The Lexicographic Treatment of Days in Sepedi, or When Mother-Tongue Intuition Fails*

Gilles-Maurice de Schryver, *Research Assistant of the Fund for Scientific Research – Flanders (Belgium) and Department of African Languages, University of Pretoria (schryver@postino.up.ac.za), and*
B. Lepota, *Unit for Language Skills Development, University of Pretoria (lepota@postino.up.ac.za), Pretoria, Republic of South Africa*

Abstract: Since May 2001 the Sepedi National Lexicography Unit officially started the actual dictionary-writing of a pioneering Explanatory Sepedi Dictionary (PyaSsaL). The compilation is undertaken within the theoretical framework of Simultaneous Feedback, is fully corpus-based, and follows an onomasiological approach to the Sepedi lexicon. In this article the various compilation aspects are examined and illustrated by means of one onomasiological sub-field, namely the days of the week. It is shown how a balanced combination of mother-tongue intuition, data from existing (bilingual) dictionaries, fieldwork results, corpus queries, and grammarians' conjectures can and should lead to a sound lexicographic treatment.

Keywords: SEPEDI NATIONAL LEXICOGRAPHY UNIT (NLU), SIMULTANEOUS FEEDBACK (SF), PRETORIA SEPEDI CORPUS (PSC), ONOMASIOLOGICAL APPROACH, DAYS OF THE WEEK, MOTHER-TONGUE INTUITION, FIELDWORK, QUESTIONNAIRES, INTER VIEWS, GRAMMARIANS' CONJECTURES, NOUN CLASSES, REPETITIVE INSERTED TEXTS

Kakaretšo: Tshekatsheko ya matšatši polelong ya Sesotho sa Leboa go ya ka dinyakwa tša go ngwala Pukuntšu, goba ge dikakanyo tša mmoledi wa polelo ye di sa atlege. Go thoma ka Mei 2001 Yuniti ya Bosetšhaba ya Pukuntšu ya Sesotho sa Leboa e thomile semmušo go ngwala Pukuntšutlhaloši ya Sesotho sa Leboa (PyaSsaL). Thulaganyo ya yona e laolwa ke teori ya *Simultaneous Feedback*, e theilwe godimo ga sešegontšu (khophase), gape e latela tsela ya go hlaloša mantšu a Sesotho sa Leboa go ya ka direrwa tše di fapanego. Mo taodišwaneng ye go tsinkelwa mekgwa ya go fapanafapana ya thulaganyo gape mekgwa yeo e šupetšwa ka karolwana e tee ya direrwa tša PyaSsaL, e lego matšatši a beke. Go bontšhwa ka fao tekanelo ya kopanyo ya dikakanyo tša mmoledi wa Sesotho sa Leboa, tshedimošo go tšwa go di-pukuntšu (tša malemepedi) tšeo di šetšego di le gona, dipoelo tša nyakišišontle, diphatišišo ka gare ga sešegontšu, le dikakanyo tša borapopapolelo di ka kgonago ebile di swanetše go fihliša tshekatshekong ya kgonthe go ya ka dinyakwa tša go ngwala pukuntšu.

Mantšu a bohlokwa: YUNITI YA BOSETŠHABA YA PUKUNTŠU (YBP) YA SESOTHO SA LEBOA, SIMULTANEOUS FEEDBACK (SF), KHOPHASE YA SESOTHO SA LEBOA YA

* An earlier version of this article was presented at the Sixth International Conference of the African Association for Lexicography, organised by the Sepedi National Lexicography Unit, University of the North, Pietersburg, 2-4 July 2001.

TSHWANE (KST), DIRERWA, MATŠATŠI A BEKE, KAKANYO YA MMOLEDI WA SESOTHO SA LEBOA, NYAKIŠIŠONTLE, MANANEOPOTŠIŠO, DIPOLEDIŠANO, DIKAKANYO TŠA BORAPOPAPOLELO, MAGORO A MAINA, KHII YA GO TSENYWA KA GO BOELETŠWA

1. South Africa's 'golden opportunity'

In a recent publication Gouws (2000: 114) refers to the process of establishing a new lexicographic dispensation in South Africa as a 'golden opportunity' and a unique occasion in international terms. This golden opportunity is materialising at this very moment, with the Pan South African Language Board (PANSALB) having established National Lexicography Units (NLUs) for each of the official South African languages in 2000, and the actual compilation of dictionaries already being undertaken by some of them. In May 2001 the Board of the Sepedi NLU appointed two full-time mother-tongue lexicographers who promptly started their activities. In addition, the Board appointed one Ph.D. student as part-time lexicographer, and accepted the offer from another one to act as facilitator. During the first three months, the Head Office joined the Branch Office at the University of Pretoria (UP), where the lexicographic team was supplemented by two part-time corpus builders from UP's Department of African Languages.[1]

2. The theoretical framework and computational support

An extensive discussion of the methods underlying the current compilation procedures and all the facets of the computational support will be described elsewhere. In short, however, we can point out that the compilation is undertaken within the theoretical framework of Fuzzy Simultaneous Feedback (cf. e.g. De Schryver and Prinsloo 2001), which can be considered as the electronic continuation of the concept of Simultaneous Feedback (SF) (cf. e.g. De Schryver and Prinsloo 2000, 2000a). Briefly, SF can be understood as entailing a dictionary-making method in terms of which the release of several small-scale Parallel Dictionaries triggers off feedback that is instantly channelled back into the compilation process of a Main Dictionary.

It is well-known that "[t]he line function of a unit, as stated by PANSALB, should eventually be the compilation of a comprehensive monolingual explanatory dictionary" (Gouws 2000: 111), and the Sepedi NLU is giving heed to this. From the start, the facilitator trained the team members in the writing of explanatory definitions, sketched the main features of the dictionary to be compiled and the structure of the articles, and approached the querying of the electronic corpus hands-on. Within two weeks a first little test dictionary was produced and circulated among mother-tongue speakers. Feedback was retrieved, fed back into the project, and the compilation adapted accordingly. Within two months the First Parallel Dictionary was printed, and the cycle repeated. This procedure will be part and parcel of the entire compilation process.

As far as the computational support is concerned, we can briefly observe that the data are entered in the Onoma Lexical Workbench, a software package developed by Lexilogik in Sweden.[2] At all times the facilitator has been — and will be for some time into the future — in close contact with the software developers in order to adapt Onoma to the requirements of the Unit. SQL[3], the database underlying Onoma, has been stored on the facilitator's computer (the server side), whilst the computers of the three mother-tongue lexicographers (the client side) have been linked to the server through a network. Another crucial facet of the computational support is the electronic corpus. In this respect the Sepedi NLU is rather fortunate, as it can make free use of the Pretoria Sepedi Corpus (PSC), a corpus that was painstakingly assembled during the past decade by D.J. Prinsloo and G.-M. de Schryver. Currently, PSC stands at 5.8 million running words.[4] At the moment, PSC has not been integrated into Onoma. Rather, PSC is analysed with WordSmith Tools, a versatile corpus query software package developed by Mike Scott in the UK.[5]

3. An onomasiological approach to dictionary compilation

With PSC at hand (or better: 'on screen'), the compilation of the *Pukuntšu-tlhaloši ya Sesotho sa Leboa* (PyaSsaL) 'Explanatory Sepedi Dictionary', is fully corpus-based. For every compilation aspect — from the selection of the lemma signs up to the writing of the dictionary articles themselves — the corpus is queried. In De Schryver and Prinsloo (2000c) the different steps one needs to follow in order to compile a corpus-based macrostructure have been reviewed. One starts by extracting a word-frequency list from the corpus, this list is then transformed into a lemmatised frequency list, after which the latter is turned into a lemma-sign list. However, taking the lemma-sign list and working through the alphabet from A to Z, rarely results in a sound end product. Indeed, such an approach is more often than not marred by inconsistencies and poor definitions. Instead, the compilation of PyaSsaL follows an onomasiological approach to the Sepedi lexicon. Different semantic fields are chosen by each lexicographer, and each then tries to cover all the basic/frequent items from that field (where the selection is based on PSC). One lexicographer will then read all the definitions from a certain field to the others to see whether or not the others can pinpoint the correct lemma sign. If not, the definition must be adapted or rewritten. Such an approach has already proven to have many advantages, foremost among them the fact that the compilers are forced to differentiate every item from every other one and to make sure circular definitions are avoided.

The latter point can be illustrated with an example taken from the *Collins COBUILD English Dictionary* (COBUILD2, Sinclair 1995[2]) — one of four reference works many lexicographers consider to be among the best learners' dictionaries available for English. The first definition (and in most cases the only one users read) for *minute* in COBUILD2 is shown in (1).

(1) A **minute** is one of the sixty parts that an hour is divided into. People often say '**a minute**' or '**minutes**' when they mean a short length of time.

As can be seen, the definition of *minute* is based on *hour*. A user who does not know the meaning of *hour* will have to consult that item in order to understand the meaning of *minute*, only to find (2) as the first definition.

(2) An **hour** is a period of sixty minutes.

Here, the definition of *hour* is based on *minutes*. Such circular definitions are unacceptable. By forcing the mother-tongue compilers to work within fields, such circularity is combated. Even though the compilation follows an onomasiological approach, the editors will have the choice to present the data stored in the database in a semasiological way (meaning that the lemma signs are listed in an alphabetical sequence) or in an onomasiological way (thus as a dictionary with a thematic character). Yet the onomasiological-compilation approach also enables the lexicographers to transcend the paper dimension. As a matter of fact, every sense of each lemma sign of a particular onomasiological field is labelled with the same 'classifier'. These data are entered in a hidden slot in Onoma — 'hidden' in that it is not shown in the printed version and can be hidden in the electronic version of PyaSsaL. Ultimately every lemma sign will contain several classifiers, and the idea is to facilitate searches in the electronic version where users go from concept to word rather than from word to concept. A user will be enabled to input some classifiers, keywords, style labels, Boolean operators, etc. after which the software will run through the multi-indexed data to suggest the item(s) the user is seeking (cf. also Geeraerts 2000).

So far, several dozen of onomasiological fields, together with their classifiers, have been treated: COLOURS, DISEASES, TIME, CROPS, VEHICLES, KITCHENWARE, CLOTHES, etc. etc. The field TIME proved to be a particularly hard one, and was classified into numerous sub-fields. DAYS is just one of these many sub-fields, and, taken at face value, one could assume that it can be treated in just a few hours. However, this is not so, and this sub-field will be taken as an example to illustrate how the different compilation aspects can and must be combined in order to arrive at a sound lexicographic treatment.

4. When mother-tongue intuition fails

Just as in English, the days are nouns in Sepedi. This immediately implies that they belong to a certain gender, i.e. a fixed singular + plural class — crucial information as all the syntactical concords follow suit. Not all, but the great majority of nouns have both a singular and a plural form. See for instance (3) for some examples (with the gender indicated after the equal-signs).

(3) *moeta* 'vessel' / *meeta* 'vessels' = 3/4
letšhollo 'diarrhoea' = 5/0
lepai 'cotton blanket' / *mapai* 'cotton blankets' = 5/6
maloba 'the day before yesterday' = 0/6
sehla 'season' / *dihla* 'seasons' = 7/8
hlogo 'head' / *dihlogo* 'heads' = 9/10

Most mother-tongue speakers know the form (where applicable) of the singular and the plural of a particular noun. However, the fact that *not all* speakers have this intuitive knowledge, and certainly not for rare or borrowed words, forces the compilers to guide the future users of PyaSsaL. It was therefore decided to include gender information for every noun, with the full treatment by default at the singular (and the reverse only where the plural is more frequent). The plural contains, apart from the comment on form, no more than a cross-reference to the singular (or where the plural is more frequent, a cross-reference from the singular to the plural). The gender information is given with the generally accepted numbering system illustrated in (3), where the class number of the treated word is printed in boldface. In addition, a repetitive inserted text at the bottom of the page (in the paper version) or a pop-up window (in the electronic version) briefly summarises the meaning and the various concords. (For a more extensive argumentation of the followed procedure when lemmatising nouns, see Prinsloo and De Schryver (1999), and for more details on the repetitive inserted text, see De Schryver and Prinsloo (2000: 200-203).)

With this procedure, the treatment of the sub-field DAYS seems simple: (a) write down all the items belonging to the closed set DAYS, including the variants, (b) query the corpus to retrieve frequency information and example sentences, and (c) treat the items accordingly, with cross-references from the lesser-used to the more-frequent ones. A sample of the first attempt at this procedure (here for Thursday) — retrieved from an early draft — is shown in (4).

(4) **Labone** *leina* ***5/6, 5/10, 5/2a***
Ke letšatši la bone la beke, le thoma ka morago ga Laboraro gomme la fela ge go thoma Labohlano: *O nteleditše mogala ka Labone*
Malabone?? *leina* ***5/6*** BONA **Labone**
Dilabone *leina* ***5/10*** BONA **Labone**
Bolabone?? *leina* ***5/2a*** BONA **Labone**

What at this point worried the mother-tongue lexicographer who had entered this into the database, was that the plurals of most days (here *Malabone*, *Dilabone* and *boLabone*) had ended up in up to *three different classes* — a rather surprising result. After consulting with the other members of the team, the compilers realised that their intuition let them down, for they could not agree on the genders of the days. A reanalysis of PSC revealed that the corpus seemed to contain only a few occurrences of the plurals of any day. At this stage, the

lexicographers concluded that plurals of days do not seem to be frequently used in *written* language (as PSC does not contain any oral components so far). At the same time, however, all lexicographers agreed that plurals *must* exist, as they themselves *do* use them. The question only was: which ones are correct?

5. Days in existing (bilingual) Sepedi dictionaries

The next step was to consult all the latest versions of the existing (bilingual) Sepedi dictionaries. This revealed that out of the nine dictionaries, only one (Kriel et al. 1989[4]) consistently includes class information for every singular day, implicitly telling the user how the plural should be formed. For ease of reference, all the relevant articles have been transcribed verbatim and are listed in Appendix A, and a summary of the data is shown in (5).

(5) Days in existing (bilingual) Sepedi dictionaries

Days	**Mon.**				**Tue.**	**Wed.**	**Thu.**	**Fri.**		**Sat.**	**Sun.**		
Alterna-tives	Mošupologo	Mošupulogo	Mošopologo	Mantaga	Labobedi	Laboraro	Labone	Labohlano	Foreitaga	Mokibelo	Sontaga	Lamorena	Lamodimo
Freq.	*144*	*8*	*3*	*0*	*61*	*68*	*59*	*127*	*0*	*219*	*163*	*63*	*1*
Z & M	–	–	–	–	✓	✓	✓	✓	–	–	✓	✓	–
S–A/E	–	–	–	–	–	–	–	–	–	–	✓	–	–
New E	✓	–	–	✓	✓	✓	✓	✓	–	✓	–	–	✓
S–E	–	–	–	–	–	–	–	–	–	–	–	–	–
New E	✓	–	–	✓	✓	✓	✓	✓	–	✓	✓	–	✓
E–S	–	–	–	–	–	–	–	–	–	–	–	–	–
Klein	–	–	–	–	✓	✓	✓	✓	–	–	✓	✓	–
S–A/E	–	–	–	–	–	–	–	–	–	–	✓	–	–
Klein	✓	–	–	–	✓	✓	✓	✓	–	✓	✓	–	–
A/E–S	–	–	–	–	–	–	–	–	–	–	–	–	–
T & O	✓	–	–	–	✓	✓	✓	✓	–	✓	–	✓	–
E–A/S	–	–	–	–	–	–	–	–	–	–	–	–	–
V Wyk	–	✓	✓	–	✓	✓	✓	✓	–	✓	✓	✓	✓
S–A	–	✓	✓	–	✓	✓	✓	✓	–	✓	✓	✓	✓
V Wyk	–	–	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	–
A–S	–	–	–	–	–	–	–	–	–	–	–	–	–
Prinsloo	✓	–	–	–	✓	✓	✓	✓	–	✓	✓	✓	–
S–E/A	–	–	–	–	–	–	–	–	–	–	–	–	–
Prinsloo	✓	–	–	–	✓	✓	✓	✓	–	✓	✓	✓	–
E/A–S	–	–	–	–	–	–	–	–	–	–	–	–	–
Popular	✓	✓	–	–	✓	✓	✓	✓	–	✓	✓	✓	✓
S–E	–	–	–	–	–	–	–	–	–	–	–	–	–

Popular	✓	–	–	✓	✓	✓	✓	✓	–	✓	✓	–	✓
E–S	–	–	–	–	–	–	–	–	–	–	–	–	–
Gerber	✓	–	✓	–	✓	✓	✓	✓	–	✓	✓	✓	✓
A–S	–	–		–	–	–	–	–	–	–	–	–	–
LEGEND	**S:** Sepedi / **E:** English / **A:** Afrikaans **Z & M:** Ziervogel and Mokgokong 1975 / **New E:** Kriel 1976[4] / **Klein:** Ziervogel and Mokgokong 1988[4] / **T & O:** Departmental Northern Sotho Language Board 1988[4] / **V Wyk:** Kriel et al. 1989[4] / **Prinsloo:** Prinsloo et al. 1996, 1997 / **Popular:** Kriel et al. 1997[4] / **Gerber:** Gerber 2000												

From (5) one can see that all dictionaries taken together offer 13 alternatives for the seven days of the week. For every dictionary, the top line shows the inclusion (✓) or exclusion (–) of the singular, and the bottom line whether or not the plural form is suggested (through the indication of class or gender information). Since we are primarily concerned with the formation of the plurals, we will first look into this aspect. Ziervogel and Mokgokong suggest, both in their comprehensive (1975) and their pocket (1988[4]) dictionary, that the plural of *Sontaga* 'Sunday' is *Disontaga*. Van Wyk, in his revision of the third edition of Kriel's *Pukuntšu* (Kriel 1983[3]), added class information to all the singular days (Kriel et al. 1989[4]). If a (non-proficient?) user follows the cryptic rules in the front matter of Van Wyk's revision, this user might derive the plurals listed in (6).

(6) Class information and derived plurals in Van Wyk's *Pukuntšu* (Kriel et al. 1989[4])

Day	**Singular**	**Class information**	**Derived plural**
Monday	Mošupulogô	n. dev. class 3	Mešupulogô
	Mošopologô	n. dev. class 3	Mešopologô
Tuesday	Labobêdi	poss. conc. class 5 + n.	(Mabobêdi / Malabobêdi / ?)
Wednesday	Laboraro	poss. conc. class 5 + n.	(Maboraro / Malaboraro / ?)
Thursday	Labonê	poss. conc. class 5 + n.	(Mabonê / Malabonê / ?)
Friday	Labohlano	poss. conc. class 5 + n.	(Mabohlano / Malabohlano / ?)
Saturday	Mokibêlô	n. dev. class 1a	bôMokibêlô
Sunday	Sôntaga	n. class 1a	bôSôntaga
	Lamorêna	poss. conc. class 5 + n.	(Mamorêna / Malamorêna / ?)
	Lamodimo	poss. conc. class 5 + n.	(Mamodimo / Malamodimo / ?)

The lexicographers were rather doubtful at this point, and the next step, fieldwork, was the consequence. However, it is necessary further to look at the lexicographic treatment of the days of the week in the above-mentioned dictionaries.

When describing the onomasiological approach to dictionary compilation in par. 3, we pointed out two crucial aspects leading to better dictionary articles when compiling within this framework: (a) the combat against circularity, cf. extracts (1) and (2), and (b) the avoidance of inconsistencies. Even a cursory

glance at any of the sets of days shown in Appendix A confirms the need for the latter. No set is consistent throughout. On a first level, one notes the inconsistencies in punctuation and layout, the random inclusion or omission of parts of speech, and the haphazard presence or absence of loanword labels. On a second level, and more problematic, one observes the variation in orthography between the different sections of these dictionaries (e.g. *La Morêna* in one direction, *Lamorêna* in the other). Furthermore, numerous discrepancies can be found among the various alternatives given in the different sections (e.g. *Mošupulogo* and *Mošopologo* in one direction, yet *Mošopologo* and *Mantaga* in the other). Luckily, present-day computational support helps to avoid many of these problems. There is however a third level that is wholly unacceptable when dealing with a closed set such as the days of the week. The supposedly most complete dictionary available for Sepedi, the *Comprehensive Northern Sotho Dictionary* (Ziervogel and Mokgokong 1975), treats only five days of the week, totally neglecting the existence of Monday and Saturday. To make things even worse, these two days belong to the three most-frequently used days (cf. (5) for frequencies in the 5.8 million PSC). Even the little pocket derivation of this comprehensive dictionary does better, as it contains those two days in the direction Afrikaans/English to Sepedi (1988[4]). Finally, but this cannot be seen from Appendix A, some items (e.g. *Sôntaga* in Kriel et al. 1989[4]) are even in the wrong alphabetical place in the dictionary, making it truly hard for users to find any day at all.[6] Sometimes the user is confronted with a case of 'impossible to find' (or dead reference), when the item in question has not been included in the dictionary. The treatment of Sunday in the *Comprehensive Northern Sotho Dictionary* is shown in (7).

(7) **SÓN'TAGA, (se-)/di- (Sôntaga) (< Afr.), cf. LÁMODÍMO,** Sondag // Sunday

Yet, when trying to follow up the cross-reference, the user will find that *Lamodimo* is not lemmatised.[7] Again, proper alphabetical order and sound cross-references are aspects where basic software can and should assist the modern lexicographer.

A tenth dictionary that is available to the team is *Basic English–N.Sotho* (Hartshorne et al. 1984). This dictionary is essentially a monolingual English learners' dictionary, with an appendix containing translation equivalents in Sepedi. A numeric-alphabetic reference system links these two sections. As such, and in dictionary-typological terms, this reference work is a bilingualised (or semi-bilingual) dictionary (cf. e.g. Laufer and Melamed 1994). In the front matter to Hartshorne et al.'s dictionary (1984: introduction), one reads:

> Basic English has been compiled for *the pupil* who is studying English as a second or foreign language. [...] In the first place a comprehensive range of textbooks in English and English literary works, frequently used by pupils, was subjected to a *computer analysis* in order to identify the vocabulary being used at this level. [...] This initial list was then checked against existing authoritative

> international basic word lists, in order to make sure that all those words which are frequently used in English were included.

Nevertheless, although much effort seemingly went into the compilation of this dictionary, one will look in vain for days of the week. The closest one comes to them is in the article for the lemma sign *week**. This article is repeated verbatim in Appendix B, together with the corresponding data from that dictionary's appendix. One cannot but fail to see that four of the seven days (namely Monday, Friday, Saturday and Sunday) are used to 'define' and 'illustrate' *week**. This brings another section from the front matter (Hartshorne et al. 1984: introduction) to mind:

> The final stage was to include [...] certain words which were needed in this dictionary's definitions and illustrative sentences.

It is therefore surprising that the names of the week were omitted, especially since COBUILD2 assigns four black diamonds to each day of the week, meaning that the names of the days belong to the top 1 900 most-frequent words of the English language. The only other major English learners' dictionary including frequency information is the *Longman Dictionary of Contemporary English, Third Edition* (LDOCE3, Summers 1995[3]). Yet, although the top-3 000 items in both spoken and written language have been marked in LDOCE3, none of the days of the week was given a frequency marker. The reason might lie in the fact that, in LDOCE3, insufficiently word-like items, including "numbers, closed sets such as nationalities and currencies, non-standard forms, and variants" (Kilgarriff 1997: 142), were not given a frequency annotation. However, this procedure of not marking, for instance, closed sets, jeopardizes the basic aim whereby frequency data are entered because "[a] central fact about a word is how common it is. The information is particularly valuable for language learners, as it immediately indicates how important it is to learn a word" (Kilgarriff 1997: 135).

6. Fieldwork: questionnaires and interviews

So far we have seen that both the mother-tongue lexicographers and the existing (bilingual) dictionaries could not provide the necessary data for a sound treatment of (the plurals of) the days. It was therefore decided to carry out some fieldwork, mainly aimed at retrieving mother-tongue speakers' suggestions for the plurals of the days. To do so, numerous techniques exist, the one having more flaws than the other. From the various alternatives, the utilisation of an informal questionnaire was chosen. Precautions were taken on two complementary levels in order to be able to 'trust' the results. On the one hand, feedback was retrieved among mother-tongue speakers in the Northern Province, Mpumalanga and Gauteng, and this in two distinct ways. 80% of the questionnaires were printed and distributed, whereby the speakers were asked

to write down their suggestions on the spot. For the remaining 20%, the paper version of the questionnaire was simulated through face-to-face interviews, whereby the interviewers tried to *trigger* spontaneous use of the plurals of the days. These interviews were recorded unknowingly. Following each interview, the interviewees were informed and their input only kept (and later transcribed) if they agreed. On the other hand, two versions of the questionnaire/interview were set up. In version 1 the participants were asked to say something about themselves first, after which they were presented with some singulars for which they had to provide plurals. Then followed a few general questions about their background. In version 2 however, the personal questions were moved to the end, and one made-up day (**Lamathomo* for Monday) was included in the list, together with more (low-frequency) day alternatives. Also, the order of the days was slightly permutated. With these various procedures, the very integrity of the methodology was tested. The two paper versions of the questionnaires can be found in Appendixes C and D.

In total, exactly 100 opinions were collected. In the preliminary analysis, the written input was kept separate from the oral input, and within these two groups, the answers for versions 1 and 2 were differentiated. Yet, after a careful study of the four types of feedback, it was noticed that there were no statistically relevant differences between them. The four types will therefore be discussed together below. That there were no such differences was surprising. On the one hand, one could have expected that participants react differently when filling in a questionnaire compared to natural spoken use. On the other hand, it is generally accepted that the retrieval of personal questions at the start of a questionnaire/interview skews the results. Fortunately, not too many participants tried to force a plural onto **Lamathomo* — indicating that they were not answering a battery of questions 'on automatic'.

The analysis of the questionnaires/interviews could best be started with Saturday, since this is the day with the least variation. The results of the fieldwork are shown in (8)(a), together with the corresponding frequency in PSC for each suggestion.

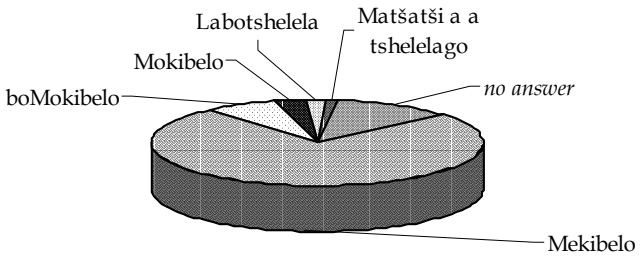
(8)(a) Fieldwork results: Saturday

Singular	*Freq.*	**Plural**	*Freq.*	**Option 1**	**Option 2**	**Option 3**
Mokibelo	*219*	Mekibelo	*10*	75	7	0
		boMokibelo	*2*	7	5	0
		Mokibelo	*219**	3	0	0
		Labotshelela	*0*	2	0	0
		Matšatši a a tshelelago	*0*	1	0	0
		Dimekibelo	*0*	0	1	2
		no answer		*12*	*87*	*98*

The participants could suggest as many possibilities as they wanted, yet no one offered more than three. Since 75% of the mother-tongue speakers opted for *Mekibelo* as the first and another 7% as the second possibility, it is clear from

the fieldwork that this noun belongs to gender 3/4, and therefore classes *mo-* for the singular and *me-* for the plural. Option 1 is shown graphically in (8)(b).

(8)(b) Graphic presentation: Saturday



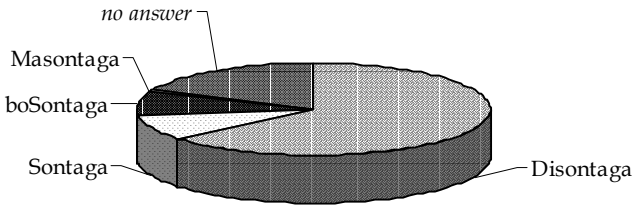
As regards Sunday, one can see from (5) that the most-frequently used form is a loanword (*Sontaga*), and not the term prescribed by the language board (*Lamorena*, see Appendix A). Likewise, as far as the suggested plurals are concerned, there is also more agreement on the former than on the latter, cf. (9)(a).

(9)(a) Fieldwork results: Sunday

Singular	*Freq.*	**Plural**	*Freq.*	**Option 1**	**Option 2**	**Option 3**
Sontaga	*163*	Disontaga	*12*	64	4	0
		Sontaga	*163**	9	0	0
		boSontaga	*1*	8	6	1
		Masontaga	*0*	1	3	1
		Sontageng	*1**	0	1	0
		boDisontaga	*0*	0	0	1
		no answer		*18*	*86*	*97*
Lamorena	*63*	Malamorena	*1*	24	3	0
		boLamorena	*0*	14	8	0
		Amorena	*0*	13	0	2
		Dilamorena	*0*	7	5	0
		Lamorena	*63**	6	0	0
		Malaborena	*0*	3	0	0
		A Lamorena	*0*	1	0	0
		Ga Morena	*206**	0	1	0
		Ba Morena	*71**	0	0	1
		Lamerena	*0*	0	0	1
		no answer		*32*	*83*	*96*
Lamodimo	*1*	Lamodimo	*1**	14	0	0
		Malamodimo	*0*	8	2	0
		Amodimo	*0*	4	0	0
		boLamodimo	*0*	2	2	0
		Alamodimo	*0*	2	0	0
		Melamodimo	*0*	2	0	0
		no answer		*68*	*96*	*100*

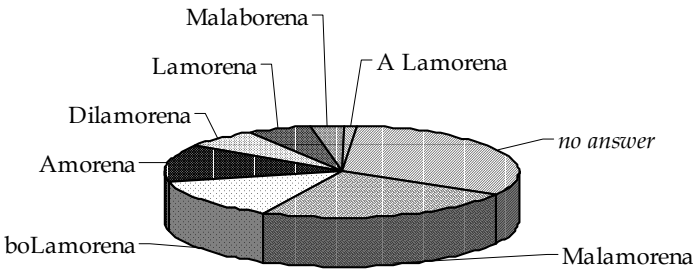
From the data in (9)(a), one must conclude that the fieldwork shows that *Sontaga* follows the pattern of most loanwords, i.e. it ends up in gender 9/10, taking *di*+ as the plural prefix. Option 1 for the plurals of *Sontaga* is shown graphically in (9)(b).

(9)(b) Graphic presentation: Sunday (i)



The other alternative for Sunday which should also be treated in a dictionary according to the fieldwork is *Lamorena*, which takes the prefix of class 6 *ma*+ for the formation of the plural, i.e. *Malamorena*. Option 1 for the plurals of *Lamorena* is shown graphically in (9)(c).

(9)(c) Graphic presentation: Sunday (ii)



Whilst Saturday is the most-frequently and Sunday the second most-frequently used day in Sepedi, Monday is the third. The summary of the fieldwork for Monday is shown in (10)(a).

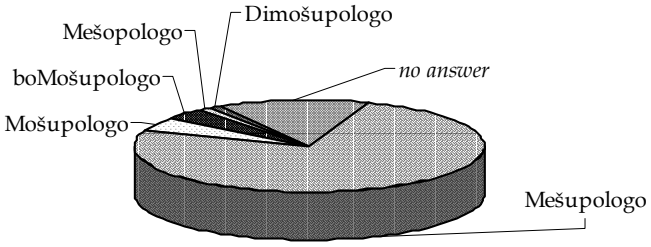
(10)(a) Fieldwork results: Monday

Singular	*Freq.*	**Plural**	*Freq.*	**Option 1**	**Option 2**	**Option 3**
Mošupologo	*144*	Mešupologo	*0*	75	3	0
		Mošupologo	*144**	5	0	0
		boMošupologo	*0*	4	6	0
		Mešopologo	*0*	1	2	0
		Dimošupologo	*0*	1	1	0

		boMešupologo	*0*	0	1	0
		Mešupologong	*0*	0	1	0
		boMedišupologo	*0*	0	0	1
		no answer		*14*	*86*	*99*
Mantaga	*0*	Dimantaga	*0*	25	6	1
		Memantaga	*0*	14	2	1
		Mebantaga	*0*	12	2	0
		boMantaga	*0*	6	12	1
		Mantaga	*0*	6	0	0
		Mantageng	*0*	1	0	0
		no answer		*36*	*78*	*97*
*Lamathomo	*0*	Lamathomo	*0*	14	0	0
		Amathomo	*0*	6	2	0
		Malamathomo	*0*	6	0	0
		boLamathomo	*0*	2	2	0
		Dilamathomo	*0*	2	2	0
		Memathomo	*0*	0	0	2
		no answer		*70*	*94*	*98*

As can be seen from (10)(a), it is reassuring to notice that up to 70% of the participants did not even try to suggest a plural for the made-up day **Lamathomo*, whilst another 14% simply repeated the singular. As this item was merely inserted to test the validity of the followed methodology, we should of course not discuss it in the context of the search for 'plurals of days'. As far as the methodology itself is concerned, we must conclude that it is valid given a fair margin of error. Furthermore, just as for the loanword *Sontaga* the fieldwork shows that the loanword *Mantaga* ends up in gender 9/10, yet with much less persuasiveness. Given the fact that this item has a zero frequency in PSC, we will not discuss it any further. Conversely, the mother-tongue speakers are rather unanimous when it comes to the plural of *Mošupologo*. Option 1 for the plurals of *Mošupologo* is shown graphically in (10)(b).

(10)(b) Graphic presentation: Monday



Just as for *Mokibelo*, the fieldwork shows that *Mošupologo* belongs to gender 3/4.

As far as the other days of the week are concerned, the analysis of the fieldwork shows a clear pattern, except for yet another loanword, *Foreitaga,* 'Friday'. Since PSC has again not a single occurrence of this alternative, we will therefore not discuss it below. The fieldwork data for the days Tuesday to Friday have been summarised in (11) to (14).

(11) Fieldwork results: Tuesday

Singular	*Freq.*	**Plural**	*Freq.*	**Option 1**	**Option 2**	**Option 3**
Labobedi	*61*	Malabobedi	*0*	31	6	1
		boLabobedi	*0*	14	9	0
		Dilabobedi	*0*	13	6	1
		Labobedi	*61**	6	0	0
		Labobedi a mabedi	*0*	1	0	0
		Labobeding	*0*	1	0	0
		Mabobedi	*0*	1	0	0
		Matšatši a mabedi	*96**	1	0	0
		Melabobedi	*0*	0	0	2
		Abobedi	*0*	0	0	1
		boMedilabobedi	*0*	0	0	1
		no answer		*32*	*79*	*94*

(12) Fieldwork results: Wednesday

Singular	*Freq.*	**Plural**	*Freq.*	**Option 1**	**Option 2**	**Option 3**
Laboraro	*68*	Malaboraro	*0*	29	6	1
		boLaboraro	*0*	16	10	0
		Dilaboraro	*0*	12	6	0
		Laboraro	*68**	8	0	0
		Laboraro a mararo	*0*	1	0	0
		Maboraro	*0*	1	0	0
		Matšatši a mararo	*126**	1	0	0
		Melaboraro	*0*	0	0	2
		Aboraro	*0*	0	0	1
		boMedilaboraro	*0*	0	0	1
		no answer		*32*	*78*	*95*

(13) Fieldwork results: Thursday

Singular	*Freq.*	**Plural**	*Freq.*	**Option 1**	**Option 2**	**Option 3**
Labone	*59*	Malabone	*0*	27	4	1
		boLabone	*0*	17	10	0
		Dilabone	*0*	14	5	1
		Labone	*59**	8	0	0
		Amane	*45**	1	0	0

Labone a mabedi	*0*	1	0	0
Labone ye mebedi	*0*	1	0	0
Melabone	*0*	0	0	2
Abone	*3**	0	0	1
boMedilabone	*0*	0	0	1
no answer		*31*	*81*	*94*

(14) Fieldwork results: Friday

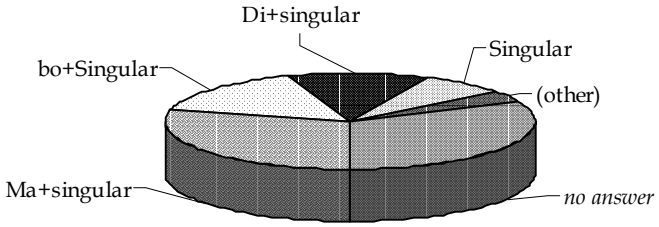
Singular	*Freq.*	**Plural**	*Freq.*	**Option 1**	**Option 2**	**Option 3**
Labohlano	*127*	Malabohlano	*0*	27	5	1
		boLabohlano	*0*	17	9	0
		Dilabohlano	*0*	11	4	2
		Labohlano	*127**	8	0	0
		Labohlano a mabedi	*0*	2	0	0
		Amahlano	*0*	1	0	0
		Labohlano ye mebedi	*0*	1	0	0
		Matšatši a mahlano	*13**	0	1	0
		Melabohlano	*0*	0	0	2
		Abohlano	*0*	0	0	1
		boMedilabohlano	*0*	0	0	1
		no answer		*33*	*81*	*93*
Foreitaga	*0*	Diforeitaga	*0*	40	4	2
		Foreitaga	*0*	8	0	0
		boFreitaga	*0*	4	4	0
		Maforeitaga	*0*	2	0	0
		no answer		*46*	*92*	*98*

Even though we are dealing with up to 10 or 11 possible plurals for one form of the singular — and hence, even though *single* mother-tongue intuition definitely fails —, the *average* mother-tongue intuition indicates that the preferred strategy is to prefix these days with *ma*+, followed by the strategies to prefix with *bo*+ and *di*+. Actually, the sequence and the percentages for these four days are so similar that one can safely make an average of the strategies for the days in the range Tuesday to Friday. These averages are shown in (15).

(15)(a) Fieldwork results: Tuesday to Friday

Plural strategy	**Tue.**	**Wed.**	**Thu.**	**Fri.**	**Average**
Ma+singular	31	29	27	27	28.5
bo+Singular	14	16	17	17	16
Di+singular	13	12	14	11	12.5
Singular	6	8	8	8	7.5
(other)	4	3	3	4	3.5
no answer	*32*	*32*	*31*	*33*	*32*

(15)(b) Graphic presentation: Tuesday to Friday



7. The value of the corpus

In par. 4 we reviewed the lexicographers' failure to pinpoint the plurals of the days of the week; in par. 5 we showed clearly that all existing (bilingual) dictionaries of Sepedi poorly treat the singulars of the days, and are even more vague (and partly wrong) when it comes to the plurals; and in par. 6 we saw how balanced fieldwork can bring a solution. Yet, there is a fourth and even a fifth level that must be taken into account. Level four consists of results from corpus queries, and level five of facts derived from the existing scientific literature. The former will be discussed in this paragraph, the latter in the next paragraph.

Calzolari (1996: 4) adequately summarises why data culled from corpora differ from the results obtained by means of other types of information retrieval:

> Carefully constructed, large written and spoken corpora are essential sources of linguistic knowledge if we hope to provide extensive and adequate descriptions of the concrete use of the language in real text. These types of descriptions certainly remain impossible if we only rely on introspection and native speaker intuition [...].

As noted in par. 2, PSC currently stands at 5.8 million running words. In building PSC, sections were sampled from several hundreds of written sources and a corpus of 5.8 million words roughly corresponds to the equivalent of 300 books. In building PSC, utmost care was taken to structure the corpus in such a way as to keep it stable. (For more information on how to build African-language corpora, see De Schryver and Prinsloo (2000b), and for an extensive discussion of the notion of corpus stability, see Prinsloo and De Schryver (forthcoming).)

Compared to intuition, informant elicitation and grammatical conjectures, the corpus has the big advantage in that it shows *real* language use. Any corpus query shows the 'attested and authentic usage *average*' of several hundreds of mother-tongue speakers. In (5) we have listed the PSC frequencies for the

various singular-day alternatives. For a corpus-based dictionary like PyaSsaL, this immediately implies that the loanwords *Mantaga* and *Foreitaga* will not be included in the dictionary, nor the low-frequency variant *Mošopologo* for Monday or the alternative *Lamodimo* for Sunday. The fieldwork satisfactorily supports this, as the mother-tongue speakers were very unsure when it came to the plurals of these items, or they simply disregarded them. Conversely, the situation with the loanword *Sontaga* is totally different. Indeed, as the second most-frequently used day of the week, it must not only be included, but must also be given a comprehensive lexicographic treatment.

Together with the fieldwork results presented in (8) through to (15), the PSC frequencies were also indicated. These frequencies are telling indeed, yet one must also carefully consider the context in order to see whether the mother-tongue suggestions are truly plurals of the days of the week. Wherever there is a severe problem with this, a star follows the frequency. As a matter of fact, the corpus shows that *Mokibelo*, *Sontaga*, *Lamorena*, *Lamodimo*, *Mošupologo*, *Labobedi*, *Laboraro*, *Labone* and *Labohlano* are all always referring to a singular. Furthermore, other suggestions have nothing to do with a plural, such as *Ga Morena* 'at the place of God', *Ba Morena* '(people) of God', *Amane* 'involves' and *Abone* 'when he/she saw'. And finally, PSC shows that still other suggestions for plurals are simply counts of days, such as *Matšatši a mabedi* 'two days', *Matšatši a mararo* 'three days' and *Matšatši a mahlano* 'five days'.

In a multicultural and multilingual environment like South Africa, it is crucial to make sure whether or not words from the (former) dominant languages are not perhaps used instead of indigenous ones. In the entire PSC, in a search for all possible singulars and plurals in English and Afrikaans, only two English days occur. These are shown in (16).

(16) *Ke* Friday today, *ge re fetša fela mo o name o tsebe. Go a iwa.*
'It is *Friday today*, once we finish here, you must know, we go.'
Karabo yona ya re, "Ka Sontaga — next Sunday, today a week — at 7 a.m."
'The answer was: "On Sunday — next Sunday, today a week — at 7 a.m."'

We can therefore safely assume that, at least in the written language, the Sepedi forms are used. One would have to see whether or not an oral corpus component would alter this finding, especially since informal observation indicates that code switches are much more frequent in spoken than in written language. Frequency markers derived from the corpus should therefore include a label differentiating between written and spoken frequencies — as is done in LDOCE3.

A corpus can also be queried with the use of wildcards, and this reveals one single instance of a plural that was *not* discovered through the fieldwork. This is shown in (17).

(17) *Gona fao ka matšatši a boSatertaga, le boSontaga go fetoga Sodoma ye nnyane, ka ge botagwe le tšohle tše mpe di direga.*

'There, days such as Saturdays and Sundays are turned into a small Sodom because drunkenness and other bad things occur.'

Single occurrences are however not enough to base conclusions on.

Finally, the loanword *Sabatha* is used in the Bible and just one other book. Here the problem is that this item has no distribution across a variety of different sources to warrant inclusion in PyaSsaL. Compare in this regard Knowles (1983: 188) who claims that "a word must occur evenly in a large number of the stratified sub-samples rather than excessively often in a small number of them, given that these two very different cases could show identical 'total-corpus' frequencies".

8. The grammarians' point of view compared to the results thus far

Unfortunately, as far as the days of the week are concerned, a systematic trawl through existing textbooks, journals and monographs did not reveal anything substantial. To make up for this, we consulted with two Sepedi grammarians, Dr E. Taljard and Prof. L.J. Louwrens. Both scholars were presented with the results of the study presented above, and their comments will now be interwoven with an overview of those results.

Firstly, Louwrens (*personal communication*, 15 June 2001) argues that the accepted way to form plurals in the field of TIME is by affixing *bo*+ at the beginning of a word that already has a class prefix. To illustrate this point, he puts forward the examples listed in (18) as 'acceptable'.

(18)	*Kresemose*	>	***bo****Kresemose*
	'Christmas day'		'Christmas days'
	Go bunwa mabele ka Julae.	>	*Go bunwa mabele ka* ***bo****Julae.*
	'Maize is harvested in July.'		'Maize is harvested around July.'
	Ba goroga ka iri ya ***le****some.*	>	*Ba goroga ka iri ya* ***bo****lesome.*
	'They arrive at the hour of ten.'		'They arrive at around the hour of ten.'
	Mo*šupologo*	>	***bo****Mošupologo*
	'Monday'		'Mondays'
	Ke letšatši la Labohlano ...	>	*Ke matšatši a* ***bo****Labohlano ...*
	'It is (the day of) Friday ...'		'It is (the days of) Fridays ...'
	Ke letšatši la ***Mo****kibelo ...*	>	*Ke matšatši a* ***bo****Mokibelo ...*
	'It is (the day of) Saturday ...'		'It is (the days of) Saturdays ...'

This would imply that every single day of the week can take the prefix *bo*+ to form some kind of plural. The fieldwork supports this.

Secondly, Louwrens emphasises that the plural prefix *di*+ is productively used for the formation of plurals of loanwords, and gives the examples shown in (19).

(19)	*tafola* 'table'	>	***di****tafola* 'tables'
	khomphuta 'computer'	>	***di****khomphuta* 'computers'

This would substantiate the finding that the plural of the loanword *Sontaga* is *Disontaga*.

Thirdly, Louwrens has his doubts as to the formation of plurals by prefixing *ma*+ to the singulars of days that are non-loanwords.

Taljard (*email*, 18 June 2001) agrees that *bo*+ can be prefixed to all days of the week, yet she remarks:

> There is a distinct semantic difference between boSontaga (Bosontaga?) and Disontaga, but this will only become clear when these forms are tested within context. The prefix *bo*+ is often used to indicate associative plurality. Usually, ordinary plurals express distributive plurality, thus *monna* 'one man', but *banna* 'many men'. When one uses *bo*+ it can indicate the same kind of plurality, but it can also indicate associative plurality. Thus *botate* can mean either 'fathers' in the sense of one, two, three, four fathers (e.g. These kids all have different fathers), but it can also mean 'father and company' i.e. other people who are associated with him on the basis of certain shared semantic features (e.g. father and his brothers, father and the other members of the *kgoro*, etc.). The prefix *bo*+ very often expresses associative plurality when used together with adverbs, which is a function frequently fulfilled by the days of the week. [...] (I have used Sontaga as an example, but it will also be valid for the other days of the week.)

Further, Taljard also agrees with Louwrens when it comes to the formation of the plurals of loanwords, i.e. through the prefixing of *di*+.

Finally, from a strictly grammatical point of view, she suggests the analogy shown in (20).

(20)

LEtšatši **LA** bobedi	>	**MA**tšatši **A** bobedi
'The second day'		'The second days'

Therefore, a "logical" conclusion could be as follows:

Labobedi	>	**A** bobedi	>	Abobedi
'Tuesday'				'Tuesdays'
Laboraro	>	**A** boraro	>	Aboraro
'Wednesday'				'Wednesdays'
Labone	>	**A** bone	>	Abone
'Thursday'				'Thursdays'
Labohlano	>	**A** bohlano	>	Abohlano
'Friday'				'Fridays'

The fieldwork brought up each of these options, yet each only once and in each case only as the third option. In addition, in PSC only *Abone* occurs, yet with the meaning 'when he/she saw', which should have been spelt as two words, namely *A bone*. Both the fieldwork and PSC therefore do not support Taljard's grammatical speculation. What is clear from (20) however, is the way in which the days Tuesday to Friday were formed — as the second, third, fourth and

fifth day, where the word *letšatši* 'day' was dropped and the remainder written conjunctively. In this context, the invented **Lamathomo* (**< (Letšatši) la mathomo* 'The first (day)') for Monday (cf. par. 6) was not so eccentric. Furthermore, it is also clear from (20) that the singulars of the days Tuesday to Friday belong to class 5. This is confirmed in PSC, as all concords for these days are class 5 concords. This same method of looking at concords can also be used to pinpoint the classes of the plurals. For instance, the fieldwork suggests that the plural of *Lamorena* 'Sunday' (< *(Letšatši) la morena* 'the Lord's (day)') is *Malamorena*, thus gender 5/6. The concords in PSC confirm this too, as can be seen in (21).

(21) *Ke lemoga lebaka leo bjale ka gore e šetše e le **Ma**lamorena **a ma**raro a go hlomagana, o sa thiše kerekeng.*
'I am aware of that, now that it has been three consecutive Sundays that you have not absented yourself from church.'

The fieldwork and the corpus therefore indicate that Tuesday to Friday (*Labobedi*, *Laboraro*, *Labone* and *Labohlano*), and Sunday (ii) (*Lamorena*) belong to gender 5/6. The fieldwork and the corpus are also unanimous when it comes to assigning gender 3/4 to Monday (*Mošupologo*) and Saturday (*Mokibelo*), and gender 9/10 to Sunday (i) (*Sontaga*). All this information will have to be indicated in PyaSsaL.

The only type of plural that needs further study at this point is *bo*+Singular. Van Wyk, in his revision of Kriel's *Pukuntšu*, suggests as plurals for *Mokibelo* and *Sontaga*, *boMokibelo* and *boSontaga* respectively. This possibility is confirmed in PSC, with both associative and distributive meanings.[8] Furthermore, the fieldwork indicates that all days of the week can take *bo*+ as prefix, and also Louwrens and Taljard suggest this possibility. The question arises however, whether the possibility of the plural *bo*+Singular should be indicated in PyaSsaL. In an enlightening article, Van Wyk (1987: 34) claims that "the morpheme *bo*- [...] can be used as a pluralizer and a nominalizer with an almost unlimited range of nouns, other parts of speech, phrases, and sentences". With this knowledge, it is obviously not a good idea to tell the dictionary user at every day of the week (and at almost every noun, for that matter) that one can add *bo*+ to the singular to form some kind of plural. As far as the days of the week are concerned, Van Wyk (1987: 37) himself gives the example shown in (22).

(22) *ba bantši ba hwile ka bo-labone*
'many died on or around Thursday'

It is now appropriate to bring all the data together, i.e. (a) the information one can find in existing (bilingual) dictionaries, (b) the results from the fieldwork, (c) the occurrences in the corpus, and (d) the grammarians' input. As explained in the previous paragraph, cross-comparing corpus frequencies for the singulars of days with the data from the fieldwork, compels us to keep only one pos-

sibility per day, except for the variant for Monday and the alternative for Sunday. Also, we will only focus on the truly frequent suggestions.

(23) Cross-comparing the various data sources for the plurals of the days of the week

Days	Singulars	PSC 5.8M	Plurals	Z&M 1975	V Wyk 1989	Field. 2001	PSC 5.8M	Louwr. 2001	Taljard 2001
Mon.	Mošupologo	✓	Me-	–	–	✓	–	✓	✓
			bo+	–	–	✓	–	✓	✓
	Mošupulogo	✓	Me-	–	✓	–	–	✓	✓
			bo+	–	–	–	–	✓	✓
Tue.	Labobedi	✓	Ma+	–	(✓)	✓	–	?	–
			bo+	–	–	✓	–	✓	✓
Wed.	Laboraro	✓	Ma+	–	(✓)	✓	–	?	–
			bo+	–	–	✓	–	✓	✓
Thu.	Labone	✓	Ma+	–	(✓)	✓	–	?	–
			bo+	–	–	✓	–	✓	✓
Fri.	Labohlano	✓	Ma+	–	(✓)	✓	–	?	–
			bo+	–	–	✓	–	✓	✓
Sat.	Mokibelo	✓	Me-	–	–	✓	✓	✓	✓
			bo+	–	✓	✓	✓	✓	✓
Sun.	Sontaga	✓	Di+	✓	–	✓	✓	✓	✓
			bo+	–	✓	✓	✓	✓	✓
	Lamorena	✓	Ma+	–	(✓)	✓	✓	?	–
			bo+	–	–	✓	–	✓	✓

With all this information it is now possible to treat the closed set 'days of the week' in such a way that the average of all the approaches is reflected. This will be done in the next section.

9. Towards a sound lexicographic treatment of days in Sepedi

A full treatment of all the days listed in (23) can be found in Appendix E. Firstly, compared to major learners' dictionaries for English such as COBUILD2 and LDOCE3, one will notice that circular definitions within this closed set were combated. Secondly, as PyaSsaL is fully corpus-based, all example sentences were culled directly from PSC, and this with minimal editing. Where attested in PSC, both singular and plural example sentences were included. Thirdly, moving to the macrostructure, one notices that the frequent items have been marked much as is done in LDOCE3, with **N** (*mo polelong ye e ngwadilwego*) '**W** (in the written language)' introducing frequencies in written Sepedi. A ① indicates that the lemma sign belongs to the top 1 000 items, a ② that it belongs to the top 2 000, and a ③ that it belongs to the top 3 000. As an illustration, the data for Saturday and Sunday are shown in (24) and (25) respectively, together with an approximate English translation.

(24) Towards a sound lexicographic treatment of Saturday

N② **Mokibelo** [Mokibêlô] *leina* ***3/4***
Ke letšatši la boselela la beke, le thoma ka morago ga Labohlano gomme la fela ge go thoma Sontaga: *Ka Mokibelo re tsogela pitšong; O e apare fela ka Mekibelo le Disontaga, le gona ge go na le phathi; Ka boMokibelo ba inwela malwa le bašomi ba bangwe ka fao lokheišeneng*

Mekibelo [Mekibêlô] *leina 3/****4***
BONA **Mokibelo**

W② **Saturday** /ˈsætədeɪ || -ər-/ *noun, singular*
The sixth day of the week, starting after Friday and ending when Sunday begins: *On Saturday we go to the meeting; You must wear it only on Saturdays and Sundays, and only when there is a party; On Saturdays they drink beers together with other workers in the township*

Saturdays /ˈsætədeɪz || -ər-/ *noun, plural*
SEE **Saturday**

(25) Towards a sound lexicographic treatment of Sunday

N ③ **Sontaga** [Sôntaga] *leina* ***9/10***
Ke letšatši la bošupa la beke, le thoma ka morago ga Mokibelo gomme la fela ge go thoma Mošupologo {LEHLALOŠETŠAGOTEE **Lamorena**}: *O be a fela a etla ga gešo gagolo ka Sontaga ka morago ga kereke; Baruti ba theogela ka Disontaga*

Disontaga [Disôntaga] *leina 9/****10***
BONA **Sontaga**

Lamorena [Lamorêna] *leina* **5**/6
Ke letšatši la bošupa la beke, le thoma ka morago ga Mokibelo gomme la fela ge go thoma Mošupologo {LEHLALOŠETŠAGOTEE **Sontaga**}: *Ge e le Lamorena ga re reke felo; Ke lemoga lebaka leo bjale ka gore e šetše e le Malamorena a mararo a go hlomagana, o sa thiše kerekeng*

Malamorena [Malamorêna] *leina 5/****6***
BONA **Lamorena**

W ③ **Sunday** /ˈsʌndeɪ/ *noun, singular*
The seventh day of the week, starting after Saturday and ending when Monday begins {SYNONYM **the Lord's Day**}: *He used to come to my place mostly on Sunday after church; Priests work on Sundays*

Sundays /ˈsʌndeɪz/ *noun, plural*
SEE **Sunday**

the Lord's Day /ˈðə lɔːds || lɔːrds deɪ/ *noun, singular*
The seventh day of the week, starting after Saturday and ending when Monday begins {SYNONYM **Sunday**}: *On the Lord's Day we have nowhere to buy; I am aware of that, now that it has been three consecutive Lord's Days that you have not absented yourself from church*

the Lord's Days /ˈðə lɔːds || lɔːrds deɪz/ *noun, plural*
SEE **the Lord's Day**

In the paper version of PyaSsaL, the bottom of every left-hand side or even page displays the even repetitive inserted text (ERIT) shown in (26), whilst the bottom of every right-hand side or odd page displays the odd repetitive inserted text (ORIT) shown in (27).[9]

(26) Even repetitive inserted text (ERIT) in the paper version of PyaSsaL

Lešala	**1** 🧍	**1** 🧍🧍	**2** 🧍	**2** 🧍🧍	**1**	**2**	**1a**	**2a**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**14**	**15**	**16**	**17**	**18**
Hlogo	nna	rena	wena	lena	mo-	ba-	–	bo-	mo-	me-	le-	ma-	se-	di-	N-	diN-	bo-	go-	fa-	go-	mo-
🧍(🧍)	ke	re	o	le	o	ba	o	ba	o	e	le	a	se	di	e	di	bo	go	go	go	go
🚲(🚲)	N-	re	go	le	mo	ba	mo	ba	o	e	le	a	se	di	e	di	bo	go	go	go	go

(27) Odd repetitive inserted text (ORIT) in the paper version of PyaSsaL

1 🧍	Botee bja mmoledi	1 🧍🧍	Bontši bja mmoledi	①, ② le ③	mantšu a 1 000, 2 000 le 3 000 a mathomo
2 🧍	Botee bja mmoledišwa	2 🧍🧍	Bontši bja mmoledišwa	**N**	mo polelong ye e ngwadi-lwego
🧍(🧍)	Lekgokasediri	🚲(🚲)	Lekgokasedirwa	**B**	mo polelong ye e bolelwago
Dinomoro tše di tšwelelago ka morago ga 'leina' di bontšha gore leina leo le hlaloswago ke la legoro lefe.					

The gender information that accompanies every noun is linked with ERIT and ORIT. With ERIT and ORIT, dictionary users are — at a single glance — not only provided with data on how the plural can be derived from the singular (or *vice versa*), but also on how to go about concordial relationships and concordial references. It should further be obvious, although the data have been grouped in (24), (25) and Appendix E, that singulars and plurals are scattered in a semasiological dictionary. The data for Sunday(s), for instance, will be found under the alphabetical categories **D**, **L**, **M** and **S**. In the electronic version of PyaSsaL, ERIT and ORIT are replaced with grammatical pop-up windows, whilst the actual 'placement' of the data is of less concern to the electronic-dictionary users.

10. Retrieving feedback on the followed lemmatisation approach

In line with the theoretical framework of Simultaneous Feedback (SF, cf. par. 2), feedback was retrieved on the lemmatisation approach illustrated in Appendix E, and (26) and (27). PyaSsaL 1.0, i.e. *PyaSsaL's First Parallel Dictionary* (De Schryver 2001), contained an earlier version of the data grouped in Appendix E. During a special session at the Sixth International Conference of the African Association for Lexicography, PyaSsaL 1.0 was presented and distributed among some 50 conference attendees. A questionnaire had also been prepared in Sepedi and English, and the aim was to retrieve a first impression of PyaSsaL 1.0 from the mixed audience of both seasoned lexicographers and

lexicographers-to-be, and from mother-tongue speakers, second- and third-language speakers, as well as from those for whom Sepedi is unknown. Of the 15 questions, only one (question 6) explicitly dealt with plurals of the days. This question is shown in (28).

(28) *Go ya ka pukuntšu ye **Labobedi** ka bontši ke eng mo polelong ya Sesotho sa Leboa?*
'According to this dictionary, what is the plural of **Labobedi** (Tuesday) in Sepedi?'

It can be expected that the brief presentation of PyaSsaL 1.0, together with the information (both in Sepedi and English) in the front matter of this dictionary, would be sufficient a basis for arriving at the correct answer. The analysis of the answers to (28) is shown in (29).

(29) Retrieval of feedback on the followed lemmatisation approach (here for *Tuesdays*)

Suggested plurals of Labobedi	**Mother-tongue speakers**	**2nd/3rd-language speakers**	**Foreign-language speakers**
Malabobedi	29 %	50 %	28 %
Bolabobedi	–	–	6 %
Di-	–	10 %	–
Labobedi	–	10 %	–
Mabobedi	–	–	22 %
Amabedi	–	–	6 %
Kalabobedi	–	–	6 %
wrong explanation	57 %	20 %	13 %
no answer	14 %	10 %	19 %
	100 %	100 %	100 %

The analysis presented in (29) is in line with previous rounds of feedback retrieval in South Africa (cf. De Schryver and Prinsloo 2000: 205-208). Firstly, it is not surprising that foreign-language speakers struggle with the conventions in a monolingual Sepedi dictionary. Only 28% of them pinpointed the correct plural using the gender information together with ERIT and ORIT. Secondly, up to half of the second- and third-language speakers were able to decode the conventions. Here one must bear in mind that many attendees in this group are actually lecturers of Sepedi (at university level), and are thus quite familiar with the use of dictionaries. Lastly, the mother-tongue speakers performed the worst of all. From par. 4 we know however that mother-tongue intuition fails in this context. The inability to answer correctly, especially considering the fact that the respondents had only just received PyaSsaL 1.0, can be attributed to a presumed lack of dictionary culture. From a metalexicographic perspective, this has been pointed out e.g. by Gouws (1999: 7, 11), while Atkins (1998: 3) has observed: "The speakers of African languages have not in their formative years had

access to dictionaries of the richness and complexity of those currently available for European languages. They have not had the chance to internalize the structure and objectives of a good dictionary, monolingual, bilingual or trilingual."

In this context it is interesting to compare the answers to question 9 in the same questionnaire. This question is shown in (30).

(30) *Ke ka lebaka lang ge dinomoro tše dingwe di swiswaditšwe mola go tše dingwe go se bjalo, mohlala* ***dinku*** *leina 9/**10*** ?
'Why are some numbers in bold and others not, for example **dinku** *leina 9/**10*** ?'

Here only 43% of the mother-tongue speakers answered correctly, while up to 60% of the second- and third-language speakers, and 56% of the foreign-language speakers knew the answer. The presumed lack of dictionary culture among mother-tongue speakers of Sepedi might again explain this finding.

Imbedded in SF is the fact that potential users continuously guide the compilers during the entire compilation process. The unabated retrieval of feedback can therefore be considered as the main strength of the methodology. According to Atkins and Varantola (1997: 1), "[t]here are two direct routes to more efficient dictionary use: the first is to radically improve the dictionary: the second is to radically improve the users". If one looks at the percentages in (29), one sees that if one would want to make PyaSsaL more accessible to foreign-language speakers, one should envisage a more explicit guidance when it comes to the formation of the plurals. Instead of just *leina 5/6* under **Labobedi**, one could — in order to improve the dictionary — for instance consider *leina 5/6 (Ø/Ma+)*, hereby explicitly telling the user that the plural is formed by *adding* Ma+ to the singular (hence **Malabobedi**, and not **Mabobedi**). Nonetheless, PyaSsaL is a dictionary aimed at mother-tongue speakers, and for them the percentages in (29) clearly indicate that one should walk the second route suggested by Atkins and Varantola, i.e. 'radically improve the users'. Therefore, while the Second Parallel Dictionary is being compiled, important efforts are directed towards the explicit and systematic teaching of dictionary skills in the classroom, as e.g. suggested by Chi (1998: 566). Formulated differently, while the Main Dictionary is being compiled, the future users are simultaneously being trained in using it, as early feedback (here only exemplified for the plural of one day of the week) indicated that there is a great need for this.

11. In conclusion

In this article we have examined the various compilation aspects to be confronted by the lexicographers active within the Sepedi NLU. We have seen that a sound treatment of the lexicon might require more than introspection and the tools present in a standard dictionary-compilation office, i.e. existing (bilingual) dictionaries, large electronic corpora, and grammar books. Indeed, for

many semantic (sub-)fields, additional fieldwork might well turn out to provide the most reliable data. It is the interplay between all these different types of data that ultimately enables the mother-tongue lexicographers to compile sound (i.e. reliable and truly representative) dictionary articles.

We have also shown clearly that an onomasiological approach to the lexicon combats circularity between dictionary definitions, and enables more consistency. Working within the framework of Simultaneous Feedback has the added advantage that the work can be brought out to the future target users, and amended if need be while the Main Dictionary is still under compilation.

Finally, this article has further indicated how practical lexicographers can be brought to engage in fundamental scientific research, if they are truly willing to provide the future users with the state of the art of their own language. As far as the days of the week are concerned, the original research reported on here has pinpointed the genders of these days for the very first time.

Endnotes

1. The two full-time lexicographers are M.P. Mogodi and M.C. Mphahlele, the part-time lexicographer is B. Lepota, the two corpus builders are S. Nong and B.P. Sathekge, and the facilitator is G.-M. de Schryver.
2. For more information on Onoma, the home page of *Lexilogik* can be consulted at: http://www.lexilogik.se
3. Structured Query Language.
4. Actually, corpora for all South African languages have been built at UP's Department of African Languages. The sizes of these are in constant evolution. For the latest developments, the home page of *ELC for ALL* (Electronic Corpora for African Language and Linguistics) can be visited at: http://www.up.ac.za/academic/libarts/ afrilang/elcforall.htm
5. For more information on *WordSmith Tools*, the home page of Mike Scott can be consulted at: http://www.lexically.net (or its mirror site: http://www.liv.ac.uk/~ms2928).
6. A thorough study has indicated that the *Pukuntšu* as revised by Van Wyk (Kriel et al. 1989[4]) contains an average of two errors in alphabetical order per page.
7. The *Comprehensive Northern Sotho Dictionary* being a stem-based dictionary, this is particularly unsatisfactory. Indeed, the user first tries to find *Lamodimo*. Upon realising that this item has not been included as such, the user will try to find *Lamodimo* under *-modimo*, then under *-dimo*, and finally under *-mo* – all to no avail.
8. In the corpus, just as in grammar books, the morpheme *bo+* is written in small letters and prefixed to whatever form, whether that form starts with a capital or not. Hence, for instance *boSontaga* and not *Bosontaga*.
9. **Lešala** = 'Pronoun'
   Hlogo = 'Prefix'
   1 = First person singular
   1 = First person plural
   2 = Second person singular
   2 = Second person plural
   ( ) = Subject concord
   ( ) = Object concord

①, ② le ③ = first 1 000, 2 000 and 3 000 items
N = in the written language
B = in the spoken language

Dinomoro tše di tšwelelago ka morago ga 'leina' di bontšha gore leina leo le hlaloš wago ke la legoro lefe.
'The numbers appearing after the part of speech "noun" indicate the gender of the word that is being treated.'

References

Atkins, B.T. Sue. 1998. *Some Discussion Points Arising from Afrilex-Salex '98*. Unpublished course evaluation document. Pretoria: University of Pretoria.

Atkins, B.T. Sue and Krista Varantola. 1997. Monitoring Dictionary Use. *International Journal of Lexicography* 10(1): 1-45.

Calzolari, Nicoletta. 1996. Lexicon and Corpus: A Multi-faceted Interaction. Gellerstam, Martin, Jerker Järborg, Sven-Göran Malmgren, Kerstin Norén, Lena Rogström and Catarina R. Papmehl (Eds.). 1996. *Euralex '96 Proceedings I-II, Papers submitted to the Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden*: 3-16. Gothenburg: Department of Swedish, Göteborg University.

Chi, Man L.A. 1998. Teaching Dictionary Skills in the Classroom. Fontenelle, Thierry, Philippe Hiligsmann, Archibald Michiels, André Moulin and Siegfried Theissen (Eds.). 1998. *ACTES EURALEX'98 PROCEEDINGS, Communications soumises à EURALEX '98 (Huitième Congrès International de Lexicographie) à Liège, Belgique / Papers submitted to the Eighth EURALEX International Congress on Lexicography in Liège, Belgium*: 565-577. Liège: English and Dutch Departments, University of Liège.

Departmental Northern Sotho Language Board. 1988[4]. *Northern Sotho Terminology and Orthography. No. 4.* Pretoria: Government Printer.

De Schryver, Gilles-Maurice (Ed.). 2001. *Pukuntšutlhaloši ya Sesotho sa Leboa 1.0 (PyaSsaL's First Parallel Dictionary)*. Pretoria: (SF)[2] Press.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2000. Dictionary-Making Process with "Simultaneous Feedback" from the Target Users to the Compilers. Heid, Ulrich, et al. (Eds.). 2000: 197-209.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2000a. The Concept of "Simultaneous Feedback": Towards a New Methodology for Compiling Dictionaries. *Lexikos* 10: 1-31.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2000b. The Compilation of Electronic Corpora, with Special Reference to the African Languages. *Southern African Linguistics and Applied Language Studies* 18(1-4): 89-106.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2000c. Electronic Corpora as a Basis for the Compilation of African-Language Dictionaries, Part 1: The *macrostructure. South African Journal of African Languages* 20(4): 291-309.

De Schryver, Gilles-Maurice and D.J. Prinsloo. 2001. Fuzzy SF: Towards the Ultimate Customised Dictionary. Lee, Sangsup (Ed.). 2001. *ASIALEX 2001 PROCEEDINGS, Asian Bilingualism and the Dictionary*: 141-146. Seoul: Center for Linguistic Informatics Development, Yonsei University.

Geeraerts, Dirk. 2000. Adding Electronic Value. The Electronic Version of the *Grote Van Dale*. Heid, Ulrich, et al. (Eds.). 2000: 75-84.

Gerber, Hendri H. 2000. *Woordeboek Afrikaans – Noord-Sotho / Pukuntšu Seburu – Sesotho sa Leboa*. Eldoraigne: Arbeidsprestasie BK.

Gouws, Rufus H. 1999. *Die maatskaplike gerigtheid van die metaleksikografie in 'n meertalige samelewing*. Inaugural lecture. Stellenbosch: Universiteit van Stellenbosch Drukkery.

Gouws, Rufus H. 2000. Toward the Formulation of a Metalexicographic Founded Model for National Lexicography Units in South Africa. Wiegand, Herbert E. (Ed.). 2000. *Wörterbücher in der Diskussion IV. Vorträge aus dem Heidelberger Lexikographischen Kolloquium*: 109-133. (Lexicographica Series Maior 100.) Tübingen: Max Niemeyer.

Hartshorne, Kenneth B., J.H.A. Swart and Edgar Posselt. 1984. *Dictionary of Basic English – N.Sotho across the Curriculum*. Johannesburg: Educum Publishers.

Heid, Ulrich, Stefan Evert, Egbert Lehmann and Christian Rohrer (Eds.). 2000. *Proceedings of the Ninth EURALEX International Congress, EURALEX 2000, Stuttgart, Germany, August 8th-12th, 2000*. Stuttgart: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart.

Kilgarriff, Adam. 1997. Putting Frequencies in the Dictionary. *International Journal of Lexicography* 10(2): 135-155.

Knowles, Frank. 1983. Towards the Machine Dictionary, "Mechanical" Dictionaries. Hartmann, Reinhard R.K. (Ed.). 1983. *Lexicography: Principles and Practice*: 181-193. London: Academic Press.

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

Kriel, Theunis J. 1983[3]. *Pukuntšu woordeboek, Noord-Sotho – Afrikaans / Afrikaans – Noord-Sotho*. Pretoria: J.L. van Schaik.

Kriel, Theunis J., Egidius B. van Wyk and S.A. Makopo. 1989[4]. *Pukuntšu woordeboek, Noord-Sotho – Afrikaans / Afrikaans – Noord-Sotho*. Pretoria: J.L. van Schaik.

Kriel, Theunis J., D.J. Prinsloo and Bethuel P. Sathekge. 1997[4]. *Popular Northern Sotho Dictionary, Northern Sotho – English / English – Northern Sotho*. Cape Town: Pharos.

Laufer, Batia and Linor Melamed. 1994. Monolingual, Bilingual and "Bilingualised" Dictionaries: Which are More Effective, for What and for Whom? Martin, Willy, Willem Meijs, Margreet Moerland, Elsemiek ten Pas, Piet van Sterkenburg and Piek Vossen (Eds.). 1994. *Euralex 1994 Proceedings, Papers submitted to the 6th EURALEX International Congress on Lexicography in Amsterdam, The Netherlands*: 565-576. Amsterdam: Vrije Universiteit.

Prinsloo, D.J. and Bethuel P. Sathekge. 1996. *New Sepedi Dictionary, English – Sepedi (Northern Sotho) / Sepedi (Northern Sotho) – English.* Pietermaritzburg: Shuter and Shooter.

Prinsloo, D.J., Bethuel P. Sathekge and Lizeth Kapp. 1997. *Nuwe Sepedi Woordeboek, Afrikaans – Sepedi (Noord Sotho) / Sepedi (Noord Sotho) – Afrikaans.* Pietermaritzburg: Shuter and Shooter.

Prinsloo, D.J. and Gilles-Maurice de Schryver. 1999. The Lemmatization of Nouns in African Languages with Special Reference to Sepedi and Cilubà. *South African Journal of African Languages* 19(4): 258-275.

Prinsloo, D.J. and Gilles-Maurice de Schryver. Forthcoming. Monitoring the Stability of a Growing Organic Corpus, with Special Reference to Sepedi and Xitsonga. *Dictionaries: Journal of the Dictionary Society of North America* 22.

Sinclair, John M. (Ed.). 1995[2]. *Collins COBUILD English Dictionary*. London: HarperCollins.

Summers, Della. (Director). 1995[3]. *Longman Dictionary of Contemporary English, Third Edition*. Harlow: Longman Dictionaries.

Van Wyk, Egidius B. 1987. Proclitic *bo* of Northern Sotho. *South African Journal of African Languages* 7(1): 34-42.

Ziervogel, Dirk and Pothinus C.M. Mokgokong. 1975. *Pukuntšu ye kgolo ya Sesotho sa Leboa, Sesotho sa Leboa – Seburu/Seisimane / Groot Noord-Sotho-woordeboek, Noord-Sotho – Afrikaans/Engels / Comprehensive Northern Sotho Dictionary, Northern Sotho – Afrikaans/English.* Pretoria: J.L. van Schaik.

Ziervogel, Dirk and Pothinus C.M. Mokgokong. 1988[4]. *Klein Noord-Sotho woordeboek, N.-Sotho – Afrikaans – English/Afrikaans – N.-Sotho/English – N.-Sotho*. Pretoria: J.L. van Schaik.

Appendix A: Days in existing (bilingual) Sepedi dictionaries

Z & M: Ziervogel and Mokgokong 1975

LÁBÓBEDÍ (< **tšatši la bobêdi**) (**Labobêdi**) Dinsdag // Tuesday; (< **lentšu la bobedi) (labobêdi)** altstem // alto (voice)
LÁBÓRÁRO (< **letšatši la boraro**) (**Laboraro**) Woensdag // Wednesday; (< **lentšu la boraro**) tenoor(stem) // tenor (voice)
LÁBÓNE (< **tšatši la bone**) (**Labonê**) Donderdag // Thursday; (< **lentšu la bone**) (**labonê**) bas(stem) // bass (voice)
LÁBÓHLÁNO (< **tšatši la bohlano**) Vrydag // Friday
SÓN'TAGA, (se-)/di- (Sôntaga) (< Afr.), cf. LÁMODÍMO, Sondag // Sunday
LÁMORENA (**Lamorêna**) (< **letšatši la Morêna**) Sondag // Sunday

New E: Kriel 1976[4]

Mošupo'logo, n., Monday.
'Labo'bedi, n., Tuesday.
'Labo'raro, n., Wednesday.
'La'bone, n., Thursday.
'Labo'hlano, n., Friday.
Mokibelo, Mo'ki'bêlô, n., Saturday.
'Lamo'dimo, n., Sunday.

Monday, Mošupologo, Mantaga (A.).
Tuesday, n., Labobedi.
Wednesday, Laboraro.
Thursday, Labone.
Friday, Labohlano.
Saturday, n., Mokibelo.
Sunday, Sontaga, Lamodimo.

Klein: Ziervogel and Mokgokong 1988[4]

Lábóbedí (Labobêdi) Dinsdag/Tuesday
Lábóráro Woensdag/Wednesday
Lábóne (**Labonê**) Donderdag/Thursday
Lábóhláno Vrydag/Friday
Sóntaga (se- kl./cl.) Sondag/Sunday
La'morena (Lamorêna) (< letšatši la Morêna) Sondag/Sunday

Máandag Mošupologô
Dínsdag Labobêdi
Wóensdag Laboraro
Dónderdag Labone
Vrýdag Labohlano
Sáterdag Mokibêlô
Sóndag Sôntaga

Monday Mošupologô
Tuesday Labobêdi
Wednesday Laboraro
Thursday Labonê
Friday Labohlano
Saturday Mokibêlô
Sunday Sôntaga

T & O: Departmental Northern Sotho Language Board 1988[4]

Days of the week / Dae van die week:
Moš. (Mošupologo)
Lbb. (Labobedi)
Lbr. (Laboraro)
Lbn. (Labone)
Lbh. (Labohlano)
Mok. (Mokibelo)
Lam. (Lamorena)

Matšatši a beke:
Moš (Mošupologo)
Lbb (Labobedi)
Lbr (Laboraro)
Lbn (Labone)
Lbh (Labohlano)
Mok (Mokibelo)
Lam (Lamorena)

V Wyk: Kriel et al. 1989[4]

Mošupulogô, snw. dev. kl 3, LLLLL: Maandag.
Mošopologô, snw. dev. kl 3, LLLLL: Maandag
Labobêdi, bsk. kl 5 + snw. HHLH: Dinsdag.

Laboraro, bsk. kl 5 + snw. HLHL: Woensdag.
Labonê, bsk. kl 5 + snw. HHL: Donderdag.
Labohlano, bsk. kl 5 + snw. HLHL: Vrydag.
Mokibêlô, snw. dev. kl 1a, LHHL: Saterdag.
Sôntaga, snw. kl 1a, HLLL: Sondag.
Lamorêna, bsk. kl 5 + snw. HHLL: Sondag, dag van die Here.
Lamodimo, bsk. kl 5 + snw. HLHL: Sondag, dag van die Here.

Maandag, Mošopologô, Mantaga.
Dinsdag, Labobêdi.
Woensdag, La boraro.
Donderdag, Labonê.
Vrydag, La bohlano, Forêitaga.
Saterdag, Mokibêlô.
Sondag, La Morêna, Sôntaga.

Prinsloo: Prinsloo et al. 1996

Mošupologô Monday
Labobêdi Tuesday
Laboraro Wednesday
Labonê Thursday
Labohlano Friday
Mokibêlô Saturday
Sôntaga Sunday
Lamorêna Sunday

Monday Mošupologô
Tuesday Labobêdi
Wednesday Laboraro
Thursday Labonê
Friday Labohlano
Saturday Mokibêlô
Sunday Sôntaga, La Morêna

Prinsloo: Prinsloo et al. 1997

Mošupologô Maandag
Labobêdi Dinsdag
Laboraro Woensdag
Labonê Donderdag
Labohlano Vrydag
Mokibêlô Saterdag
Sôntaga Sondag
Lamorêna Sondag

Maandag Mošupologô
Dinsdag Labobêdi
Woensdag Laboraro
Donderdag Labonê
Vrydag Labohlano
Saterdag Mokibêlô
Sondag Sôntaga, La Morêna

Popular: Kriel et al. 1997[4]

Mošupologô Monday
Mošupulogô Monday
Labobêdi Tuesday
Laboraro Wednesday
Labonê Thursday
Labohlano Friday
Mokibêlô Saturday
Sôntaga Sunday
Lamorêna Sunday
Lamodimo Sunday

Monday Mošupologô, Mantaga
Tuesday Labobêdi
Wednesday Laboraro
Thursday Labonê
Friday Labohlano
Saturday Mokibêlô
Sunday Sôntaga, Lamodimo

Gerber: Gerber 2000

Maandag Mošupologô, Mošopologô
Dinsdag Labobêdi
Woensdag Laboraro
Donderdag Labonê
Vrydag Labohlano
Saterdag Mokibêlô
Sondag Lamorêna, Lamodimo, Sôntaga

Appendix B: Week* in *Basic English – N.Sotho* (Hartshorne et al. 1984)

W46 **week** (n) the seven consecutive days which begin with Monday. *There are roughly four* **weeks** *in each month and 52* **weeks** *in a year***.** 2 **weekday** (n) one of the days from Monday to Saturday. 3 week-end (n) the end of the week, usually from Saturday to Sunday. 4 **long week-end** *When a holiday falls on a Monday a* **long week-end** *is usually from Friday afternoon until Monday evening***.** 5 **weekly** (adj) once a week or every week. **Weekly** *newspapers are usually published on Sundays but some are published on other days of the week***.**

W46 beke 2 letšatši la beke 3 mphelabeke 4 mphelabeketelele 5 ka beke

Appendix C: Questionnaire to retrieve the form of day plurals, version 1

MATŠATŠI A BEKE A SESOTHO SA LEBOA

Thaetlele (Mna/Mtšana/Moh./Ngaka/Prof./ bj.bj.): _____
Sefane: _____
Leina: _____
Bong (Monna/Mosadi): _____
Mengwaga (Mohlala: 20): _____

HLOKOMELA: PAMPIRI YE E NA LE DIPOTŠIŠO TŠE HLANO FELA. O KGOPELWA GO ARABA DIPOTŠIŠO TŠE KA MOKA.

1. Ngwala bontši bja mantšu a a latelago ka mo tlase. Ge go kgonega, efa dikarabo tša go feta e tee.

Mohlala: **Kgomo > (a) Dikgomo (b) Magomo**

BOTEE	**BONTŠI (a)**	**BONTŠI (b)**	**BONTŠI (c)**
Lamorena			
Sontaga			
Mošupologo			
Mantaga			
Labobedi			
Laboraro			
Labone			
Labohlano			
Mokibelo			

2. Ke lengwalo lefe la thuto le o nago nalo? (Swaya ka X)

Praemari	
Matriki	
Kholetšheng	
Tekhnikon	
Yunibesithi	

3. O šoma eng? (Hlaloša ka boripana)

_____
_____
_____
_____

4. O dula kae? (Swaya ka X)

Polaseng	
Nagalegaeng	
Lokheišeneng	
Toropong	

5. Ke polelo goba dipolelo dife tše o di šomišago ka gae, mošomong, le ge o na le bagwera? (Swaya ka X)

Polelo	**Gae**	**Mošomong**	**Bagwera**
Sesotho sa Leboa			
Setswana			
Sesotho sa Moshoeshoe			
Setebele			
Sezulu			
Seswatsi			
Sethosa			
Setšhakane			
Sevenda			
Seisemane			
Seafrikanse			

RE LEBOGA TEMA YA GAGO MOŠOMONG WO!!!

Appendix D: Questionnaire to retrieve the form of day plurals, version 2

MATŠATŠI A BEKE A SESOTHO SA LEBOA

HLOKOMELA: PAMPIRI YE E NA LE DIPOTŠIŠO TŠE HLANO FELA. O KGOPELWA GO ARABA DIPOTŠIŠO TŠE KA MOKA.

1. Ngwala bontši bja mantšu a o tsebago a kgona go tšwelela ka bontšing. Mo go kgonegago, efa dikarabo tša go feta e tee.

Mohlala: **Kgomo > (a) Dikgomo (b) Magomo**

BOTEE	**BONTŠI (a)**	**BONTŠI (b)**	**BONTŠI (c)**
Mošupologo			
Mantaga			
Lamathomo			
Labobedi			
Laboraro			
Labone			
Labohlano			
Foreitaga			
Mokibelo			
Lamorena			
Sontaga			
Lamodimo			

2. Ke lengwalo lefe la thuto le o nago nalo? (Swaya ka X)

Praemari	
Matriki	
Kholetšheng	
Tekhnikon	
Yunibesithi	

3. O šoma eng? (Hlaloša ka boripana)

4. O dula kae? (Swaya ka X)

Polaseng	
Nagalegaeng	
Lokheišeneng	
Toropong	

5. **Ke polelo goba dipolelo dife tše o di šomišago ka gae, mošomong, le ge o na le bagwera? (Swaya ka X)**

Polelo	**Gae**	**Mošomong**	**Bagwera**
Sesotho sa Leboa			
Setswana			
Sesotho sa Moshoeshoe			
Setebele			
Sezulu			
Seswatsi			
Sethosa			
Setsonga			
Sevenda			
Seisemane			
Seafrikanse			

Re kgopela gore o re botše ka bowena ka boripana gomme tshedimošo ye e ka se fiwe motho yo mongwe.

Thaetlele (Mna/Mtšna/Mdi./Ngaka/Prof./ bj.bj.): _____
Sefane: _____
Leina: _____
Bong (Monna/Mosadi): _____
Mengwaga (Mohlala: 20): _____

RE LEBOGA TEMA YA GAGO MOŠOMONG WO!!!

Appendix E: A sound lexicographic treatment of the days of the week based on the research presented in this article

Monday

N③ **Mošupologo** [Mošupologô] *leina* ***3****/4*
Ke letšatši la mathomo la beke, le thoma ka morago ga Sontaga gomme la fela ge go thoma Labobedi: *Ka Mošupologo ge re fihla sekolong morutiši a re bitša ka o tee ka o tee*
Mešupologo [Mešupologô] *leina* 3/**4**
BONA **Mošupologo**
Mošupulogo [Mošupulogô] *leina* ***3****/4*
BONA **Mošupologo**
Mešupulogo [Mešupulogô] *leina 3/***4**
BONA **Mošupulogo**

Tuesday

Labobedi [Labobêdi] *leina* ***5****/6*
Ke letšatši la bobedi la beke, le thoma ka morago ga Mošupologo gomme la fela ge go thoma Laboraro: *E be e le Labobedi ge sefatanaga sa Ariel se ikgaša gare ga mpa ya Tshwane*
Malabobedi [Malabobêdi] *leina 5/****6***
BONA **Labobedi**

Wednesday

Laboraro *leina* ***5****/6*
Ke letšatši la boraro la beke, le thoma ka morago ga Labobedi gomme la fela ge go thoma Labone: *Ka Laboraro mo gare ga beke ye nngwe ba bolela gore ba tla tsoga ba eya Tshwane*
Malaboraro *leina 5/****6***
BONA **Laboraro**

Thursday

Labone [Labonê] *leina* ***5****/6*
Ke letšatši la bone la beke, le thoma ka morago ga Laboraro gomme la fela ge go thoma Labohlano: *O nteleditše mogala ka Labone*
Malabone [Malabonê] *leina 5/****6***
BONA **Labone**

Friday

Labohlano *leina* ***5****/6*
Ke letšatši la bohlano la beke, le thoma ka morago ga Labone gomme la fela ge go thoma Mokibelo: *O tla swara taxi ge a seno tšwa sekolo ka Labohlano*
Malabohlano *leina 5/****6***
BONA **Labohlano**

Saturday

N② **Mokibelo** [Mokibêlô] *leina* ***3****/4*
Ke letšatši la boselela la beke, le thoma ka morago ga Labohlano gomme la fela ge go thoma Sontaga: *Ka Mokibelo re tsogela pitšong*; *O e apare fela ka Mekibelo le Disontaga, le gona ge go na le phathi*; *Ka boMokibelo ba inwela malwa le bašomi ba bangwe ka fao lokheišeneng*
Mekibelo [Mekibêlô] *leina 3/****4***
BONA **Mokibelo**

Sunday

N③ **Sontaga** [Sôntaga] *leina* ***9****/10*
Ke letšatši la bošupa la beke, le thoma ka morago ga Mokibelo gomme la fela ge go thoma Mošupologo {LEHLALOŠETŠAGOTEE **Lamorena**}: *O be a fela a etla ga gešo gagolo ka Sontaga ka morago ga kereke*; *Baruti ba theogela ka Disontaga*
Disontaga [Disôntaga] *leina 9/****10***
BONA **Sontaga**
Lamorena [Lamorêna] *leina* ***5****/6*
Ke letšatši la bošupa la beke, le thoma ka morago ga Mokibelo gomme la fela ge go thoma Mošupologo {LEHLALOŠETŠAGOTEE **Sontaga**}: *Ge e le Lamorena ga re reke felo*; *Ke lemoga lebaka leo bjale ka gore e šetše e le Malamorena a mararo a go hlomagana, o sa thiše kerekeng*
Malamorena [Malamorêna] *leina 5/****6***
BONA **Lamorena**