

Honderd jaar *Afrikaanse Woordelys en Spelreëls* – ’n oorsig en waardering. Deel 2: Die gebruiker in fokus

*One hundred years of Afrikaanse Woordelys en Spelreëls – An overview and evaluation.
Part 2: The user in focus*

ELSABÉ TALJARD

Departement Afrikatale
Universiteit van Pretoria
E-pos: elsabe.taljard@up.ac.za



Elsabé Taljard



Danie Prinsloo

DANIE PRINSLOO

Departement Afrikatale
Universiteit van Pretoria
E-pos: danie.prinsloo@up.ac.za

NERINA BOSMAN

Departement Afrikaans
Universiteit van Pretoria
E-pos: nerina.bosman@up.ac.za



Nerina Bosman

ELSABÉ TALJARD is professor in die Departement Afrikatale, Universiteit van Pretoria. Haar spesialistaal is Noord-Sotho en haar navorsingsareas is Noord-Sothotaalkunde, leksikografie, terminologie en menslike taaltegnologie. Haar navorsing is oorwegend korpusgebaseerd. Sy dien op die bestuur van AFRILEX (Tesourier), is lid van die Suid-Afrikaanse deel van die ISO/TC37-komitee en dien ook op die kundigheidspaneel van mensliketaaltegnologie van die Departement van Kuns en Kultuur. Sy is die outeur of mede-outeur van 35 artikels in geakkrediteerde tydskrifte en lewer gereeld referate by nasionale en internasionale kongresse. Sy is mede-projekbestuurder van die *Open Education Resource Term Bank*, ’n multi-miljoenrand projek, befonds deur die Departement van Hoër Onderwys vir die daarstel van ’n

ELSABÉ TALJARD is professor in the Department of African Languages, University of Pretoria. Her language of specialization is Northern Sotho and her research interests are linguistics, lexicography, terminology and human language technology. Her research is mainly corpus based. She serves on the Board of AFRILEX (Treasurer), is a member of the South African version of the ISO/TC37 committee and also serves on the expert panel of human language technology of the Department of Arts and Culture. She has authored or co-authored 35 articles in accredited journals and regularly reads papers at national and international conferences. She is co-manager of the Open Education Resource Term Bank, a multi-million rand project funded by the Department of Higher Education for the creation of a multilingual, online term bank for South

<p>veeltalige, aanlyn termbank vir Suid-Afrikaanse studente. Sy is 'n NRF-gegradeerde navorser.</p>	<p>African students. She is an NRF-rated researcher.</p>
<p>DANIE PRINSLOO is professor in Afrikatale aan die Universiteit van Pretoria. Hy is 'n voormalige voorsitter van AFRILEX (African Association for Lexicography), Hoof van die Departement Afrikatale en voorsitter van die Skool vir Tale. Hy is nasionaal en internasionaal bekend vir sy baanbrekersnavorsing in korpusgebaseerde Afrikataal-leksikografie en is die outeur of mede-outeur van meer as 100 publikasies en 'n soortgelyke aantal kongresreferate oor die onderwerp. Die huidige fokus van sy navorsing is op die ontwikkeling van hulpmiddele vir elektroniese woordeboeke en die koppeling van elektroniese woordeboeke met verwerkte korpusdata. Hy is die ontvanger van verskeie toekennings as Uitmendende Akademiese Presteerder en as 'n Denkleier van die eeu van die Universiteit van Pretoria asook van 'n PanSAT-toekenning vir sy bydrae tot meertaligheid.</p>	<p>DANIE PRINSLOO is Professor of African Languages at the University of Pretoria. He is a former Chairperson of AFRILEX (African Association for Lexicography), Head of the Department of African Languages and Chair of the School of Languages. He is well known nationally and internationally for his groundbreaking research in corpus-based African language lexicography and is the author or co-author of more than 100 publications and an equal number of conference papers on this subject. The current focus of his research is on the development of support tools for electronic dictionaries and linking electronic dictionaries with processed corpus data. He received several awards as Exceptional Academic Performer, and as a Centenary Leading Mind from the University of Pretoria, as well as a PanSALB award for his contribution to multilingualism.</p>
<p>NERINA BOSMAN is medeprofessor in die Departement Afrikaans aan die Universiteit van Pretoria. Sy doseer voor- en nagraadse kursusse in Afrikaanse Taalkunde en nagraadse kursusse in Toegepaste Taalstudies. Haar navorsingsgebiede is in die eerste plek die Afrikaanse leksikologie en die leksikale semantiek. Binne die teoretiese raamwerk van die konseptuele metafoorteorie is sy tans besig met 'n projek oor eet- en drinkmetafore in Afrikaans en ander tale. Danksy 'n navorsingstoekenning deur die Nasionale Navorsingstigting, is daar in 2014 onder haar leiding begin met 'n groot projek oor aanspreekvorme in Afrikaans, nie net in Suid-Afrika nie, maar ook in Namibië. In 2017 verskyn, onder haar en prof. Wannie Carstens se redakteurskap, die tweede uitgawe van <i>Kontemporêre Afrikaanse Taalkunde</i>.</p>	<p>NERINA BOSMAN is an Associate Professor in the Department of Afrikaans at the University of Pretoria. She teaches undergraduate and postgraduate courses in Afrikaans Linguistics and postgraduate courses in Applied Language Studies. Her research areas are in the first instance Afrikaans lexicology and lexical semantics. Within the theoretical framework of conceptual metaphor theory, she is currently working on a project on eat and drink metaphors in Afrikaans and other languages. Courtesy of a research grant by the National Research Foundation, a big project on forms of address in Afrikaans was launched under her leadership. This project will involve participants in the whole of South Africa as well as in Namibia. In 2017 the second edition of <i>Kontemporêre Afrikaanse Taalkunde</i>, co-edited by prof. Wannie Carstens, was published.</p>

ABSTRACT

One hundred years of Afrikaanse Woordelys en Spelreëls – An overview and evaluation. Part 2: The user in focus

This article is the second in a series of two concerned with the word list section of the Afrikaanse Woordelys en Spelreëls (AWS). The online version of the word list section was released to users in 2010, shortly after the tenth version of the AWS was published in 2009. This version of the 2009 Woordelys is the focus of the current article. We investigate the Woordelys from a user perspective and contextualise our study against the background of the rapid development of e-lexicography, and the possibilities opened up by the availability of online lexicographic

reference sources. We start our investigation with a critical evaluation from the perspective of the user of the *Woordelys*. In our evaluation we refer to the implications of the unavailability of the Afrikaans spelling rules to the online user, the inadequacy of the software underpinning the online version of the *Woordelys*, and lastly, we comment on the fact that no typographical distinction is made within the articles between different data types. Secondly, we discuss the status of the *Woordelys* as lexicographic information source. In the third section we focus on the user of the *Woordelys* viewed against the background of the information needs of the user in the information era. Analysis of logfiles is a relatively new technique for obtaining information on the real information needs of the user. Based on such analysis, we conclude that the needs of the user of the *Woordelys* are not restricted to seeking information on spelling of specific words only – users of the *Woordelys* have far more sophisticated needs than those pertaining exclusively to spelling. By compiling a corpus of unedited writing by Afrikaansspeaking students, we were able to identify words which are often misspelled and which consequently are suitable candidates for inclusion in future versions of the *Woordelys*. We conclude that the current online version of the *Woordelys* – despite being a valuable reference work – does not provide adequately for the information needs of the modern user. In terms of the function theory – one of the modern theories of lexicography – lexicographic information sources should not only provide cognitive information, but should also provide clear guidance with regard to text production. Provision of this kind of information should be a consideration when a revised version of the *Woordelys* is planned. Apart from making a number of recommendations with regard to a more functional lemma selection, we also recommend the use of a lexicographic ruler, although the relevance of such a ruler for the *Woordelys*, which could be seen as a kind of special dictionary, has not yet been ascertained. Currently, there is a good correlation between the predictions made by the ruler for Afrikaans and the actual alphabetical stretches covered in the *Woordelys*. We conclude with a number of suggestions which could assist in the creation of an improved online version of the old familiar *Woordelys*, which could serve its users well in the next hundred years.

KEY WORDS: *Afrikaanse Woordelys en Spelreëls*, e-lexicography, user perspective, spelling guidance, log file analysis, information needs, lexicographic ruler, lemma selection, user guidance, online dictionaries

TREFWOORDE: *Afrikaanse Woordelys en Spelreëls*, e-leksikografie, gebruikersperspektief, spelleiding, loglêeranalise, inligtingsbehoefte, leksikografiese linaal, lemmakeuse, gebruikersleiding, aanlyn woordeboeke

OPSOMMING

Hierdie artikel vorm die tweede deel van 'n reeks van twee oor die *Woordelys*gedeelte van die *Afrikaanse Woordelys en Spelreëls (AWS)*. Die aanlyn weergawe van die *Woordelys*gedeelte van die *AWS* is in 2010, net nadat die tiende uitgawe van die *AWS* in 2009 die lig gesien het, aan gebruikers beskikbaar gestel. Hierdie weergawe van die 2009-*Woordelys* word in die artikel vanuit 'n gebruikersperspektief beskou en wel teen die agtergrond van die snelle ontwikkeling van e-leksikografie en die moontlikhede wat aanlyn beskikbaarheid aan leksikografiese naslaanwerke bied. Ons begin met 'n kritiese evaluering vanuit 'n gebruikerperspektief van die *Woordelys*. Tweedens kyk ons na die status van die *Woordelys* as leksikografiese inligtingsbron. In die derde plek stel ons die gebruiker van die *Woordelys* in fokus, en wel teen die breër agtergrond van die inligtingsbehoefte van die woordeboekgebruiker in die inligtingsera. Gebaseer op ons

ondersoek, wat onder andere gebruik gemaak het van loglêers en ’n korpus van ongeredigeerde studenteskryfwerk, kom ons tot die gevolgtrekking dat die huidige elektroniese *Woordelys* – alhoewel dit ’n waardevolle naslaanwerk is – nouliks aan die moderne gebruiker se inligtingsbehoefes voldoen. In terme van die funksieteorie – een van die moderne leksikografiese teorieë – moet leksikografiese inligtingsbronne byvoorbeeld naas kognitiewe inligting ook sterk leiding bied ten opsigte van teksproduksie. Die gee van hierdie tipe leiding is iets wat die samestellers van die *Woordelys* in die toekoms miskien ook in gedagte sou kon hou. Ten opsigte van lemmakeuse stel ons voor dat die gebruik van ’n leksikografiese liniaal bykomende hulp sou kon verleen. Ons bied wenke aan oor hoe ’n beter aanlyn weergawe van die ou bekende *Woordelys* by die aanvang van die tweede honderd jaar van die *AWS* daar sou kon uitsien.

1. INLEIDING

Die aanlyn weergawe van die woordelysgedeelte van die *AWS*, tesame met ’n elektroniese gebruikersgids, is in 2010 aan gebruikers beskikbaar gestel (Odendaal 2012:311). In hierdie artikel – die tweede in ’n reeks van twee – word hierdie weergawe van die woordelys (hierna *Woordelys*) (<http://www.woordelys.co.za/>) vanuit ’n gebruikersperspektief beskou. Dit word gedoen teen die agtergrond van die snelle ontwikkeling van e-leksikografie en die moontlikhede wat aanlyn beskikbaarheid aan leksikografiese naslaanwerke bied. Ons begin met ’n kritiese evaluering vanuit ’n gebruikersperspektief van die *Woordelys*. Tweedens kyk ons na die status van die *Woordelys* as leksikografiese inligtingsbron. In die derde plek stel ons die gebruiker van die *Woordelys* in fokus, en wel teen die breër agtergrond van die inligtingsbehoefes van die woordeboekgebruiker in die inligtingsera. Ons sluit af met aanbevelings vir ’n nuwe benadering tot lemmakeuses en spellingleiding.

2. KRITIESE EVALUERING VAN DIE *WOORDELYS*

Die aanlyn weergawe van die *Woordelys* is ’n tipiese voorbeeld van wat Tarp (2011:59) ’n “faster horse” noem: ’n leksikografiese naslaanwerk waarin die samestellers wel van nuwe beskikbare tegnologie gebruik maak, “but only in a very restricted way in order to provide quicker access to their data by means of links or search strings, (...) and which might recognize not only inflectional forms but also parts of words, orthographic variants and even mistakes”. Die *Woordelys* bied dus wel vinniger toegang tot leksikografiese data, maar verken nie volledig die moontlikhede wat deur die elektroniese medium gebied word nie. Die data is steeds gerangskik in tradisionele en statiese inskrywings, wat presies ooreenstem met die inskrywings in die papierweergawe – uit die oogpunt van die gebruiker kan dit as ’n onkritiese oornamie van die papiermodel beleef word. Enkele aspekte van die onderbenutting van die moontlikhede van die e-medium word vervolgens bespreek.

Die eerste ooglopende frustrasie vir die gebruiker is die verwysing na spelreëls waartoe die gebruiker nie aanlyn toegang het nie. Trouens, die onervare gebruiker sal waarskynlik geen benul hê van die betekenis van die syfers wat ná sommige van die inskrywings gegee word nie, terwyl die logiese oplossing gewoon is om die betrokke spelreël deur middel van ’n klikbare skakel aan gebruikers beskikbaar te stel. Sou gebruikers die gebruikersgids raadpleeg deur op die skakel *Belangrike gebruiksinligting*¹ oor die *Woordelys* te klik, vind hulle die volgende verduideliking onder punt 2:

¹ Let op dat die bladsy wat oopmaak, egter die opskrif “Oriëntasie” het.

Die primêre oogmerk van die *Woordelys* is om spelleiding te gee, en dit moet dus in samehang met die Spelreëls gebruik word. Vir dié doel word die voorbeelde wat ter illustrasie by reëls gebruik word, in die *Woordelys* saam met die toepaslike reëlverwysing(s) aangebied. ’n Reëlnommer bestaan uit die hoofstuknommer gevolg deur die volgnummer van die reël, byvoorbeeld 5.1. Indien daar na ’n bepaalde opmerking by ’n reël verwys word, volg die opmerkingnommer ná die reëlnommer, byvoorbeeld 5.1 (a).

Die gebruikers van die *Woordelys* baat geensins by die verwysing na die spelreëls waartoe hulle geen toegang het nie – dit kan lei tot erge frustrasie aan die kant van die gebruiker. Dit is onwaarskynlik dat gebruikers wat die elektroniese weergawe raadpleeg, ’n papierkopie van die *AWS* waarin die spelreëls vervat is, ter hand sal hê – hulle raadpleeg waarskynlik die e-weergawe, juis omdat hulle nie die papierweergawe byderhand het nie. Die opmerking in die Veel Voorkomende Vrae (VVV) dat “(d)ie lyste op hierdie webblad dien as aanvulling tot die *AWS* en nie as ’n vervanging van die *AWS* nie” bied nie veel troos nie – aanvulling impliseer dat bykomende inligting verskaf word – iets wat nie die geval is nie. Verderaan onder VVV word aangedui dat ’n volledige elektroniese weergawe van die *AWS*, die *e-AWS*, met die spelreëls ingesluit, op ’n betaalbasis in die toekoms beskikbaar gestel gaan word. Daar word aangedui dat meer inligting oor die *e-AWS* in 2011 bekend gemaak sal word. Hierdie belofte het egter tot nog toe nie tot uitvoer gekom nie.

In die tweede plek is die rekenaaromatige onderbou van die *Woordelys* gebrekkig wat daartoe lei dat soektogte vreemde en soms verwarrende resultate oplewer. Enkele lukraak gekose voorbeelde word vervolgens bespreek.

Wanneer *voltooide* byvoorbeeld as soekwoord ingetik word, is daar ’n enkele soekresultaat, naamlik *volt.dw.* as lemma, gevolg deur die verklaring “voltooide deelwoord”. Vergelyk die volgende skermskoot:

The screenshot shows the PHAROS website interface. At the top, it says 'PHAROS' with navigation links for 'Tuis | pharos.co.za » | PharosAanlyn » | Veel Voorkomende Vrae'. Below this is the title 'Afrikaanse Woordelys en Spelreëls'. There is a search bar with a magnifying glass icon, a 'Soek' button, and a 'Kies Lyste »' button. Below the search bar is an alphabetical index 'A B C D E F G H I J K L M N O P Q R S T U V W X Y Z | Woordelys' and a link 'Belangrike gebruiksinligting oor die Woordelys »'. The current search term is 'voltooide' and the result count is '1'. A list of search results follows, including 'VNHKV Verenigde Nasies se Hoekommissariaat vir Vluchteling', 'vnu. voornaamwoord; voornaamwoordelik(e)', 'vob. of VOB vragontvangsbewys', 'VOC Generale Verenigde Nederlandsche Geootroyeerde Oostindische Compagnie', 'voegw. of vgw. voegwoord', 'voëlk. voëlkunde; voëlkundig(e)', 'v.o.g. of VOG vry op goederewa', 'VOK of v.o.k. vry op kaal', 'vok. 1. vokaal, 2. vokatief', 'vol. volume', 'volkek. volkekunde; volkekundig(e)', 'volksetim. volksetimologie; volksetimologies(e)', 'volksk. volkskunde; volkskundig(e)', 'volm. volmaak(te)', 'volt. voltooid(e)', 'volt.dw. voltooid deelwoord', and 'volt.teenw.t. voltooid teenwoordige tyd'. On the right side of the results, there are navigation arrows and the text '1. volt.dw. 58% Afkortingslys'.

Die soekresultaat is om verskeie redes vreemd. In die eerste plek verskyn daar in die blokkie aan die regterkant van die skerm 'n lewendige skakel *volt.dw*. Wanneer daarop geklik word, egter, verleen dit nie toegang tot enige nuwe inligting nie. Tweedens word daar in dieselfde blokkie die syfer 58% verstrekk, maar nóg in die gebruikersgids, nóg in die Veel Voorkomende Vrae word enige aanduiding gegee van wat die betekenis van dié syfer is. In die derde plek is dit ook nie duidelik waarom die soekresultaat as *volt.dw* aangegee word nie, terwyl *volt*. voltooid(e) vroeër in die alfabetiese strek voorkom. Die soekresultate is egter irrelevant vir die gebruiker se soektog. Dit wil dus voorkom of die programmering wat die soekfunksie onderlê nie meer gesofistikeerde soektogte ondersteun nie – die soekfunksie is soortgelyk aan dié wat in woordverwerkingspakkette aangetref word, met dié verskil dat hierdie funksie in die *Woordelys* nie aan die gebruiker die opsie bied om byvoorbeeld heelwoorde of slegs gedeeltes van woorde te soek of kassensitiewe soektogte uit te voer nie. Die program kan ook nie 'n kombinasie van soekfunksies hanteer nie. Die soekfunksie is afgestem op die blote masjienmatige herkenning van ortografiese eenhede wat tot onbevredigende soekresultate kan lei.

Bogenoemde leemte in die sagteware word verder duidelik wanneer “tersprake” as soekwoord ingetik word. Dit lewer geen resultate op nie, wat die gebruiker op een van twee maniere kan interpreteer: dit kan naamlik impliseer dat die woord korrek geskryf is, maar nie in die *Woordelys* opgeneem is nie, of dit kan beteken dat die skryfwyse foutief is. Geen verdere gebruikersleiding word verskaf nie. Sou die gebruiker vermoed dat die korrekte skryfwyse *ter sprake* is en dit as soekwoord intik, is daar nie minder nie as 84 resultate. Van die 84 resultate bevat 15 die voorsetsel *ter*, maar *ter sprake* is nie in die resultate te vinde nie. Die gebruiker word verder onkant betrap wanneer die lemma *prompt* onder die soekresultate verskyn. Wanneer daarop geklik word, blyk dit dat die verboë vorm en trappe van vergelyking as “*prompt*, *-te*, *-ter*, *-ste*” aangegee word. Hoewel die soektog dus 84 resultate opgelewer het, bly die gebruiker se inligtingsbehoefte onbevredig. 'n Pluspunt is dat die sagteware wel soektogte met behulp van die sogenaamde Booleaanse soekwysigers (*Boolean Search Modifiers*) * en “” ondersteun, maar selfs hierdie funksie is nie sonder frustrasie vir die gebruiker nie. Sou die gebruiker byvoorbeeld **tjie* intik om 'n lys van alle verkleinwoorde te kry wat op *-tjie* eindig, word aangedui dat daar 258 resultate is. Van die 285 kan die gebruiker egter net die eerste 10 sien en daar is skynbaar geen manier waarop toegang tot die res van die soekresultate verkry kan word nie.

'n Derde tekortkoming in die aanlyn *Woordelys* is dat daar nie (tipografies) onderskei word tussen verskillende inligtingstipes wat deurgaans in hakies en kursiewe druk ná die lemma gegee word nie: soms is dit woordsoortaanduiding, soms is dit betekenisleiding, soms dui dit 'n bepaalde vakgebied of gebruikskonteks aan, soms is dit 'n sinoniem van die soekwoord en soms word 'n vaste uitdrukking of idioom gegee waarin die soekwoord tipies voorkom. Enkele voorbeelde word ter illustrasie aangebied.

mag (ww.): woordsoortaanduiding

maerman (soort plant): betekenisleiding

souffleer (drama): vakgebied

madras (Islamities): gebruikskonteks, moontlik selfs bedoel as etimologiese inligting?

maermerrie (skeen): sinoniem

baat (- vind by): vaste uitdrukking

Daar kan nie sonder meer aanvaar word dat gebruikers die nodige diskriminasie aan die dag sal lê met die interpretasie van sodanige inligting nie. Daar is verwarring oor of daar etikette gebruik word en of ander inligting gegee word. By die lemma *magie* byvoorbeeld, word

“toorkuns” in hakies verskaf. Dit is nie duidelik of “toorkuns” ’n sinoniem vir *magie* is, en of *magie* ’n woord is wat tot die terrein van die toorkuns behoort nie en dit dus gebruikskonteks aandui nie. Dat enige addisionele inligting wat verstrekkend word, potensieel nuttig is, val nie te betwyfel nie, maar alvorens hierdie inligting op ’n beter gestruktureerde en gebruikersvriendelike wyse aangebied word, kan die gebruiker nie maksimaal daaruit voordeel trek nie. Vanuit ’n sagtewareperspektief is een of ander vorm van tipografiese diskriminasie tussen verskillende inligtingstipes ’n relatief eenvoudige prosedure.

In die vierde plek is daar heelwat interne mikrostrukturele inkonsekwentheid. Dieselfde inligtingstipes word naamlik verskillend hanteer. Wanneer betekenisleiding gegee word deur na die superordinaat van die betrokke lemma te verwys, word hierdie hiponimiese verhouding meestal aangetoon deur aan te dui dat lemma X ’n “soort Y” is, vergelyk byvoorbeeld *baar* (“soort seevis”), *rabas* (“soort plant”). Tog word voorbeelde aangetref waar dieselfde leksikale verhouding geld, maar dit word op ’n ander wyse gemerk, bv. *babala* (“gras”), *radioliet* (“fossiel”). Veral laasgenoemde voorbeeld is potensieel verwarrend omdat daar – soos in die voorafgaande paragraaf aangetoon – geen tipografiese onderskeid tussen verskillende inligtingstipes getref word nie. Die gebruiker kan daarom verkeerdelik aanneem dat “fossiel” ’n sinoniem vir *radioliet* is. Verder word lemmas wat tematies verwant en morfologies ooreenstemmend is, verskillend behandel. Vergelyk in dié verband die volgende voorbeelde:

Basotho (die –)

Barolong (lid van ’n bevolkingsgroep), -s

Bapedi (die –)

Volgens die gebruikersleiding moet Basotho en Bapedi saam met ’n bepaalde lidwoord gebruik word. Daar word geen aanduiding gegee ten opsigte van meervoudsvorming nie. Hoewel dieselfde reël ten opsigte van die gebruik van die bepaalde lidwoord by *Barolong* geld, word dit nie aangedui nie en word daar aangetoon dat die meervoud daarvan *Barolongs* is. Benewens die inkonsekwente behandeling van gelyksoortige lemmas is die korrektheid van laasgenoemde meervoudsvorm debatteerbaar.

Daar is egter wel gebruikersvriendelike elemente in die *Woordelys*. Een hiervan is die voorspellende teks. Sodra die gebruiker ’n soekwoord begin intik, verskyn ’n lys woorde wat met die letters wat deur die gebruiker ingetik word, begin. Hierdie eienskap ondervang tot ’n mate die probleem van soekwoorde waarvan die spelling foutief is. Sou die gebruiker bv. *paralel* as soekwoord intik, verskyn die korrekte spelling (*parallel*) wel in die lys woorde wat outomaties “voorspel” word, en kom die gebruiker tog by die korrekte spelling uit. Dit is egter nie altyd die geval nie: sou die gebruiker *teroris* intik, bied die program geen keusemoontlikhede aan nie, en die gebruiker kry slegs ’n nul-resultaat. Verdere verfyning van hierdie funksie is dus nodig.

Dit is dikwels die geval dat gebruikers nie weet hoe om deeltokens of kappies rekenaarmatig aan te dui nie. Die soekfunksie maak hiervoor voorsiening: ’n soektog na “beindruk” lei die gebruiker wel na die korrekte spelling, naamlik “beindruk”.

Vir die *Woordelys* om in navolging van Tarp (2011) se metafoor van ’n “faster horse” na ’n Model T-Ford en uiteindelik tot ’n Rolls-Royce te vorder, is ’n paar modifikasies nodig. In die eerste plek moet die moontlikheid ondersoek word om inskrywings dinamies te maak, dit wil sê om die gebruiker die geleentheid te bied om sy eie profiel te skep en die aktiwiteit of funksie waarvoor hy die *Woordelys* wil gebruik, te spesifiseer. Leksikografiese Model T-Fords bied gebruikers ook die opsie om deur middel van lewendige skakels toegang tot bestaande inligting op die internet te kry om sodoende die gebruiker se spesifieke inligtingsbehoefte te

bevredig. Ons dink byvoorbeeld hier aan skakels na die aanlyn grammatika van Afrikaans soos dit op die taalportaal van die ViVA-webblad (<http://viva-afrikaans.org/index.php>) verskyn. Hoewel Tarp (2011:61) daarop wys dat die kategorie van leksikografiese Rolls-Royces tans nog 'n leë een is, is dit die toekoms van leksikografiese inligtingsbronne. Hierdie toekomstige inligtingsbronne sal pasgemaakte oplossings vir individuele gebruikers in spesifieke gebruikssituasies bied; dit sal toegang tot geselekteerde data in 'n databasis met internetsoektoegte kombineer en sodoende dinamiese oplossings, gebaseer op 'n herbewerking en hervoorstelling van die data in die databasis, aan gebruikers bied.

3. DIE *WOORDELYS* AS (LEKSIKOGRAFIESE) INLIGTINGSBRON

Die huidige status van die *AWS* is dié van 'n preskriptiewe leksikografiese inligtingsbron, wat 'n aanslag reflekteer wat kontrasteer met die moderne leksikografiese benadering – veral wat betref algemene woordeboeke, waar die oorheersende benadering deskriptief van aard is. Spesiale woordeboeke oftewel vakwoordeboeke is egter tot 'n groter mate preskriptief, juis vanweë die leidende rol wat hulle ten opsigte van normering en standaardisering speel. Juis in terme van hierdie rol van die *AWS* sou geargumenteer kon word – soos Odendaal (2016) inderdaad doen – dat die *AWS* 'n woordeboek vir spesiale doeleindes is. Dit word naamlik spesifiek gebruik om spellingleiding aan gebruikers te verskaf en kan derhalwe as 'n spel- of ortografiese woordeboek beskou word wat as verwysingspunt vir korrekte spelling gebruik word. Hoewel algemene woordeboeke ook dikwels geraadpleeg word vir korrekte spelwyses, lê die verskil daarin dat 'n ortografiese woordeboek geen ander leksikografies relevante inligting verstrek nie.

Teenoor Odendaal se mening dat die *AWS* 'n woordeboek is, haal ons aan uit die gebruikersgids van die aanlyn weergawe van die *woordelys*, die eerste punt: “Die *Woordelys* is **nie** 'n volledige lys van alle bestaande Afrikaanse woorde nie. Dit is ook nie 'n woordeboek nie en moet dus nie as een gebruik of beskou word nie.” Die gebruiker word dus eksplisiet gewaarsku om nie enige inligting benewens leiding ten opsigte van spelling en ander ortografiese aangeleenthede in die *Woordelys* te verwag nie. Belangriker as die tipologiese status van die *Woordelys* as spesiale woordeboek al dan nie, is die vraag of die tipe inligting wat in die *Woordelys* verskaf word noodwendig beperk moet word tot spellingleiding. Spellingleiding kan kwalik as 'n geïsoleerde leksikografiese aktiwiteit gesien word – dit kan dikwels nie geskei word van ander tipes gebruikersbehoefte nie, soos verderaan deur die analise van die loglêers uitgewys word. Die sentrale rol wat die gebruiker en sy inligtingsbehoefte in die moderne leksikografie speel, kom vervolgens onder die loep.

4. DIE GEBRUIKER IN FOKUS

4.1 Die gebruikersperspektief in die inligtingsera

Dit is duidelik dat die *Woordelys* – waarskynlik as gevolg van historiese redes en die feitlik uitsluitlik preskriptiewe en normerende aard daarvan – nie die gebruiker en dié se behoeftes ten opsigte van inligting en die verkryging daarvan vooropstel nie. Dit blyk duidelik uit die instruksies wat in die gebruikersleiding van die *Woordelys* gegee word. Die onus word dikwels op die gebruikers geplaas om te sorg dat die inligting wat hulle uit die *Woordelys* herwin het, korrek en gepas vir 'n bepaalde situasie is: “Die lys bevat ook heelwat geografiese name wat om verskeie redes as histories beskou kan word en van tyd tot tyd in publikasies opduik.

Sulke name en afleidings word nie gemerk nie en *dit berus by die gebruiker om te verseker dat die gekose naam of afleiding gepas is vir die betrokke konteks*". (Ons kursivering) S6 'n benadering is uit pas met 'n moderne leksikografiese benadering waar klem op die gebruiker geplaas word en dit as die leksikograaf se plig geag word om toe te sien dat gebruikers so vinnig as moontlik die korrekte inligting verkry ten einde in hul inligtingsbehoefte(s) te voorsien.

Ons wil juis in ons herwaardering die *AWS* en spesifiek die *Woordelys* vanuit 'n gebruikersperspektief benader. Ons argumenteer dat die normerende en standaardiserende funksie van die *AWS* en 'n sensitiwiteit vir die behoeftes van sy gebruikers nie noodwendig onderling uitsluitend hoef te wees nie. 'n Deeglike kennisname van die sentrale rol wat die gebruiker en sy inligtingsbehoefte speel, is nodig indien 'n hersiening van die aanlyn *AWS* en spesifiek die *Woordelys* oorweeg sou word.

In die afgelope ongeveer twee dekades het die gebruikersperspektief, en spesifiek die gebruiker se inligtingsbehoefte, sterk na vore gekom en die oorheersende kriterium vir die beoordeling van die kwaliteit van woordeboeke en by implikasie enige leksikografiese inligtingsbron geword.

The user-perspective, so prevalent in modern-day metalexigraphy, compels lexicographers to compile their dictionaries according to the needs and research skills of well-defined target user groups. The dominant role of the user has had a definite effect on the compilation of dictionaries as well as on the evaluation of their quality. Good dictionaries do not only display a linguistically sound treatment of a specific selection of lexical items. Good dictionaries are products that can be used as linguistic instruments by their respective target user groups. The better they can be used, the better dictionaries they are. (Gouws & Prinsloo 2005:39)

A good dictionary is one in which you can find the information you are looking for – preferably in the very first place you look. (Haas 1962:48)

Die gebruiker en spesifiek sy inligtingsbehoefte is nie net van kardinale belang vir woordeboeke nie – dit is belangrik vir die saamstel van enige leksikografiese inligtingsbron, vergelyk Bergenholtz en Gouws (2010:106) in dié verband:

For the user the type of information source is not important. Important is that he/she retrieves the exact required information as quickly as possible.

In die lig hiervan is dit ook nie verbasend dat die toonaangewende leksikografievereniging *Euralex* se kongrestema in 2016 *The user in focus* was nie.

4.2 Die woordeboekgebruiker in die inligtingsera

In die inligtingsera is gebruikers toenemend daarop ingestel om inligting maklik en vinnig aanlyn te bekom. Daar word dikwels hierna verwys as die eenmalige gebruiker ("on the fly user") se behoefte aan die onmiddellike verkryging van die korrekte inligting ("information on demand"). Inligting oor spelling en ander skryfprobleme kan deesdae ook byvoorbeeld deur middel van skakels na ondersteunende webwerwe of programme soos skryfhelp en verwerkte korpusdata verkry word.

The e-dictionary could, for example, link to a writing assistant in order to solve text production problems of users, and to provide specific information on demand. (Bothma et al. 2016:110)

In terme van die funksieteorie – een van die moderne leksikografiese teorieë – moet leksikografiese inligtingsbronne naas kognitiewe inligting ook sterk leiding bied ten opsigte van teksproduksie. Daar kan nie vandag sonder meer aanvaar of vereis word dat gebruikers lang teksgedeeltes in 'n woordeboek of grammatika moet deurwerk alvorens hul inligtingsbehoefes bevredig is nie.

Die vereistes van die inligtingseeu en die toename in die omvang en kwaliteit van inligtingsbronne op die internet moet in gedagte gehou word by toekomstige samestellings van die *Woordelys*. Soos met ander naslaanbronne loop woordelyste en spelreëls die gevaar om irrelevant te raak indien gebruikers die inligting wat hulle benodig makliker en vinniger kan bekom deur bloot te google. 'n Gesofistikeerde spelkontroleerder vir Afrikaans en 'n aanlyn vertalingsprogram soos *Google Translate* kan byvoorbeeld 'n groot persentasie spellingleiding wat in die verlede slegs deur die *AWS* voorsien is, ondervang en vir die gebruiker maklik die voorkeursopsie bô die gebruik van die aanlyn *AWS* word. Toekomstige weergawes van die *AWS* moet na ons mening hulleself dus as hulpmiddels onderskei ten opsigte van veral dié inligtingstipes wat nie deur spelkontroleerders of ander hulpmiddels soos aanlyn vertaalpakkette, korpuskonkordansies en ander skryfhulpmiddels gebied word nie.

Die voorgestelde verandering in die benaderingswyse tot die samestelling van veral die *Woordelys* wat in hierdie artikel voorgehou word, stel dus die gebruikersfokus, en in besonder die behoeftes van dié gebruiker, wat vinnig en direk die regte antwoord verlang, voorop.

4.3 Inligtingsbehoefes van die *Woordelys*gebruiker

Die aard van die *AWS* en spesifiek dié van die *Woordelys* dui daarop dat dit nou afgestem is op teksproduksie, een van die kommunikatiewe funksies van leksikografiese inligtingsbronne. Die funksie van enige leksikografiese inligtingsbron word vanuit die perspektief van die samesteller daarvan gedefinieer. Wat egter belangrik is, is dat daar belyning tussen die bedoelde funksie van so 'n naslaanwerk en die inligtingsbehoefte van die teikengebruiker is. Tarp (2009: 283) identifiseer 'n aantal strategieë deur middel waarvan gebruikers se inligtingsbehoefes bepaal kan word, en een hiervan is die analise van loglêers² – 'n relatief nuwe prosedure wat moontlik gemaak is deur die beskikbaarheid van die elektroniese medium.³

Daar is verskeie voordele verbonde aan die ontleding van loglêers. Eerstens beïnvloed dié tipe waarneming nie die gebruiker of die leksikografiese konsultasieproses nie. Dit voorsien verder die navorser van 'n groot hoeveelheid data wat die werklike gebruikersgroep verteenwoordig, en op grond van empiriese data kan afleidings gemaak word ten opsigte van die werklike inligtingsbehoefes van die gebruiker. Die kwantitatiewe analise van die data is relatief eenvoudig. 'n Analise van die loglêers van die gebruikers van die *Woordelys* kan dus meehelp om 'n betroubare aanduiding te gee van die inligtingsbehoefes van die gebruiker. Vir die doel van hierdie artikel is die loglêers van die *Woordelys* van 4 Junie 2010 tot 14 Maart 2012 ontleed. 'n Totaal van 41 653 soektogte is vir hierdie tydperk gedokumenteer.

Bergenholtz en Johnsen (2007:2) toon aan dat een van die voor die hand liggendste gebruike van die loglêeranalise is om gapings in die lemmaseleksie in woordeboeke te identifiseer en aan te vul deur woorde wat opgesoek is, maar nie gevind is nie, in te sluit. De

² 'n Loglêer is 'n lêer wat alle soektogte aanteken. Dit stoor onder andere inligting rakende die IP-adres van die gebruiker, die tyd waarop die soektog gedoen is en inligting oor die soektog self.

³ Ons dank aan die Taalkommissie wat die loglêers van die aanlyn *Woordelys* aan ons beskikbaar gestel het.

Schryver en Joffe (2004) rapporteer byvoorbeeld dat die opname van woorde waarvoor daar onsuksesvol gesoek is in die SeDiPro-projek (’n aanlyn, tweetalige Noord-Sotho-Engelse woordeboek) die suksesvolle soektogte van 67% na 75% opgestoot het. Soos Bergenholtz en Johnsen (t.a.p.) tereg opmerk, hang die insluiting van lemmas af van onder meer die funksie van die woordeboek, die teikengebruiker en die werklike doel van die naslaanwerk. ’n Eerste, oppervlakkige analise van die soektogte soos in die loglêers van die *Woordelys* opgeneem, bring aan die lig dat 14 431, d.w.s. 34.6% uit die totaal van 41 653 soektogte ’n nulresultaat opgelewer het. Die bespreking verderaan fokus op hierdie onsuksesvolle soektogte, omdat ons van mening is dat ’n behoorlike analise van dié data die persentasie suksesvolle soektogte beduidend kan verhoog en so waarde tot die *Woordelys* kan toevoeg.

Uit die loglêeranalise is dit duidelik dat daar verskillende permutasies ten opsigte van die aard van die soektog, die resultaat daarvan en die opname al dan nie in die *Woordelys* is:

- soekwoord korrek gespel, maar nie in die woordelys opgeneem nie
- soekwoord verkeerd gespel, korrekte vorm in die woordelys opgeneem
- soekwoord verkeerd gespel, korrekte vorm nie in die woordelys opgeneem nie

Ten opsigte van die eerste kategorie is dit belangrik om in gedagte te hou dat die soektog onsuksesvol was en die gebruiker se inligtingsbehoefte gevolglik onbevredig gelaat is. Die blote feit dat die woord opgesoek is, is ’n aanduiding dat die gebruiker óf onseker is oor die korrekte spelling, óf op soek was na bevestiging dat die woord wel as Standaardafrikaans beskou word. Die gebruiker se inligtingsbehoefte en die gevolglike soektog is dus in lyn met die gestelde funksies van die *Woordelys*. Om klakkeloos alle woorde wat deur gebruikers opgesoek word, in te sluit, is waarskynlik nie haalbaar of sinvol nie en vra vir die oordeel van die samestellers. ’n Oorweging vir die opname van sulke woorde sou byvoorbeeld kon wees hoeveel keer die betrokke woord (deur verskillende gebruikers) opgesoek is. Ter illustrasie: die woorde *sexy* en *postapartheid*, is sewe en vier maal onderskeidelik opgesoek, en sou op grond hiervan vir opname oorweeg kon word, al is dit dan in die geval van *sexy* met ’n aanduiding dat dit nie Standaardafrikaans is nie. ’n Verdere oorweging sou ook die algemene gebruiksfrekwensie van die woord wees. Met die beskikbaarheid van groot datakorpusse is dit relatief eenvoudig om vas te stel wat die gebruiksfrekwensie van enige woord is. ’n Soektog in die 48 miljoen Taalkommissiekorpus (TK) dui byvoorbeeld daarop dat “*sexy*” ’n gebruiksfrekwensie van 349 het. Die samestellers van die woordelys sou argumentsonthaltwe op ’n minimumfrekwensie in die korpus kon besluit waaraan kandidaatwoorde moet voldoen om te kwalifiseer vir opname.

In gevalle waar die soekwoord verkeerd gespel is, en die korrekte vorm wel in die *Woordelys* opgeneem is, is die soektog ook nie altyd suksesvol nie. Die teksvoorspellingsfunksie lei die gebruiker soms na die korrekte woord, maar dit is slegs van nut wanneer die foutiewe spelling en die korrekte weergawe na aan mekaar binne dieselfde alfabetiese strek val. In die meeste gevalle is daar geen leiding wat die gebruiker by die korrekte spelwyse uitbring nie; trouens, soms kan die teksvoorspelling misleidend wees: “nogsteeds” is byvoorbeeld 16 keer opgesoek, en sodra die gebruiker “nog” inlees, verskyn die woorde “nogmaals” en “nugtans” in die kieslys. Na analogie hiervan kan die gebruiker die foutiewe aanname maak dat “nogsteeds” inderdaad een woord is. Verdere voorbeelde sluit in: “aantuiging” (sewe maal gesoek), “soesjie” (nege maal gesoek), “bestanddeel” (ses keer gesoek) en “ensomeer” (11 keer gesoek). Voorbeelde soos hierdie is goeie kandidate vir ’n geïntegreerde outomatiese speltoetsers wat alternatiewe aanbied indien die soekreeks nie gevind word nie – die tipiese “*Did you mean...?*”-funksie. So ’n funksie sou die gebruiker na die korrekte inskrywing begelei en so die gebruiker se inligtingsbehoefte bevredig.

Die sterkste kandidate vir opname in die woordelys, is waarskynlik dié wat tot die laaste kategorie behoort, veral indien dit blyk dat so 'n lemma herhaaldelik opgesoek is. So is *onverskoning* agt keer en *omverskoning* sewe keer opgesoek, maar die korrekte vorm verskyn nie in die *Woordelys* nie, *vuurmaak* is vier keer gesoek, maar *vuur maak* is nie opgeneem nie. Ook hier behoort gebruiksfrekwensie 'n rol te speel in die besluit of so 'n woord opgeneem behoort te word of nie.

Benewens die vul van gapings ten opsigte van lemmaseleksie, kan die analise van die loglêers ook 'n aanduiding gee van die aard van die inligtingsbehoefes van die gebruiker. 'n Analise van die loglêers dui daarop dat die gebruiker ook van die *Woordelys* verwag om in kognitiewe inligtingsbehoefes te voorsien. Hierdie tipe inligtingsbehoefte verwys na gevalle waar die gebruiker kennis oor 'n spesifieke verskynsel wil inwin. Soektogte soos die volgende bevestig dat die inligtingsbehoefes van die gebruiker van die *Woordelys* verder as die soeke na spellingleiding vir teksproduksie gaan:

deelteken
wanneer een woord of twee
wanneer gebruik jy *y* of *ei*
verkleinwoorde spelreëls
koppelteken reëls
betekenis van webtuiste
antoniem *spyt*
ander woord *skurk*

Die feit dat die gebruiker *deelteken* as soekwoord intik, dui daarop dat hy inligting nodig het oor die gebruik van die deelteken, en nie soseer op soek is na die korrekte spelling van 'n spesifieke woord waarin daar 'n deelteken voorkom nie. Die aard van die inligtingsbehoefte van die gebruiker in die laaste drie gevalle spreek vanself.

Die besluit tot watter mate daar aan inligtingsbehoefes soos blyk uit die soektogte hierbo gelys, aandag gegee gaan word in toekomstige weergawes van die *Woordelys*, is uiteraard vir die diskresie van die samesteller(s) van die *Woordelys*. 'n Baie belangrike vraag hier is: kan daar aangeneem word dat 'n toekomstige besluit oor die aard en omvang van 'n aanlyn woordelys nog steeds die tradisionele doel en aard van die lys, soos uiteengesit in Deel 1, in ag behoort te neem?

5. 'N NUWE BENADERING TOT LEMMAKEUSE EN SPELLINGLEIDING

Die benadering tot lemmakeuse vir 'n woordelys moet verskil van die lemmakeuse van 'n gewone woordeboek. By laasgenoemde val die klem op die oeroue probleem rakende watter woorde ingesluit en watter weggelaat kan word. By gewone woordeboeke val die klem ook sterk op lemmatisering, dit wil sê die keuse van 'n lemma wat in werklikheid 'n abstraksie is en wat 'n bepaalde paradigma (moet) verteenwoordig. So, byvoorbeeld, word *hond* (die ongemerkte vorm) as lemma gekies vir die paradigma *hond, honde, hondjie, hondjies*. Om ekonomiese redes gebruik die gewone woordeboek dikwels ook sekere konvensies, byvoorbeeld deur slegs die enkelvoudsvorm op te neem en meervoudsaanduiding op 'n verkorte wyse te gee, *man(s)* of *man, ~s*. Die gebruiker van 'n gewone woordeboek of woordelys kan nie verwag om alle afleidings van 'n spesifieke woord in die naslaanbron, veral nie as lemmas nie, te vind nie.

Die klem in 'n aanlyn, elektroniese woordelys behoort egter anders geplaas te word. In teenstelling tot betekenisinligting as hoofdoel in 'n gewone woordeboek, (hetsy deur betekenis-

omskrywing in 'n verklarende woordeboek of vertaalekwivalente in 'n vertalende woordeboek), staan spellingleiding voorop in 'n woordelys. Betekenisinligting moet egter gegee word in dié gevalle waar verskillende spellings van 'n woord verskillende betekenis impliseer, byvoorbeeld *wyer* (breër) versus *weier* (toestemming weerhou).

Tweedens moet die woordelys sensitief wees vir woorde wat gereeld verkeerd gespел word en kategorieë van problematiese spelaspekte. Voorbeelde van sulke kategorieë is *ei* versus *y* (*wyer/weier*), *f* versus *v* (*fel/vel*), met of sonder deelteken of kappie, (*finansieel/finansieël*), enkel- of dubbelletters (**plakaat/plakkate*), ens. 'n Uiteers belangrike aspek van inligtingsoektogte is die versekering/bevestiging ("reassurance/confirmation") dat die gebruiker die woord reg gespел het en dat die gebruiker korrek gespelde woorde fisies in die lys moet kan sien. Daarvoor word 'n omvattende woordelys van byvoorbeeld die mees gebruikte, korrek gespelde 50 000 Afrikaanse woorde in die elektroniese weergawe benodig.

Só 'n woordelys moet dan deur 'n spelkontroleerder ondersteun word. Die gebruiker wat wil seker maak en die woord *onmiddellik* intik, vind die woord in die lys en kry daardeur bevestiging en vertrou dat die woordelys hom daarop attent sou maak indien daar 'n ander spelling was met 'n verskil in betekenis. Die gebruiker wat **onmiddelik* of **onmidellik* intik word deur die spelkontroleerder die opsies *onmiddellik*, *onmiddellike*, ensovoorts aangebied. In gevalle waar verskillende spellings op verskillende betekenis dui, moet daar klikbare skakels wees na 'n teks met meer betekenis- en produksie-inligting. So, byvoorbeeld, moet gebruikers wat *wyer* opsoek daarop bedag gemaak word dat hulle ook rekening moet hou met *weier* wat 'n totaal ander betekenis het en bowendien 'n werkwoord is.

6. GEBRUIKERSLEIDING VIR DIE KORREKTE SPELLING VAN WOORDE

Vir die doel van hierdie artikel is 'n Afrikaanse korpus van 1.3 miljoen woorde saamgestel uit taakopdragte van voorgraadse Afrikaanssprekende studente en is ondersoek ingestel na woorde wat tipies verkeerd gespел word.⁴ Enkele voorbeelde is uitgesonder vir bespreking hieronder. (Die frekwensietelling van woorde in die korpus word tussen hakies aangegee.)

Dit is opvallend dat woorde dikwels op meer as een wyse verkeerd gespел word. So, byvoorbeeld, kom drie foutiewe spellings vir *plakkate* en twee elk vir *professioneel* en *interessante* voor.

plakkate (164)	*plakate (18)	*plakaate (4)	*plakkaate (9)
plakkaat (143)		*plakaat (54)	
professioneel (8)	*profesioneel (4)	*proffesioneel (1)	
professionele (20)	*profesionele (4)		
interessante (80)	*interessante (1)	*interesaante (1)	
interessant (48)	*interesaant (1)		

Die *Woordelys* gee die volgende resultate:

⁴ Ons dank aan kollegas wat die data wat in die korpus opgeneem is, aan ons beskikbaar gestel het: prof. Rinelle Evans, me. Annelie Botha en me. Linette van der Merwe.

plakkaat, ..kate
 plakkaatboek, -e

Huidige soektog: **plakkaat** | Resultate: **0**

Huidige soektog: **interessant** | Resultate: **0**

Die konvensie om die meervoud aan te dui as ..kate kan op sigself misleidend wees vir relatief onervare woordeboekgebruikers. Dit mag wees dat hulle nie seker is of ..kate vervangend optree vir -kkaat of -kaat nie en gebruikers kan moontlik tot die foutiewe gevolgtrekking kom dat die meervoud van *plakkaat*, **plakate* is as -kkaat met ..kate vervang word.

Bloot prakties of leksikografies gesproke begaan die *Woordelys* geen fout om nie foutiewe vorme soos **plakkaate* en **profesioneel* op te neem nie, maar gegewe die groot aantal foutiewe spellings in die korpus sou dit raadsaam wees om **plakkaat* en selfs **plakate* op te neem met 'n waarskuwing en riglyn na die korrekte spelling, byvoorbeeld **plakkaat**: **foutiewe spelling van plakkaat**, of **plakkaat**: **korrekte spelling = plakkaat**. Hierdie beginsel is nie onbekend in woordeboeke nie. In die *Oxford Bilingual School Dictionary: Northern Sotho and English* (De Schryver 2007) word vorme wat gereeld verkeerd gespел word, in die lemmalys opgeneem met 'n verwysing na die korrekte vorm. Vergelyk byvoorbeeld:

waka Correct spelling = *wa ka*

yaba Correct spelling = *ya ba*

Die lemma **plakkaat* kan byvoorbeeld in die alfabetiese posisie *plak* gelemmatiseer word met verwysing na die korrekte spelling, *plakkaat*. Na ons mening is dit 'n tekortkoming in die woordelys dat dit nie gebruikersleiding voorsien waar die gebruiker die verkeerde spelling in gedagte het en in die woordeboek opsoek nie. 'n Gebruiker wat vermoed dat die korrekte spelling **plakkaate* is, sal meer as een poging moet aanwend om by die korrekte spelling uit te kom. In hierdie geval sou die gebruiker meer gebaat het by 'n goeie spelkontroleerder van Afrikaans. Gebruikers kan byvoorbeeld *plakkate* op enige manier verkeerd spel soos **plakate*, **plakkaate*, of **plakkaate* soos hierbo en die spelkontroleerder sal onmiddellik korrekte opsies aanbied of die spelling outomaties korrigeer.

Meer problematies is gevalle waar verkeerde spelling en alternatiewe spellings met verskillende betekenis gelyktydig voorkom soos *weier/wyer*, *verassing/verrassing*, *fel/vel*, *lei/ly*, *voorkeer/voorkeur*, *waardeer/waardeur*, ens. Tabel 1 gee 'n aanduiding van die korrekte versus foutiewe spelling in die studentekorpus van enkele willekeurig gekose woorde wat binne konteks beoordeel is vir korrektheid al dan nie.

TABEL 1: Reg gespelde versus verkeerd gespelde woorde in die ongeredigeerde studentekorpus

ly	4 uit 14 keer verkeerd gespel
lei	4 uit 100 keer verkeerd gespel
weier	1 uit 13 keer verkeerd gespel
wyer	7 uit 19 keer verkeerd gespel
voorkeer	3 uit 3 keer verkeerd gespel
voorkeur	0 uit 17 keer verkeerd gespel
waardeer	4 uit 39 keer verkeerd gespel
waardeur	6 uit 29 keer verkeerd gespel

In die geval van *wyer* (breër) is die woord in meer as 40% van die gevalle verkeerd gespel omdat die gebruiker *weier* (toestemming weerhou) bedoel het. Vergelyk die volgende uittreksel van konkordansielyne uit die korpus:

N Concordance

- 1 te bewerkstellig, en in die wyer /I sin die struktuur te ver
- 2 partheidsregering het egter wyer gestrek as bloot net die aa
- 3 ksamen toetse. Hier moet ek wyer dink om my les te beplan omr
- 4 **die wat is nou fout, hoekom wyer jy om jou skool klere aan t**
- 5 rbeeld: lank, kort, langer, wyer ensovoorts. Lewensvaardighe
- 6 **toe dit by tandeborsel kom wyer Trudie om haar tande te bor**
- 7 nder vakke. Dit het my laat wyer dink aan moontlikhede wat g
- 8 s handel oor om leerders 'n wyer perspektief te gee oor die
- 9 iese kanon, en ook binne 'n wyer intertekstuele konnteks'' (Litn
- 10 houding te kry, asook om 'n wyer perspektief te kry. Liggaam
- 11 swart leerlinge, dit 'n wyer veld van leer oopgemaak 'n
- 12 wees om die situasie uit 'n wyer perspektief te beskou en ai
- 13 **staan vir skool. Die oggend wyer Trudie nie om wakker te wor**
- 14 ngrik omdat dit ons leer om wyer te soek vir interessante kl
- 15 ngrik omdat dit ons leer om wyer te soek vir interessante kl
- 16 **er staan Trudie op, maar sy wyer om haar skool klere aan te**
- 17 **t die bed te klim nie en sy wyer ook nie om haar skool klere**
- 18 **eet nie lekkers nie. Trudie wyer toe om op te staan en sy ho**
- 19 **beur dieselfde ding, Trudie wyer om skool toe te gaan as haa**
- 20 hul Engels as 'n taal van wyer kommunikasie aanleer. Onde

Die *Woordelys* gee wyer as deel van die afleidingsparadigma vir *wyd*.

Huidige soektog: **wyer** | Resultate: **1**

wrywing, -e of -s
 wrywingkrag of wrywingskrag, -te
 wuf, -te; -ter, -ste
 wuftheid, ..hede
 wuif of wuiwe, ge-
 wuifgroet, ge-
 wulfeniet
 wulps, -e; -er, -ste
 wurg, ge-
 wurggreep, ..grepe
 wurm, -s
 wurmpie, -s
 WVK-agtig, -e 12.10
 WVK-sitting, -s 12.12
 wy (*plegtig in diens stel*), ge- 6.1
 wyd, wye; wyer, -ste

Die gebruiker word nie daarop gewys dat die risiko bestaan dat *wyer* met *weier* verwar kan word nie. Omdat die betekenis van “wyd” of “wyer” nie hier gegee word nie en daar ook nie ’n kruisverwysing na “weier” is nie, kan die gebruiker wat “weier” bedoel maklik konkludeer dat daar net een spelling, naamlik “wyer” is. Gebruikersleiding is hier noodsaaklik by wyse van ’n kruisverwysing, byvoorbeeld “sien ook *weier*”, óf deur ’n hiperskakel na ’n opspringkieslys (“pop-up menu”) óf direkte konteksleiding, *wyer* (breër) vergelyk *weier* (toestemming weier). Die feit dat *weier* glad nie in die *Woordelys* opgeneem is nie, is beslis ’n leemte.

Huidige soektog: **weier** | Resultate: **0**

Gebruikers wat *weier* in die sin van “toestemming weerhou” wil gebruik maar onseker is oor die spelling daarvan, kan byvoorbeeld probeer om albei vorme *wyer* en *weier* op te soek. Die feit dat hulle dan slegs *wyer* teëkom, kan veroorsaak dat daar verkeerdelik aangeneem word dat *wyer* die korrekte spelling vir *weier* is.

Vergelyk ook *verassing* (verbranding) versus *verrassing* (onverwagte gebeurtenis). *Verassing* of ’n afleiding van hierdie woord kom tien keer in die studentekorpus voor, naamlik *verassing* (3), *verassings* (3) en *?verassende* (4).

Die gebruiker wat *verassing* opsoek vind die resultaat: *verassing*, -e of -s. Indien hy egter *verrassing* opsoek is daar geen resultate nie:

Huidige soektog: **verrassing** | Resultate: **0**

In al tien gevalle waar studente *verassing/verassings/?verassende* gebruik het, was die bedoeling ’n handeling wat *verras* en nie die verbranding van ’n lyk nie.

- 1 sy wou graag weet wat die **verassing** is. Bennie vertel haar t
- 2 die manier versier geen **verassings** nie en alles wat gebeur
- 3 ders gepak het. Die groot **verassing** van die dag was die Towe
- 4 Bennie vir Lettie van ’n **verassing**. Dit het Lettie baie nuu
- 5 Wolf vra: “Hou julle van **verassings**?” Thandi en Nina se ja ’n
- 6 Verseboek is steeds vol **verassings** en sal altyd ’n digbunde

- 7 van die karakter. “Die **verassende** kompleksiteit word bygev
- 8 n kom die skrywer met ’n **verassende** einde. Dit bevat element
- 9 se aandag te hou met ’n **verassende** oplossing waar die skuld
- 10 karakters en kan lei tot **verassende** aspekte van die karakter

In ’n koerantkorpus is *verassing* 50 keer in die eerste honderd voorkomste, dit wil sê in 50% van die gevalle verkeerd gebruik, naamlik om na *verrassing* te verwys, en selfs in die korpus van die Taalkommissie self is *verassing* 3 uit 31 (10%) keer verkeerd gebruik.

Statistiek soos hierdie beklemtoon die omvang van die probleem en die dringende behoefte aan duidelike en sterk gebruikersleiding in ’n woordelys, en spesifiek dat daar kruisverwysings tussen *veras* en *verras* en hulle afleidings gegee moet word.

Daar kan geargumenteer word dat die *Woordelys* wel voorsiening maak vir *verassing* en *verrassing* in die sin dat dit uit die alfabetiese konteks afgelei kan word met die inskrywings vir *veras*, *verras* en hulle afleidings.

veras, -te
veras (*verbrand*), het –
verassing, -e of -s

verras, het –
verras, -te
verrassend, -e
verrassinkie, -s

Daar kan egter nie van die gebruiker wat inligting onmiddellik en direk verlang, verwag word om die alfabetiese kontekste vir *veras* en *verras*, wat nie onderling van eksplisiete kruisverwysings voorsien is nie, raak te sien nie. Die gevaar hier is dus dat gebruikers wat onseker is oor die spelling van die veelgebruikte woord *verrassing*, en dit direk opsoek, (*verrassing* en nie via *verras* nie), nie gewaarsku word dat hulle by die laefrekwensie-woord wat “lykverbranding” beteken in die *Woordelys* beland het nie. Die voorkomste van *verras* versus *veras* in die Taalkommissiekorpus dui daarop dat die frekwensie van *verras* 30 keer hoër is as *veras*, vergelyk *verras* (950) versus *veras* (31) en in die koerantkorpus 5 445 keer teenoor 352. Die *Woordelys* moet sodanige direkte toegang bied vir die eenmalige gebruiker (“casual” of “on the fly user”), dit is die gebruiker wat nie noodwendig die woordvormings-/spellingreëls wil bestudeer of selfs bereid is om tipiese leksikografiese lemmatiseringsreëls of selfs die alfabetiese konteks te bestudeer nie. Gebruikers wil bloot weet of hulle *verrassing*, in die betekenis “onverwagte gebeurtenis”, reg spel. In hierdie geval is die grootste leemte dus dat *verrassing* nie in die *Woordelys* gelemmatiseer is nie. Die gebruiker kan dus verkeerdelik aflei dat *verassing* die korrekte spelling vir *verrassing* is.

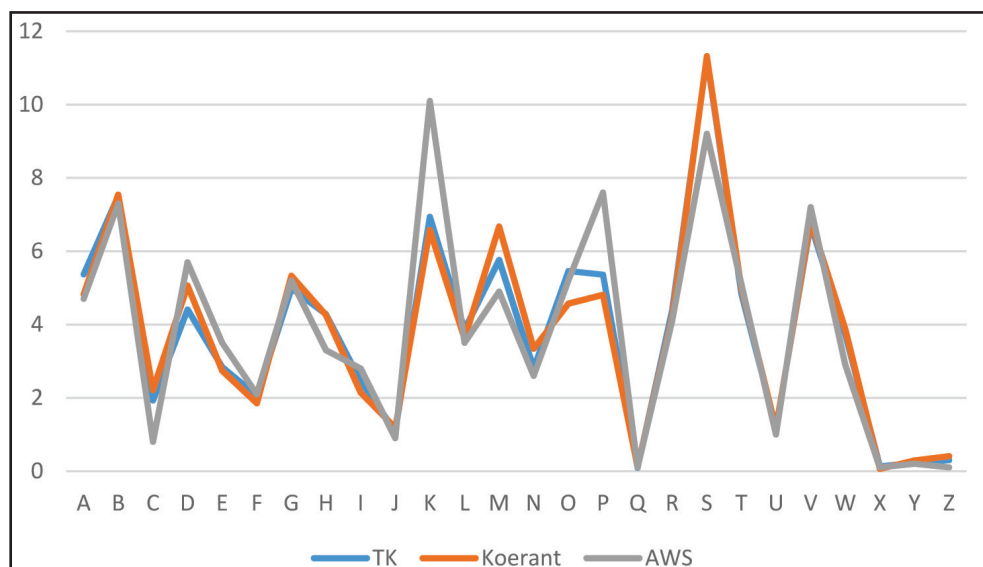
Die gebruiker sou in hierdie spesifieke geval meer baat gevind het by *Google Translate* wat vir *verrassing* die vertaalekwivalent *surprise* aanbied en vir *verassing* *cremation*. Gebruikers benodig egter ’n betroubare inligtingsbron wat hulle nie sal mislei of soms die verkeerde antwoord gee nie. *Google Translate* sit byvoorbeeld die pot mis ten opsigte van *wyer* en *weier* wat albei met *refuse* vertaal word, *voorkeer* met *for once* en *lei* met *slate*. ’n Hersiene *Woordelys* kan hierdie leemte vul.

Die krag van die *Woordelys* lê in sy gesag as normeerder vir korrekte spelling in Afrikaans. Dit moet beslag vind in reëls maar ook ondersteun word deur ’n omvattende woordelys van korrek gespelde woorde waardeur gebruikers seker kan wees dat hulle nie alleen die reël reg toegepas het nie, maar dat die eindproduk wel korrek gespel is en die toepaslike spelling is vir die betekenis wat hulle in gedagte gehad het. ’n *Woordelys* van byvoorbeeld die mees

gebruikte 50 000 Afrikaanse woorde met hiperskakels waar toepaslik, sou uitnemend hieraan voldoen. 'n Moontlike benadering sou wees om frekwensielyste handmatig deur te werk en met 'n spelkontroleerder te vergelyk.

7. DIE *WOORDELYS* IN VERGELYKING MET LEKSIKOGRAFIESE LINIALE VIR AFRIKAANS

Heelwat navorsing is deur Prinsloo en De Schryver (2002, 2003) oor leksikografiese liniale vir Afrikaanse woordeboeke gedoen. Hierdie liniale gee 'n aanduiding van die relatiewe grootte van Afrikaanse alfabetiese strekke in verhouding tot mekaar. So, byvoorbeeld, is die kategorieë B, K, S en V relatief groot terwyl C, J, Q, X en Z relatief klein is. Vir die doel van hierdie artikel is 'n leksikografiese liniaal vir die *AWS* saamgestel en met liniale vir die TK korpus en 'n koerantkorpus vergelyk. Daar is bevind dat die drie liniale tot 'n groot mate ooreenstem en dat die *AWS* gevolglik in terme van hierdie liniale die nodige balans ten opsigte van woorde in die verskillende alfabetiese kategorieë gehandhaaf het.



Figuur 1: *Leksikografiese liniale vir TK, Koerantkorpus en die AWS*

TABEL 2: Persentasie woorde per alfabetiese kategorie in TK, Koerantkorpus en die *AWS*

	TK	Koerantkorpus	AWS
A	5.37	4.82	4.7
B	7.51	7.54	7.3
C	1.93	2.21	0.8
D	4.41	5.06	5.7
E	2.84	2.74	3.5
F	2.06	1.85	2.1
G	4.99	5.33	5.2
H	4.28	4.26	3.3
I	2.45	2.14	2.8
J	1.09	1.18	0.9
K	6.94	6.58	10.1
L	3.86	3.62	3.5
M	5.76	6.67	4.9
N	2.8	3.34	2.6
O	5.46	4.57	5.2
P	5.36	4.81	7.6
Q	0.09	0.12	0.1
R	4.38	4.2	4.1
S	11.28	11.32	9.2
T	4.83	4.99	5.1
U	1.18	1.16	1
V	6.74	6.83	7.2
W	3.74	3.87	2.9
X	0.13	0.06	0.1
Y	0.23	0.29	0.2
Z	0.3	0.41	0.1
	TK	Burger&Beeld	AWS

Daar kan uiteraard nie sonder meer verwag word dat daar 'n eksakte korrelasie tussen woorde in 'n korpus en die woorde wat geselekteer is met die oog op spellingleiding in die *AWS* sal wees nie. By gebrek aan ander duidelike riglyne vir die opname al dan nie van woorde in die *Woordelys*, sou so 'n liniaal wel één van die kriteria kon wees. 'n Mens sou dan kon verwag dat 'n groter kategorie soos S ook meer woorde sal bevat waarvoor spellingleiding nodig is as 'n kleiner kategorie soos J.

8. SAMEVATTING

Dat die *AWS* en spesifiek die *Woordelys* van groot waarde vir die Afrikaanssprekende en -skrywende is, val nie te betwyfel nie. Vir 'n eeu reeds dien dit as gesaghebbende bron vir standaardisering en normering van Standaardafrikaans. Deur 'n aanlyn weergawe van die woordelysgedeelte aan gebruikers beskikbaar te stel, gee die samestellers reeds 'n eerste tree om te voorsien in die behoeftes van die gebruikers van leksikografiese inligtingsbronne wat in die inligtingsera daaraan gewoon geraak het om korrekte inligting direk en vinnig te bekom. Om in die snel ontwikkelende era van e-leksikografie relevant te bly, staan die samestellers van die *Woordelys* onder 'n verpligting om die moontlikhede wat deur die elektroniese media gebied word, maksimaal te ontgin. Ideaal gesproke, behoort die *Woordelys* 'n kombinasie te wees van 'n standaardlys van die 50 000 mees frekwente, korrek gespelde Afrikaanse woorde, aangevul met woorde wat dikwels verkeerd gespeld word, soos deur korpusondersoek uitgewys en woorde waar verskillende spellings verskillende betekenis impliseer. Dié lys kan ook verder uitgebrei word deur woorde wat introspektief deur die samestellers as problematies geïdentifiseer is.

Aangesien die Standaardafrikaanse uitspraak sedert 1917 pertinent genoem word as 'n rede vir die insluiting van woorde, stel ons laastens voor dat toekomstige weergawes van die *AWS* meer uitspraakleiding verskaf. Die aanlyn webwerf kan dan ook klanklêers aanheg van problematiese woorde soos byvoorbeeld *weliswaar*, *sinister* en *veras*.

BIBLIOGRAFIE

- Bergenholtz, H. & Gouws, R. H. 2010. A new perspective on the access process. *Hermes. Journal of Language and Business Communication*, 44:103-127.
- Bergenholtz, H. & Johnsen, M. 2007. Log files can and should be prepared for a functionalistic approach. *Lexikos*, 17:1-20.
- Bothma, T.J.D., Gouws, R.H. & Prinsloo, D.J. 2016. *The role of e-lexicography in the confirmation of lexicography as an independent and multidisciplinary field*. Kongresverrigtinge van die 17de Internasionale EURALEX Kongres, 6-10 September 2016. Tbilisi. Georgië.
- De Schryver, G-M. 2007. *Oxford Bilingual School Dictionary: Northern Sotho and English/Pukuntšu ya Polelopedi ya Sekolo: Sesotho sa Leboa le Seisimane. E gatišitšwe ke Oxford*. Cape Town: Oxford University Press Southern Africa.
- De Schryver, G-M & Joffe, D. 2004. On How Electronic Dictionaries are Really Used. Williams, G. & S. Vessier (eds). 2004. *Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004, Lorient, France, July 6–10, 2004*:187-196. Lorient: Faculté des Lettres et des Sciences Humaines, Université de Bretagne Sud.
- Fuertes-Olivera, P. & Bergenholtz, H. (eds). 2011. *e-Lexicography: the internet, digital initiatives and lexicography*. London: Continuum International Publishing Group.
- Gouws, R.H. & Prinsloo, D.J. 2005. *Principles and practice of South African lexicography*. Stellenbosch: African Sun Media.
- Haas, M.R. 1962. What belongs in the bilingual dictionary? In: Householder, F.W. & S. Saporta (eds). *Problems in Lexicography*. Bloomington: Indiana University. 1962:45-50.

- Odendaal, B.G. 2012. Die herstandaardisering van Afrikaans: 'n Praktiese benadering met die *AWS* as gevallestudie. Ongepubliseerde proefskrif. Universiteit van Stellenbosch.
- Odendaal, G. 2016. Die rol van die US se Departement Afrikaans en Nederlands in die ontwikkeling van die Afrikaanse leksikografie: die Nederlands-Afrikaans-stryd gedurende die aanvangsjare. *Tydskrif vir Geesteswetenskappe*, 56(1):257-276.
- Pharos. 2010. *Afrikaanse woordelys en spelreëls*. [Aanlyn]. Beskikbaar: <http://www.woordelys.co.za/> [11 November 2016].
- Prinsloo, D.J. & de Schryver, G-M. 2002. Designing a Measurement Instrument for the Relative Length of Alphabetical Stretches in Dictionaries, with special reference to Afrikaans and English. Tenth Euralex International Congress, Copenhagen, Denmark, August 2002: 483-494.
- Prinsloo, D.J. & de Schryver, G-M. 2003. Effektiewe vordering met die Woordeboek van die Afrikaanse Taal soos gemeet in terme van 'n multidimensionele Liniaal. In Botha, W. (red.). 2003. *'n Man wat Beur: Huldigingsbundel vir Dirk van Schalkwyk*, pp. 106-126.
- Tarp, S. 2009. Reflections on lexicographic user research. *Lexikos*, 19:257-296.
- Tarp, S. 2011. *Lexicographical and other e-tools for consultation purposes: towards the individualization of needs satisfaction* in Fuertes-Olivera, P. & Bergenholtz, H. (eds), pp. 54-70.
- ViVA Virtuele instituut vir Afrikaans <http://viva-afrikaans.org/index/php>