
African Association for Lexicography



Department of African Languages
University of Pretoria
Pretoria, 0002
South Africa

Tel.: +27-12-420-2494

Fax: +27-12-420-3163

E-mail: etaljard@postino.up.ac.za

Internet: <http://www.up.ac.za/academic/libarts/afrilang/homelex.html>

AFRILEX Newsletter Nr. 7 – February 2004 – Compiled by E. Taljard

TshwaneLex in use at several of South Africa's *National Lexicography Units*

Dear AFRILEX Members,

As announced in the PanSALB Newsletter of June 2003, David Joffe and Gilles-Maurice de Schryver, with hands-on advice from Salmina Nong and D.J. Prinsloo, have created a brand-new software application for the compilation of modern dictionaries and terminology lists. The program is known as **TshwaneLex**, derived from *Tshwane* 'Pretoria' and *lexicography*. TshwaneLex is already in use for the compilation of mono- and bilingual dictionaries at several of South Africa's *National Lexicography Units*, and copies have been acquired by Macmillan, as well as by a number of lexicography projects based in Europe. For the benefit of those NLUs and private dictionary compilers who are still considering whether or not to store their lexicographic gems in TshwaneLex, we would like to briefly introduce the software here.

From the outset it is good to know that TshwaneLex supports **Unicode** throughout (see Annex 1), which means that it can handle virtually all of the world's languages. In the South African context this means that the so-called 'special characters' for Tshivenda can simply be inserted by means of straightforward keystrokes. Specific **customisations for the African languages** have also been implemented, such as provisions to deal with noun class systems. TshwaneLex further includes features such as immediate article preview, customisable fields, automatic cross-reference tracking, automated lemma reversal,

online and electronic dictionary modules, export to Microsoft Word format, the ability to link sound recordings to dictionary fields, dictionary error/integrity checks, automated backups, a filter function (see Annex 2), and teamwork (network) support. TshwaneLex also efficiently handles **large databases**, and a search function allows the entire database to be rapidly searched for some given text. Search options such as case-sensitivity and 'find whole word only' are available. More advanced 'regular expressions' may also be used in the search function.

A primary design goal has been to produce a **user-friendly** tool that is easy to learn and use, and allows lexicographers to get up and running quickly with a new dictionary project. Lexicographers work with a **familiar abstraction**, namely that of dictionary articles, rather than at a level where they are exposed to unnecessary technical implementation details. Thus, lexicographers do not need to have an advanced level of computer literacy in order to perform the general day-to-day tasks of compiling a dictionary. This makes dictionary compilation more accessible to less technically literate lexicographers, and also has the benefit of **lowering training time**.

Among the more advanced features are the 'linked view mode' and the 'dictionary compare/merge function'. When in **linked view mode**, the language window for the target language automatically displays those lemmas that are related to the selected lemma in the source language. With this feature the *consistency* across the two sides of a bilingual dictionary can be safeguarded. The **dictionary compare/merge function** allows different

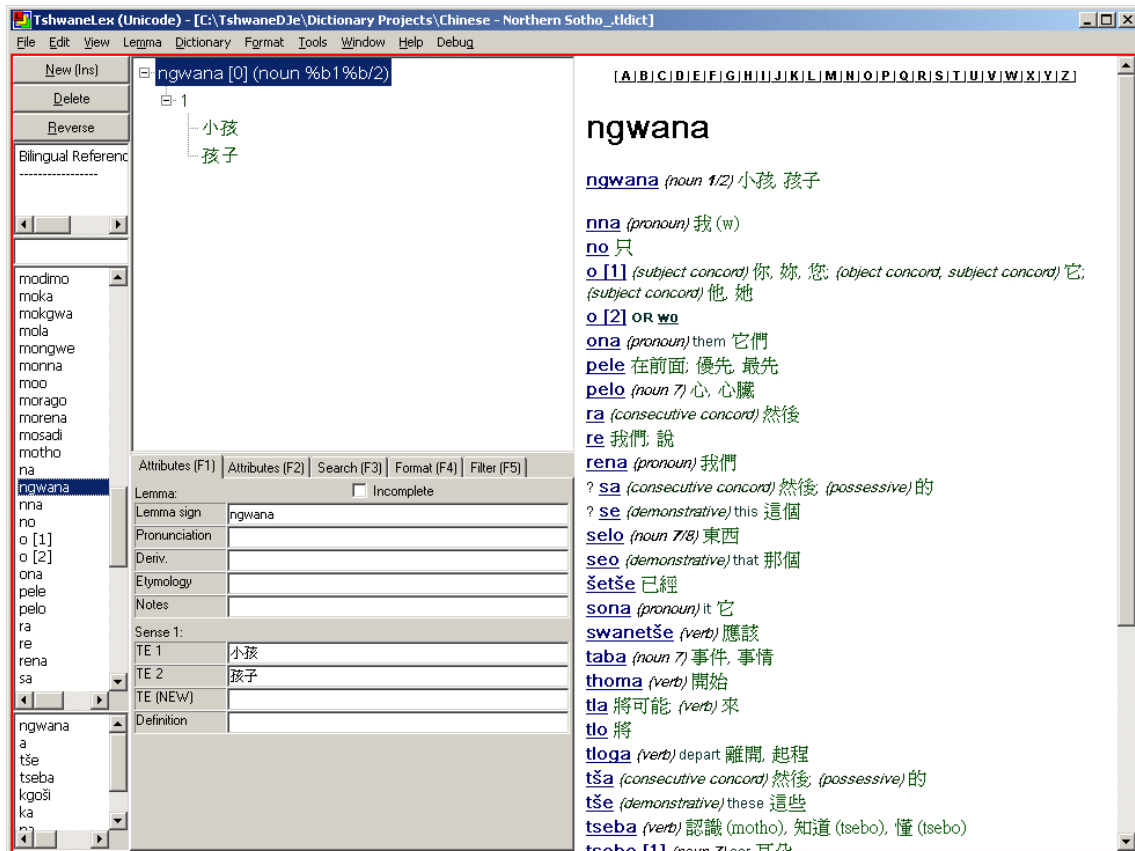
versions of a dictionary database to be compared with one another. This function is especially useful in situations where lexicographers are split up geographically, where it may not be possible to have a high-speed network connection to the main database. Changes may then be made ‘offline’, and periodically merged back into the main database.

TshwaneLex supports the creation of the three major types of dictionary output media: **hardcopy** (paper), **electronic** (CD-ROM), and **online** (Internet). Two basic methods are provided for placing the database contents online. The first is to generate static output, where the dictionary is placed online as a pre-generated file (e.g. HTML, XML, RTF, PDF, Microsoft Word, etc.). It is also possible to export to XML format from TshwaneLex. An XML stylesheet transform may then be applied to generate an HTML page. The

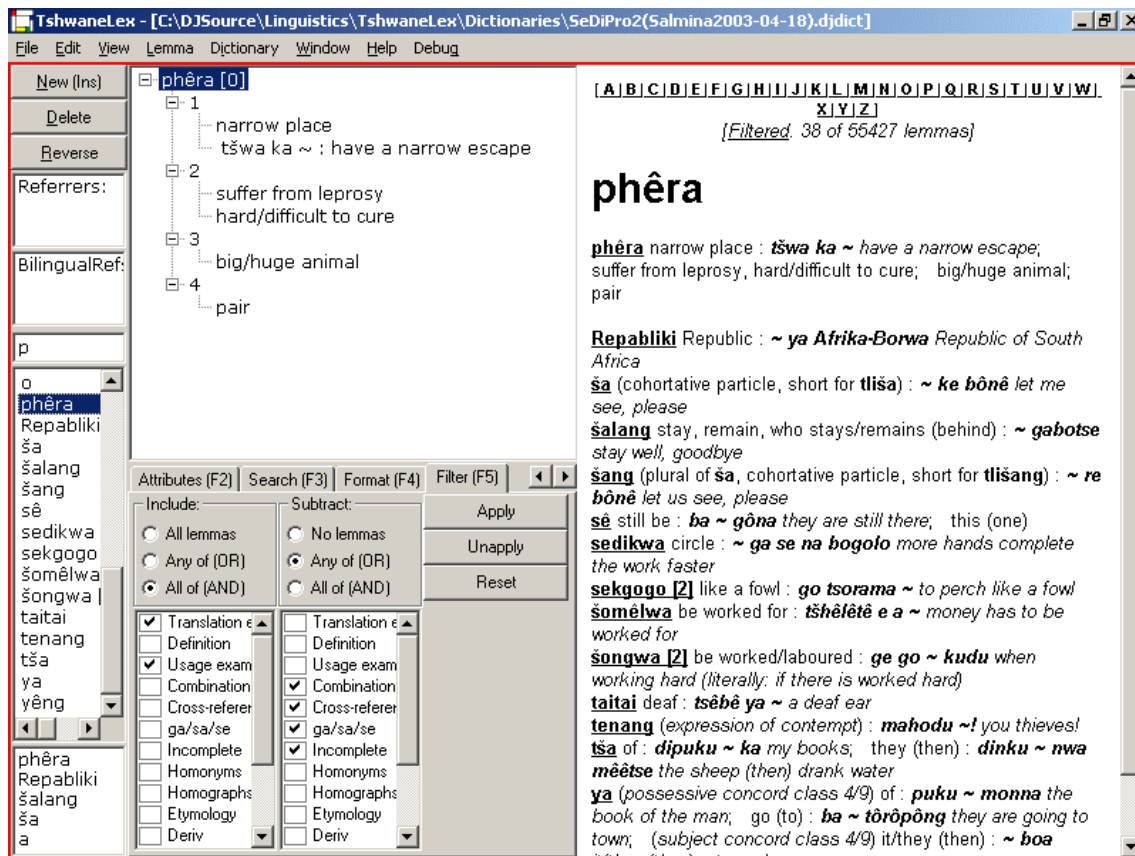
second method, using the **online dictionary software module**, dynamically generates output, and provides far greater flexibility and functionality, such as advanced dictionary usage tracking and analysis. These dictionary usage statistics can be used in many ways to improve the dictionary; for example, one can quickly see what the most frequent not-found are.

TshwaneLex allows new data **exporters and importers** to be created via a plugin system. A plugin for importing databases stored in the *Shoebox / Toolbox* dictionary format is under development, while one for the now discontinued *Onoma Lexical Workbench* has already been created.

For more detailed information, a large selection of screenshots and scientific articles on TshwaneLex, we would like to refer you to <http://tshwanedje.com/tshwanelex/>



Annex 1: TshwaneLex screenshot showing the use of **Unicode** in a Sesotho sa Leboa – Mandarin Chinese dictionary (here the parts of speech are in English, yet this is easily changed to a ‘label set’ in any language; this feature thus allows for true customisation of the printed, electronic or online version of the database)



Annex 2: TshwaneLex screenshot demonstrating the **lemma filter** in a Sesotho sa Leboa – English dictionary (*here in order to view all complete articles which have both a translation equivalent and a usage example, but which have no combinations, no cross-references and no ga/sa/se fields; of which there are only 38 out of 55 427 in this database*)

— Contributed by David Joffe & Gilles-Maurice de Schryver

David Joffe has a BSc Computer Science (1998) degree from the University of Pretoria. He has five years full-time C/C++ software development and project management experience, and over ten years software development expertise. He has managed several complex simulator software systems, and was manager for such projects as the Haul Truck and Shovel Mining Training Simulators for AngloPlatinum, and a flight simulator visualisation system for the South African Air Force.

Gilles-Maurice de Schryver has an MSc in microelectronics (1995), an MA in African-language linguistics (1999), and is currently finalising a PhD revolving around Fuzzy Simultaneous Feedback applied to the South African lexicographic landscape (Ghent University, Belgium). He is the author or co-author of several books and peer-reviewed articles dealing with African-language corpus and dictionary topics.

TshwaneLex

Ye ke porokerama ya khomphuthara ya go hlama dipukuntšu. Nna bjalo ka yo mongwe wa ba mathomo ba go diriša porokerama ye, ke bona e le ye kaone kudu ge ke e bapetša le tše dingwe tše ke ilego ka di diriša peleng. Ke porokerama ya mohuta wa yona o nnoši yeo e fanago ka dinyakwa ka moka tšeo mongwadi yo mongwe le yo mongwe wa pukuntšu a ka ratago go ba le tšona. E sa le go tloga ke thoma go diriša porokerama ye, mošomo wa ka ke bona o tšwela pele ka lebelo le le makatšago ka lebaka la dilo tše dintši tšeo di humanegago mo porokerameng ye. Go ya ka maitemogelo a ka, nka rata go eletša bangwalapukuntšu ba bangwe le bona gore ba diriše porokerama ye. Porokerama ye ga se ya direlwa fela leleme la Sesotho sa Leboa eupša e ka šomišwa go ngwala dipukuntšu tša maleme a mangwe a mantši ao a lego gona mo lefaseng.

Translation

TshwaneLex

This is a computer program for compiling dictionaries. As one of the first users of this program, I find it very user-friendly as compared to other programs that I have used before. It is a unique program that offers everything that a lexicographer might look for. Since I started using this program, my work seems to go faster because of the features of this program. Based on my experience, I would advise other lexicographers to use this program too. The program is made not only for Sesotho sa Leboa, it can be used for almost all the languages in the world.

— Contributed by Salmina Nong

Die Etimologiewoordeboek van Afrikaans (EWA) in gedrukte en elektroniese formaat

Die *Etimologiewoordeboek van Afrikaans (EWA)* in gedrukte en elektroniese formaat is gedurende 2003 deur die Buro van die WAT op Stellenbosch bekend gestel.

Die EWA in gedrukte formaat is op 13 Junie 2003 bekend gestel. Dit is die eerste Afrikaanse etimologiewoordeboek sedert 1967. Die herkoms van meer as 8 000 Afrikaanse woorde word op 'n gebruikersvriendelike wyse aangebied. Die lemmas is só gekies dat dit die ryk, multikulturele oorsprong van Afrikaanse woorde illustreer.

Die woordeboek is 'n leesfees vir elkeen wat meer van die herkoms van Afrikaanse woorde te wete wil kom. Lesers sal 'n goeie aanduiding kry van wat sedert 1652 met die Nederlandse woordeskat in Suid-Afrika gebeur het, en ook van dit wat Afrikaans sonder die toedoen van Nederlands bygekry het. So byvoorbeeld word die invloed van Engels en die inheemse tale van Suid-Afrika op Afrikaans aangedui deurdat talle leenwoorde en leenvertalings uit hierdie tale in die woordeboek opgeneem is. Die wisselwerking tussen Afrikaans en ander tale word verder geïllustreer deurdat daar aangetoon word watter woorde op hulle beurt deur die betrokke tale aan Afrikaans ontleen is. Verder sal die leser ook 'n goeie aanduiding kry van watter woorde of betekenisonderskeidings van woorde in Afrikaans self ontstaan of ontwikkel het.

Die EWA het 'n maklik leesbare karakter deurdat verskillende soorte inligting duidelik van mekaar geskei word. Die woordeboek tree vernuwend op deur ruim aandag aan betekenis-aanbieding te bestee, interessante benoemingsmotiewe aan te bied, en die onderskeid tussen verskillende tipes etimologiese inligting tipografies van mekaar te onderskei.

Die projek is reeds in 1995 deur prof. Piet van Sterkenburg van die *Instituut voor Nederlandse Lexicologie (INL)* geïnisieer, maar het eers in 2001 werklik goed op dreef gekom. Aanvanklik is prof. Johan Combrink en dr. Helena Liebenberg as outeurs aangewys. Na prof. Combrink se afsterwe in Julie 1999 is die redaksie hersaamgestel, met mnr. Gerhard van Wyk as tegniese redakteur om die manuskrip te redigeer en tegniese te versorg, en met dr. Liebenberg, prof. Johan Lubbe, mev. Alet Cloete en mnr. Anton Jordaan as outeurs.

Prof. Fons Moerdijk, oudhoofredakteur van die *Woordenboek der Nederlandsche Taal (WNT)* en tans hoofredakteur van die *Algemeen Nederlands Woordenboek*, het die redaksie van die EWA in die etimologie opgelei. Die EWA is in sy geheel saamgestel met fondse wat uit Nederland bekom is.

Die woordeboek in gedrukte formaat beslaan 612 bladsye en kan teen R195 by die Buro van die WAT bestel word.

Die EWA in elektroniese formaat is op 5 September 2003 bekend gestel. Dit bied 'n gebruikersvriendelike en effektiewe soek-fasiliteit waarmee met die druk van 'n knoppie onder andere vasgestel kan word wat die herkomstaal of herkomswoord van 'n Afrikaanse woord is, wat die datum van 'n herkomswoord is, watter woordvormingsproses of klank- en vormveranderingsproses 'n rol gespeel het in die ontstaan van 'n Afrikaanse woord, wat die benoemingsmotief of datum van eerste optekening van 'n Afrikaanse woord is, wat die verwyderde etimologie van 'n Afrikaanse woord is, of watter Afrikaanse woorde deur ander tale ontleen is.

Met die EWA op CD het die gebruiker nou maklike toegang tot die etimologiese inligting wat daarin opgesluit lê en is die herkoms van Afrikaanse woorde nou toegankliker as wat dit ooit in die verlede was.

Die woordeboek in elektroniese formaat kan teen R175 by die Buro van die WAT bestel word. Alle bestellings moet gerig word aan:

Die Hoofredakteur, Buro van die WAT
Posbus 245, Stellenbosch, 7599.
Telefoon: 021-8873113, Faks: 021-8839492,
E-pos: wat@sun.ac.za

Summary

The Etimologiewoordeboek van Afrikaans (EWA) in paper and electronic format

The Etimologiewoordeboek van Afrikaans (EWA) in paper and electronic format was released by the Bureau of the WAT in Stellenbosch during 2003.

The EWA in paper format is the first Afrikaans etymology dictionary since 1967. The origin of more than 8 000 Afrikaans words is presented in a user-friendly manner. The words were selected to illustrate the rich, multicultural heritage of Afrikaans words. The reader will get a good indication as to what happened with the Dutch lexicon in South Africa since 1652, as well as what influence other languages such as English and the indigenous South African languages had on Afrikaans.

The EWA in electronic format consists of an effective and user-friendly search form that enables the user to get easy access to the etymological information contained within the dictionary. By using the search form in the EWA the origin of Afrikaans words is now more accessible than ever before.

Both products can be ordered from the Bureau of the WAT. See details above.

— Contributed by the Bureau of the WAT

Isihlathululi-mezwi sesiNdebele as fully recognized Dictionary Unit

IsiHlathululi-mezwi sesiNdebele (isiNdebele Dictionary Unit) is well on track with its activities. The Unit was started informally in July 1999 and currently has four staff members, appointed by the Board of Directors. Until recently, Mr B. Skhosana filled the position of Acting Editor-in-Chief, since the Unit did not have a permanent Editor-in-Chief. In August 2003, Ms Katjie Sponono Mahlangu was appointed in this position. Ms Mahlangu had been working for the Department of Education as an educator for 18 years. She started at Mbalenhle High School in 1986, two

years later moving to Khanyisa Primary School where she was appointed as Head of the Department of Languages. While teaching, she furthered her studies in African languages, obtaining a BA degree from UNISA and an Honours degree from the University of Pretoria. She also completed a Computer Literacy Diploma in 1998 and is currently busy with the research paper for her MA in African Languages. Because of her interest in the Ndebele culture, she also studied ‘Special art’ in isiNdebele.

The isiNdebele Dictionary Unit is following a corpus-based approach in their lexicographical activities. They have already succeeded in building an electronic corpus containing two million words, which is used for the compilation of a mono- and a bilingual dictionary. For the compilation itself, they are making use of the new software application *TshwaneLex*. The Unit plans to produce both a mono- and a bilingual dictionary by 2005.

— Contributed by Katjie Sponono Mahlangu

Online Dictionary Interface Sesotho sa Leboa – English

On 22 April 2003 an online dictionary interface Sesotho sa Leboa – English, the first of its kind in South Africa, was placed on the **Internet** at <http://africanlanguages.com/sdp/> by **TshwaneDJe**, a Human Language Technology development team. Two months later, on 20 June, an official academic launch followed at the University of Pretoria, and media releases started appearing from July onwards. Ten months after the first upload of the database to the Internet the popularity remains overwhelming. In ten months’ time as many as 34 000 searches were made by 4 000 unique users. This dictionary is currently also the largest for any African language on the Internet, with approximately 25 000 Sesotho sa Leboa articles and 28 000 items in the English index. The contents of this online dictionary were originally brought together by Gilles-Maurice de Schryver and are currently being revised and expanded by Salmina Nong, who uses David Joffe’s dictionary compilation software **TshwaneLex** (cf. pp. 1–3 above). David Joffe also created the online dictionary

software module, which can be seen as an extension of TshwaneLex.

Linked to the main reference work there is also a linguistics terminology list. The interface of the latter contains several **innovative features** and even a **world's first**, namely the customisation of the output of part-of-speech (POS) tags, usage labels and cross-references depending on the language chosen. The language of the dictionary interfaces can effectively be set in an African language, also a world's first. Those who could not attend the academic launch might be interested in reading about these features and some of their underlying principles in the following publications: *The Compilation of Electronic Dictionaries for the African Languages* (by D.J. Prinsloo, in Lexikos 11), *Lexicographers' Dreams in the Electronic-Dictionary Age* (by G-M de Schryver, in the International Journal of Lexicography 16.2), *Online Dictionaries on the Internet: An Overview for the African Languages* (by G-M de Schryver, in Lexikos 13), and *On How Electronic Dictionaries are Really Used* (by G-M de Schryver & D. Joffe, in the EURALEX 2004 Proceedings).

The TshwaneDJe team is currently collecting more terminology lists and also working towards **CD-ROM** and **paper** versions of the dictionaries. The team further wishes to produce similar online, electronic and hardcopy lexicography tools for numerous **other languages**, and hereby invites interested dictionary makers in South Africa and beyond to contact TshwaneDJe in this regard.

— Contributed by the TshwaneDJe team (<http://tshwanedje.com/>)

Pukuntšu ya Inthanete

Kamego ya ka mo pukuntšung ye ke go badišiša le go godiša diteng tša yona. Ke amegile gape ka go fetolela interface ya wepesaete ya pukuntšu ya inthanete go Sesotho sa Leboa. Ke šomiša porokerama ya *TshwaneLex* go dira diphetogo mo pukuntšung ya inthanete. Pukuntšu ye ya inthanete e na le foromo yeo batho ba ka e tlatšago gomme ba re romela ditshwaotshwao, tšeo re ka di dirišago go kaonafatša pukuntšu ye. Ke hwetša ditshwaotshwao tše di nthuša kudu ka ge le

nna le ge ke le mmoledi wa polelo ye nka se tsebe dilo ka moka malebana le polelo ye. Ke rata go hlohleletša batho gore ba dirišane le rena gore re kgone go kaonafatša pukuntšu ye, gape ba botše ba bangwe ka sedirišwa se.

Translation

Online dictionary

My involvement with the dictionary is to revise and extend the contents. I was also involved in the translation of the interface of the online dictionary web site. I use the TshwaneLex program to make the changes to the online dictionary. The online dictionary has a form that visitors can use to send comments, and we use these comments to improve the dictionary. I find these very useful because even though Sesotho sa Leboa is my home language, I can't know everything about the language. I would like to encourage people to work together with us to improve this dictionary, and to tell others about this resource.

— Contributed by Salmina Nong

News from the Xitsonga National Lexicography Unit

Yuniti ya Rixaka ya Tidikixinari ta Xitsonga yi kumeka eTivumbeni Multi-purpose Centre, tikhilomitara ta kwalomu ka 15 evuxeni bya doroba ra Tzaneen eLimpopo. Xikongomelonkulu xa Yuniti, ku ya hi milawu leyi yi lawulaka, i ku tsala dikixinari yo angarhela ya ririmin'we ya Xitsonga. Kambe yi nga tumbuluxa ni tidikixinari ta tinxaka tin'wana to hambana. Vatirhi va Yuuniti eka nkarhi wa sweswi i: Prof. N.C.P. Golele – Muhlerinkulu; Man. W.V. Mtebule – Mulekhsikhografi; Tat. J.D. Baloyi – Mufambisi wa Hofisi; Tat. M.J. Mongwe – Mulekhsikhografi. Tanihleswi swi nga swa nkoka ku tiva swilaveko eka rixaka, Yuniti ya Xitsonga yi ringeta hi matimba ku fikelela vini va ririmi ni ku va katsa eka migingiriko ya Yuniti. Ku fikelela mhaka leyi ya nkoka, Yuniti yi endla leswi landzelaka.

Ku tirhisana ni Huvo ya Rixaka ya Ririmi ra Xitsonga (XNLB) Leswi i swa nkoka tanihleswi Huvo yi nga yona mulanguteri wa timhaka hinkwato leti khumbaka ririmi ra Xitsonga. Ya letela eka ku ringanisa ririmi, ni ku tlhela yi hlela ntirho lowu endliwaka hi Yuniti.

Ku tirhisana na xitici xa rhediyo xa Xitsonga xa Munghana Lonene FM Leswi i swa nkoka swinene, ngopfu-ngopfu eka ririmi leri ra ha hluvukaka, tanihi Xitsonga, tanihleswi rhediyo yi fikelelaka vanhu vo tala swinene eku tiviseni ka mahungu, ku tlula tindlela tin'wana. Timhaka ta nkoka leti Yuniti yi tsakelaka ku ti tivisa hi leti landzelaka.

Ku tivisa rixaka ku va kona ka Yuniti Leswi swi humelerile hi siku ra 5 Mhawuri 2003. Xitici xa Munghana Lonene FM xi nyikile nkarhi wo ringana awara ni hafu eka nongoloko wa ninhlikanhi. Yuniti yi komberile Huvo yo Angarhela ya Tindzimi ta Afrika-Dzonga (PanSALB) ku va kona eka nongoloko lowu, leswaku yi tivisa rixaka hi tlhelo ra xiyimo ni matumbuluxelo ya Yuniti. Nongoloko wu vuye wu famba hi ndlela leyi: Mbulavulo hi mufambisi wa Hofisi ya Huvo yo Angarhela, Prof. C.N. Marivate; Mbulavulo hi mutirhi wa le ka Huvo yo Angarhela mayelana ni lekhsikhografi, Tat. H.T. Mashele; Mbulavulo hi mutshama-xitulu wa Huvo ya Vafambisi ya Yuniti, Tat. S.E. Mushwana; Mbulavulo hi mukhomeri wa muhlerinkulu, Prof. N.C.P. Golele.

Ku thya Yuniti vito, ni ku yi kumela mfungho Loko Yuniti yi vonile ku fanela ka ku thya vito, ku twananiwile leswaku leswi swi endlwa hi ndlela ya mphikizano eka rixaka. Munghana Lonene FM u pfumerile ku haxa mphikizano lowu, kutani Yuniti yi veka sagwati ra R1000.00 eka loyi a nga ta hlula. Mavito ni mimfungho swi rhumeriwile; ku sale ntsena leswaku swi hleriwa.

Xitumbuluxiwa xo sungula xa Yuniti xi le ku hleriweni I dikidxinari yitsongo ya Xinghezi-Xitsonga. Ku languteriwa leswaku yi ta nyiketwa eka nhlengeletano ya lembe ya vini va ririmi hi n'hweti ya Hukuri, leyi nga ta katsa ni ku khanguriwa ka Yuniti.

Vuxokoxoko byin'wana bya Yuniti byi kumeka eka papila-hungu ra PanSALB ra Dzivamisoko-Khotavuxika 2003.

Summary

The Xitsonga Lexicography Unit is located at the Tivumbeni Multi-purpose Centre, about 15 km east of Tzaneen in Limpopo. The highlights of the Unit at present are as follows.

Close cooperation with the Xitsonga National Body, which is imperative, is well

established, as well as cooperation with the Xitsonga radio station, Munghana Lonene FM.

On the 5th of August 2003 the announcement of the existence of the Unit was aired by Munghana Lonene FM radio station in a 90 minutes afternoon programme. Prof. C.N. Marivate spoke in general on the mandate of PanSALB and its structures. Mr H.T. Mashele spoke specifically on the lexicography programme of PanSALB, followed by Mr S.E. Mushwana, Chairperson of the Board of Directors, and Prof. N.C.P. Golele, Acting Editor-in-Chief.

Munghana Lonene FM also aired a name and logo competition for the Unit, with a prize of R1000.00 provided by the Unit.

The first product of the Unit, a small English – Xitsonga dictionary is at present being edited, to be presented at the end-of-the-year stakeholder meeting and launch of the Unit.

— Contributed by the Xitsonga NLU

Van ossewa na sneltrein: die ontwikkeling van die Elektroniese WAT

Die evolusie van die *Woordeboek van die Afrikaanse Taal* het onlangs 'n hoogtepunt bereik met die voltooiing en verskyning van die *Elektroniese WAT* op CD-ROM en die Internet. Hierdie proses is met *ContentLot / Van Schaik Electronic* (VSE) as vennoot onderneem en is moontlik gemaak deur 'n skenking van die Universiteit van Pretoria.

Die omskakeling van papierweergawe na elektroniese teks was egter geen maklike taak nie. Aanvanklik is besluit om die projek in te deel in twee fases. Fase 1 sou voltrek word met die produksie, binne 17 weke, van 'n elektroniese woordeboek met voltekssoek-kapasiteit. Fase 2 sou dan die omvattende markerings van die teks behels, om dit sodoende versoenbaar te maak vir insluiting by 'n leksikografiese databasis soos *Onoma*, wat deur die Buro gebruik word. Dit het egter spoedig duidelik geword dat veral twee aspekte problematies was. In die eerste plek moes volumes I–VIII (A–K) wat nog nooit vantevore elektronies geredigeer is nie, omgeskakel word na elektroniese teks; in die tweede plek moes dele IX–XI aangepas word om in te skakel by die voorafgaande dele. Ten opsigte van die omskakeling van volumes I–VIII, het onakkurate skandering tot grootskaalse spel- en formateringsfoute gelei

en die proefleesproses wat na die skandering gevolg het, was ook nie na wense nie. Nadat 'n proses van kwaliteitsbeheer in plek gestel is, is hierdie probleme egter vinnig uit die weg geruim. Die grootste struikelblok in die aanpassing van dele IX–XI was die inkonsekwente toepassing van merkers in die bronteks. Daar is nie in die saamstel van dele IX–XI van databasisgedrewe sagteware gebruik gemaak vir die skep van die woordeboekartikels nie (sagteware kompleks genoeg om die Buro se stelsel te akkommodeer, was bloot nog nie beskikbaar nie). Die merkers is deur mense ingevoeg en derhalwe was daar inkonsekwentheid wat die koderingsproses bemoelijk het. VSE het egter wondere verrig om nie net hierdie merkers uiteindelik binne die dele eenvormig te maak nie, maar ook om die kodering relatief eenvormig te maak met die vorige dele.

Die eindproduk van hierdie bloedsweet is 'n omvattende elektroniese naslaanbron met gevorderde soekfunksionaliteit. Dit bevat nie slegs die omvangryke inligting wat in 11 WAT-dele (A–O) vervat is nie, maar ook 'n kerntesourus van Afrikaans (*Woordkeusegids*).

Die *Elektroniese WAT* se nagenoeg 200,000 trefwoorde (lemmas) reflekteer die verskillende variëteite van Afrikaans, asook woordeskat uit geskrewe en gesproke taal. Hierdie inligting is uit omvattende sitaatversamelings en korpora verkry. Die *Elektroniese WAT* bied onontbeerlike hulp vir enige persoon vir wie effektiewe kommunikasie in Afrikaans 'n saak van erns is en werk, volgens 'n Nederlandse kritikus, "als een trein, heel erg goed".

Die *Elektroniese WAT* is beskikbaar op CD-ROM teen 'n koste van R450 en kan by die Buro van die WAT, Posbus 245, Stellenbosch, 7599 (e-pos: wat@sun.ac.za) bestel word. Gebruikers kan as alternatief inteken op die Internet-weergawe, teen 'n koste van R150 per jaar, by <http://www.woordeboek.co.za>

Summary

The Bureau of the Woordboek van die Afrikaanse Taal recently reached an important milestone with the completion and publication of the Elektroniese WAT on CD-ROM and the Internet. This process was undertaken in partnership with ContentLot /

Van Schaik Electronic and was made possible by a generous donation from the University of Pretoria.

The conversion was by no means an easy process. Two aspects were particularly problematic. Firstly, volumes I–VIII (A–K), which were not previously edited electronically, had to be converted to electronic text. This encompassed an onerous process of scanning, proofreading and coding. Secondly, volumes IX–XI (L–O), the text of which was already available between customised SGML-type tags, had to be prepared to be published electronically. The tags had, however, been inserted manually, which led to inconsistencies that hampered the coding process.

The final product of all this hard labour is a comprehensive electronic reference work with advanced search functions. It not only presents the data contained in the eleven volumes of WAT, but also includes a core thesaurus of Afrikaans.

— Contributed by the Bureau of the WAT

Lexicom 2003 & Lexicom 2004

'Learning how to' rather than 'learning about' is the now well-tried approach of the annual *Lexicom Workshop in Lexicography and Lexical Computing*. Last year's workshop was held in Brighton, England, from July 13–18, 2003. Each seminar introduced and discussed a key topic, which was then further explored in practical exercises at the computer. As well as the basics of practical lexicography, the programme covered corpus design and annotation, the extraction of information from corpus data to build a dictionary entry, a brief look at various types of dictionary databases, and an introduction to frame semantics as a practical approach to corpus analysis. Participants received complete documentation of all the seminars in a bound copy of the Course Notes. The course was led by Sue Atkins, Adam Kilgarriff and Michael Rundell, of the *Lexicography MasterClass Ltd*.

This year's workshop will be held in Brighton, England, from June 6–11, 2004. Participation is limited to 24 people. The programme and general approach will be substantially the same, with a few enhancements suggested by this year's group. Pre-registration is already taking place on our web site: <http://www.lexmasterclass.com>

— Contributed by Sue Atkins