pressure as reasons for not honouring their appointments. The only pharmacist that was initially mentioned was offered a position elsewhere before the scheduled date of appointment.

Unfortunately, the whole issue of people not honouring their appointments showed a negative attitude towards the language issue in general. This was noticed in the

comments that potential participants made. On being pressed to provide reasons for their lack of enthusiasm for the project, typical comments were:

- Do you think that this is possible? (i.e. to have chemistry terms in Northern Sotho)
- How long will it take before the use of Northern Sotho chemistry terms can be implemented?
- Do you think our own language will match the world's standard in terminology?

Although the number of participants selected is too small to make a general conclusion regarding the attitudes of Northern Sotho speakers towards the development of technical terminology in their language, it would seem that the terminological development of their language in order to enable it to become a language of scientific discourse does not enjoy priority amongst subject field experts. What these experts do not realize, is that lack of proper technical terminology hampers the effective transfer of knowledge from the English language community to the Northern Sotho language community.

After the unsuccessful attempt to involve subject experts in providing translation equivalents for the remaining 51 terms, other means of finding translation equivalents for these remaining terms had to be found. The only option left at this stage is the coining of TEs. There are several strategies used by terminologists in coining new terms which can be viable in the case of Northern Sotho. These techniques and their applications are discussed in the next chapter.

CHAPTER 4 TERM FORMATION STRATEGIES

4.1 Introduction

In chapter 1 it was stated that terminologists distinguish between two major types of term formation strategies, viz primary and secondary term formation strategies. These two concepts are briefly explained, and from the explanation it will become clear that for this particular project, secondary term formation is the relevant strategy. Therefore, a detailed discussion of secondary term formation is provided, illustrated where possible, with examples taken from the chemistry subject field.

It should be clearly stated that terms cannot be formed in a vacuum. A particular environment is needed where people feel the need to express new ideas and concepts. That environment (situation) may be a laboratory, design office, a warehouse, etc., or anywhere else where new ideas may emerge.

The international organization for standardization (ISO) provides some guidelines for the creation of new terms.

1. Terms should be created systematically with respect their morphological, syntactic, semantic and pragmatic characteristics.
2. Terms should conform to the morphology, spelling and pronunciation convention of the language for which it is intended.
3. Once a term has gained a general acceptance, it should not be changed without compelling reason and a strong certainty that the new term will be accepted as a full substitute.

4. If a new term succeeded only partially in replacing an existing term, the confusion may become worse as this would amount to deliberate synonym creation. In this case, it is preferable to introduce a new term. (Sager 1990:89)

4.2 Term formation strategies

Two mayor types of term formation strategies are distinguished, viz. primary and secondary term formation strategies.

4.2.1 Primary term formation strategy

Primary term formation takes place when a terminologist identifies a new concept which needs to be labeled (given a term). This term formation strategy should follow the guidelines as provided by ISO (Sager 1990:85). This approach is termed the onomasiological (naming) approach. Primary term formation has no linguistic precedent and accompanies concept formation, therefore it is often monolingual. Probably, all languages have an onomasiological origin. Every item which forms part of a specific language group was named according to the language rules of that particular group. The amount of richness in terminology always correlates with the complexity of the life style of that particular language group. The more advanced, educated, adventurous and sophisticated the nation becomes the more terminology it will acquire.

4.2.2 Secondary term formation strategies

This strategy is used when a new term is created for a known concept. This can happen because of knowledge transfer from one linguistic community to another, or when there is a revision of terminology. In secondary term formation strategy there is always the precedence of an existing term with its own motivation. Guidelines are needed when applying the secondary term formation strategy. For this specific project, secondary term

formation is relevant, resulting from the need to transfer knowledge from the English language community to the Northern Sotho community.

The following strategies form part of secondary term formation and a thorough knowledge of them is needed in order to coin appropriate terms.

4.2.2.1 Language internal term-formation processes

This includes:

a. Semantic transfer

Semantic transfer is the process of attaching new meanings to existing words by modifying their semantic content. During this process, a Language for General Purpose word (LGP word) can be changed to a Language for Specific Purpose (LSP) term. The shift is in the reference rather than the sense. What usually happens in these cases is that the original LGP word co-exists with the LSP term; the word being used in general language and the term being used in its specific subject field. In Northern Sotho the word 'element' is translated as *setho* 'member, part (of)'. In the subject field of Chemistry it refers to a group of atoms combined together. Semantic transfer is divided into two categories i.e. **semantic specialization** and **semantic generalization**.

Semantic specialization is the opposite of generalization and it is applied when a word from the LGP attains the status of an LSP term by acquiring an additional, more technical meaning. Compare the following examples:

'current' *moela* general meaning = something that flows; stream
specialized meaning = electric charges in motion

'particle' *setsekana, lerathana* general meaning = a piece; a crumb
specialized meaning = the smallest building block of an organism

'negative' *ganetšago* general meaning = opposing; that which denies

specialized meaning = opposite of being positive.

‘**element’** *setho* general meaning = member; part of
specialized meaning = substance whose atoms all have the same atomic number

Semantic generalization is in a sense the opposite of semantic specialization and entails an upward shift in the semantic hierarchy. It is also known as extension in the sense that the term receives a broader realm of meaning than it originally possessed. This means that the semantic features of an item are extended. Examples that are often cited are cases where brand names are generalized. Within the South African context, the brand name ‘Colgate’ is often used to refer to any kind of toothpaste, regardless of the brand; ‘Omo’ refers to kind of washing powder, etc. In this specific project, semantic generalization did not feature as a term formation strategy.

b. Paraphrasing

Paraphrasing is a way of explaining or describing a concept by making use of a phrase or even a sentence. Although paraphrases can be a very effective term formation strategy, indiscriminate use within the context of a translation could give rise to complicated grammatical constructions that are difficult to interpret due to lack of coherence. The advantage of paraphrases is that they are often self-explanatory and therefore transparent. Compare for example:

‘carbonated water’ *meetse ao a tšhetšwego khapone*. (lit.meaning) ‘water to which carbon has been added’
‘dissolved substance’ *selo seo se tološitšwego*. (lit.meaning) ‘a thing that has been melted / dissolved’
‘charged particles’ *dikgawana tše di maatlafaditšwego*. (lit.meaning) ‘particles which have been energized’
‘electronegative’ *dieleketerone tše di ganetšago*. (lit. meaning) ‘electrons that oppose’

The main disadvantage of paraphrasing is that although it is transparent, it is not economical, especially if that particular term has to be used repeatedly, cf.

Carbonated water can be good for watering plants. **Carbonated water** can also be used by human beings.

Translation:
Meetse ao a tšhetšwego khapone a loketše go nošetša dimela. Meetse ao a tšhetšwego khapone a ka šomišwa le ke batho.

Therefore, although paraphrases are transparent, they are not ideal in terms of economy.

c. Compounding

During compounding, a new term is formed by joining more than one existing words or lexical items. There seems to be a natural chronological link between paraphrasing and compounding, whereby paraphrases tend to be compacted to become compounds, probably because of frequency of use. Examples are the following:

'atomic' *seka-athomo,* lit. meaning 'like an atom'
'electrochemical' *mohlagasekhemikhale,* lit. meaning 'electric + chemical' < *mohlagase wa khemikhale*.
'time unit' *motšonako,* lit. meaning 'unit of time' < *motšo wa nako.*
'trivial name' *leinatlwaelo* lit. meaning 'name of habit' < *leina la tlwaelo.*
'magnetic field' *lepatlelokgogedi* lit. meaning 'field of attraction' < *lepatlelo la kgogedi.*

Deleting the possessive concord *(la, wa, sa*, etc) from so-called possessive constructions in order to form compound nouns, seems to be a productive strategy of term formation in Northern Sotho.

4.2.2.2 Language external term formation processes

a. Borrowing

Borrowing is the process whereby linguistic elements are taken over from one language to another. Different types of borrowing are:

- Loan words/ foreign words

When terms are borrowed they are taken as wholes, both morphologically and phonologically and their meanings are retained intact. Compare for example music terms, e.g. pianissimo, allegro, forte. Although these terms are of Italian origin, they are internationally in music scores. In this particular project, symbols representing chemical elements and units of meaning measurement were, according to international practice, retained as is. Compare the following examples:

Debye (unit of measurement, named after Dutch physicist P.J.W Debye)
H (symbol) > H (*seka*)
E (symbol) > E (*seka)*
HCI (symbol) > HCI (*seka)*
CI (symbol) > CI (*seka)*

- Transliterations/Adoptives

Adopted words are completely adapted to the language system of the borrowing language: syntactically, morphologically, phonologically and tonologically. When terms

are translated from English / Afrikaans to Northern Sotho, one such an adaptation is the adaptation of the syllable structure. In Northern Sotho the preferred syllable structure is CV-CV-CV. This requires that consonant clusters in the source language be separated by inserting vowel(s) between them. Examples of transliterated terms are the following:

electrode > *e-le-ke-te-ro-te*
titration > *ta-e-te-re-i-še-ne*
carbon > *kha-po-ne*
endothermic > *e-n-to-the-mi-ki*
hydrogen > *ha-e-te-ro-tše-ne*

Transliteration has both advantages and shortcomings. Its main advantage is that it is readily available, since it entails little more than an adaptation of the phonological and morphological structure of the source term. Secondly, it retains the visual relatedness between the source term and the target term. Its disadvantage is that it gives no assistance if the target user is not familiar with the conceptual content of the source term. The other disadvantage of transliteration is that foreign terms may dominate the language to such an extent that it runs the risk of losing its character.

With specific reference to Northern Sotho, inconsistency in spelling is a serious problem when using transliterations. In some cases terms have multiple TEs, which differ only with regard to spelling, The Northern Sotho equivalent for 'electron' is a case in point: no less than three spelling variations are in existence, viz **electron** => *elektrone* or *eleketerone* or *eleketrone*.

According to Taljard (2008) the inconsistency can be contributed to inadequate Northern Sotho spelling rules. The official Terminology and Orthography provides little help to someone who is faced with all these spelling options, especially if that person is not a mother tongue speaker of Northern Sotho. The word list or terminology list in the Terminology and Orthography (T&O) is preceded by an explanation of the spelling rules /orthography of Northern Sotho. Generally speaking, the purpose and function of

formalized spelling rules is to enable language users, be they students, academics, or terminologists, to correctly spell words which do not necessarily appear in dictionaries and / or term lists. The applicability of spelling rules to as yet unstandardized words and / or terms is especially relevant for the terminologists, since they are primarily responsible for the creation and standardization of new terms.

Spelling of all new terms, regardless of whether these have been formed by language internal processes such as semantic transfer, compounding or paraphrasing, or by means of language external processes such as transliteration or borrowing must be in accordance with the recognized spelling rules of any particular language. Terminologists invariably run into trouble when these rules are not (a) clear, (b) consistent and (c) phonologically and linguistically sound. Even though a section in the 1988 T&O is devoted to the expression of new words and ideas, the guidelines provided in this regard are to a large extent inadequate and the terminologist is largely left to his / her own devices as regards for example the adaptation of syllable structure and the resultant spelling of especially transliterations. This has direct implications for any terminological activity, since the use of transliterations or adoptives is a very frequently used strategy in the formation of terms. In the Terminology and Orthography a distinction is made between adoptives / transliterations "that could be adapted to the Northern Sotho syllable structure" (paragraph 14.2.1 T&O 1988:22) and "those words which as international terms are not adapted according to the Northern Sotho syllable structure" (paragraph 14.2.2 T&O 1988:22). No specific guideline is given as to which words should be adapted to the Northern Sotho syllable structure and which ones should not be. The impression is inadvertently created that it is up to the user to make this decision. The examples that are given also do not serve to clarify the issue, but only to confuse it. As examples of words that could be adapted to the Northern Sotho syllable structure, the following terms are *inter alia* given:

themometara cf. 'thermometer / thermometer'
barometara cf. 'barometer'
sekolo cf. 'school / skool'

Under words which are regarded as international terms and which should not be adapted, only two examples are given, i.e. *'krediti'* (intead of *kerediti*) and '*moprofeta'* (instead of *moporofeta*) (T&O 1988:22). It is not clear whether these two terms are the only ones to be regarded as international terms, or whether they should be viewed as being representatives of a larger class of terms. In the case of the latter possibility, not enough guidance is given to enable the reader to identify the class of words represented by these two examples. It is not clear why *krediti* 'credit' should be regarded as an 'international term', but not *barometara* 'barometer' or *themometara* 'thermometer'. It is clear that despite attempts by the compilers of T&0 to provide guidelines regarding phonological adaptations and resultant spelling of transliterations, the hapless terminologist is no closer as to whether the syllable structure of any transliteration should be adapted or not. In 2008, these rules were revised by the Sesotho sa Leboa National Language Body and published by PanSALB. Unfortunately, existing problems were not addressed in the revised version; if anything, the revised rules are even more confusing than the ones in the 2008 version of the T & O. As a result, confusion reigns as regards the spelling of especially transliterated terms. Nevertheless, transliteration is a very productive term formation strategy, despite the sometimes vociferous resistance from academics.

4.3 Coining of new chemistry terms

As was indicated in chapter 3, after harvesting TEs from both standardized and non-standardized sources, 51 terms without TEs were left. Consequently, TEs were coined for these 51 terms, making use of a variety of the term formation strategies discussed above.

TABLE 9 **Terms with coined translation equivalents**

Term	**Translation equivalent**
TRANSLITERATIONS	
actinides	diakethenaete

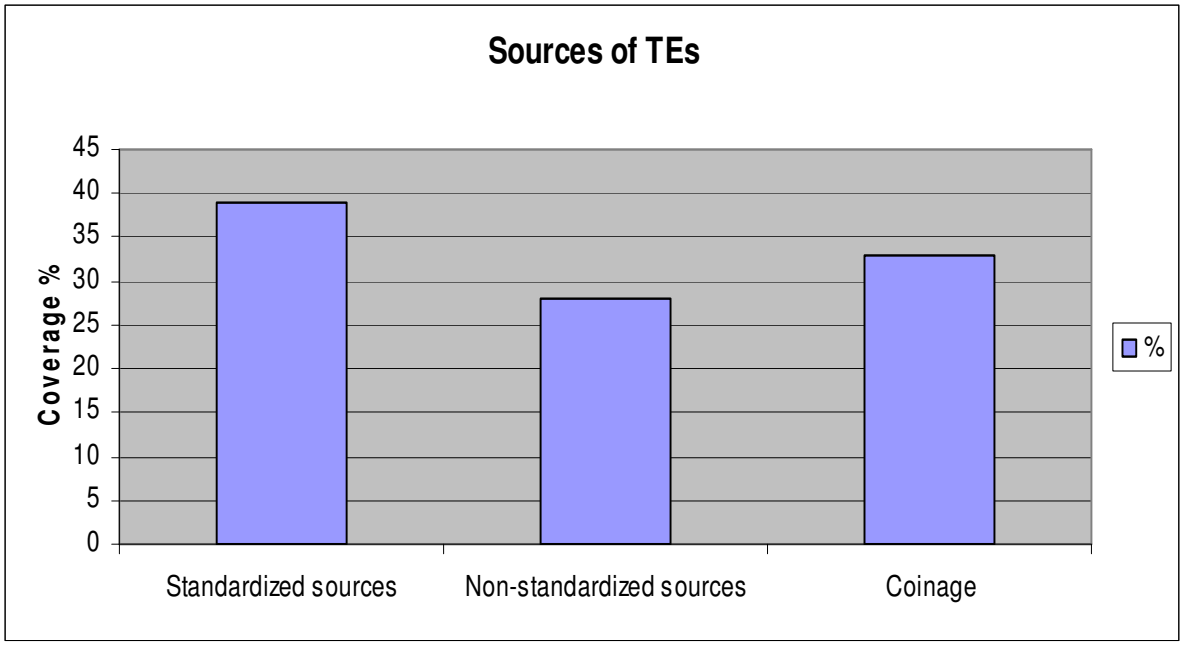
coke	khoukhu
debye	tepaye
diamagnetic	taemakenethiki
diamagnetism	setaemakenethiki
dimer	taemara
diprotic	taephorothiki (phorothone-pedi)
electrolysis	eleketerolaesese
electrolyte	eleketerolaete
electrolytic	eleketerolaetiki
enthalpy	entalipi
ethoxy ethane	etheini ya ethokisi
ethylene	ethelini
lanthanides	dilanthenaete
magnetite	makenetaete
metallic element	elemente ya metale
radioactive	radio-ekethifi
titration	taetereišene
COMPOUNDS	
control rods	melangwanataolo
diatomic	seathomopedi
dipole	ntlhapedi
electromotive	mohlagase-sepedi
end point	ntlhaphetšo
equivalent point	ntlhatekatekano
homogeneous mixture	motswakotshwano
resultant mixture	motswakopheletšo
PARAPHRASES	
charged particles	dikgawana tše maatlafaditswego
dipole moment	lebaka la ntlhapedi
electric circuit	lepatlelo la mohlagase

electrolytic cell	sele (lelahle) ya eleketerolaete
electromotive force	maatla a mohlagase-sepedi
electron configuration	peakanyo ya dieleketerone
electron density	pitlagano ya dieleketerone
electronegative	dieleketerone tše di ganetšago
electroplating	pharo ka mohlagase
empirical formula	fomula ye e netefaditšwego
enthalpy change	phetogo ya entalipi
visible lights	dietša tše di bonalago
nuclear reactors	difetoši tša nutliliya
orbital	se dukulogago
organic compound	tlhakantšhetšo ya tše di phelago
propagated	gašanywa
SEMANTIC SPECIALIZATION	
device	sešomišwa
dissolution	motološo
donating	aba (fana)
emission	ntšha (lokolla)
emulsify	phaphamala
homogeneous	se swanago
molten	tološitšwego
overall	kakaretšo
via	ka (tsejana)

4.4 Conclusion

Terminological processing of the source text, i.e the 50 chemistry terms and their definitions resulted in the compilation of a bilingual term list, containing 148 source terms and their equivalents in Northern Sotho. These equivalents were sourced from a variety of sources, as can be seen from figure 1 below:

Figure 1 *Sources of term equivalents*



It can clearly be seen that less than 40% of the harvested TEs come from standardized sources. This surely shows how inadequate standardized terminology sources are in Northern Sotho. With regard to the use of non-standardized sources, it needs to be acknowledged that these sources need to be treated with the necessary circumspection. However, in this particular instance, these sources provided well-formed and appropriate TEs for many of the source terms.

CHAPTER 5 USERS' PREFERENCES AS STANDARDIZATION MECHANISM: A CASE STUDY

5.1 Introduction

After having completed the coining of equivalents for those source terms for which no equivalents could be found, the first deliverable of the initial terminological processing of the source text was available, i.e. a bilingual term list, containing all source terms isolated from the source text, followed by their Northern Sotho equivalents. However, the problem of multiple term equivalents for 32 of the 148 source terms still persisted. Ideally, such a list of source terms and their multiple equivalents should be submitted to an official standardization body for formal standardization, a process during which a preferred term from amongst multiple translation equivalents is identified, again in consultation with the subject field experts. In practice this is rarely feasible, due to the time pressure under which translators normally operate. Furthermore, as was earlier pointed out, the standardization process of terminology in South Africa seems rather flawed; one of the main problems being the dissemination and general accessibility of standardized terms. As a result, translators use their own discretion in deciding on appropriate term equivalents for the source terms. This practice does however not solve the issue of the multiplicity of term equivalents and may even contribute to the unnecessary proliferation of terms.

For this particular project, it was decided to use preferences of the target users of the terminology as a guideline for internal standardization. Consequently, a small case study was conducted in three secondary schools in the Limpopo province where chemistry forms part of the curriculum. The aim of this case study was to determine the feasibility of involving target users in the standardization process, even if is it only a preliminary and internal standardization. It could be argued that the views of Grade 12 learners can hardly be a substitute for the input of the experts, and the researcher was well aware of the possible pitfalls of such an approach. However, the value of the involvement of the target users in the terminological process should not be under-estimated. Being in grade

12, it can be assumed that participants (learners) would already have internalized the basic chemistry concepts and therefore do have a contribution to make. Secondly, involving the target users in the development of terminology may in the long run encourage them to take ownership of not only terminology development in the home language, but also to actively use the terminology. Thirdly, on more general level, it may help to dispel the myth that the indigenous languages are incapable of producing technical and scientific terminology.

5.2 Procedural preparations

Science teachers and science students were identified as potential participants in the case study. The learners were allowed to participate only if they were 18 years and above and doing Chemistry or Physical science as one of their school subjects.

Educators were easy to approach and the whole procedure was easily understood. Time and space were the only obstacles. The only time that educators are available is during working hours (during the day) and the only place is the work place (at school). Bearing this in mind, it is not possible to interview educators without the approval of the principal, who also needs approval from circuit manager. The circuit manager needs to have permission from the district manager, who in turn needs approval from the Provincial level supervisor or MEC of Education. Obviously, permission should first be requested from the Department of Education (Provincial Level), and then with approval at hand, one can approach the circuit and school managers.

A request to do research by means of interviews at school level was made to the Department. Circuit managers and school Principals were given copies of approval. Interviews were scheduled for specific high schools. Three high schools were targeted. Two science educators and 20 learners from each school were identified as potential participants. Because of the lack of more science students, only ten learners and one educator from each school were interviewed.

5.3 Challenges encountered when trying to liaise with educators and learners.

As was explained in the previous paragraph, all the necessary permission was requested from all authorities involved. Even though the proper procedure was followed, it was not easy to get an interview with either educators or learners. Educators were either busy teaching or marking their learners' work, and the learners were either attending classes or writing tests.

Some students were reluctant to participate, and simply went through the motions of just filling in the forms; some copied the information from their "clever" fellow students, without really committing themselves to the task at hand. This was discouraging because there was no sign of commitment from their side. This caused a setback because new forms were to be provided and this was expensive.

Obtaining a venue for administering the questionnaire was also a challenge. Most schools in our region do not have extra classrooms or halls for activities such as conducting of interviews. In two of the three schools visited, interviews had to be conducted outside, in the shade of some trees on the school ground. This was rather disturbing, since participants were constantly interrupted by fellow learners who were asking questions about the research activities. Controlling the inputs given by participants in order to do some quality control was rather difficult under such circumstances.

Some learners were reluctant to fill in their personal details as they thought that the purpose of the interview was to encourage them to follow a career in translation. A lot of effort had to be made to convince them that it was just for research purposes.

Lastly, most learners were convinced that there was some form of remuneration to be made to the researcher and like-wise as participants, felt they should also be remunerated. It is not easy to convince an eighteen year old that there is no remuneration involved for the interviewer, and that everything is from the interviewer's pocket.

On the other hand, teachers were also not cooperative. One would expect great support from teachers who are also mother tongue speakers of Northern Sotho. However, they showed little interest in the research. Some thought it was a waste of money and time. Sometimes, teachers would leave the researcher alone with the learners. During the absence of their teachers, learners tend to be uncontrollable. It was frustrating to try to restore order in a class of learners who barely know you. That consumed time and energy and caused frustration on the side of the interviewer.

5.4 Structure of questionnaire

During the survey 30 questionnaires were distributed among three high schools in the Tsimanyane circuit, Dennilton Area, Greater Sekhukhune District in Limpopo. The high schools are:

1. Mahlontebe high school
2. Ngwanakena high school
3. Mphahlele high school

Ten learners and one educator from each school participated in the survey. The criteria for the selection of learners were amongst others that the learner should be above eighteen years of age. Secondly, the learner should be doing one of the science subjects, preferably physical science or chemistry. The learners should be in grade 12 most preferably or grade 11.

The aims of the questionnaire were as follows:

- To investigate language attitude, especially the use of Northern Sotho as mother tongue in the teaching of Chemistry as a subject.
- To do preliminary standardization of a selected number of terms for which multiple equivalents were found during the initial data collection. The users' preference is thus used as a preliminary standardization tool.
- To investigate the users' preference regarding the use of transliterated equivalents versus the so-called indigenous Northern Sotho terms.

- To investigate the users' preference for the adaptation of the syllable structure on the transliterated terms.

These aims were also reflected in the structure of the questionnaire, with each of the four sections of the questionnaire dealing with each of the aims.

Even though one should be careful to generalize on the basis of these findings in this survey, the results of the survey do shed some light on people's attitude towards and perception about their language. In involving the target users in the decision making process, it is no longer the sole responsibility of the terminologist/researcher to make the decision with regard to all terminological issues.

5.4.1 Attitudes towards Northern Sotho as the language of teaching and learning of chemistry

Section A of the questionnaire addresses the question of language attitude, specifically among learners whose mother tongue is Northern Sotho, but who are taught and who learn through the medium of English. The following is an extract from the questionnaire as administered to the learners:

- Is it easy to learn Chemistry in English?
- If *no*, in which language do you prefer to be taught?
- Is the language chosen (in 5) above your mother tongue?
- If *yes* what do you think are the advantages of learning a subject in your mother tongue?
- If all subjects were taught in learners' mother tongue, what do you suggest could be a pass rate in matric results in percentages?

The following questions were addressed to educators

- Was it easy to learn science subjects like Chemistry in a language other than your mother tongue?
- In your opinion, which language should be used to teach subjects like Chemistry?
- Should the language chosen in 5 above be the mother tongue?
- If yes, what do you think are the advantages of learning a subject in one's mother tongue?
- If Chemistry was taught in learners' mother tongue, what do you suggest could be a pass rate in matric?

The responses of the participants are consequently analyzed, in order to determine their (general) attitude towards the use of Northern Sotho, which is their mother tongue, as the language of teaching and learning. From 30 learners, 24 (80%) think it is easy to learn Chemistry in English. Only 6 perceive it as difficult, and their alternative language for learning Chemistry is Northern Sotho. These six respondents indicate that they grew up speaking Northern Sotho and that they understand it much better than English. All six chose Northern Sotho not only because they can speak the language, but because it is their mother tongue. On the other hand, all 30 learners believe that if teaching and learning were done in their mother tongue, a 80-100% pass rate will be achieved. The learners' response to these two questions therefore presents an interesting contradiction, indicating that these learners have an ambivalent attitude with regard to the value of their mother tongue as the medium of teaching and learning. On the one hand, 80% of learners indicate that the use of English as the medium of instruction does not present any difficulty to them, but they do seem to value their mother tongue as a tool towards success. They seem to realize that they can express themselves easier and clearer in their home language, and that if given a chance to learn and be taught in their mother tongue, they would perform better than they are currently doing.

All three educators gave more or less the same response. They indicated that it was not easy for them to learn Chemistry in English and they would have preferred to learn it in

their mother tongue which is Northern Sotho. They also believe that if the mother tongue is used for teaching and learning, there would be a pass rate of higher than 80%, because terminology would have evolved and developed over time. It is significant that teachers are of the opinion that the use of mother tongue in the teaching and learning would bring such a great change in the pass rate of learners. This view can possibly be attributed to teachers' experience in the classrooms. They seem to understand that most learners cannot express themselves in English or interpret statements written in English. From the teachers' responses it is clear that they realize the importance of proper terminology when using Northern Sotho as the medium of teaching and learning. From the responses of both learners and educators, it is clear that their attitude towards Northern Sotho as the language of teaching and learning is not a negative one, and that there is some understanding of the fact that learning a difficult subject in a language other than a home language/mother tongue may have a negative impact on the mastering of the subject.

5.4.2 Selection of the preferred term from multiple translation equivalents

The aim of section B of the questionnaire was to identify a preferred term, based on users' preferences, for a number of source terms for which multiple term equivalents were found. These 11 source terms were selected in such a way that the (multiple) TEs for any given ST represented the same term formation strategy. As can be seen from the questions below, respondents were provided with all possible TEs for any specific term, and asked to choose the term which they prefer. The preferred terms would then be regarded as being internally standardized for the purpose of this project. Respondents were also provided with a definition of every source term, in order to make sure that they select the term which is conceptually the closest match to the source term. Compare the following excerpt from the questionnaire:

SECTION B

In this section a definition will be provided together with possible terms for that definition, yours is to tick the term which you think is best suitable for the definition.

A. Bond: A thing or force that unites or restrains.

Term 1: tlemagano

Term 2: kamanyo

Term 3: setlamo

Term 4: pofo

B. Decomposition: A reaction involving the chemical separation of a given compound into two or more simple compounds or substances e.g. $\underline{2H_2O \rightarrow 2H_2+O_2}$

Term 1: polo

Term 2: tlharamollo

Term 3: kamologano

C. Dehydration: The removal of water from a substance e.g. $\underline{CuSO_4.5H_2O}$

(Hydrated copper sulphate) $\rightarrow CuSO_4 + 5H_2O$ (dehydrated copper sulphate + water)

Term 1: komo

Term 2: tšhwabo

Term 3: meetsefatšollo

D. Density: Mass per unit of volume e.g. the density of mercury is 13,5g.cm^3

Term 1: pitlagano

Term 2: kitlano

Term 3: teteano

Term 4: kitlagano

Term 5: pitlagantšho

E. Dispersion: To go in different directions or to scatter.

Term 1: phatlalatšo

Term 2: tšitlano

Term 3: phatlalalo

F. Dissociation: The separation of compounds or atoms, e.g. the dissociation of acetic acid in water to form H+ ions and acetate ions.

Term 1: tlogelano

Term 2: kgaogano
Term 3: tšhwalalano

G. Energy: The potential to do work or to transfer heat.
Term 1: mafolofolo
Term 2: mooko
Term 3: maatla

H. Diffusion: The process during which a substance moves from a higher to a lower concentration.
Term 1: phatlalatšo
Term 2: kgašano

I. Particle: A very small bit or piece of a substance.
Term 1: sekgwana
Term 2: karolwana
Term 3: lerathana
Term 4: seripana
Term 5: tsekana

J. Separation: Stop being combined, to remove elements from each other.
Term 1: karogano
Term 2: kgaogano
Term 3: tlogelano

K. Reaction: Chemical change produced by two or more substances acting upon each other.
Term 1: phetogo
Term 2: kgohlagano

Table 10 **ANALYSIS OF RESULTS FOR LEARNERS**

	ST and TEs	**Mphahlele**	**Ngwanakwena**	**Mahlontebe**	**Total**	**%**

ST	**BOND**					
TE1	tlemagano	7	10	7	24	**80**
TE2	kamanyo	-	-	1	1	**3**
TE3	setlamo	-	-	1	1	**3**
TE4	pofo	-	-	-	-	**-**
ST	**DECOMPOSITION**					
TE1	polo	1	6	8	15	**50**
TE2	tlharamollo	7	2	2	11	**37**
TE3	kamologano	2	2	-	4	**13**
ST	**DEHYDRATION**					
TE1	komo	8	5	4	17	**57**
TE2	tšhwabo	-	1	1	2	**6**
TE3	meetsefatšolla	2	4	5	11	**37**
ST	**DENSITY**					
TE1	pitlagano	1	1	2	4	**13**
TE2	kitlano	1	4	1	6	**20**
TE3	teteano	1	1	3	5	**17**
TE4	kitlagano	6	2	2	10	**33**
TE5	pitlagantšho	1	2	2	5	**17**
ST	**DISPERSION**					
TE1	phatlalatšo	1	3	5	9	**30**
TE2	tšitlano	3	6	4	13	**43**
TE3	phatlalantšho	6	1	1	8	**24**
ST	**DISSOCIATION**					
TE1	tlogelano	4	2	3	9	**30**
TE2	kgaogano	3	6	4	13	**43**
TE3	tšhwalalanyo	3	2	3	8	**27**
ST	**ENERGY**					
TE1	mafolofolo	10	17	4	21	**70**

TE2	mooko	-	-	-	-	**-**
TE3	maatla	-	3	6	9	**30**
ST	**DIFFUSION**					
TE1	phatlalatšo	9	6	7	22	**73**
TE2	kgašanyo	1	4	3	8	**27**
ST	**PARTICLE**					
TE1	sekgwana	-	-	1	1	**3**
TE2	karolonyana	1	-	-	1	**3**
TE3	lerathana	6	1	6	13	**43**
TE4	seripana	1	5	2	8	**27**
TE5	tsekana	-	1	1	2	**7**
TE6	karolwana	2	3	-	5	**17**
ST	**SEPARATION**					
TE1	karogano	2	3	1	6	**20**
TE2	kgaogano	1	3	4	8	**27**
TE3	tlogelano	7	4	5	16	**53**
ST	**REACTION**					
TE1	phetogo	7	6	6	19	**63**
TE2	kgohlagano	3	4	4	11	**37**

As can be seen from the table, in some cases preferences were very clear – 80% of respondents preferred the term *tlemagano* as equivalent for 'bond'; 70% preferred *mafolofolo* as equivalent for 'energy'. In other cases, preferences were not so clear-cut. Preferences for the equivalent for the equivalent of the source term 'dispersion' were as follows*: phatlalatšo* (30%*), tšitlano* (43%), and *phatlalantšho* (24%). Nevertheless, it was possible to identify a preferred term for all the source terms.

On comparing the preferences of learners and educators, it is clear that there is to a large extent correlation between the preferences of the learners and those of the educators. The two notable exceptions are the preferred TEs for 'dispersion' and 'particle'. In the case of

'dispersion' the majority of learners (43%) prefer *tšitlano* whereas only one of the teachers regarded this as the most suitable term. With regard to the TE for 'particle', none of the teachers gave preference to *lerathane* whereas this was the preferred term for learners. Given the fact that the important issue is the learners' cognition, it was decided to use the learners' preferences as a guiding principle.

5.4.3 The use of transliteration versus indigenous terms

The purpose of section C of the questionnaire is to ascertain whether there is a preference amongst teachers and learners for either transliterations or indigenous terms as equivalents for chemistry terms. The use of transliterations is a much debated issue amongst academics, but thus far, no investigation has been made into the preferences of the target users.

Learners and educators were again given a choice between two translation equivalents for 12 source terms, one being an indigenous term, the other being a transliterated term. Compare the relevant section from the questionnaire:

In this section a term will be given together with two possible translation equivalents (TE). One TE is a Sepedi coined term, the other is a transliteration. Choose which TE you prefer.

Tick with √ next to the chosen TE.

	Indigenous word	**Transliteration**
1. acid	sedilana	esiti
2. cell	lelahle	sele
3. compound	tlhakantšhetšo	khomphaonte
4. dimer	phokotšaetša	timara
5. element	setho	elemente
6. formula	kaelo	fomula
7. gas	moya	gase

8. negative	ganetšago	nekethifi
9. mercury	tshipimeetse	mekhuri
10. melt	nyaoga	meleta
11. layer	tlhatlagano	llaga
12. copper	mpshiri	koporo

The results of Section C are tabulated in table 11 below.

Table 11 **Indigenous words vs transliterations**

	Mphahlele		**Ngwanakwena**		**Mahlontebe**		**Total**	**Total**
TERM	**IND**	**TRANS**	**IND**	**TRANS**	**IND**	**TRANS**	**IND**	**TRANS**
Acid	3	7	4	6	3	7	10	20
Cell	5	5	4	6	5	5	14	16
Compound	7	3	5	5	4	6	16	14
Dimer	5	5	4	6	5	5	14	16
Element	5	5	6	4	4	6	15	15
Formula	3	7	2	8	1	9	6	24
Gas	6	4	7	3	2	8	15	15
Negative	6	4	5	5	7	3	18	12
Mercury	5	5	6	4	1	9	12	18
Melt	3	7	4	6	2	8	9	21
Layer	7	3	6	4	8	2	21	9
Copper	8	2	8	2	7	3	23	7
TOTAL PERCENTAGE							**48%**	**52%**

Results for this section indicated that 52% of the preferred equivalents were transliterations, the rest (48%) being indigenous terms. The results obtained from the educators present an interesting contrast to those of the learners: and analysis of the educators' preferences indicated that only 28% of their preferred equivalents are

transliterations, the rest being indigenous terms. Although we acknowledge the fact that the preferences of target users, who can at most be regarded as semi-experts, cannot be the final criterion in the selection of a standardized term from multiple equivalents, it surely needs to be taken into consideration that target users seem to have no serious objection to the use of transliterations.

5.4.4 Adaptation of syllable structure

One of the problems with regard to the use of transliterations in Northern Sotho is the phonological adaptation and resultant spelling of these items. The preferred syllable structure in Northern Sotho is the CV-structure, which implies that whenever a word is borrowed from English of Afrikaans, its phonological structure and consequently its spelling needs to be adapted to conform to this requirement. However, this rule is not applied consistently, resulting in multiple equivalents which differ on orthographical level. In the last section of the questionnaire, respondents were requested to choose between one of two variants, the one displaying an adapted syllable structure and spelling corresponding to the syllable structure, the other equivalent being the non-adapted variant. A second aspect that complicates the spelling of transliterations is the indication of aspiration, specifically with regard to the three voiceless plosives [p], [t] and [k]. Speakers often differ with regard to the pronunciation of these sounds, which consequently leads to differences in spelling, cf *molekhule* vs *molekule*. Four items were included in the list of terms where respondents had to choose between a version where aspiration was indicated, and one where it was not. The official spelling rules of Northern Sotho provide no guideline in this regard.

Analysis of target users' preferences revealed a clear bias towards those forms where the orthographical representation reflects the adapted phonological structure. Compare Section D from the questionnaire.

SECTION D:

In this section, a term will be given together with two possible term equivalents. The two term equivalents (TEs) are the same, but spelled differently. Choose the TE that you prefer. Tick with √ next to the chosen TE.

Term	**TE1**	**TE2**
compound	khompaonte	khomphaonte
electrode	eleketerote	elektrote
electron	eleketerone	elektrone
hydrogen	haeterotšene	haetrotšene
isotope	aesothopo	isothopo
molecule	molekhule	molekule
positive	phosethifi	phosetifi
titration	taetereišene	taetreišene
atom	athomo	atomo

Compare Table 12 below for an analysis of the results of Section D of the questionnaire.

Table 12 **Adaptation of syllable structure**

	Mphahlele school		**Ngwanakwena school**		**Mahlontebe school**		**TOTAL**	
TERM	**TE 1**	**TE 2**	**TE 1**	**TE 2**	**TE 1**	**TE 2**	**TE1**	**TE2**
Compound	-	10	1	9	2	8	3	27
Electrode	10	-	8	2	3	7	21	9
Electron	8	2	7	3	6	4	21	9
Hydrogen	6	4	5	5	6	4	17	13
Isotope	6	4	7	3	6	4	19	11
Molecule	9	1	10	-	9	1	28	2
Positive	6	4	7	3	6	4	19	11
Titration	6	4	6	4	5	5	17	13
Atom	9	1	9	1	8	2	26	4
Percentage	**66,6%**	**33,3%**	**66,6%**	**33,3%**	**56,6%**	**43,3%**	**63**	**37**

As can be seen from the table above, an analysis of target users' preferences reveal a clear bias towards those forms where the orthographical representation reflects the adapted syllable structure. Furthermore, there is also a clear preference for the forms where aspiration of the plosives [p], [t] and [k] are indicated. However, this sample is much too small to base any final conclusion on, but it surely is an issue that needs further investigation.

The responses of the educators mostly coincided with those of the learners, except for one case where an educator preferred *khompaonte* over the generally preferred *khomphaonte*.

By going through the various steps as described above, a total number of 148 terms were provided with translation equivalents. The procedure described in this chapter represents proper terminological groundwork that should be done prior to any project regarding technical translation. The final product still contains multiple term equivalents for 25 of the terms. It is not the task of the translator\ terminologist to make a choice or preference among different TEs.

5.5 Selection of term equivalents through informal discussions with semi-experts

It is always important to take into consideration the view of people involved in the subject field before taking the final stand. It was decided to undertake an informal discussion with people who have knowledge of science and are also mother tongue speakers of Northern Sotho, to help in selecting the most preferred term among the following terms as was done with other terms in the questionnaire.

Table 13 **Acquisition of TEs through informal discussions**

	ST AND TEs	**PART1**	**PART2**	**PART3**	**PART4**	**PART5**	**TOTAL**	**%**
ST	**ACID**							
TE1	acid						0	-

TE2	sedilana	✓	✓	✓	✓	✓	5	100
ST	**CELL**							
TE1	sele	✓	✓				2	40
TE2	okwane			✓	✓	✓	3	60
TE3	lelahle						0	-
ST	**CHEMICAL**							
TE1	sekhemise						0	-
TE2	khemikhale	✓	✓	✓	✓	✓	5	100
ST	**DISSOLVED**							
TE1	tologa	✓	✓	✓	✓	✓	5	100
TE2	nyaoga						0	-
TE3	nyaoša						0	-
ST	**DISTILLATION**							
TE1	tlhwekišo	✓					1	20
TE2	hlwekišo							
TE3	tshekišo		✓	✓	✓	✓	4	80
ST	**ELEMENT**							
TE1	elemente				✓		1	20
TE2	setho		✓	✓		✓	3	60
TE3	tokollo	✓					1	20
ST	**FORMULA**							
TE1	fomula	✓					1	20
TE2	kaelo		✓	✓	✓	✓	4	80
ST	**GAS**							
TE1	gase						0	-
TE2	kese	✓		✓	✓	✓	4	80
TE3	moya		✓				1	20
ST	**MAGNETIC FIELD**							

TE1	lepatlelokgogedi	✓	✓	✓	✓	✓	5	100
TE2	karolokgogedi							
ST	**MOTION**							
TE1	tšhikinyego	✓	✓				2	40
TE2	tšhišinyego			✓	✓	✓	3	60
ST	**MOVEMENT**							
TE1	tshepelo	✓	✓	✓	✓	✓	5	100
TE2	tšhikinyego						0	-
TE3	tšhišinyego						0	-
ST	**NEGATIVE**							
TE1	ganetšago	✓	✓	✓	✓	✓	5	100
TE2	latolago						0	-
ST	**PERIODIC**							
TE1	lebaka			✓			1	20
TE2	paka						0	-
TE3	nako	✓			✓		2	40
TE4	sebaka						0	-
TE5	ka mabaka		✓			✓	2	40
ST	**POINT**							
TE1	ntlha	✓	✓	✓	✓	✓	5	100
TE2	khutlo						0	-
ST	**REGION**							
TE1	selete	✓	✓	✓	✓		4	80
TE2	tikologo					✓	1	20
ST	**SOLVENT**							
TE1	setološi	✓		✓			2	40
TE2	setološa		✓		✓	✓	3	60
ST	**SUBSTANCE**							

TE1	selo	✓	✓	✓		✓	4	80
TE2	nto				✓		1	20
ST	SYMBOL							
TE1	sešupo	✓					1	20
TE2	seka		✓	✓	✓	✓	4	80
ST	**SYNONYM**							
TE1	lehlalološetšagotee	✓	✓	✓	✓	✓	5	100
TE2	sinonimi						0	-
ST	**UNIT**							
TE1	motšo	✓	✓	✓		✓	4	80
TE2	botee				✓		1	20

As has been mentioned earlier, the users' preferences are, for this project, regarded as internally standardized. It should also be mentioned that the above TEs may not be the only TEs for a particular term. It is also important to mention that even though one TE may be regarded as the preferred one, usually the context is the determining factor as to which TE to use.

5.6 Summary

Chapter 5 dealt mainly with the actual interviews which included teachers and learners who; directly or indirectly teach and learn Chemistry respectively. The problems encountered during interviews are all listed in 5.3 and are not directly related to the interviews. Some are social problems which only surfaced during interviews. The questionnaire was structured in such a way that it addresses the whole problem but using only the samples. The first part of the questionnaire dealt with preferences between multiple TEs (see Table 10). The second section dealt with the use of transliterations versus indigenous terms. According to the survey 48% preferred indigenous terms whereas 52% preferred transliterated terms (see Table 11). The third section of the questionnaire dealt with the adaptation of syllable structure .A term is written both as it is uttered and according to the traditional syllable structure of Northern Sotho i.e c-v-c-c.

The results show that both learners and educators prefer the traditional way of spelling which is the c-v-c-v structure (see Table 12). The informal interviews held, helped to recognize the preferred term between multiple terms for the remaining terms which had multiple term equivalents (see Table 13).

CHAPTER 6 CONCLUSION

The main aim of this research was to give a detailed account of the terminological processing of a technical text – in this case a text on chemistry – prior to translation. This entailed an investigation into the different techniques for extracting terms from the source text in order to compile a bilingual term list, consisting of source terms with their translation equivalents. Such a list is a prerequisite for translations which are done into a language such as Northern Sotho, in which there is a dearth of standardized sources on terminology. It is especially valuable in cases where a team of translators is working on a project, since access to such a list will facilitate consistency in the final translation.

In this particular instance a total of 148 source terms were isolated from the source text. The first technique that was utilized for term extraction is a corpus-based one, which requires the querying of electronic corpora. For this particular project, 50 English definitions of chemistry terms were regarded as our special purpose corpus whereas the Pretoria English Internet Corpus (PEIC) is our general or reference corpus. By using Wordsmith Tools KeyWord function, 89 single and multiword terms, i.e. 60% of the total of 148 terms were thrown up by means of a KeyWord search. This confirms the results obtained by Taljard and De Schryver (2002) in a similar experiment. Semi-automatic term extraction was followed by manual term excerption, which added 59 source terms to the list.

The next step in the terminological processing of the source text is the provision of Northern Sotho term equivalents. As a result of the lack of standardized chemistry terminology in Northern Sotho, the terminologist has no choice but to make use of both standardized and non-standardized sources. Two bilingual dictionaries and Terminology and Orthography of Northern Sotho were consulted as standardized references in search for translation equivalents. From the total of 148, only 57 terms got their translation equivalents from these formal reference works. During the search for term equivalents another problem presented itself, i.e. that of multiple equivalents: 32 of the 57 terms have multiple TEs. This is the result of the current inadequate standardization procedure.

During our investigation it was found that proper structures for the standardization of terminology are indeed in place in South Africa, but that these structures are not at the moment functioning properly. This is seriously hampering terminology development of the African languages.

Consultation of non-standardized sources of terminology indicated that these sources, although not standardized, can make a valuable contribution: term equivalents for an additional 19 terms were found. The information gained when searching for the TEs from the formal sources, helped in coining an additional 21 term equivalents, bringing the total of term equivalents found to 97.

In chapter 4, term formation strategies are discussed in order to ultimately coin the translation equivalents for the remaining 51 terms. Different strategies are discussed in order to try to match the suitable one for the translation of the individual terms. Term equivalents for 18 source terms were coined by means of transliteration, 8 by means of compounding, 16 by means of paraphrasing and 9 by means of semantic specialization. Even though regarded as a somewhat dubious term formation strategy by academics, transliteration seems to be a productive method, although by no means the only one for coining terms.

During the search for TEs from formal and informal sources, a number of terms acquired multiple TEs. By way of experiment, it was decided to involve the target users of the final product in the preliminary standardization of the bilingual term list. Grade 12 learners in three schools, together with their educators were approached to fill in a questionnaire. All learner participants are mother tongue speakers of Northern Sotho and take chemistry as a subject in school. The main aim of the questionnaire was to, among others:

1. determine the attitude of the interviewees towards Northern Sotho as medium of teaching and learning Chemistry as a subject. (Section A)
2. identify the preferred term equivalent from multiple term equivalents. (Section B)

3. determine the target users' (students and educators) preferences with regard to indigenous Northern Sotho terms and transliterations. (Section C)
4. identify the syllable structure they best prefer. (Section D)

The results obtained from the analysis of the responses to the questionnaire were used to preliminary standardize the bilingual term list. The aim is to submit the final product to the standardization committee of Northern Sotho for official standardization. The finalized list is included as Addendum A.

BIBLIOGRAPHY

Alberts, M. 2000. *Terminology Management at the National Language Service.* Lexikos 10:

Bowker, L. and J. Pearson. 2002. *Working with specialized language: A practical guide to using corpora*. London: Routledge.

Department of Education and Training. 1988. *Northern Sotho terminology and orthography no.4*. Pretoria: The Government Printer.

Gauton, R. 2000. *Introduction to translation (Volume 1) (TRL 810).* Unpublished study material, University of Pretoria

Gauton, R. 2000. *Translation Equivalence and Modern Translation Models (TRL 851).* Unpublished study material, University of Pretoria.

Librarius. 1985. *Learner English-Northern Sotho Dictionary*. Librarius Felicitas (PTY)LTD.

Mtintsilana, P.N. and R. Morris. 1988. Terminology in African Languages in South Africa. *South African Journal of African Languages* 8(4):109-113.

Pare, P. (n.d) *English-Sepedi-Afrikaans chemistry term list*. Unpublished: personal collection.

Pearson, J. 1998. *Terms in context*. Amsterdam: John Benjamins.

Reader's digest. 1991. *Multi-Language Dictionary and Phrase book*. The Reader's digest Association South Africa (PTY) Limited Cape Town

Sager, J.C. 1990. *A practical Course in Terminology Processing.* Amsterdam: John Benjamins.

Scott, M. 1997. *PC Analysis of Key Words-and Key Key Words*. System 25(1):236

Taljard, E and G.-M. de Schryver . 2002. *Semi-automatic Term Extraction for the African Languages, with Special Reference to Northern Sotho.* LEXIKOS 12. Afrilex-series 12:2002. Stellenbosch.

Taljard, E. 2004. *Terminology AFT 810*. Unpublished study material: University of Pretoria.

Taljard, E. 2008. *Issues in scientific terminology in African Languages*. In Lafon, M and V. Webb (eds.) IFAS Working Paper Series, p. 88 – 92.

ADDENDUM A Bilingual term list

Note:

In cases of multiple term equivalents, the preferred term is listed first

Source term	**TE1**	**TE2**	**TE3**	**TE4**
accelerate	akgofiša			
acid	sedilana	esiti		
actinides	diakethenaete			
anion	ayone			
anode	anode			
aqueous	sa meetse			
atom	athomo	atomo		
atomic	seka-athomo	[-]go ba le athomo		
atomic number	athomopalo			
bond	tlemagano	kamanyo	pofo	setlamo
carbon	khapone			
carbon monoxide	monokside wa khapone			
carbonated water	meetse ao a tšhetšwego khapone			
cathode	khatote			
cation	khateayone			
cell	okwane	sele		
charged particles	dikgawana tše maatlafaditšwego			
chemical	khemikhale	sa khemise		
chemical compound	tlhakantšhetšo ya sekhemise			
chemical reaction	kgohlagano-khemikhale			
Cl (symbol)	Cl (seka)			
coke	khoukhu			

colloidal	thulano			
compound	tlhakantšhetšo			
concentration	motswakoti			
conductor	sesepediši			
control rods	melangwanataolo			
copper sulphate	salfate ya koporo			
current	moela			
cyclotron	sesesedi			
debye	tepaye			
decompose	bola			
decomposition	polo	kamogolano	popolo	tlharamollo
decomposition reaction	phetogo ya polo			
density	kitlagano	kitlano	teteano	pitlagantšho
device	sešomišwa			
diamagnetic	taemakenethiki			
diamagnetism	setaemakenethiki			
diatomic	seathomopedi			
diffusion	kgašanyo	phatlalatšo		
dilute	hlaphola			
dimer	taemara			
dipole	ntlhapedi			
dipole moment	lebaka la ntlhapedi			
diprotic	taephorothiki	phorothone-pedi		
dispersion	tšitlano	phatlalatšo	phatlalantšo	
dissociation	kgaogano	tlogelano	tšwalalanyo	
dissolution	motološo			
dissolve	tologa	nyaoša	nyaoga	
dissolved substance	selo se se tološitšwego	setološwa		
distillate	hlwekiša	sekiša		
donating	aba	fana		
E (symbol)	E (seka)			

electric	mohlagase			
electric circuit	lepatlelo la mohlagase			
electric current	moela wa mohlagase			
electrochemical	mohlagasekhemikhale			
electrochemical cell	lelahle la mohlagasekhemikhale			
electrochemical reaction	phetogo ya mohlagasekhemikhale			
electrode	eleketerote			
electrolysis	eleketerolaesese			
electrolyte	eleketerolaete			
electrolytic	eleketerolaetiki			
electrolytic cell	sele / okwane ya eleketerolaete			
electromotive	mohlagase-sepedi			
electromotive force	maatla a mohlagase-sepedi			
electron	eleketerone	elektrone		
electron configuration	peakanyo ya dieleketerone			
electron density	pitlagano ya dieleketerone			
electronegative	dieleketerone tše di ganetšago			
electroplating	pharo ka mohlagase			
element	setho	elemente	tokollo	
emission	ntšha	lokolla		
empirical formula	fomula ye e netefaditšwego			
emulsify	phaphamala			
end point	ntlhaphetšo			
endothermic	entothemiki			
energy	mafolofolo	enetši	mooko	maatla
enthalpy	entalipi			
enthalpy change	phetogo ya entalipi			
equivalent point	ntlhatekatekano			
ethoxy ethane	etheini ya ethokisi			

ethylene	ethelini			
formula	kaelo	fomula		
gas	kese	gase	moya	
gas exchange	neeletšano ya moya			
H (symbol)	H (seka)			
HCl (symbol)	HCl (seka)			
homogeneous	[se] swanago			
homogeneous mixture	motswakotshwano			
hydration	meetsefatša			
hydrogen	haetrotšene			
hydrogen chloride	tleloraete ya haeterotšene			
immersed	inetše			
ion	ayone			
iron	tshipi			
lanthanides	dilanthenaete			
layer	leyara			
liquid	seela			
magnetic field	lepatlelokgogedi	karolokgogedi		
magnetite	makenetaete			
metallic element	elemente ya metale			
molecule	molekhule	molekule		
molten	tološitšwego			
monoxide	monokosaete			
motion	tšhišinyego	tšhikinyo		
movement	tshepelo	tšhikinyego	tšhišinyego	
negative	ganetšago	latolago		
nuclear reactors	difetoši tša nutliliya			
orbital	se dukologago			
organic compound	tlhakantšhetšo ya tše di phelago			
overall	kakaretšo			
oxidation	okositšenefatšo			

particle	lerathana	karolonyana	seripana	sekgawana
periodic	nako	ka mabaka	lebaka	paka
periodic table	papetla ya pheriotiki			
point	ntlha	khutlo		
process	tshepetšo			
propagated	gašanywa			
property	seeng			
protons	thoto			
radio waves	maphotho a radio			
radioactive	radio-ekethifi			
reaction	phetogo			
reduce	fokotša			
region	selete	tikologo		
resultant mixture	motswakopheletšo			
separation	tlogelano	karogano	kgaogano	
solute	setologi			
solution	setološwa			
solvent	setološa	setološi		
spin	dikologa	dikološa		
stable	se sa fetogego			
strength	maatla			
substance	selo	nto		
sulphate	salefeiti			
symbol	seka	sešupo		
synonym	lehlalošetšagotee	sinonime		
time unit	motšonako			
titration	taetereišene			
trivial name	leinatlwaelo			
unit	motšo	botee		
via	ka (tsejana)			
visible lights	dietša tše di bonalago			
volume	bolumo			

water	meetse			
x-ray	eksrei			
zero	lefela			