Reversing an African-language lexicon: the *Northern Sotho Terminology and Orthography No. 4* as a case in point[1]

DJ Prinsloo*

Department of African Languages, University of Pretoria, Pretoria 0002, South Africa
E-mail: prinsloo@postino.up.ac.za

Gilles-Maurice de Schryver

Department of African Languages, University of Pretoria, Pretoria 0002, South Africa
Research Assistant of the Fund for Scientific Research — Flanders (Belgium)
E-mail: gillesmaurice.deschryver@rug.ac.be

April 2002

The aim of this article is to give a perspective on the required strategies and the potential difficulties involved in reversing a unidirectional bilingual dictionary with an African language as the target language. This will be done by means of a full-scale case study of the (hypothetical) reversal of the *Northern Sotho Terminology and Orthography No. 4*. It will be pointed out that the African languages do indeed pose unique problems — difficulties that emanate directly from the structure of those languages. It will also be shown that the use of the forward slash causes particular additional complications. The discussion is preceded by an overview of the main issues involved in the compilation of (bidirectional) bilingual dictionaries with the focus on the different types of equivalence relations on the one hand and the reversibility principle on the other.

Senaganwa

Maikemišetšo a sengwalwana se ke go fa ponelopele go maano ao a nyakegago le go mathata ao a ka rarollwago a amanego le go fetošetša pukuntšu ya maleme a mabedi yeo e amago leleme le tee go pukuntšu yeo e tla lebelelago leleme la Afrika ge go fetošwa. Se se tlo dirwa ka go badišiša le go lekolla *Sesotho sa Leboa mareo le mongwalo No.4* yeo e fetošitšwego ka botlalo. Go tla laetšwa gore maleme a Afrika a na le mathata a a itšego, ao a tšwago go sebopego sa ona. Go tla bontšhwa gape gore tšhomišo ya leswao le (/) e tliša tšharakano ye e itšeng godimo ga mathata ao a lego gona. Pele ga poledišano go tla ba le tlhalošo ye kopana ya ditaba tše bohlokwa tšeo di amanago le go hlama pukuntšu ya maleme a mabedi (e lebeletše maleme ao ka bobedi) ka tsinkelo go mehuta yeo e fapanego ya kamano ya tekantšho gape le taba ya phetošo ya pukuntšu ya maleme a mabedi.

Brief theoretical conspectus

Tomaszczyk (1988:289) describes the purpose of a bilingual dictionary as follows:

The basic task of a bilingual dictionary is to provide L2 equivalents of L1 items in the L1–L2 part and L1 equivalents of L2 items in the

---

* Author to whom correspondence should be addressed.

L2–L1 part. The equivalents should be of an insertable kind, i.e. capable of being used in actual texts and, preferably, monolexemic (Akhmanova 1975:127). Moreover, the equivalents proposed should be carefully selected closest possible ones rather than cross-linguistic (near) synonyms "freely thrown about" (Liberman 1984:285).

Through the ages the selection of suitable translation equivalents (L2) for source-language items (L1) proved to be a major challenge to the lexicographer due to the complex network of semantic and equivalence relations that exist between source- and target-language items. Any attempt towards addressing these complexities lies beyond the scope of the present article and the reader is referred to pioneering studies on the compilation of bilingual dictionaries such as Zgusta (1971), Al-Kasimi (1977), Hartmann (1983), Landau (1984; 2001), Gouws (1989), Kromann *et al.* (1991) and Svensén (1993).

The following discussion will be limited to the basic issues relevant to the process of *reversing* a unidirectional bilingual dictionary in order to obtain a bidirectional bilingual dictionary. In the words of Tomaszczyk (1988:289):

A two-language dictionary is monodirectional if it serves the needs of the native speakers of one of the two languages. It is bidirectional if it attends to the needs of the speakers of both languages. Thus, the L1–L2 part of a bidirectional dictionary would be a reading dictionary (for decoding texts in the FL) for the native speakers of L2 and a writing dictionary (for encoding texts in the FL) for the speakers of L1. The L2–L1 part, in turn, would be a reading dictionary for speakers of L1 and a writing dictionary for speakers of L2 (cf. Steiner 1984:173).

Hartmann & James (1998), in their *Dictionary of Lexicography*, offer the following definitions of the key concepts *equivalent* and *equivalance* and rightfully comment at the same time on the intrinsic difficulties underlying the selection of suitable translation equivalents:

(1)   **equivalent**
A word or phrase in one language which corresponds in MEANING to a word or phrase in another language, e.g. English *mystery tour* and German *Fahrt ins Blaue*. Because of linguistic and cultural ANISOMORPHISM, translation equivalents are typically partial, approximative, non-literal and asymmetrical (rather than full, direct, word-for-word and bidirectional). Their specification in the BILINGUAL DICTIONARY is therefore fraught with difficulties, and recourse must be had to surrogate EXPLANATORY EQUIVALENTS.

(2)   **equivalence**
The relationship between words or phrases, from two or more languages, which share the same MEANING. Because of the problem of ANISOMORPHISM, equivalence is 'partial' or 'relative' rather than 'full' or 'exact' for most contexts. Compilers of bilingual dictionaries often struggle to find and codify such translation EQUIVALENTS, taking into account the directionality of the operation. In bilingual or multilingual TERMINOLOGICAL DICTIONARIES, equivalence implies interlingual correspondence of DESIGNATIONS for identical CONCEPTS.

In addition to *full* ~ (or *absolute* ~) versus *partial* ~ (or *relative equivalence*); *convergence*, *divergence* and *lexical gaps* (or *zero equivalence*) form the basis or point of departure for the compilation of a bilingual dictionary (Baunebjerg Hansen, 1990:13; Rettig, 1985). Consider, once again, the definitions offered by Hartmann & James (1998), with illustrations from Northern Sotho ⟷ English throughout:

(3a)   **convergence**
In cases of partial TRANSLATION equivalence, the rendering of two or more words in one language by a single word in the other language. Thus the meanings of the two English words *slug* and *snail* are covered by the single Dutch word *slak*. [...] The BILINGUAL DICTIONARY must allow for such asymmetrical relations, in conjunction with the problem of the user orientation.

(3b) *Northern Sotho → English*
khurumela close (*by putting a lid on*)
tswalela close (*of e.g. a gate*)

(3c) *English → Northern Sotho*
adultery bootswa
fornication bootswa
immoral conduct bootswa

(4a) **divergence**
In cases of partial TRANSLATION equivalence, the rendering of a word in one language by two or more words in the other language. Thus the meaning of the English word *aunt* is expressed in Danish by two words, *moster* 'maternal aunt' and *faster* 'paternal aunt'.

(4b) *Northern Sotho → English*
sehlano five; five shillings; fifty cents; quintet

(4c) *English → Northern Sotho*
aunt *(maternal aunt)* mmamogolo; *(paternal aunt)* rakgadi

(5a1) **lexical gap**
The absence of a word to express a particular meaning, e.g. lack of translation equivalents for CULTURE-SPECIFIC VOCABULARY terms across languages.

(5a2) **culture-specific vocabulary**
The words and phrases associated with the 'way of life' of a language community. In translation and the BILINGUAL DICTIONARY, these lexical items cause particular problems of EQUIVALENCE.

(5a3) **explanatory equivalent**
In the translation of CULTURE-SPECIFIC VOCABULARY, the explanation of a word or phrase by means of a surrogate PARAPHRASE in the target language rather than a one-to-one EQUIVALENT […]

(5b)[2] *Northern Sotho → English*
maganagodiša [*cattle herder who refuses to look after the cattle*]

(5c) *English → Northern Sotho*
illiteracy [*go se kgone go bala le go ngwala, go se tsebe ditlhaka*]

In (5b–c) a lexical gap exists in L2 and the lexicographer has to revert to other options such as *explanatory equivalents* (also known as *surrogate equivalents*), pictorial illustrations, glosses, etc. See in this regard especially Zgusta (1984; 1987), Schnorr (1986), Benson (1990), Duval (1991) and Rey (1991). For a detailed discussion of equivalence relations in bilingual dictionaries with Northern Sotho as L1 or L2, see Lekganyane (2001: 58–70).

Turning to the issue of *reversibility* — it can be stated that lexicographers generally agree on the basic underlying principle, as e.g. formulated by Tomaszczyk (1988:290) in oversimplified terms:

> everything that appears on the right-hand side of the L1–L2 part should reappear – as far as the structure of the two lexicons allows – on the left-hand side of the L2–L1 part.

Gouws (1989:162) is somewhat more specific in stating that all lexical items presented as lemma signs or translation equivalents in the X–Y section of a dictionary should respectively be translation equivalents and lemma signs in the Y–X section of the dictionary:

> 'n Leksikale item A wat aan die X-kant van 'n vertalende woordeboek as vertaalekwivalent verskyn in die artikel van 'n leksikale item B as lemma, moet aan die Y-kant van daardie woordeboek 'n lemma wees met minstens leksikale item B (die lemma aan die X-kant) as vertaalekwivalent. 'n Leksikale item B wat aan die X-kant van 'n vertalende woordeboek as lemma verskyn met 'n leksikale item A as vertaalekwivalent moet aan die Y-kant van daardie woordeboek as vertaalekwivalent verskyn in die artikel waarin leksikale item A (die vertaalekwivalent aan die X-kant) die lemma is.

Consider the following examples, taken from Kriel's (1976[4]) *New English – Northern Sotho Dictio-*

nary, illustrating these two conditions:

(6)   **A** = raloka, **B** = play
      Side **X:** *English → Northern Sotho*
      B (lemma sign)   A (translation equivalents)
      play, ... v., …   raloka, bapala, letša

      Side **Y:** *Northern Sotho → English*
      A (lemma sign)   B (translation equivalents)
      raloka, ... v., …   play, run about, romp

Thus the translation equivalent **raloka** in side X becomes the lemma sign in side Y, while the lemma sign **play** in side X becomes a translation equivalent of **raloka** in side Y. To the conditions stipulated in Gouws (1989:162), Gouws (1996:80) adds:

> Although the formulation of [the reversibility principle] only deals with translation equivalents with a lemmatic address, the principle must be interpreted as also referring to translation equivalents with a non-lemmatic address, provided that the address has lexical item status.

This reflects an *ideal situation* on the assumption that all lemma signs (including the sublemmas with lexical-item status) and their translation equivalents will in turn be suitable as respectively translation equivalents and lemma signs in the reverse side. However, this is *not* what is generally reflected in bidirectional bilingual dictionaries for a number of reasons which will be briefly outlined below.

Tomaszczyk (1988:290) refers to a series of conditions where the reversibility principle is *not* followed. These are:

- Suitable lexical items do exist but the lexicographer simply fails to apply the reversibility principle consistently: '…inconsistencies of the kind *mapmaking = Kartographie* but *Kartographie = cartography* (Liberman 1984:285)'.
- Zero equivalence: 'The principle is said to be inapplicable in the case of equivalentless lexis (Gold 1982:250 n.2; see however Gold 1985:319 n.2 and Tomaszczyk 1983:48 ff.)'.

- Frequency-of-use considerations: 'It may also not be followed when the L2 equivalent of an L1 item is much less frequent (Liberman 1984:285, Gold, op cit.)'.
- Prescriptiveness: 'Finally, entries are not reversed when one part of the dictionary (obviously monodirectional) is meant to be more prescriptive than the other (Gold, op. cit. both references). In such a dictionary e.g. the four-letter words etc. could be entered in the L2–L1 part but their equivalents could be euphemized, and they would not be entered in the L1–L2 part (cf. Dennis 1985:317)'.

Consider the following examples for a Northern Sotho ⟷ English dictionary:

(7a)   *Northern Sotho → English*
       kgaetšedi (younger) sister (of a brother); (younger) brother (of a sister)
       mogolo elder brother; elder sister
       moratho younger brother; younger sister
       morwarra brother; son of father's eldest brother

(7b)   *English → Northern Sotho*
       brother morwarra; **elder ~** mogolo; **younger ~** moratho; **younger ~ of a sister** kgaetšedi
       sister, **elder ~** mogolo; **younger ~** moratho; ...; **younger ~ of a brother** kgaetšedi

In these cases of rather complicated Northern Sotho kinship terminology [3], the reversibility principle is honoured — even though one is dealing with several lexical gaps. The word **kgaetšedi**, for instance, as well as '(younger) sister (of a brother)' and '(younger) brother (of a sister)', are all meticulously catered for in the Northern Sotho → English side and in the reverse English → Northern Sotho side.

However, in the case of the lemma signs and/or translation equivalents **bapala**, **raloka** and **play** Kriel (1976[4]) was only partially successful:

(8a)   **'bapala,** v., to play; **go bapala ka motho,** to make fun of a person.
(8b)   **play,** ... v., raloka, bapala, letša.

(8c) **raloka,** ... v., play, run about, romp.
(8d) **romp,** v., tlolaka, bapala, raloka.
(8e) **tlo'laka,** v., int., jump about, frolic, hop.

The lemma sign **bapala** and translation equivalent **play** in (8a) are presented as translation equivalent and lemma sign respectively in (8b). Likewise, **raloka** and **play** in (8c) are presented in (8b). Thus in (8a-c) the reversibility principle for **bapala**, **raloka** and **play** is successfully honoured. In spite of this, both **bapala** and **raloka** are given as translation equivalents for **romp** (8d), but **romp** is only included in the translation equivalent paradigm of **raloka** and not of **bapala**, thus violating the reversibility principle. The verb **romp** should have been included in the translation equivalent paradigm of **bapala** as well. Besides **bapala** and **raloka**, also **tlolaka** is given as translation equivalent of **romp**. Yet, the article for **tlolaka** does not include **romp** (8e), again violating the reversibility principle.

A more difficult decision in terms of the reversibility principle is whether an equivalent such as **run about** (8c) should perhaps not be entered as a lemma sign. The verb **raloka** does not appear anywhere in the translation equivalent paradigms of the lemma signs **run** or **about**. It can be argued that **run about** should have been treated in one way or another in the article of **run** and/or in the article of **about**, as has for example been done in the article of **about** in COBUILD3 (Sinclair, 2001[3]):

(9) **about** …
     [7] If someone or something moves **about**, they keep moving in different directions. *The house isn't big, what with three children running about.*

In (9) the fact that **about** is one of the typical collocates of **run** is taken care of in the illustrative example sentence.

Reversing a unidirectional bilingual dictionary

So far a brief and *over-simplified* overview was given of the main issues involved in the compilation of (bidirectional) bilingual dictionaries. The focus has been on the different *types of equivalences* on the one hand and

the *reversibility principle* on the other. It is clear that there is a direct correlation between the two. If a full equivalence exists between a certain lemma sign and its translation equivalent, then the reversal of that particular article will be non-problematic. Conversely, when one is dealing with a lexical gap, or thus with a zero equivalence between lemma sign and translation equivalent, then the reversal cannot be anything else but problematic. In-between these two extremes lie those articles for which a relation of partial equivalence exists; both convergence and divergence are applicable in those cases. Landau is thus entirely correct when he observes that 'the bilingual lexicographer cannot just reverse the direction of translations to obtain a word list for a companion volume' (2001: 10). If compiling a sound 'reversed macrostructure' is already problematic, it stands to reason that honouring the reversibility principle in the entire microstructure of the reversed side is even more ambitious. In any case, the reversed macrostructure must first be compiled, before one can even start thinking about the more demanding microstructural aspects. In the remainder of this article, the *macrostructure* of a reversed bilingual dictionary will therefore receive most attention.

In today's digital age, it should not come as a surprise that more and more 'software systems' are designed that enable the (semi-)automatic conversion of one dictionary type to another, and this drive includes the reversal of bilingual dictionaries. To name but a few examples from the 1990s' scholarly literature: Honselaar & Elstrodt (1992) presented a computer programme for the electronic reversal of a Dutch – Russian dictionary; Martin & Tamm (1996) discussed OMBI, i.e. a non-directional but linkable bilingual database from which databases and/or dictionaries in both directions can be automatically derived; and Newmark (1999) reported on his experiences with the automatic conversion of an Albanian – English dictionary.

It will be wise for any *new* African-language bilingual-dictionary project to take cognisance of the computational methods currently utilised. However, it might very well be the case that one will rather wish to start by reversing existing unidirectional African-language lexica, before embarking on technically complex projects. In this regard Newmark's observations, phi-

losophy, methods and aims for the conversion of his Albanian – English dictionary are of *utmost importance* for the African languages. These languages also have, as Newmark puts it, 'limited worldwide commercial importance' (1999:37). This means that it is unlikely that the commercial world will invest large amounts of money in dictionary compilation for African languages. Countless examples of one-way bilingual African-language dictionaries can be quoted which are out of print and/or outdated and which have never been reversed. What is needed for many such unidirectional bilingual dictionaries is very similar to what Newmark describes as 'a useful bilingual dictionary with the reverse orientation by automatic conversion of the entries in the data files from which the first dictionary was generated' (1999:37). The relative ease with which automatic conversion could be achieved will determine whether such projects will be undertaken or not. If such conversions could be done relatively quickly, with limited financial input, with few sophisticated computational and programming requirements, and with limited human intervention, the more likely it will be that such reversing activities will be undertaken.

The purpose of the remainder of this article is precisely to look into some of the potential difficulties prospective 'reversing African-language lexicographers', with only limited computational support at hand, will encounter in the future. This will be done by means of a full-scale case study of the (hypothetical) reversal of one particular dictionary for Northern Sotho. It will be pointed out that the African languages do indeed pose unique problems — problems that emanate directly from the structure of those languages. It will also be shown that the use of the forward slash causes particular additional complications.

Case Study: Reversing the *Northern Sotho Terminology and Orthography No. 4* (T&O)[4]

As a case study one can look into the reversal of the Departmental Northern Sotho Language Board's *Northern Sotho Terminology and Orthography No. 4* (1988[4]), henceforth T&O. The current T&O exists only in the direction English → Afrikaans → Northern Sotho. For many decades both speakers and learners of Northern Sotho have been in need of a Terminology & Orthography list that would make it possible to look up the Northern Sotho terms themselves; hence what is needed, as a first step, is a *'reversed T&O'* with Northern Sotho as the source language and, say, English as the target language (Northern Sotho → English). The mere fact that, for the past 14 years, no effort has been made to revise T&O, not to mention reversing it, is deplorable. This warrants an investigation into the required *degree of human intervention* assuming that: (i) the current T&O would be accessible electronically (this could e.g. be obtained by scanning, followed by optical character recognition (OCR)), and (ii) a simple software program would be available that could reverse the data in the English and Northern Sotho sections in such a way that the reversibility principle is honoured.

Although the line function of each of the recently established South African *National Lexicography Units* (NLUs) *'should eventually be the compilation of a comprehensive monolingual explanatory dictionary'* (Gouws, 2000: 111), the present authors were asked to do a preliminary study of the difficulties that would arise if the trilingual T&O were reversed in such a (semi-)automatic way. The value of such a Northern Sotho → English list for the Northern Sotho NLU can hardly be overestimated. Despite its numerous deficiencies, the fact of the matter is that T&O is the only general resource of its kind available for Northern Sotho. It is thus defendable to analyse T&O with a view to assist with the monumental task awaiting the Northern Sotho NLU. Taljard & Gauton (2001:208), for example, studied the formation of abbreviations in T&O, and based on their study they proposed abbreviations for the Northern Sotho *'parts of speech'* (2001:201–3). With the present study we aim to propose *guidelines* for the successful reversal of T&O. Not only will these guidelines be useful for any other Northern Sotho lexicon that needs to be reversed, but also for the reversal of any other unidirectional lexicon with an African language as the target language.

Before looking into the aspects revolving around the reversal itself, however, it is sensible to consider some general aspects of T&O. Firstly, T&O is an all-purpose terminology list, i.e. it does not deal with a

particular subject field, but rather covers a wide range of them. The compilers of T&O state in the foreword:

> Terms included in this list, are intended in the first place for use in the *primary school* and have mainly been taken from the syllabuses for the various subjects of the primary school. […] Terms that could be useful in training schools where instruction in some subjects such as Religious Instruction is given through the medium of Northern Sotho have also been included. Likewise, terms that might be required by writers and translators of school handbooks as well as general terms for which an increasing demand exists outside the school are included in this list. […] The list of terms was compiled with all the recognised African languages in South Africa in mind. Because comprehensive dictionaries do not exist for all the languages, some words which in the strict sense are not terms, have also been included. […] The terms are arranged alphabetically according to English, Afrikaans, with Northern Sotho in the third column without regard to the subject or course of study in which it is used. Where necessary the subject concerned is shown in brackets after the term. (Departmental Northern Sotho Language Board, 1988[4]:1)

As a result of this broad spectrum, the distribution of T&O's macrostructure is more akin to a general dictionary's macrostructure than to the macrostructure of a typical LSP dictionary. This can best be illustrated with the graph shown in Figure 1.

In Figure 1, the exact number of English lemma signs for each alphabetical category in T&O (expressed as a percentage of the total) is compared with a ruler devised by Thorndike for general-purpose dictionaries with an English-language macrostructure.[5] From Figure 1 one sees that the alphabetical breakdown of T&O is very similar to that of an LGP dictionary. The correlation coefficient *r* is as high as 0.977. As a comparison, *The American Heritage® Dictionary of the English Language, Fourth Edition* (2000[4]) is compared to the Thorndike distribution in Figure 2.[6]

In Figure 2 the correlation coefficient *r* is as high as 0.980. The patterns of the graphs in Figures 1 and 2 stand in sharp contrast with the typical pattern of an LSP dictionary. Figure 3 compares the *BioTech Life Science Dictionary* (1998), a typical LSP dictionary, with Thorndike.[7]

Although it is obvious from Figure 3 that one is still dealing with the same language (the stretches **C**, **P** and **S** remain the top categories), the LSP BioTech distribution does not follow the LGP Thorndike distribution as closely as was the case in Figures 1 and 2. The correlation coefficient *r* is only 0.788.

Secondly, T&O contains a total of 11,411 articles. The macrostructure (the English lemma signs) is to be found in a first column, the two microstructures (the translation equivalents in Afrikaans and Northern

Figure 1 Alphabetical stretches in T&O versus the Thorndike ruler
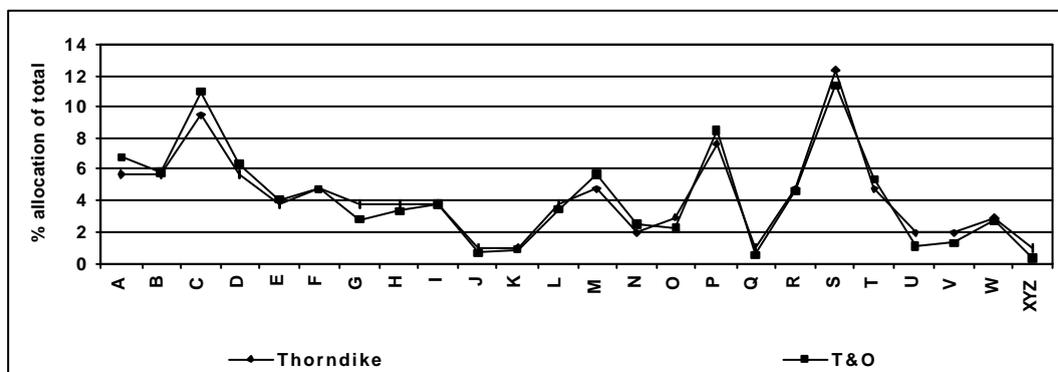
Figure 2 Alphabetical stretches in Heritage (2000[4]) versus the Thorndike ruler
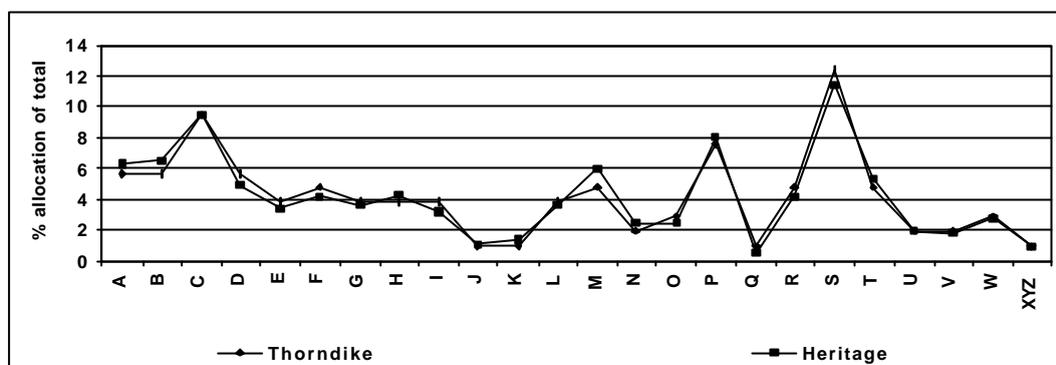
Figure 3 Alphabetical stretches in BioTech (1998) versus the Thorndike ruler
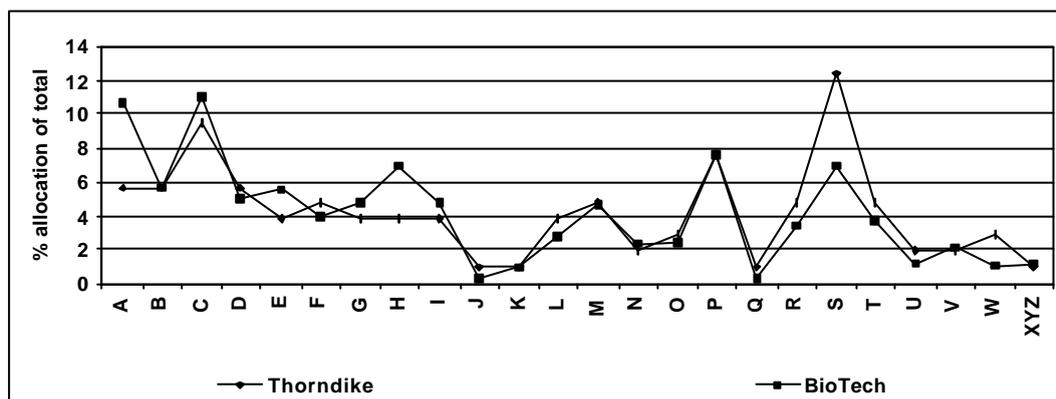
Sotho) in two subsequent columns. Most articles fit on just one line (spread over three columns). If one studies the appearance of the Northern Sotho translation equivalent paradigms (the third column), one sees that there are four main types (Types I to IV). These types are listed in Table 1, each with its respective abstract formula, a description, and the number of occurrences.

From Table 1 one can see that roughly two-thirds of the Northern Sotho translation equivalent paradigms consist of just one group of words (i.e. one word or one single string of words), possibly with bracketed parts (Type I). A small third is basically a concatenation of Type 1 paradigms, separated by commas (Type II). One out of twenty-five translation equivalent paradigms (3.9%) contain one or more forward slashes in

one way or another (Type III). Finally, a small number of equivalents can be considered as special cases, since they include colons, equal signs, quotes, etc. (Type IV).

We can now turn to the issue of the reversal of T&O. Taken at face value, reversing Types I and II should not be problematic, as the equivalent(s) could be used as lemma sign(s) in the reversed list. For Type I a simple software program can be instructed to reverse lemma sign and translation equivalent, for Type II each of the translation equivalents should become a lemma sign in the reversed list, with the English term as translation. As the translations are separated from one another by commas in the current T&O, this is easy to compute. Nonetheless, besides the question whether or not 'groups of words (possibly with brack-

Table 1 Main types of Northern Sotho translation equivalent paradigms in T&O

T	Formula	Description	#	%
I	K	A 'group of words (possibly with bracketed parts)' (**K**).	7,668	67.2
II	K, L, M, …	Two or more 'groups of words (possibly with bracketed parts)' (**K**, **L**, **M**, …) separated from one another by commas.	3,280	28.7
III	α: X/Y/Z/… : β *or* α : Bx/y/z/…E	Two or more 'groups of words' (**X, Y, Z**, …) separated by forward slashes, preceded by 'a group of words' (**α**) and/or followed by 'a group of words' (**ß**). *or* The beginning of a word (**B**) or the end of a word (**E**) combines with various parts (**x, y, z**, …) separated by forward slashes, preceded by 'a group of words' (**α**).	441	3.9
IV	special	Special cases ((**…**) *or* **:** *or* **bj.bj.** *or* = *or* **;** *or* " " etc.).	22	0.2
SUM			**11,411**	**100.0**

eted parts)' (K, L, M, …) can truly function as lemma signs — cf. the problems discussed in the theoretical conspectus above regarding surrogate paraphrases for lexical gaps, etc. — one also immediately notices that such an automated reversal is not always possible for very practical reasons in African-language lexicography. Indeed, many equivalents start with a word or concord that may not be included in the reversed list, for it would make that reversed list notably user-unfriendly.

The fact that reversing Types III and IV automatically is not straightforward either should be obvious, the more so that these types may, in addition, also have words or concords at the start of the equivalents that cannot, for practical reasons, appear in the reversed list. An overview of the number of such 'start problems' (or 'lemma-sign initial problems') per type is shown in Table 2.

As can be seen from Table 2, the 'start problems' are more serious for Types III and IV than for Types I and II. We will now first review a few typical instances of Types I and II, to illustrate the various formulae, after which we will look into the 'start problems'. Only then will instances of Types III and IV be studied.

Table 2 'Start problems' per type when reversing T&O

T	Formula	Total number per type		'Start problems' per type	
		#	%	#	%
I	K	7,668	67.2	292	3.8
II	K, L, M, …	3,280	28.7	202	6.2
III	α : X/Y/Z/… : β *or* α : Bx/y/z/…E	441	3.9	40	9.1
IV	special	22	0.2	9	40.9
SUM/Average		**11,411**	**100.0**	**543**	**4.8**

Type I – K (i.e. k1 k2 k3 …)

In this type, which accounts for 67.2% of all translation equivalent paradigms, one is dealing with a 'group of words (possibly with bracketed parts)'. The discussion of this type can successfully be broken down into a number of levels. On a first level, whenever the translation equivalent consists of just one single word (E → k1; with E the English lemma sign and k1 the single-word Northern Sotho equivalent), the automatic reversal is non-problematic. Examples from T&O are shown in (10), with the reverse (k1 → E) in (11).

(10a)	cardinal numeral	lebalapalo
(10b)	ordinal numeral	lebalatatelano
(10c)	Moslem	Lemoslem

(11a)	**lebalapalo**	cardinal numeral
(11b)	**lebalatatelano**	ordinal numeral
(11c)	**Lemoslem**	Moslem

In many cases, however, the equivalent consists of a group of words (E → k1 k2 k3 …). This is the second level. Examples are shown in (12).

(12a)	act of parliament	molao wa palamente
(12b)	mint (v)	bopa tšhelete
(12c)	trial method	mokgwa wa maitekelo
(12d)	tribal authority	pušo setšhaba

In such cases the reversal is still not problematic, yet, once the software has reversed these instances, the compilers should decide whether or not the equivalent has lemma-sign status. Even if the latter is the case, it might be wise to include such multi-word units under all (or some) of the words making up the multi-word unit (i.e. under k1, k2, k3, …). The dictionary policy and the intended target user group should of course guide the compilers in this regard. Reversing (12) might for instance result in (13), (14) or (15) depending on the type of dictionary and the intended target user group. Software can easily take care of each of these options.[8]

(13a)	**molao wa palamente**	act of parliament
(13b)	**bopa tšhelete**	mint (v)
(13c)	**mokgwa wa maitekelo**	trial method
(13d)	**pušo setšhaba**	tribal authority

(14a)	**palamente, molao wa ~**	act of parliament
(14b)	**tšhelete, bopa ~**	mint (v)
(14c)	**maitekelo, mokgwa wa ~**	trial method
(14d)	**setšhaba, pušo ~**	tribal authority

(15a)	**molao wa palamente**	act of parliament
	palamente, molao wa ~	act of parliament
(15b)	**bopa tšhelete**	mint (v)
	tšhelete, bopa ~	mint (v)
(15c)	**mokgwa wa maitekelo**	trial method
	maitekelo, mokgwa wa ~	trial method
(15d)	**pušo setšhaba**	tribal authority
	setšhaba, pušo ~	tribal authority

In (13) it is assumed that one is dealing with a dictionary which allows the inclusion of multi-word units as lemma signs. In (14) the opposite is true, i.e. only single words are entered as lemma signs. In (15) each of the components of K (i.e. k1, k2, k3, …) has been entered in the reverse list with E as translation equivalent, i.e.:

(16a)	**k1 k2 k3 … → E**
(16b)	**k2, k1 ~ k3 … → E**
(16c)	**k3, k1 k2 ~ … → E**
(16d)	etc.

The latter is definitely the most user-friendly approach, yet very space-consuming. It should also be noted, even though it is trivial, that the concords (**wa** in these examples) should not be singled out. It would for example be absurd, even though technically correct, to include articles such as (17).

(17a)	**wa, molao ~ palamente**	act of parliament
(17b)	**wa, mokgwa ~ maitekelo**	trial method

A third level is when a part of the equivalent is bracketed. Examples are shown in (18).

(18a)	bench-stop	thibedi (ya panka)
(18b)	embroidery	mokgabišo (wa moroko)
(18c)	high-relief carving	mmetlopopego (godimo)

In such cases the compilers must manually check the automated output and decide whether the bracketed part could form part of the lemma sign of a reversed list, or whether the bracketed part should rather

be taken care of in the microstructure. Compare (19) and (20) respectively.

(19a)	**thibedi (ya panka)**	bench-stop
(19b)	**mokgabišo (wa moroko)**	embroidery
(19c)	**mmetlopopego (godimo)**	high-relief carving
(20a)	**thibedi** …; **~ (ya panka)**	bench-stop
(20b)	**mokgabišo** …; **~ (wa moroko)**	embroidery
(20c)	**mmetlopopego** …; **~ (godimo)**	high-relief carving

On a fourth level one observes that there are several instances in T&O of articles in which an attempt is made to delimit the exact sense by means of collocates or labels. Examples include:

(21a) knob (of knitting needle)
          hlogwana (ya nalete ya go loga)
(21b) list (of names, articles, etc.)
          lenano (la maina, dilo, bj.bj.)

When reversing such articles, the bracketed information belongs to the microstructure, as shown in (22).

(22a) **hlogwana** *(ya nalete ya go loga)*
          knob *(of knitting needle)*
(22b) **lenano** *(la maina, dilo, bj.bj.)*
          list *(of names, articles, etc.)*

As software cannot be programmed to differentiate between levels three and four, manual intervention will be required in these cases.

On a fifth level, some articles contain extra information, often the full form of an abbreviation, which also has lemma-sign status in a reversed list. Manual intervention will again be necessary. Examples from T&O are shown in (23), with a possible reverse in (24).

(23a)	h (height)	bg. (bogodimo)
(23b)	T.C.P.	T.C.P. (Thisiphi)
(24a)	**bg.** *(bogodimo)*	h *(height)*
	bogodimo *(bg.)*	height *(h)*
(24b)	**T.C.P.** *(Thisiphi)*	T.C.P.
	Thisiphi *(T.C.P.)*	T.C.P.

Unfortunately, as a last sixth level, there are numerous instances of the use of brackets where one should have used commas instead. Such instances obviously need human intervention. Examples include:

(25a)	exponent (arith.)	sematlafatši (eksponente)
(25b)	hairpin bend	morarelo wa nkgopo (kgopamo ya nkgopo)
(25c)	sting	lebolela (lebola)

With commas, examples such as (25) belong to Type II, topic of the next section.

Type II – K, L, M, … (i.e. k1 k2 k3 …, l1 l2 l3 …, m1 m2 m3 …, …)

Since Type II is basically only a concatenation of Type I formulae, separated by commas, the present discussion can be brief. As mentioned above, software can be programmed to use the comma as a delimiter between potential lemma signs for the reverse list. Once broken up, each of the parts may be handled as an instance of Type I. Some typical examples of levels one to four include:

(26a) explorer
          moutolli, mohlohlomiši, mohlotletši
(26b) susceptibility
          tseno malwetši, peko, tšhabelelo
(26c) sweet-oil
          makhura (mohlware), sutuoli
(26d) iris (of the eye)
          leratladi (la leihlo), aerise

In a reversed list, these could be included as:

(27a)	**moutolli**	explorer
	mohlohlomiši	explorer
	mohlotletši	explorer
(27b)	**tseno malwetši**	susceptibility
	peko	susceptibility
	tšhabelelo	susceptibility
(27c)	**makhura (mohlware)**	sweet-oil
	sutuoli	sweet-oil
(27d)	**leratladi** *(la leihlo)*	iris *(of the eye)*
	aerise	iris *(of the eye)*

Variants for the first reverse shown in (27b) and the first reverse shown in (27c) are respectively:

(28a)	**malwetši, tseno ~**	susceptibility
(28b)	**makhura** …; **~ (mohlware)**	sweet-oil

'Start problems' when reversing T&O

Before examining the problems involved when reversing Types III and IV, it is appropriate to firstly study the 'start problems' when reversing T&O, as these occur for the four types, and only become more complex when they come on top of the intrinsic problems of Types III and IV.

In Table 2 it was indicated that there are a total of 543 Northern Sotho translation equivalent paradigms out of 11,411 (or thus 4.8%) which contain initial words or concords that should not appear as initial words or concords in a reversed list. In the discussion of Type II it was shown that a reversed list 'grows' in size, as each of the equivalents (K, L, M, …) must be considered as a candidate for inclusion as lemma sign in the reversed list. Even reversing Type I items can result in a 'growing' list, cf. (15) and (24). As will be shown below, also the reversal of Types III and IV results in a 'growing' list. More in particular, the original 543 equivalents with 'start problems', add up to 687 potential articles in a reversed list, an increase of 26.5%. All types taken together, the various kinds of words and concords causing problems are shown in Tables 3 to 6, grouped in logical sets, with an indication of their occurrence frequencies.

Table 3   Possessive concords as 'start problems'

Possessive concords	#	%
wa ~	10	1.5
ba ~	3	0.4
ya ~	47	6.8
la ~	5	0.7
a ~	3	0.4
sa ~	57	8.3
tša ~	6	0.9
(wa) ~	3	0.4
(ya) ~	47	6.8
(la) ~	2	0.3
(a) ~	1	0.1
(sa) ~	24	3.5
(tša) ~	3	0.4
(ga) ~	1	0.1
SUM	**212**	**30.9**

Table 4   Other concords, morphemes and pronouns as 'start problems'

Other concords, morphemes and pronouns	#	%
e ~	25	3.6
se ~	23	3.3
(ye) ~	5	0.7
(ye) e ~	1	0.1
(se) ~	5	0.7
yo ~	3	0.4
o ~	1	0.1
(yo a) di ~	1	0.1
SUM	**64**	**9.3**

Table 5   Infinitive class prefix and particles as 'start problems'

Infinitive class prefix and particles	#	%
go ~	184	26.8
go ba ~	3	0.4
go fa ~	3	0.4
go wa ga ~	1	0.1
go wa (ga ~)	1	0.1
go yo di mo ~	1	0.1
(go ya ka) ~	1	0.1
ka ~	156	22.7
(ka) ~	5	0.7
SUM	**355**	**51.7**

Table 6   Dashes and brackets as 'start problems'

Dashes and brackets	#	%
-~ *or* - ~	47	6.8
(~)	9	1.3
SUM	**56**	**8.2**

Translation equivalents which include initial concords, relative pronouns, infinitive prefixes, negative morphemes, etc. cannot be lemmatised as such, namely in the alphabetical stretch according to the alphabetical category to which the concords, pronouns, etc. belong. Thus none of the initial elements listed in Tables 3 to 5 which occur in T&O as initial element to translation equivalents, qualify as alphabetised initial parts of lemma signs. Likewise, a reverse list cannot include

lemma signs that start with erroneous dash-initial items or bracketed parts, cf. Table 6. The different categories of these initial elements will now be evaluated and discussed in more detail in order to determine how such translation equivalents should be presented as lemma signs in a reversed dictionary. It should already be obvious that these start problems can hardly be handled successfully automatically, and that human editing is the rule for these cases.

POSSESSIVE CONCORDS as 'start problems' (Table 3)

The following possessive concords are used as the first elements of translation equivalents:

(29)   **wa ~, ba ~, ya ~, la ~, a ~, sa ~, tša ~, (wa) ~, (ya) ~, (la) ~, (a) ~, (sa) ~, (tša) ~, and (ga) ~**

Examples from T&O include:

(30a) astute   wa matšato go hlalefa
(30b) manual (adj.)   (wa) diatla

The absence or presence of brackets, such as in (30a) versus (30b), is merely the result of inconsistent treatment and has no specific conventional function. The bracketed **(sa)**, on the other hand, is consistently used as possessive concord but the unbracketed **sa** is once again inconsistently used as either possessive concord or negative morpheme. This is illustrated in (31a) and (31b) respectively.

(31a) annual (n) (bot.)   sa ngwaga
(31b) unglazed   - sa phadimišwago ...

Note further that **sa** (just like many other items, cf. Table 6) is sometimes inconsistently preceded by a dash without any specific conventional function.

In the reverse side of the dictionary the noun should be lemmatised followed by a convention representing the possessive concord *paradigm* (and not the current haphazard choice for this or that concord) and a tilde '~', e.g. for (31a):

(32)   **ngwaga, ya/tša/wa/... ~**   annual (n) (bot.)

OTHER CONCORDS, MORPHEMES AND PRONOUNS as 'start problems' (Table 4)

Apart from possessive concords, a variety of other concords/morphemes/pronouns and combinations of these are used as the first elements of translation equivalents:

(33)   **e ~, se ~, (ye) ~, (ye) e ~, (se) ~, yo ~, o ~ and (yo a) di ~**

Examples from T&O include:

(34a) corrosive   (ye) e jago, e senyago
(34b) definite   (ye) itšego
(34c) exciting   e huduago, thakgatšago

In (34a-c) direct relative constructions (Class 9) are used as translation equivalents. However, in (34a) both relative pronoun and subject concord are given for the first equivalent, but only the subject concord for the second equivalent, in (34b) only the relative pronoun is included, and in (34c) only the subject concord is given for the first equivalent, and neither of the two for the second equivalent. One is thus dealing with *four* different ways of presenting the same construction. Furthermore, this presentation creates the impression that the translation equivalent is restricted to Class 9. In fact, depending on the subject, any of the concords of the other noun classes can be generated. This, once again, underlines the need for a user-friendly but accurate convention to represent the entire *paradigm* of e.g. subject concords, object concords, relative pronouns, etc. (cf. also Prinsloo & Gouws, *forthcoming*). One could argue that reflecting the concords of just one noun class is defendable in a severely limited number of translation equivalents, e.g.:

(35)   antiseptic (adj.)   (se) šitišago ... phero

In this example **sehlare** 'medicine' or **setlolo** 'ointment' seem the most likely subjects and both are in Class 7, thus generating **se**. Yet, since **se** is a required element at that point, the brackets around it would have to be removed. Unfortunately, assuming that the most probable antecedent for **šitišago** would be a Class

7 noun is not borne out by the corpus. One is thus compelled to assume that if a construction can take one concord, it can theoretically take them all.

An example was found where no less than *three* concords are used as the first elements of a translation equivalent, complicating the selection of a lemma sign for the reverse side of the dictionary:

(36)   adept (adj.)   (yo a) di bonego

In the reverse side of the dictionary the verb should be lemmatised followed by a convention representing the relative pronoun paradigm and/or the subject concord paradigm and/or the object concord paradigm, and then a tilde '~', e.g. for (34b) and (36) respectively:

(37a) **itšego, yo/ye/tše/… ~**
       definite
(37b) **bonego, (yo/ye/tše/… a/e/di/…) di ~**
       adept (adj.)

The negative morphemes **ga**, **sa** and **se** are also used in many translation equivalents as the first elements:

(38a) no entry   ga go tsenwe
(38b) no overtaking   ga go fetanwe

(38c) non-ruminant (adj.)   sa otleng, sa otlego
(38d) uncorrected   sa ntšhwago phošo

(38e) absent (v)   … se be gona
(38f) non-poisonous   - se nago mpholo

When reversing a terminology list, the selection of a negative morpheme is fairly restricted (as either **ga** or **sa** or **se**) and in the reversed section the relevant negative morpheme should be presented following the lemma sign. Sound reversals are thus:

(39a) **tsenwe, ga go ~**   no entry
(39b) **fetanwe, ga go ~**   no overtaking

(39c1)**otleng, sa ~**   non-ruminant (adj.)
(39c2)**otlego, sa ~**   non-ruminant (adj.)
(39d) **ntšhwago phošo, sa ~** uncorrected

(39e) **gona, se be ~**   absent (v)
(39f) **nago mpholo, se ~**   non-poisonous

Note however that when a *general bilingual dictionary* is reversed, verb stems such as **tsena**, **fetana** and **otla** in (39a) to (39c) above will take **ga**, **sa** or **se** depending on the mood of the verb. In such cases it will be advisable to use the **..ga/sa/se..~** convention presented in Prinsloo & Gouws (1996).

INFINITIVE CLASS PREFIX AND PARTICLES as 'start problems' (Table 5)

The infinitive prefix (Class 15), with 26.8% of all 'start problems', is the most frequently used first element of translation equivalents in T&O. Examples include:

(40a) germination   go mela …
(40b) perusal   go badišiša …

In the reversed section such nouns should be presented as:

(41a) **mela, go ~**   germination
(41b) **badišiša, go ~**   perusal

In the case of fixed expressions lemmatisation under **go** can, however, be considered. In (42) an example from T&O is shown, with (43a) and (43b) sound reversals.

(42)   to whom it may concern go yo di mo amago

(43a) **go yo di mo amago**   to whom it may concern
(43b) **amago, go yo di mo ~** to whom it may concern

Depending on the dictionary policy and the intended target user group, one could opt for either (43a) or (43b), or for both (43a) and (43b) simultaneously. Lemmatising **to whom it may concern** in T&O under **to** (cf. (42)), although inconsistent with the other lemma signs with initial part **to**, is justifiable for such a fixed expression. However, to lemmatise **bear a flower**, **bear fruit**, **circle**, **magnetise a knitting needle** and **purify water** respectively under **to bear a flower**, **to bear fruit**, **to circle**, **to magnetise a**

knitting needle and **to purify water** makes no sense — yet this was done in T&O. Even if these so-called 'terms' were meant to be translations of labels for merchandise, e.g. *a product* **to magnetise a knitting needle**, or *a product* **to purify water**, where **to** thus indicates the purpose or intention of an action, then the infinitive marker **to** should at least have been translated in Northern Sotho as **go** as well. Yet, T&O has (44) instead of (45).

(44a) to magnetise a knitting needle
　　　 maknethaesa nalete ya go loga
(44b) to purify water
　　　 sekiša meetse

(45a) to magnetise a knitting needle
　　　 go maknethaesa nalete ya go loga
(45b) to purify water
　　　 go sekiša meetse

Reverse articles for translation equivalents with the potential or with particles such as the instrumental **ka** are straightforward. Such elements should be presented following the noun or verb with the tilde '~'; thus (46), for instance, could be lemmatised as (47) on the reverse side.

(46a) active　　　　　　(ka) mahlahla …
(46b) methodically　　　ka tsela … *(instrumental)*
(46c) underground　　　ka tlase ga mobu …
　　　　　　　　　　　　 (locative)

(47a) **mahlahla, ka ~**　　 active
(47b) **tsela, ka ~**　　　　 methodically
(47c) **tlase ga mobu, ka ~** underground

Unfortunately, as can be seen from (46a) versus (46b) and (46c), brackets around **ka** are wrongly added in some cases.

DASHES AND BRACKETS as 'start problems' (Table 6)

In (31b) we already saw that dashes are often used without any conventional function. In other cases, however, they do have a function, e.g. to mark that an adjective takes on a class prefix in an adjectival con-

struction or that a relative takes on a relative pronoun and/or a subject concord:

(48a) minor (lessor) (adj.)　　 -nyane
(48b) questionable　　　　　　 -belaetšago

The use of the dash in cases such as (48b) is, however, inconsistent with the explicit display of a concord paradigm, such as in for instance (34b). Examples (34b) and (48b) are both relative constructions, yet two different strategies were used in T&O to indicate this. Automatically reversing these different strategies will simply perpetuate the inconsistencies, which is obviously not desirable.

Finally, translation equivalents that start with bracketed parts cannot be reversed blindly either, e.g.:

(49)　 mimic　　　　(ketšaetšane, mankekišane) ekiša

Such instances will be discussed in detail under the discussion of Type IV.

On the whole it should be clear from the above discussion that all the mentioned 'start-problems' cannot really be handled in an automatic way by a computer program. It will be much safer to check all these instances manually.

Type III —— α : X/Y/Z/… : ß *or* α : Bx/y/z/…E

In a different article (Prinsloo & De Schryver, 2002b), the use of the forward slash in T&O's Northern Sotho translation equivalent paradigms is thoroughly analysed. There it is shown that the forward slash is used in eight different environments, five of which can be considered 'main types' (Types III.1 to III.5). A brief overview of these environments, with their respective formulae, is shown in Table 7.

Types III.1 to III.3 deal with forward slashes on multi-word level, whilst Types III.4 and III.5 deal with forward slashes on word-level. In Prinsloo & De Schryver (2002b) it is pointed out that an unambiguous use of slashes for Types III.1 to III.3 requires that the forward slashes separate single words only; and that a correct use of slashes for Types III.4 and III.5 could be to utilise both vertical and forward slashes, or else to separate the alternative words with commas.

Types III.2 and III.3 could be collapsed with Type

Table 7   Types of forward slash uses in T&O's Northern Sotho equivalents

T	Formula	Description	#	%
III.1	$\alpha : X/Y/Z/\ldots : \beta$	Two or more 'groups of words' (**X, Y, Z**, …) separated by forward slashes, preceded by 'a group of words' (α) and/or followed by 'a group of words' (β).	344	78.0
III.2	$C : \alpha : X/Y/Z/\ldots$	A concord (paradigm) (**C**) and 'a group of words' (α) precede two or more 'groups of words' (**X, Y, Z**, …) separated by forward slashes.	13	2.9
III.3	$X/Y/Z/\ldots : *C : \beta$	Two or more 'groups of words' (**X, Y, Z**, …) separated by forward slashes, followed by a partly erroneous concord (***C**), followed by 'a group of words' (β)	12	2.7
III.4	$\alpha : Bx/y/z/\ldots$	The beginning of a word (**B**) combines with various parts (**x, y, z**, ...) separated by forward slashes, preceded by 'a group of words' (α).	19	4.3
III.5	$\alpha : x/y/z/\ldots E$	The end of a word (**E**) combines with various parts (**x, y, z**, ...) separated by forward slashes, preceded by 'a group of words' (α).	6	1.4
III.6	special /	A special case with /.	1	0.2
III.7	error	Forward slashes and commas are used without any logic.	41	9.3
III.8	?	Impossible to decode.	5	1.1
SUM			**441**	**99.9**

III.1 ($\alpha : X/Y/Z/\ldots : \beta$), since the former two can be considered special cases of the latter. Furthermore, Types III.4 and III.5 could also be collapsed into one new type ($\alpha: Bx/y/z/\ldots E$). Especially with reverse purposes in mind, however, working with the mentioned five different main types (and not with two 'bigger' main types) is extremely relevant. Reversing these five main types, and the three subsidiary types (Types III.6 to III.8), can now be reviewed one by one.

From a sound lexicographic point of view, reversing Type III.1 ($\alpha: X/Y/Z/\ldots : \beta$) will only result in an unambiguous reading when X, Y, Z, ... are each a single word (symbolically: X, Y, Z, … = 1). Consider the following example:

(50)   public works
           mediro ya setšhaba/mang le mang

As X and Y are not both = 1 in (50), there is no way that (basic) software can differentiate automati-

cally between the two options: does the forward slash stand for 1/1 **setšhaba** *or* **mang**, 2/2 **ya setšhaba** *or* **mang le**, 3/3 **mediro ya setšhaba** or **mang le mang**, etc. (actually it stands for 1/3 **setšhaba** or **mang le mang**).

Even when X, Y, Z, … = 1, there are still three options. Firstly, when α is present (with β present or not), the reversal is non-problematic:

(51)   Department of Public Works
           Kgoro ya Mešomo/Mediro ya Mmušo

(52)   **Kgoro ya Mešomo/Mediro ya Mmušo**
           Department of Public Works

Secondly, when α is not present but β is, then the reverse list must contain articles under X, Y, Z, ..., at which point the forward slashes disappear, and/or lemmatisation is done under β. Examples before and after conversion are shown in (53) and (54) respectively.

(53a)   Citrus Board
        Boto/Lekgotla la senamune/Sitrase
(53b)   omnipresent
        lego/bego gohle

(54a)   **Boto la senamune/Sitrase**
        Citrus Board
        Lekgotla la senamune/Sitrase
        Citrus Board
(54b1) **lego gohle**
        omnipresent
        bego gohle
        omnipresent
(54b2) **gohle, lego/bego ~**
        omnipresent

Thirdly, when both α and β are absent, the reverse list *must* contain articles under X, Y, Z, ...:

(55)   wild dog   lekanyane/lehlalerwa

(56a)   **lekanyane**   wild dog
(56b)   **lehlalerwa**   wild dog

Even so, in African-language lexicography, there are examples of exceptions to the metalexicographic notion that *forward slashes* should not separate different words *in lemma signs*. In Botne & Kulemeka's (1995) *A Learner's Chichewa and English Dictionary*, for instance, nouns are entered in the format '**singular/plural**'. Examples of such lemma signs are:

(57a)   **munthu/anthu** *n* person [agr = a*-/a-].
(57b)   **mwendo/miyendo** *n* leg [agr = u-/i-].
(57c)   **phwando/maphwando** *n* feast [agr = li-/a-].

Reversing Type III.2 (C : α : X/Y/Z/...) is in principle not feasible, as this would imply the inclusion of the same microstructure under each of the concords (cf. discussion above). Each concord would moreover take up numerous pages to cater for each instance of this type. A simple solution, however, is to enter such items in the reversed lexicon under α as shown below:

(58)   α**, C   ~   X/Y/Z/...**

In order to convey the correct linguistic informa-

tion, **C** must be a concord paradigm (such as e.g. **ya/tša/wa/...** as discussed above), or a symbol for it (such as e.g. **lk** for **lekgokamong** 'possessive concord') with matching front- (or back-)matter guidance. Since such lemma signs actually have lemma-sign status, one should also consider marking such entries with a unique non-typographical structural marker. In the extract shown below, taken from the *Beknopt woordenboek Ciluba — Nederlands* (BCN) (De Schryver & Kabuta, 1998[2]), that marker is the square '▣':

(59)   **lufù** [11/4 < **-fwà**; cf spw 14] de dood
        ▣ **-à ~** [cn adj] **1** dodelijk; **2** gevaarlijk

In Ciluba **lufù** 'death' is a noun with gender 11/4, is derived from the verb **kufwà** 'die, pass away', and belongs to the top 400 lemma signs (frequency band 2). The construction 'possessive concord + **lufù**' means '**1** deadly, mortal, lethal; **2** dangerous'. Firstly, to indicate that this construction actually has lemma-sign status — or from a metalexicographic point of view: to point out that the macrostructure has been 'inserted' into the microstructure (cf. De Schryver & Kabuta, 1998[2]: xiii) — this construction is preceded by the square '▣' in the Extra Column (on the left in BCN). Secondly, the 'symbol' for a possessive concord paradigm is '-à' in BCN. (The possessive concords themselves are formed by prefixing the pronominal prefixes to '-à', resulting in e.g. **bàà** for Class 2, **dyà** for Class 5, **kàà** for Class 12, etc.)

Also for this type X, Y, Z, ... should each be a single word in order to enable an unambiguous reading. If α is absent, then the reversed lexicon must include articles under X, Y, Z, ..., at which point the forward slashes disappear.

Reversing Type III.3 (X/Y/Z/... : *C : β) is in principle unacceptable, unless articles are provided under X, Y, Z, ..., at which point the forward slashes disappear, and/ or lemmatisation is done under β. This can only be done unambiguously when X, Y, Z, ... are a single word each, and when they all take the same concord. Examples before and after (a partly failed) conversion are shown in (60) and (61) respectively.

(60a)   hagiography
        histori/ditiragalo tša bakgethwa

(60b)   jockey club
         klapo/mokgatlo wa bakatišapere …

(61a1)   **histori *tša bakgethwa**
         hagiography
         ditiragalo tša bakgethwa
         hagiography
(61a2)   **bakgethwa, histori/ditiragalo *tša ~**
         hagiography
(61b1)   **klapo *wa bakatišapere**
         jockey club
         mokgatlo wa bakatišapere
         jockey club
(61b2)   **bakatišapere, klapo/mokgatlo *wa ~**
         jockey club

Reversing Type III.4 (α: Bx/y/z/...), whether or not α is present, is, in principle, unacceptable. Examples before and after (a partly failed) conversion are shown in (62) and (63) respectively.

(62a)   executor   moabalefa/bohwa
(62b)   key industries   diintasterikhii/theo/kgolo

(63a)   **moabalefa**   executor
        *bohwa   executor
(63b)   **diintasterikhii**   key industries
        *theo   key industries
        *kgolo   key industries

One can, however, accept those cases where one is dealing with a variant final, such as in the following example taken from BCN:

(64)   **mwoyo/i** [3/4] **1** 't leven; **2** [vgl **mucìma**[1] 1] hart; **~ (wòmba) tukùtukù** [ud; (VD)] *met 'n kloppend hart*; **3** groet; *kwela* **~** [ud] *groeten*

In (64), **mwoyo** '**1** life; **2** heart; **3** greeting', with as variant **mwoyi**, was placed first as it is the more frequent one (cf. De Schryver & Prinsloo, 2000: 304).

Reversing Type III.5 (α : x/y/z/...E), whether or not α is present, is, in principle, unacceptable. An example before and after (a partly failed) conversion is shown in (65) and (66) respectively.

(65)   airways terminal   bogoma/bokomafofane

(66)   ***bogoma**   airways terminal
        bokomafofane   airways terminal

Nonetheless, one possible way to make the inclusion of this type acceptable has been provided by Rycroft in his *Concise SiSwati Dictionary* (1981). He consistently uses forward slashes to indicate 'or what follows (e.g. a plural prefix instead of singular)':

(67a)   **im-phúmulo /tim-** *n*. nose, snout.
(67b)   **lí-phílo /éma-** *n*. pillow-case.

This convention should thus be interpreted as **imphúmulo** for the singular *or* **timphúmulo** for the plural, and **líphílo** for the singular *or* **émaphílo** for the plural respectively.

Finally, reversing the special case **taolo ya ema/ sepela pejana** 'stop/go control ahead' (Type III.6) is an easy matter, as it can be entered as such, as a special kind of multi-word lemma sign. The microstructure should, however, explain its special status. Those instances where forward slashes and commas are used without any logic (Type III.7) should first be corrected and then be handled as one of the previous types. The paradigms that are hard to decode at present (Type III.8) must first be simplified (potentially following extra fieldwork) and then also be handled as one of the previous types.

It should be clear from the discussion of Type III that automatically reversing any translation equivalent paradigm containing forward slashes is likely to produce erroneous lemma signs. This is mainly a result of the often faulty, inconsistent and haphazard use of forward slashes in the current T&O. Each of those instances should thus be handled manually.

Type IV – Special cases

Reversing the so-called special cases requires in principle that the lexicographer must decide which word(s) should get lemma-sign status in the reverse side. The special cases are thus bound to fail if handled by a software program only.

START BETWEEN BRACKETS as a special case

It is unlikely that a general rule for reversing transla-
tion equivalents starting with a bracketed section can
be formulated since each equivalent has to be evalu-
ated in its own right. Inconsistencies and even errone-
ous use of brackets in T&O aggravate the problem:

(68a)   hedged in (reply)
       (karabo) kopana
(68b)   mental hygiene
       (maphelobjoko) tlhokomelobjoko

    In the case of (68a) no brackets should have been
used in both lemma sign and translation equivalent,
thus rendering (69) and (70) as article and reversed
article respectively.

(69)   hedged in reply   karabokopana

(70)   **karabokopana**   hedged in reply

    In the case of examples such as (68b),
maphelobjoko, literally 'health-mental', and
tlhokomelobjoko, literally 'care-mental', a comma
instead of brackets should separate the translation
equivalents. Both translation equivalents should re-
ceive lemma-sign status in the reversed section, thus
respectively:

(71)   mental hygiene
       maphelobjoko, tlhokomelobjoko

(72a)   **maphelobjoko**   mental hygiene
(72b)   **tlhokomelobjoko**   mental hygiene

COLON as a special case

If the translation equivalent includes a colon, the en-
tire translation equivalent could be lemmatised in the
reverse section, e.g. (73) and (74) respectively.

(73a)   chief: language service
       hlogo: tirelo ya polelo
(73b)   Limited hours: No parking
       Diiri tše beilwego: Ga go phakwe
(73c)   three periods: no
       dipakatharo: aowa

(74a)   **hlogo: tirelo ya polelo**
       chief: language service
(74b)   **Diiri tše beilwego: Ga go phakwe**
       Limited hours: No parking
(74c)   **dipakatharo: aowa**
       three periods: no

    Lemmatisations under other parts of the multi-
word units could however also be considered.

ETCETERA as a special case

The functionality of **bj.bj.** 'etc.' is questionable and
should only be used if a *logical series* of additional
applicable equivalents is implied. The suggested re-
verse articles for (75), for example, are given in (76).

(75a)   consist
       bopilwe ka, hlophilwe ka, bj.bj.
(75b)[9]   matching (of colours, material, etc.) (v)
       nyalanya (ya mebala) materiale, bj. bj.

(76a1)   **bopilwe ka**
       consist
(76a2)   **hlophilwe ka**
       consist
(76b)   **nyalanya** *(ya mebala, materiale, bj.bj.)*
       matching *(of colours, material, etc.)* (v)

EXPLANATION as a special case

In the following example each translation equivalent
should be considered for lemma-sign status in the re-
verse section:

(77)   in the work quoted=opere citato=op. cit
       dingwalong tše di tsopotšwego=opere
       citato=op. cit

    In fact (77) should be lemmatised under **in the
work quoted**, **opere citato** and **op. cit.** as in (78),
with the entire reversed treatment as in (79).

(78a)   in the work quoted
       dingwalong tše di tsopotšwego, opere citato,
       op. cit.

(78b) opere citato
       dingwalong tše di tsopotšwego, opere citato,
       op. cit.
(78c) op. cit.
       dingwalong tše di tsopotšwego, opere citato,
       op. cit.

(79a) **dingwalong tše di tsopotšwego**
       in the work quoted, opere citato, op. cit.
(79b) **opere citato**
       in the work quoted, opere citato, op. cit.
(79c) **op. cit.**
       in the work quoted, opere citato, op. cit.

SEMICOLON as a special case

One example was found in T&O with a semicolon in
the Northern Sotho translation equivalent paradigm:

(80)   cardinal (n)   ye kgolo; mokhadinala

     In this specific example where a semicolon sepa-
rates the translation equivalents, only the second
equivalent should be considered as a lemma sign in the
reverse section since a lemma sign **ye kgolo** with the
general meaning 'a/the big one' cannot logically trigger
cardinal as a translation equivalent.

QUOTES as a special case

The functionality of using quotes can be questioned
and can have negative alphabetical sorting implications.
Note also the inconsistency in the use of double ver-
sus single quotes in the source and target languages:

(81)   "safety-first"
       'totego pele', 'polokego pele'

     Example (81) should be given as (82) and its sug-
gested reversed articles as (83), thus without quota-
tion marks.

(82)   safety first   totego pele, polokego pele

(83a) **totego pele**   safety first
(83b) **polokego pele**   safety first

NO EQUIVALENT as a special case

Instances where no equivalents are given are simply
errors that should be rectified manually by giving one
or more suitable equivalent(s) and by selecting the
lemma sign(s) for the reverse section on the principles
outlined above:

(84)   poliomyelitis   Ø

(85)   **polio**   poliomyelitis

Degree of human intervention required in a semi-automated reversal of T&O

We can now return to the *degree of human interven-
tion* required in a semi-automated reversal of T&O.
The case study has clearly indicated that two catego-
ries (Type III with 441 items, and Type IV with 22
items) cannot be handled without human intervention.
Further, all start problems for Types I and II (292
items and 202 items respectively) should be checked
manually as well. Summating these gives 957 items.
Since there are 11,411 articles in all in T&O, this means
that, at best, circa 90% of the reversals could be suc-
cessfully taken care of by software. The real figure is
unfortunately much lower. Firstly, we also saw that
levels 3 and higher of Types I and II need to be checked
manually.

     Secondly, the resulting list needs heavy editing.
Indeed, reversing the entire 11,411-article-strong T&O
as described above, results in a reversed list with 15,828
potential lemma signs, thus an increase of 38.7%. In
this reversed list numerous identical lemma signs have
been generated as a result of convergence (in the cur-
rent T&O) now resulting in redundant divergence (in
the reverse section). Compare for example (86) where
three occurrences of **aba** as translation equivalents
generated three identical lemma signs in (87) that should
be reduced to a single lemma sign as in (88).

(86a) allocate   aba
(86b) confer   aba   *convergence*
(86c) designate   aba …

(87a) **aba**   allocate
(87b) **aba**   confer   *divergence*

(87c) **aba**   designate

(88) **aba**   allocate; designate; confer

Even though an automated reversal of T&O would only be successful for less than 90% of the articles, and even though the resulting reversed list would still need heavy editing, there can be no doubt that having such a reversed list at one's disposal would be far better than having to do the entire job manually.

Summary and conclusion

In this article we have firstly reviewed the basic notions one needs to be aware of in (bidirectional) bilingual lexicography, i.e. the main types of equivalence relations and the reversibility principle. We have also pointed out that numerous African-language lexica exist that are only available with the African languages as target languages, and that there is a dire need for reversed lexica in which the African-language items could be looked up directly.

In an attempt to evaluate the success of computationally-simple reverse software, we then presented a full-scale case study of the (hypothetical) reversal of the *Northern Sotho Terminology and Orthography No. 4* (T&O). As parallel lists exist for the other official African languages of South Africa, the direct multiplication potential is hardly imaginable. In (2) Hartmann & James were quoted as saying: 'In bilingual or multilingual TERMINOLOGICAL DICTIONARIES, equivalence implies interlingual correspondence of DESIGNATIONS for identical CONCEPTS'. The latter entails that reversing a terminology list is, by default, primarily a macrostructural exercise. The focus was therefore on the lemma signs in the reversed list, and not on the microstructures. Besides the obvious relevance for the case study proper and Northern Sotho by extension, the results are also important for African-language lexicography as a whole. We therefore summarise the various results in Table 8.

In Table 8 both the various types of Northern Sotho translation equivalent paradigms (in Roman) and the conditions for successful (semi-)automatic reversal (in italics) have been listed. Type III is especially detailed, as conditions were stipulated beyond which

the use of the forward slash cannot really be considered user-friendly anymore. This leads to a first important general conclusion, namely that the forward slash should be implemented with a much more rigid set of rules governing its use than is generally the case today.

Another important conclusion can be drawn from the discussion of the 'lemma-sign initial problems' (cf. Tables 2 to 6). These occur across all types. The result with the greatest consequence in this regard for the *African languages* is doubtless the realisation that there is a dire need for the creation of concord paradigms on the one hand, and a need for the indication of the lemma-sign status of the resulting items on the other. This topic will be further investigated in forthcoming endeavours, yet numerous ideas and practical solutions were already suggested above.

Finally, we have also pointed out that it will be necessary to go manually through the computationally-reversed list, as the original convergence gave way to divergence. This human intervention obviously also enables one to deal with and sort out inconsistencies, to correct mistakes in general, and to make important suggestions for improvement of the L1 → L2 section in view of a future revision thereof. We therefore wish to suggest that the semi-automatic reversal of African-language lexica should seriously be considered as a viable option, as it is not necessarily computationally complex and definitely faster than a purely manual reversal.

NOTES

1   Since this article is being submitted for publication in a South African journal, necessary sensitivity with regard to the term 'Bantu' languages is exercised in the authors' choice rather to use the term *African* languages. Keep in mind, however, that the latter includes more than just the 'Bantu Language Family'.

2   When dealing with lexical gaps, dictionary compilers often feel tempted to 'coin' a term. For **maganagodiša**, for instance, Kriel (1976[4]) suggests 'a boy who refuses to herd; a herd-funk' as translation-equivalent paradigm. Funk is an old-fashioned British verb meaning 'avoid doing something because one is afraid'.

Table 8   Main types of Northern Sotho translation equivalent paradigms in T&O

T	Formula	Description	#	%
I	K	A 'group of words (possibly with bracketed parts)' (**K**).	7,668	67.20
		Feasible to reverse automatically, except when there is a bracketed part (levels I.3 and higher)		
II	K, L, M, …	Two or more 'groups of words (possibly with bracketed parts)' (**K, L, M**, …) separated from one another by commas.	3,280	28.74
		Feasible to reverse automatically, except when there is a bracketed part (levels II.3 and higher)		
III.1	α : X/Y/Z/… : β	Two or more 'groups of words' (**X, Y, Z**, …) separated by forward slashes, preceded by 'a group of words' (α) and/or followed by 'a group of words' (ß).	344	3.01
		Only acceptable when X, Y, Z, … = 1 AND • *when α is present, ß is or is not present: reversible* • *when α is absent, but ß is present: reverse articles under X, Y, Z, … (forward slashes disappear) and/or under ß* • *when both α and ß are absent: reverse articles under X, Y, Z, … (forward slashes disappear)*		
III.2	C : α : X/Y/Z/…	A concord (paradigm) (**C**) and 'a group of words' (α) precede two or more 'groups of words' (**X, Y, Z**, …) separated by forward slashes.	13	0.11
		Unacceptable, unless X, Y, Z, … = 1 AND a convention for the 'concord paradigm' is developed (+ extra marker for multi-word unit with lemma-sign status) AND • *when α is present: reverse under α* • *when α is absent: reverse articles under X, Y, Z, … (forward slashes disappear)*		
III.3	X/Y/Z/… : *C : β	Two or more 'groups of words' (**X, Y, Z**, …) separated by forward slashes, followed by a partly erroneous concord (***C**), followed by 'a group of words' (ß).	12	0.11
		Unacceptable, unless X, Y, Z, … = 1 AND reverse articles under X, Y, Z, … (forward slashes disappear) + each with correct concord		
III.4	α : Bx/y/z/…	The beginning of a word (**B**) combines with various parts (**x, y, z**, ...) separated by forward slashes, preceded by 'a group of words' (α).	19	0.17
		Unacceptable, except where only a variant final; α is or is not present		
III.5	α : x/y/z/…E	The end of a word (**E**) combines with various parts (**x, y, z**, ...) separated by forward slashes, preceded by 'a group of words' (α).	6	0.05
		Unacceptable; α is or is not present		
III.6	special /	A special case with **/**.	1	0.01
		OK, yet include enough extra contextual information		
III.7	error	Forward slashes and commas are used without any logic.	41	0.36
		Correct the mistakes and handle as one of the main forward slash types (Types III.1–III.5)		

III.8	?	Impossible to decode.	5	0.04
		Simplify the paradigms (potentially following extra fieldwork) and handle as one of the main forward slash types (Types III.1–III.5)		
IV	special	Special cases ((**…**) *or* **:** *or* **bj.bj.** *or* = *or* **;** *or* " " etc.).	22	0.19
		Reverse manually, as each case is different		
SUM			**11,411**	**99.99**

3   See Prinsloo & Van Wyk (1992) for a detailed discussion of Northern Sotho kinship terminology.

4   The authors are grateful for Ms Salmina Nong's help with the decoding of various translation equivalents from the *Northern Sotho Terminology and Orthography No. 4* (T&O).

5   For a detailed discussion of the design of measurement instruments for alphabetical stretches, see Prinsloo & De Schryver, 2002a.

6   Heritage (2000[4]) is freely available over the Internet (see References for URL), and the number of lemma signs can be calculated with counts of the 'entry index'.

7   Just as Heritage (2000[4]), also BioTech (1998) is freely available over the Internet (see References for URL). The number of lemma signs can be calculated with wildcard searches. Note also that only those lemma signs beginning with letters were considered, ignoring digit-initial scientific terms.

8   Even though software can be written to easily process such instances, initial errors and/or orthography rules that have changed with time make the reversal unnecessarily complex. Example (12d), for instance, should be written as one word. An automatic reversal rendering (13d), (14d) and (15d) is thus not satisfactory.

9   The brackets in the Northern Sotho equivalent of **etc.** are in the wrong place in T&O, and the abbreviation for **bjalo-bjalo** 'etcetera' should be written without a space, thus as **bj.bj.**

References

(Where applicable, superscripts at publication dates indicate the edition number of the mentioned dictionary.)

Akhmanova, Olga. 1975. Review: Marcus Wheeler. 1972. The Oxford Russian – English Dictionary, *International Journal of Slavic Linguistics and Poetics*, 21: 127–31.

Al-Kasimi, Ali M. 1977. *Linguistics and Bilingual Dictionaries*. Leiden: E.J. Brill.

Baunebjerg Hansen, Gitte. 1990. *Artikelstruktur im zweisprachigen Wörterbuch. Überlegungen zur Darbietung von Übersetzungsäquivalenten im Wörterbuchartikel* (Lexicographica Series Maior 35). Tübingen: Max Niemeyer Verlag.

Benson, Morton. 1990. Culture-Specific Items in Bilingual Dictionaries of English, *Dictionaries: Journal of The Dictionary Society of North America*, 12: 43–54.

[BioTech 1998] *BioTech Life Science Dictionary.* Bloomington: Indiana Institute for Molecular and Cellular Biology. <http://biotech.icmb.utexas.edu/search/dict-search.html>

Botne, Robert & Kulemeka, Andrew T. 1995. *A Learner's Chichewa and English Dictionary* (Afrikawissenschaftliche Lehrbücher 9). Cologne: Rüdiger Köppe.

Dennis, Ronald D. 1985. Review: Bobby J. Chamberlain & Ronald M. Harmon. 1983. A Dictionary of Informal Brazilian Portuguese (With English Index), *The Modern Language Journal*, 69(3): 316–7.

Departmental Northern Sotho Language Board. 1988[4]. *Northern Sotho Terminology and Orthography No. 4 / Noord-Sotho terminologie en spelreëls No.4 / Sesotho sa Leboa mareo le mongwalo No.4.* Pretoria: Government Printer.

De Schryver, Gilles-Maurice & Kabuta, Ngo S. 1998[2].

Beknopt woordenboek Cilubà – Nederlands & Kalombodi-mfùndilu kàà Cilubà (Spellingsgids Cilubà), Een op gebruiksfrequentie gebaseerd vertalend aanleerderslexicon met decodeerfunctie bestaande uit circa 3.000 strikt alfabetisch geordende lemma's & Mfùndilu wa myakù ìdì ìtàmba kumwèneka (De orthografie van de meest gangbare woorden). Ghent: Recall.

De Schryver, Gilles-Maurice & Prinsloo, D.J. 2000. Electronic corpora as a basis for the compilation of African-language dictionaries, Part 1: The *macro-structure*, *South African Journal of African Languages*, 20(4): 291–309.

Duval, Alain. 1991. L'équivalence dans le dictionnaire bilingue. In Hausmann, Franz J. *et al.* eds. 1989–1991: 2817–24.

Gold, David L. 1982. Review: Carlos Castillo *et al.* 1972. The University of Chicago Spanish Dictionary, Revised Edition, A New Concise Spanish – English and English – Spanish Dictionary of Words and Phrases Basic to the Written and Spoken Languages of Today, *Dictionaries: Journal of The Dictionary Society of North America*, 4: 238–50.

Gold, David L. 1985. Review: John C. Traupman. 1966. The New College Latin and English Dictionary, *Dictionaries: Journal of The Dictionary Society of North America*, 7: 311–20.

Gouws, Rufus H. 1989. *Leksikografie.* Pretoria: Academica.

Gouws, Rufus H. 1996. Idioms and Collocations in Bilingual Dictionaries and Their Afrikaans Translation Equivalents, *Lexicographica, International Annual for Lexicography*, 12: 54–88.

Gouws, Rufus H. 2000. Toward the Formulation of a Metalexicographic Founded Model for National Lexicography Units in South Africa. In Wiegand, Herbert E. ed. 2000. *Wörterbücher in der Diskussion IV. Vorträge aus dem Heidelberger Lexikographischen Kolloquium* (Lexicographica Series Maior 100): 109–33. Tübingen: Max Niemeyer Verlag.

Hartmann, Reinhard R.K. ed. 1983. *Lexicography: Principles and Practice* (Applied Language Studies 5). London: Academic Press.

Hartmann, Reinhard R.K. & James, Gregory. 1998. *Dictionary of Lexicography.* London: Routledge.

Hausmann, Franz J., Reichmann, Oskar, Wiegand, Herbert E. & Zgusta, Ladislav. eds. 1989–91. *Wörterbücher / Dictionaries / Dictionnaires, Ein internationales Handbuch zur Lexikographie / An International Encyclopedia of Lexicography / Encyclopédie internationale de lexicographie* (Handbücher zur Sprach- und Kommunikationswissenschaft 5.1–3). Berlin: Walter de Gruyter.

[Heritage 2000[4]] *The American Heritage® Dictionary of the English Language, Fourth Edition.* Boston: Houghton Mifflin.

<http://www.bartleby.com/61/>

Honselaar, Wim & Elstrodt, Marijke. 1992. The electronic conversion of a dictionary: from Dutch – Russian to Russian – Dutch. In Tommola, Hannu *et al.* eds. 1992. *EURALEX '92 PROCEEDINGS I–II, Papers submitted to the 5th EURALEX International Congress on Lexicography in Tampere, Finland* (Studia Translatologica A/2): 229–37. Tampere: Department of Translation Studies, University of Tampere.

Kriel, Theunis J. 1976[4]. *The New English – Northern Sotho Dictionary, English – Northern Sotho, Northern Sotho – English.* Johannesburg: Educum Publishers.

Kromann, Hans-Peder, Riiber, Theis & Rosbach, Poul. 1991. Principles of bilingual lexicography. In Hausmann, Franz J. *et al.* eds. 1989–1991: 2711–28.

Landau, Sidney I. 1984. *Dictionaries: The Art and Craft of Lexicography.* New York: Charles Scribner's Sons.

Landau, Sidney I. 2001. *Dictionaries: The Art and Craft of Lexicography (2nd edition).* Cambridge: Cambridge University Press.

Lekganyane, D.N. 2001. Lexicographic Perspectives on the Use of Sepedi as a High Function Language. (Unpublished doctoral dissertation, University of Pretoria.)

Liberman, Anatoly. 1984. Review: Peter Terrel. 1980. Collins German – English, English – German Dictionary, *The German Quarterly*, 57(2): 280–7.

Martin, Willy & Tamm, Anne. 1996. OMBI: An editor for constructing reversible lexical databases. In Gellerstam, Martin *et al.* eds. 1996. *Euralex '96*

Proceedings I-II, Papers submitted to the Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden: 675–87. Gothenburg: Department of Swedish, Göteborg University.

Newmark, Leonard. 1999. Reversing a One-Way Bilingual Dictionary, *Dictionaries: Journal of The Dictionary Society of North America*, 20: 37–48.

Prinsloo, D.J. & De Schryver, Gilles-Maurice. 2002a. Designing a Measurement Instrument for the Relative Length of Alphabetical Stretches in Dictionaries, with special reference to Afrikaans and English. In Braasch, Anna & Povlsen, Claus. eds. 2002. *Proceedings of the Tenth EURALEX International Congress, EURALEX 2002, Copenhagen, Denmark, August 13-17, 2002*: 483–94. Copenhagen: Center for Sprogteknologi, Københavns Universitet.

Prinsloo, D.J. & De Schryver, Gilles-Maurice. 2002b. The use of slashes as a lexicographic device, with special reference to the African languages, *South African Journal of African Languages,* 22(1):70–91.

Prinsloo, D.J. & Gouws, Rufus H. 1996. Formulating a new dictionary convention for the lemmatization of verbs in Northern Sotho, *South African Journal of African Languages*, 16(3): 100–7.

Prinsloo, D.J. & Gouws, Rufus H. *forthcoming*. Formulating dictionary conventions for concordial paradigms in Northern Sotho.

Prinsloo, D.J. & Van Wyk, J.J. 1992. Verwantskapsterminologie van die Noord-Sotho, *South African Journal of Ethnology*, 15(2): 43–58.

Rettig, Wolfgang. 1985. Die zweisprachige Lexikographie Französisch – Deutsch, Deutsch – Französisch. Stand, Probleme, Aufgaben, *Lexicographica, International Annual for Lexicography*, 1: 83–124.

Rey, Alain. 1991. Divergences culturelles et dictionnaire bilingue. In Hausmann, Franz J. *et al.* eds. 1989–1991: 2865–70.

Rycroft, David K. 1981. Concise SiSwati Dictionary. SiSwati – English / English – SiSwati. Pretoria: J.L. van Schaik.

Schnorr, Veronika. 1986. Translational Equivalent and/ or Explanation? The Perennial Problem of Equivalence, *Lexicographica, International Annual for Lexicography*, 2: 53–60.

Sinclair, John M. ed. 2001[3]. *Collins Cobuild English Dictionary for Advanced Learners.* London: HarperCollins Publishers.

Steiner, Roger J. 1984. Guidelines for Reviewers of Bilingual Dictionaries, *Dictionaries: Journal of The Dictionary Society of North America*, 6: 166–81.

Svensén, Bo. 1993. *Practical Lexicography: Principles and Methods of Dictionary-Making* (Translated from the Swedish (1987) by John Sykes and Kerstin Schofield). Oxford: Oxford University Press.

Taljard, Elsabé & Gauton, Rachélle. 2001. Supplying Syntactic Information in a Quadrilingual Explanatory Dictionary of Chemistry (English, Afrikaans, isiZulu, Sepedi): A Preliminary Investigation, *Lexikos*, 11 (AFRILEX-reeks/series 11: 2001): 191–208.

Tomaszczyk, Jerzy. 1983. On Bilingual Dictionaries. The case for bilingual dictionaries for foreign language learners. In Hartmann, Reinhard R.K. ed. 1983: 41–51.

Tomaszczyk, Jerzy. 1988. The bilingual dictionary under review. In Snell-Hornby, Mary. ed. *ZüriLEX'86 Proceedings, Papers read at the EURALEX International Congress, University of Zürich, 9–14 September 1986*: 289–97. Tübingen: A. Francke Verlag.

Zgusta, Ladislav. 1971. *Manual of Lexicography*. The Hague: Mouton.

Zgusta, Ladislav. 1984. Translational equivalence in the bilingual dictionary. In Hartmann, Reinhard R.K. ed. 1984. *LEXeter '83 Proceedings, Papers from the International Conference on Lexicography at Exeter, 9–12 September 1983* (Lexicographica Series Maior 1): 147–54. Tübingen: Max Niemeyer Verlag.

Zgusta, Ladislav. 1987. Translational Equivalence in a Bilingual Dictionary. Bahukosyam, *Dictionaries: Journal of The Dictionary Society of North America*, 9: 1–47.