

Die aard, doel en omvang van die *Afrikaanse Woordelys en Spelreëls*. Deel 1

The nature, goal, and scope of the Afrikaanse Woordelys en Spelreëls. Part 1

GB (GERHARD) VAN HUYSSTEEN

Sentrum vir Tekstegnologie (CTeXt)

Potchefstroomkampus

Noordwes-Universiteit

E-pos: gerhard.vanhuysteen@nwu.ac.za



Gerhard van Huyssteen

GERHARD VAN HUYSSTEEN is professor in Afrikaanse taalkunde en taaltegnologie aan die Noordwes-Universiteit. Hy dien sedert 2005 op die Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns en is tans voorsitter van dié liggaam. Vanaf 2009–2014 dien hy ook op die paneel van mensetaaltegnologiesekundiges van die Departement Kuns en Kultuur. Hy is onder andere ontvanger van die Elizabeth Eybers-beurs, die CL Engelbrecht-prys vir taalkundenavorsing vir sy aandeel in die publikasie van die *Afrikaanse Woordelys en Spelreëls*, en die Stalsprys vir multi- en interdisiplinêre spanwerk. Hy het al verskeie grootskaalse, internasionale multimiljoenrandprojekte gelei, waaronder projekte vir Microsoft, MTN en die Suid-Afrikaanse regering. Van Huyssteen is veral bekend vir sy bydrae as taalkundige in die ontwikkeling van tekstegnologietoepassings en verskeie elektroniese hulpbronne vir Afrikaans, asook as konseptualiseerder, samesteller en beoordelaar van die kykNET-program *Op die Spel*.

GERHARD VAN HUYSSTEEN is professor in Afrikaans linguistics and language technology at the North-West University. Since 2005, he served in the “Taalkommissie” (Language Commission) of the “Suid-Afrikaanse Akademie vir Wetenskap en Kuns” (South African Academy for Science and Arts), and is currently chair of this commission. He has also served in the Human Language Technology Expert Panel of the Department of Arts and Culture from 2009–2014. He is the recipient of, among others, the Elizabeth Eybers Bursary, CL Engelbrecht Award for linguistic research for his contribution to the publication of the *Afrikaanse Woordelys en Spelreëls*, and the Stals Award for multi- and interdisciplinary team work. He has initiated and led numerous large-scale, international, multimillion rand projects for institutions like Microsoft, MTN and the South African government. Van Huyssteen is known for his contribution as a linguist in the development of text technology applications and various resources for Afrikaans, as well as conceptualiser and developer of the kykNET programme *Op die Spel*.

ABSTRACT***The nature, goal, and scope of the Afrikaanse Woordelys en Spelreëls. Part 1***

When the first Afrikaanse woordelys en spelreëls (AWS) was conceived more than a hundred years ago, it was not yet standard practice to explicate the nature, goal and scope (for example with regard to inclusion and elimination criteria) of dictionaries explicitly in a dictionary conceptualisation plan. Through the years, the compilers of the AWS (i.e. the Language Commission (“Taalkommissie” (TK)) of the “Suid-Afrikaanse Akademie vir Wetenskap en Kuns”) provided in the respective forewords of each new edition of the AWS indications of their specific perspectives, goals and methodologies. Until 2006, no internal (or public) dictionary conceptualisation plan of the AWS existed.

From 2006 onwards, the TK endeavoured to compile a style guide (addressing matters such as the sequence of dictionary tags, alphabetisation, sequence of disjunctively and conjunctively written forms, treatment of clippings, typography, etc.), as well as methodological guidelines for inclusion and elimination of lemmas in the word-list section of the AWS. None of these documents have been formally approved by the TK, or have been released in the public domain. The aim of this article is to organise and interpret some of the information contained in these documents, to construct an operationalisation framework, and to make this information available for scientific scrutiny.

With regard to the nature of the AWS (as a so-called spelling or orthography dictionary), it is characterised within a broader dictionary typology as a limited, monolingual, synchronic, linguistic-encyclopaedic reference work. As such, it shares certain similarities with standard monolingual dictionaries, specifically regarding the normative character of some of these dictionaries. One can therefore also say that it is the primary goal of the AWS to set norms and standards for using present-day, written, formal Standard Afrikaans.

Regarding the scope of the word-list section of the AWS, this article argues that lemmas in the AWS has a bipartite function, serving simultaneously as macro- and microstructural dictionary elements. Numerous inclusion and elimination criteria that determine the scope of the AWS are discussed and illustrated with examples from the AWS.

Macrostructural criteria include: (1) lexical items that could present potential spelling problems and problems of style; (2) the status of lexical items (taking the distribution and frequency of a lexical item in corpora into account); (3) lexical items that should be included for the sake of completing certain paradigms in the AWS (e.g. derivations of language names); (4) illustration material in the rule section of the AWS; and (5) physical limitations in the scope of a to-be-published book, set by a publisher.

Microstructural criteria include: (1) linguistic correctness/motivation of lexical items; (2) institutionalisation of lexical items in other dictionaries; (3) frequency of lexical items (and specifically variant lemmas) in a variety of corpora; and (4) the dictionary style and editorial policy of the AWS.

A few statistical models and techniques that could be considered by the TK when analysing corpus data are also discussed. The article concludes with a summative table that could serve as an operationalised frame of reference for the AWS. A few topics for further future research are also included.

KEYWORDS: Afrikaans; *Afrikaanse Woordelys en Spelreëls (AWS)*; elimination criteria; inclusion criteria; lexicography; normative linguistics; orthography; Suid-Afrikaanse Akademie vir Wetenskap en Kuns (SAAWK); Taalkommissie (TK)

TREFWOORDE: Afrikaans; *Afrikaanse Woordelys en Spelreëls (AWS)*; elimineringskriteria; leksikografie; normatiewe taalkunde; opnamekriteria; ortografie; Suid-Afrikaanse Akademie vir Wetenskap en Kuns (SAAWK); Taalkommissie (TK)

OPSOMMING

Toe die eerste *Afrikaanse woordelys en spelreëls (AWS)* meer as honderd jaar gelede gekonsipieer is, was dit nog nie standaard praktyk om die aard, doel en omvang van woordeboeke duidelik, omvattend en sistematies uit te spel in 'n woordeboekkonseptualiseringsplan nie. In 2006 het die TK begin om twee dokumente saam te stel, naamlik 'n stylgids (waarin kwessies soos die volgorde van etikette, alfabetisering, volgorde van los-, vas- en koppeltekenskrifvorme, inkortings, tipografie, en so meer aangeraak word) en riglyne vir die TK (waarin opname- en elimineringskriteria vir lemmas in die woordelysgedeelte van die AWS bespreek word). Geen van beide is tot op datum van publikasie van hierdie artikel formeel deur die TK goedgekeur of in die openbare domein beskikbaar nie. Die doel van hierdie artikel is om van daardie inligting te organiseer, daarop uit te brei en as 'n operasionaliseringsraamwerk te formuleer. Met betrekking tot die aard van die AWS word die AWS binne 'n groter woordeboektipologie gekarakteriseer as 'n beperkte, eentalige, sinchroniese, linguisties-ensiklopediese naslaanbron. Die belangrikste doel van die AWS is om norme daar te stel vir die gebruik van hedendaagse, geskrewe, formele Standaardafrikaans. Met betrekking tot die omvang van die woordelysgedeelte in die AWS word daarop gewys dat lemmas in die AWS 'n dubbele rol as woordeboekinskrywings vervul, naamlik as makro- én mikrostrukturelemente. Verskeie opname- en elimineringskriteria wat die omvang van die AWS bepaal, word bespreek en aan die hand van verskeie voorbeelde toegelig.

1. INLEIDING

In sy resensieartikel oor die negende uitgawe van die *Afrikaanse Woordelys en Spelreëls (AWS)*⁹ skryf Van Sterkenburg (2004:428): “Verder is de woordenlijst uitgebreid met nieuwe lemma's en met meer afleidingen en samenstellingen. Maar, over de wijze van selectie daarvan geen woord.” Hy beveel dan ook verder aan: “Je moet als commissie een referentiekader hebben waaraan je alles kunt toetsen” (Van Sterkenburg 2004:428). In die moderne leksikografiepraktyk word so 'n verwysingsraamwerk idealiter a priori in die leksikografiese stylgids (Bergenholtz 1990) omskryf, wat deel van die woordeboekkonseptualiseringsplan (Gouws & Prinsloo 2005:14) van enige woordeboek is.

Toe die eerste AWS meer as honderd jaar gelede gekonsipieer is, was dit nog nie standaardpraktyk om die aard, doel en omvang (byvoorbeeld opname- en elimineringskriteria vir lemmas in die woordelysgedeelte) van woordeboeke duidelik, omvattend en sistematies uit te spel in 'n konseptualiseringsplan nie. Die enigste aanduiding oor die aard van AWS¹ vind ons in die voorwoord:

Maar hoewel die spelreëls [wat op 18 September 1915 deur die Suidafrikaanse Akademie vir Taal, Lettere en Kuns goedgekeur is] die algemene beginsels bevat, is dit op sigself nie voldoende nie. Daar blij nog altoos baie gevalle van spelling oor wat veel beter in 'n aparte woordelys kan behandel word, omdat sodanige vorme daar veel makliker te vind is. Daarom het die Akademie ... opdrag gegee ... om 'n Afrikaanse woordelys saam te stel, in aansluiting bij, en ter aanvulling van die ... spelreëls. (AWS¹:IV-V)

Oor die doel – en veral die teikengebruikers – is dit duidelik dat die eerste Taalkommissie (TK) hom rig op die volwasewording van Afrikaans as (veral) onderwystaal:

Vanself dus wil ons sê dat ons taal in ieder opsig tot sy reg kom; dat dit die draer word van ons kultuur. Vandaar dat ons in die eerste plaas ons oog rig op die onderwys. (AWS¹:III)

Reeds in AWS¹ identifiseer die TK een van die belangrikste riglyne wat die omvang van die AWS oor die volgende eeu sou bepaal, naamlik dat prioriteit gegee word aan woorde wat potensieel spellingprobleme oplewer:

Wat die aantal woorde betref, is dit die samestellers nie in die eerste plaas daarom te doen gewees om 'n volledige lys te gee van al die gebruiklike Afrikaanse woorde nie. ... Daarom het die samestellers hul veral bepaal by die woorde, wat in meer of minder mate spellingmoeilikhede oplewer. Daarnaas is egter ook gestreef na 'n redelike mate van volledigheid. ... Maar die grenslyn is dikwels vaag; en hoe maklik word 'n woord nie uitgelaat as 'n lys vir die eerste keer saamgestel word nie. Die opname van vreemde woorde is met opset beperk, om die deur nie al te wijd oop te sit nie. (AWS¹:VI)

Twee ander kriteria, naamlik frekwensie en algemeenheid, word ook (hoewel indirek) reeds in die voorwoord van AWS¹ gestel, en dit is twee beginsels wat 'n eeu later – soos in hierdie artikel sal blyk – steeds geld:

Wie 'n woordelys saamstel van woorde wat nog in algemene gebruik is, en daarby probeer om die mees algemene uitspraak weer te gee ... so iemand neem 'n lastige taak op hom.... Een versekering kan die samestellers darem gee, en dit is dat hul altoos daarna gestreef het om, na hul beste wete, die gebruiklikste, dus die beste uitspraak weer te gee. Waar verskillende uitsprake ewe gewoon is – wat dikwels die geval is, veral met 'n jong taal – daar is ook verskillende vorme opgeneem. Te veel vryheid kan nie gegee word nie; dit stig net verwarring. Aan die ander kant mag ons nie ons taal in 'n dwangbuis forseer nie. (AWS¹:V-VI)

Met verloop van die jare is daar in elke voorwoord van elke nuwe uitgawe van die AWS duidelik gemaak wat die sienings en werkwyse van die TK('s) van daardie spesifieke uitgawe was (vergelyk ook McLachlan 2016). Behalwe vir oorlewering van TK tot TK of uitgawe tot uitgawe van die AWS, en behalwe vir besluite wat opgeteken is in die notules van TK-sittings (vergelyk Van Rensburg 2017), het daar tot en met 2006 geen formele stylgids of konsep-tualiseringsplan vir die AWS bestaan nie.

In 2006 het die TK begin om twee dokumente saam te stel:

- *Stylgids*: In hierdie dokument is kwessies soos die volgorde van etikette, alfabetisering, volgorde van los-, vas- en koppeltekenskrifvorme, inkortings, tipografie, en so meer aangeraak. Die doel van die dokument is veral om inkonsekwentheid in die AWS te probeer uitkakel en is aanvanklik deur Mariëtta Alberts opgestel (en daarna verder uitgebrei en onderhou deur veral Gerhard van Wyk en Herman Beyer).
- *Riglyne vir die Taalkommissie*: Daar is begin om geformaliseerde riglyne, wat as opname- en elimineringskriteria vir lemmas in die woordelysgedeelte kan dien, vir die TK neer te lê. Ek en Tom McLachlan was die aanvanklike outeurs van dié dokument.

Beide hierdie dokumente is dinamiese dokumente wat steeds organies groei; geen van beide is tot op datum van publikasie van hierdie artikel formeel deur die TK goedgekeur en as finaal

“afgeteken” nie. Soos wat nuwe gevalle hulle voordoën, of soos nuwe insigte in die werkwyse van die TK verkry word, word die dokumente dienooreenkomstig aangepas.

In hierdie artikel is die fokus laasgenoemde dokument. Die doel van die artikel is om die inligting daarin na my eie siening te organiseer, daarop uit te brei en as ’n operasionaliserings-raamwerk te formuleer. In Van Huyssteen (2017) illustreer ek die toepassing daarvan aan die hand van ’n toetsgeval. As sodanig is hierdie artikel dus my persoonlike (kritiese) beskouing van die aard, doel en omvang van die AWS en die werkwyse van die TK in die periode 2006 tot 2016. Ofskoon hierdie artikel van interne TK-dokumentasie gebruikmaak, is en bly dit my persoonlike interpretasie en verteenwoordig dit nie die amptelike standpunt van die TK nie; enige denk- of berekeningsfoute moet dus voor my deur gelê word. Ek hoop egter om hiermee van die kennis en werkwyse van die TK in die openbare domein beskikbaar te stel, sodat dit ook aan die normale praktyk van die wetenskaplike proses onderwerp kan word. Hopelik dra dit op dié wyse by tot die vakkundige diskoers oor die AWS en ander spelling- of ortografiewoordeboeke soos die AWS, ten einde uiteindelik tot nuwe insigte en metodologieë te lei.

In afdeling 2 word die aard van die AWS in oënskou geneem, met spesifieke verwysing na die AWS se plek in ’n groter woordeboektipologie en die ooreenkomste wat dit met standaard verklarende woordeboeke vertoon. Afdeling 3 sluit hierby aan waar besin word oor die doel van die AWS, ook met spesifieke aandag aan die normatiewe aard en preskriptiewe funksie van ’n bron soos die AWS. In afdeling 4 word verskeie opname- en elimineringskriteria wat die omvang van die woordelysgedeelte in die AWS bepaal, bespreek en aan die hand van verskeie voorbeelde toegelig. Die artikel word afgesluit met ’n samevattende tabel wat myns insiens as geoperasionaliseerde “referentiekader” vir die AWS kan dien (dit is deur Van Sterkenburg (2004:428) as ’n leemte uitgewys), asook met ’n lys onderwerpe wat in die toekoms verder nagevors kan word.

2. DIE AARD VAN DIE AFRIKAANSE WOORDELYS EN SPELREËLS

By gebrek aan empiriese gegewens oor gebruikers se siening van wat die AWS is of behoort te wees, word die aanname hier gemaak dat daar uiteenlopende sieninge is oor die aard van die AWS. Daar word gepostuleer dat sommige gebruikers dit sien as ’n woordeboek, ander as ’n naslaanbron, ander as ’n reëlboek; dat sommige wil hê dit moet ’n normatief-preskriptiewe handleiding wees, terwyl ander wil hê dit moet ’n sinchronies-deskriptiewe beeld van hedendaagse Afrikaans weergee. Sulke uiteenlopende verwagtinge is nie vreemd ten aansien van ’n normatiewe bron soos die AWS nie.

Hierdie verskillende verwagtinge lei veral tot verwarring oor wat van die TK verwag word om te doen en hoe dit gedoen moet word. Dié kwessie kan gedeeltelik uitgeklaar word deur ’n duidelike en eenduidige karakterisering van die aard van die AWS te gee.

Aangesien die AWS primêr uit twee dele bestaan (d.i. die woordelys en die spelreëls), het dit ’n distinktiewe karakter, sodat dit nóg as woordeboek, nóg as grammatika/handboek beskou kan word. Dit is eerder ’n hibridiese naslaanbron met eienskappe van sowel woordeboeke as taalkundige grammatikas en selfs ’n ensiklopedie, soos hier onder aangedui sal word (vergelyk die inhoudsopgawe van AWS¹¹).

Gebaseer op die bekende tipologie van Zgusta (1971), kan die AWS (en moontlik ook ander sogenaamde spelling- of ortografiewoordeboeke (Singh 1982)) geklassifiseer word as ’n beperkte, eentalige, sinchroniese, linguisties-ensiklopediese naslaanbron:

- **Beperk:** Die beperking van die AWS lê daarin dat dit
 - 'n beperkte skopus het –
 - in die woordelys word slegs 'n beperkte deel van die volledige Afrikaanse leksikon beskryf, te wete
 - leksikale items in geskrewe, formele Standaardafrikaans
 - waar daar moontlik twyfel kan wees oor die
 - skryfwyse/spelling; en/of
 - status; en
 - in die spelreëls word slegs reëls opgeneem met betrekking tot die ortografiese konvensies van Afrikaans, terwyl sintaktiese, semantiese en pragmatiese aspekte (soos woordvolgorde, betekenisuitbreiding en styl) nie in die AWS bereël word nie – en
 - 'n beperkte doel het (kyk afdeling 3 hier onder).
- **Eentalig:** Die AWS fokus slegs op Afrikaans, sonder om vertalings in ander tale aan te bied. Vir die eerste keer in honderd jaar bevat sommige van die bylae in AWS¹¹ wel 'n Engelse vertaalekwivalent, soos die Engelse landname en name van elemente in die periodieke tabel. Die insluiting van Engelse vertaalekwivalente verander egter nie aan die eentalige aard en doel van die AWS nie; dit kan eerder as ensiklopediese inligting gesien word wat toeligend is tot die Afrikaanse inskrywings (sien hier onder).
- **Sinchronies:** Die AWS fokus slegs op hedendaagse Afrikaans. Die AWS word met gereelde tussenposes hersien en bygewerk om juis veranderinge in die woordeskat en ortografie op te teken.
- **Linguisties-ensiklopedies:** Benewens die alfabetiese woordelys, bevat die AWS ook ensiklopediese inligting, waarvan die spelreëls en die bylae die belangrikste is. Die spelreëls is weliswaar linguistiese inligting, maar dit word op ensiklopediese wyse aangebied (tipies alfabeties gerangskik volgens die aspekte van die ortografie, maar met besprekings, kommentaar, opmerkings, voorbeelde, uitsonderings, en so meer). Hiernaas verskyn daar ook van tyd tot tyd bylae in die AWS, soos 'n afkortingslys (vanaf AWS⁵), lys buitelandse geografiese eiename en hulle afleidings (vanaf AWS⁹), terminologielys (vanaf AWS⁹), riglyne vir die skryf van Omgangsafrikaanse woorde (vanaf AWS⁹), riglyne vir transliterasie (vanaf AWS¹⁰), en so meer. Vanaf AWS¹¹ word nog meer ensiklopediese inligting ingesluit (dikwels met Engelse vertaalekwivalente as toeligende materiaal daarby), soos die lys met elemente in die periodieke tabel; 'n lys met plekname in die Ooste; 'n lys met landname, hulle geldeenhede en ISO-kodes; en gebruiksiglyne vir die SI-stelsel. Die ensiklopediese aard van sommige van hierdie inligting blyk onder andere uit die organisering daarvan: Die bylae met plekname in die Ooste is byvoorbeeld nie alfabeties gerangskik soos wat 'n mens in 'n tradisionele (linguistiese) woordeboek sou verwag nie, maar volgens land, provinsie/streek, ensovoorts. 'n Gebruiker sou ook AWS¹¹ kon gebruik om te sien watter lande in 2017 deur die Verenigde Nasies erken is, of wat die geldeenheid van 'n bepaalde land is. In 'n tradisionele linguistiese woordeboek sal sulke inligting nie tematies aangebied word nie, maar eerder alfabeties.

Die AWS toon verder ook bepaalde ooreenkomste met standaard verklarende woordeboeke/handwoordeboeke. Opsommend kan die volgende karakteristieke van sodanige woordeboeke wat op die AWS (in die geheel) van toepassing is, genoem word (na aanleiding van Gouws 1989:69-70):

- Vanweë die gerigtheid van sodanige woordeboeke om ’n weergawe van die standaardtaal te gee, het dit ’n duidelike **normatiewe karakter**.
- As normatiewe naslaanbron word dit as ’n **gesaghebbende bron** beskou (kyk hier onder).
- **Betekenisverklarings en lektiese variasie** (byvoorbeeld sosiolekte en dialekte, asook vak- en ander kringtale) val buite die AWS se regstreekse reikwydte (en wanneer uitsonderings hierop gemaak word, geld gebruiksfrekwensie as kriterium, en word die lemma dikwels van ’n etiket voorsien).
- Die AWS is op **die hede en die toekoms** gerig en gee dus min aandag aan etimologiese en historiese inligting of aan verouderde taalgebruik.
- Dit het primêr ’n **preskriptiewe funksie** omdat dit nie net ’n aanduiding gee van hoe die taal in die hede gebruik word nie (die deskriptiewe funksie), maar ook hoe dit na verwagting nog vir ’n tyd lank gebruik sal word.

Laasgenoemde punt is veral van kernbelang vir die AWS, aangesien dit saamgestel word deur die TK, as geïnstusionaliseerde liggaam wat die ortografie en skryfkonvensies van Afrikaans bepaal. Ofskoon die status van die TK en AWS nooit direk deur middel van wetgewing geïnstusionaliseer is nie (soos byvoorbeeld die geval in Nederland deur middel van die Spellingwet (Nederland 2005) is), het ’n formele proklamasie van die Sekretaris van Onderwys op 15 September 1932 bepaal dat die *Woordeboek van die Afrikaanse Taal* (WAT) die Taalkommissiespelling moet aanvaar. Hierdie gesag van die TK oor die spelling en skryfwyse van Afrikaans is in 1933 verder bevestig toe die toenmalige minister van Binnelandse Sake, JH Hofmeyr, opdrag aan die staatsvertaler gegee het om die TK-spelling in alle staatspublikasies te volg (Van Rensburg 2017). In die Wet op die Suid-Afrikaanse Akademie vir Wetenskap en Kuns (54 van 1959) is ook bepaal dat, ten einde sy oogmerke te bereik, die Suid-Afrikaanse Akademie vir Wetenskap en Kuns (SAAWK) “norme [kan] bepaal ten bate van die gebruik en ontwikkeling van die Afrikaanse taal in al sy funksies”.

As sodanig was dit nog altyd die TK se eksplisiete opdrag om die spellingkonvensies van Afrikaans te bepaal en te bereël. In ’n reglement van die SAAWK (hersiene weergawe goedgekeur in September 2015) word die mandaat van die TK bepaal:

Die Taalkommissie is die liggaam van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns wat hom, onderworpe aan die gesag van die Akademieraad, besig hou met die voortgesette standaardisering en beskrywing van Standaardafrikaans as gebruiksvariëteit, met inbegrip van die spelling, skryfwyse, puntuasie en grammatika daarvan, en van die variëteite van verskillende gebruikersgroepe. Hierdie werksaamhede sluit advies en voorligting oor Standaardafrikaans aan die Akademie en die breë publiek in.

Gegewe hierdie normatiewe opdrag van die TK en die preskriptiewe aard van die AWS, en gegewe die dinamiek van taal en taalgebruikers, is dit nie verbasend dat daar dikwels ’n spanning tussen (ingeligte) taalgebruikers (die “skrywende volk”, byvoorbeeld subredakteurs, terminoloë, leksikograwe, vertalers, ensovoorts) se heersende praktyk en die AWS se norme/reëls kan ontstaan nie. Ten einde hierdie spanning op te los, is dit dus noodsaaklik dat sowel die TK as eindgebruikers ’n ooreenstemmende siening moet hê van wat die doel van die AWS is.

3. DIE DOEL VAN DIE AWS

Die doel van die AWS is tweërlei van aard, naamlik om:

- (1) (preskriptiewe) norme daar te stel vir die gebruik van hedendaagse, geskrewe, formele Standaardafrikaans; en
- (2) as ekspertnaslaanbron oor skryfnorme vir gebruikers van Afrikaans te dien.

Die AWS is die resultaat van die TK se normeringswerk in Afrikaans oor meer as 'n eeu en word beskou as “dié normatiewe gesagsbron oor Afrikaanse spelling” (Carstens 2011:146). Ten spyte hiervan is daar soms steeds kritiek dat die AWS nie normatief genoeg is nie: “Ik vind dat de ‘Taalkommissie’ niet prescriptief of normatief genoeg is geweest. Er blijft veel te veel optioneel; er blijft ook veel te veel variatie toegestaan” (Van Sterkenburg 2004:430, oor AWS⁹).

Oor die verwagtinge van eindgebruikers merk Heid et al. (2004) op: “The most important criterion is the user perspective: the intended use and usergroup of a dictionary, the prior knowledge and the needs of the users.” Drie aspekte moet hiervolgens in gedagte gehou word:

- **Gebruikersgroep:** By gebrek aan gebruikersgebaseerde navorsing oor die AWS, word dit in hierdie artikel aanvaar dat die primêre gebruikersgroep volwasse moedertaalsprekers van Afrikaans is wat versorgde, geskrewe Afrikaans/ Standaardafrikaans wil gebruik (meestal binne 'n professionele omgewing).
 - **Voorkennis van gebruikersgroep:** Daar word in hierdie artikel vooronderstel dat die primêre gebruikers van die AWS 'n goeie (maar nie gevorderde nie) kennis van geskrewe Afrikaans, die Afrikaanse spelsisteen en skryfkonvensies het. Die vooronderstelling hier is dat gebruikers wat nie 'n goeie kennis van Afrikaanse skryfwyses het nie, meestal nie bewus van bepaalde spelreëls of skryfkonvensies is nie, of dat hulle hulle dikwels tot ander, meer algemene bronne soos die *Handwoordeboek van die Afrikaanse Taal* (HAT), of slegs tot die woordelysgeedeelte van die AWS wend.
 - **Behoeftes van gebruikersgroep:** Die aanname word hier gemaak dat gebruikers se eerste toegang tot die AWS die woordelysgeedeelte is (eerder as om die reëlgeedeelte te raadpleeg). Indien die opgesoekte leksikale item nie gevind word nie, sal eindgebruikers waarskynlik eers probeer om “eie reëls” af te lei uit omstaande of soortgelyke voorbeelde, alvorens hulle hulle tot die reëlgeedeelte wend. Myns insiens het gebruikers van die AWS meestal een van die volgende primêre behoeftes:
 - Leiding met betrekking tot spelling/skryfwyse (*Hoe skryf/spel ek X?*)
 - Leiding met betrekking tot erkenning/status van 'n leksikale item (*Is X 'n erkende Standaardafrikaanse woord?*) (Byvoorbeeld: Is “app” 'n erkende Standaardafrikaanse woord?)
 - Leiding met betrekking tot fleksie of afleiding van, of samestellings met 'n leksikale item (*Wat is X se Y-vorm?*) (Byvoorbeeld: Wat is “slu” se oortreffende trap?, of: Is dit “werkswinkel” of “werkwinkel”?)
- Sekondêre behoeftes sou die volgende kon insluit:
- Leiding met betrekking tot spelreëls (*Wat is die reël waarom X so geskryf word?*)
 - Leiding met betrekking tot die Afrikaanse grammatika (byvoorbeeld: *Wat is 'n massanaamwoord? Hoe tree massanaamwoorde in Afrikaans op?*)

Soos blyk uit die feit dat bostaande bloot vooronderstellings is, moet al hierdie aannames nog in die toekoms deeglik empiries nagevors word. Dit is die taak/verantwoordelikheid van die TK om in sy normeringswerk te sorg dat die behoeftes van die gebruikersgroep deeglik in ag geneem word en hanteer word op 'n vlak wat vir die eindgebruikers aanvaarbaar is.

Met betrekking tot normering (d.i. standaardisering) sover dit die ortografie van 'n taal betref, kan die volgende drie algemene uitgangspunte gestel word:

- **Normering moet gebruikergerig wees:** Gebruikers wil antwoorde hê op wat reg en wat verkeerd is (Carstens 2011:17). Daarom moet 'n woordeboek soos die AWS klinkklaar en ondubbelsinnig antwoorde aan gebruikers verskaf.
- **Normering moet kontroleerbaar wees:** By gebrek aan empiriese ondersteuning word die status van norme tans op basis van die oordele van taalkundiges en taalpraktisyns bepaal. Dit kan egter nie willekeurig bepaal word nie (Carstens 2011:18), maar moet berus op “objektiewe waarneming” deur “teoreties geskoolde taalkundiges” (volgens De Villiers 1977:2) en ervare taalpraktisyns (myns insiens) wat bevoeg is om tussen die norm en afwykings van die norm te onderskei. Voortspruitend moet die normeerder (d.i. die TK) deeglik oor die status van die bron(ne) van norme besin (byvoorbeeld korpusse, ander woordeboeke, of taalkundige navorsing) en waaraan dit gemeet word.
- **Normering moet sistematies wees:** Alle norme moet aan die taalsisteem voldoen, met inagneming van (a) hoe taal (in die algemeen) teoreties inmeekaarsteek, (b) hoe die spesifieke taal prakties gebruik word, en (c) wat die status van relevante taaldata is (Carstens 2011:18). Dit impliseer ook dat die TK kennis moet neem van moontlike veranderinge wat in die taalsisteem kan intree en op verantwoordbare wyse leiding gee oor sodanige verandering.

Hierdie drie karakteristieke van normering mag voor die hand liggend wees, maar Carstens (2011:19-20) wys tereg daarop dat dit egter in die praktyk die moeilike vraag oproep: *Wanneer en hoe kry deskripsie 'n preskriptiewe funksie?* 'n Liggaam soos die TK is veronderstel om 'n sinchroniese beskrywing te verskaf van die variëteit van Afrikaans wat algemeen in formele, geskrewe kontekste gebruik word (oftewel Standaardafrikaans), maar gebruikers verwag egter van die TK om preskriptief op te tree. Dit lei oteenseglik tot 'n metodologiese spanning waar daar 'n balans tussen beskrywing en voorskrywing gevind moet word. In die “Werkwyse”-afdeling in AWS¹¹ word dit soos volg verwoord:

Dit is dus altyd vir die Taalkommissie noodsaaklik om 'n goue middeweg te vind tussen enersyds voorskrywing (soos wat taalgebruikers dikwels van hom verwag), en andersyds beskrywing (soos wat die taalkundige dikwels van hom verwag). Daar moet dus 'n ewewig gehandhaaf word tussen wat vir die gemiddelde ingeligte gebruiker verstaanbaar en aanneemlik is, en wat vir die taalkundige aanneemlik en wetenskaplik verantwoordbaar is.

In die samestelling van 'n ortografiewoordeboek (soos die AWS) is dit myns insiens ononderhandelbaar dat besluite en uitsprake wetenskaplik begrond word deurdat dit ten minste **taalkundig verantwoordbaar** moet wees (kyk uitspraak van De Villiers (1977) hier bo). Al moet die TK dus ingevolge sy mandaat en gebruikersverwagtinge preskriptief optree, moet sodanige voorskrywings/normering op goeie wetenskaplike gronde fundeer word (wat objektiewe waarneming (byvoorbeeld van korpusdata) en subjektiewe intuitiewe oordeel gebaseer op ervaring insluit – vergelyk Tarp en Fiertes-Olivera (2016:275)). Dit moet gedoen word binne 'n raamwerk

van 'n stel riglyne wat die sistematiek en kontroleerbaarheid van sy werk verhoog. Alle besluite moet verdedigbaar wees en nie bloot idiosinkraties (of, erger, ideologies) wees nie.

In die res van hierdie artikel word slegs gefokus op die woordelysgedeelte van die AWS. Daar word eerstens spesifiek 'n stel riglyne nagevors aan die hand waarvan inskrywings in die woordelys makrostruktureel opgeneem of weggelaat kan word. Tweedens word 'n beskrywing gegee van die mikrostrukturele data wat oor opgeneemde lemmas verskaf word.

4. DIE OMVANG VAN DIE *AFRIKAANSE WOORDELYS EN SPELREËLS*: OPNAME- EN ELIMINERINGSKRITERIA

Gouws en Prinsloo (2005:65) stel dit duidelik dat “[i]t is important that the compilation of any dictionary must be dominated by well-defined principles regarding the collection and selection of lexical items to be included as lemmata.” De Schryver en Prinsloo (2001:375) sluit hierby aan: “One could say that setting up a dictionary’s lemma-sign list is the first major problem with which any lexicographer is confronted”, soos ook bevestig word deur Tomaszczyk (1983:51): “One of the basic problems of lexicography is to decide what to put in the dictionary and what to exclude.”

In die leksikografie word 'n onderskeid tussen die makro- en mikrostruktuur van 'n woordeboek getref: “Die makrostruktuur verwys ... na die lemmakeuse terwyl die mikrostruktuur verwys na die interne struktuur van elke artikel”, d.i. die tipes inligting wat by elke lemma aangebied word, soos grammatiese inligting, uitspraakleiding, betekenisverklaring, voorbeeldsinne, ensovoorts (Gouws 1989:40). Gouws (1989:75) wys daarop dat lemmas 'n dubbele rol as woordeboekinskrywings vervul, aangesien hulle as makro- én mikrostrukturelemente optree:

As makrostrukturelemente verteenwoordig die lemmas daardie versameling woordeskatelemente wat leksikografies bewerk moet word en funksioneer elke lemma as die kop van 'n woordeboekartikel. As mikrostrukturelement vorm elke lemma 'n onderdeel van 'n woordeboekartikel. Wat dié optrede betref, is 'n lemma leksikografies gelykwaardig aan die ander inskrywings van die betrokke artikel.

Dit is des te meer en by uitstek waar vir lemmas in 'n ortografiewoordeboek soos die AWS. In byvoorbeeld 'n vertalende woordeboek is 'n gebruiker primêr geïnteresseerd in die mikrostruktuur van die woordeboekartikel: H/sy wil 'n vertaalekwivalent/die vertaalekwivalente van 'n opgeneemde lemma vind; die lemma is dus eintlik die sleutel wat toegang bied tot die inligting waarna die gebruiker eintlik soek. In 'n ortografiewoordeboek bevat die lemma sêlf die inligting waarna die gebruiker primêr soek, naamlik hoe om 'n leksikale item te spel/skryf, of wat die status van 'n bepaalde leksikale item in Standaardafrikaans is; die lemma is dus die sleutel én die inligting. Die vraag oor **wat** opgeneem moet word, is dus gelyktydig ook die vraag **hoe** dit opgeneem moet word.

In elk van die elf uitgawes van die AWS (1917; 1918; 1921; 1931; 1937; 1953; 1964; 1991; 2002; 2009; 2017) gee elk van die TK's op hul eie direkte of indirekte manier rekenskap van die werkwyse wat in die samestelling van die betrokke uitgawe se woordelys gevolg is. Vergelyk byvoorbeeld die werkwyse wat gestel word in die voorwoord van AWS¹ (1917:IV; sien die aanhaling in afdeling 1 hier bo), of in die afdeling “Werkwyse” in AWS¹¹ (2017) waar gestel word:

... dat die woorde wat in die Woordelys opgeneem word aan soveel as moontlik van die volgende maatstawwe moet voldoen:

- (1) Dit moet taalkundig korrek wees in ooreenstemming met die grondbeginsels van die Afrikaanse spelling- en skryfsisteem.
- (2) Dit moet met 'n relatief hoë frekwensie in 'n verskeidenheid hoëfunksiepublikasies voorkom.
- (3) Dit moet opgeneem wees in ander gesaghebbende woordeboeke of taalkundige bronne.
- (4) Dit moet moontlik spelling- of skryfprobleme vir gebruikers inhou.
- (5) Die erkenning daarvan as deel van Standaardafrikaans is nodig, veral ook met betrekking tot woorde of frases wat unieke benoemers van (nuwe) konsepte is.
- (6) Opname is wenslik ter wille van paradigmatische volledigheid.
- (7) Dit moet hedendaagse Afrikaans wees of waarskynlik kan voorkom in hoëfunksiepublikasies wat (nog) deur die huidige geslag Afrikaanssprekendes gelees en gebruik word.
- (8) Dit moet, in die lig van die gebruik van standaardtaal vir hoër funksies, die geskrewe gebruik deur die Afrikaanse taalgemeenskap in sulke funksies weerspieël.
- (9) Dit moet sosiaal, kultureel en politiek sensitief wees.

Hierdie nege maatstawwe kan breedweg in twee interafhanklike kategorieë verdeel word, naamlik kriteria wat:

- (a) die makrostruktuur van die AWS bepaal (d.i. of 'n leksikale item enigsins oorweeg moet word vir opname/skrapping); en
- (b) betrekking het op die lemma as mikrostrukturele element in die AWS (d.i. hoe 'n leksikale item opgeneem moet word).

4.1 Makrostrukturele kriteria

Die sentrale vraag met betrekking tot die makrostrukturele kriteria is: *Moet die leksikale item vir opname (of skrapping) oorweeg word?* In dié verband is die oorkoepelende kriterium **bruikbaarheid**, wat deur Gove (1961:4a) as 'n belangrike seleksiekriterium geïdentifiseer word: "Selection is guided by usefulness, and usefulness is determined by the degree to which terms most likely to be looked for are included." Teen die agtergrond van veronderstelde gebruikersverwagtinge moet die TK dus die leksikale items en reëls opneem waarna die veronderstelde eindgebruiker waarskynlik sal soek as hy/sy die AWS raadpleeg. Die volgende kriteria geld (myns insiens in volgorde van belangrikheid vir die veronderstelde gebruiker):

- **Spelling- en skryfprobleme** (kyk punt (4) hier bo): 'n Hoëfrekwensiewoord soos *sit* sal minder geredelik opgeneem word as 'n woord soos *masjinis*, gewoon omdat laasgenoemde waarskynliker deur 'n gebruiker opgesoek sal word.
- As elimineringskriterium kan **deursigtigheid** gebruik word om die woordelysgedeelte te verklein: Indien 'n mens uit die leksikale item se onderdele kan aflei hoe dit gespel of geskryf moet word, hoef dit nie noodwendig opgeneem te word nie. Hier gaan dit oor die verhouding tussen reëls en lyste: As 'n leksikale item genoegsaam deur die reëls ondervang word, is dit minder noodsaaklik om die leksikale item op te neem. Soos die samestellers van die Nederlandse *Het Groene Boekje* tereg opmerk: "Daarom werden veel samenstellingen geschrapt waarvan de spelling blijkt uit gelijksoortige gevallen" (Instituut voor Nederlandse Lexicologie 2005).
- **Status** (kyk punt (5) hier bo): Gebaseer op informele terugvoer van eindgebruikers (soos studente), asook op eie ervaring, weet ons dat die woordelys dikwels gebruik

word om te bepaal of 'n bepaalde leksikale item 'n “erkende woord in Afrikaans” is. Iemand wat byvoorbeeld wil weet of *karnivoor* (naas *vleiseter*) 'n erkende Afrikaanse woord is, sou die woord in die woordelys kon naslaan. Hierdie kriterium geld dikwels vakterme, rekenaarterme, nuutskappings, unieke benoemers (byvoorbeeld die werkwoorde *netwerk* en *teks*; of *fyngoed* vir bokke en skape, en *grofgoed* vir grootvee en trekdier), en so meer. Dit kan ook gebruik word as motivering vir die opname van leksikale items wat dalk meer gewestelik of van spesifieke kulturele belang is (kyk punt (9) hier bo), maar wat dalk tog 'n plek in die breër Afrikaanse gemeenskap sou kon kry, byvoorbeeld *brötchen* (“soort broodjie”; word meestal in Namibië gebruik), *ghoemahare* (“spookasem”; word meer algemeen in die Wes-Kaap gebruik), of *motjie* (“bejaarde vrou”; word meestal in die Moesliemgemeenskap gebruik). Uiteraard moet 'n liggaam soos die TK uiters omsigtig te werk gaan om byvoorbeeld nie net op een variëteit te fokus nie, of om byvoorbeeld nie hipermoralisties of polities op te tree nie.

As elimineringskriterium dien hierdie kriterium ook om verouderde leksikale items, of leksikale items wat sosiaal, kultureel of politiek onsensitief is uit die AWS te skrap (kyk punt (9) hier bo). Voorbeelde van verouderde leksikale items wat in AWS¹⁰ voorgekom het, maar nie meer in AWS¹¹ nie, sluit in *fransbrandewyn*, *hoegrootheid*, *juts* en *KGB-agent*, en voorbeelde van onsensitiewe leksikale items sluit in *idiot-savant*, *Kakie*, *mongool* en *bruinman/swartman/witman*.

Om die status van 'n leksikale item in hedendaagse Standaardafrikaans te bepaal, word veral twee subkriteria in ag geneem:

- **Spreiding** (kyk punt (7) en (8) hier bo): In afdeling 2 is genoem dat die skopus van die AWS in wese beperk is tot geskrewe, formele Standaardafrikaans. Leksikale items behorende tot beperkte sosiale, geografiese of vakkundige groepe/domeine behoort nie in die **woordelys** opgeneem te word nie, behalwe waar sodanige leksikale items ook algemeen in die standaardvariëteit gebruik word (byvoorbeeld in koerante), of waar erkenning van unieke benoemers nodig is. Daarteenoor behoort die **spelreëls** wel eksplisiet voorsiening te maak vir die spelling en skryfwyses van Standaardafrikaans in gespesialiseerde kontekste, byvoorbeeld in vaktaal. Die skrywer, teksredakteur of vertaler van 'n gespesialiseerde teks moet dus steeds die AWS as rigtinggewende bron kan gebruik, selfs al word spesifieke terme en uitdrukkings nie in die woordelys opgeneem nie.
- **Frekwensie** (kyk punt (2) hier bo): Ten einde patrone te bepaal vir die formulering van spelreëls, of om op die spelling of skryfwyse van 'n bepaalde leksikale item te besluit, kan korpusfrekwensie as 'n kriterium in samehang met ander kriteria in 'n spesialiswoordeboek soos die AWS gebruik word: “In practice, corpus frequency is only one of several selection criteria” (Heid et al. 2004). Die gebruik van frekwensie as kriterium moet egter verantwoordelik hanteer word: Die feit dat 'n bepaalde vorm frekvent voorkom, waarborg nie noodwendig die taalkundige korrektheid daarvan nie (vergelyk byvoorbeeld Reynaert (2009) se navorsing oor die hoë frekwensie van spelfoute in korpusse). Die belangrikste vraag hier is: *Wanneer dikteer frekwensie van gebruik die norm?* Om die balans te vind tussen beskrywing en voorskrywing, moet daar op verantwoordbare en (sover moontlik) wetenskaplike wyse met korpusdata omgegaan word, al is dit nie altyd in praktyk volkome moontlik nie (weens, byvoorbeeld, gebrek aan hulpbronne en tyd).

Twee aspekte is in dié verband van belang. In die eerste plek is dit noodsaaklik om te bepaal wat as korpus(se) dien. Gouws en Prinsloo (2005: 65) stel die ideaal: “The lemmata have to be drawn from a representative corpus of the specific language and a vital part of the planning of any lexicographic endeavour is the development of a corpus.” Vir ’n hulpbronskaars taal soos Afrikaans is dit makliker gesê as gedaan, veral aangesien uitgewers en mediahuise dikwels huiwerig is om hulle data as (kommersiële) korpusse beskikbaar te stel. Dit word dus moeilik om ’n **betroubare, verteenwoordigende en gebalanseerde korpus** as basis vir die bepaling van frekwensietellings saam te stel. In dié verband het die TK in 2005 begin om ’n verwysingskorpus vir leksikografiese doeleindes (d.i. “... any collection of texts where lexicographers can find inspiration for completing the dictionary structures they need when making a real dictionary” (Fuertes-Olivera 2012:51) saam te stel. Die korpus (Taalkommissie 2011) bestaan uit ongeveer 57 miljoen woorde, waarvan nagenoeg 27 miljoen woorde uit niefiksie-, nieakademiese tekste soos koerante, tydskrifte, godsdienstige geskrifte, notules en ander boeke kom; ongeveer 6 miljoen woorde uit fiksie kom; en 24 miljoen woorde uit akademiese tekste soos artikels, studiegidse en proefskrifte kom (kyk TABEL 1 vir die gedetailleerde stratum). Ander korpusse van (veronderstelde) geredigeerde teks wat op die webwerf van die Virtuele Instituut vir Afrikaans (VivA) (2016) beskikbaar is, word ook geraadpleeg, tesame met soektogte in die WAT as korpus, die Afrikaanse korpus van Projekt Deutscher Wortschatz (2016; voortaan Leipzigkorpus) en ander soortgelyke hulpbronne. Desalnietemin is dié versamelings tekste dikwels steeds te klein om klinkklare afleidings of statisties betroubare gevolgtrekkings te maak, en dan word daar van Googlesoektogte gebruikgemaak om tendense buite beskikbare korpusse na te speur. Só ’n gebruik van Google (of ander soekenjins) word deur Fletcher (2005) “web hunting” genoem, d.i. om websoektogte te gebruik om seldsame voorkomste van konstruksies (wat nie noodwendig in standaardkorpusse voorkom nie) te soek. Van Huyssteen (2016) toon byvoorbeeld aan dat slegs die ortografiese vorme *Ou Testamenties* en *Ou-Testamenties* in die korpusse op VivA se Korpusportaal (2016) aangetref word, terwyl ’n allesomvattende Googlesoektog nog nege ander variante aan die lig bring. Aangesien beskikbare korpusse te klein is of neig om alleen geredigeerde teks te bevat, kan soekenjins gebruik word om ’n vollediger prentjie van die taalwerklikheid te probeer kry. Dit moet egter met groot omsigtigheid gedoen word, soos Fletcher (2007) waarsku: “The query, search and ranking optimization techniques [search engines] have adopted can either assist or sabotage a scholar’s quest ...”.

Die ander aspek wat van belang is, is dat die **beginsels en metodes** wat aangewend word om korpusse te ontgin, ook neergelê moet word. Dit sluit onder andere die volgende aspekte in:

- Word frekwensietellings in vergelykbare korpusse vir **alle leksikale items** gemaak, of net vir twyfelagtige gevalle? En hoe word **nuwe leksikale items en/of patrone** sistematies gevind (d.i. ten einde die insidentele opname van toevallighede, soos eendagsvlieë, teen te werk)? Die woordelys in die AWS is die organiese resultaat van meer as ’n eeu se werk deur verskillende TK’s. Aan die beginjare was korpuslinguistiese metodes nog nie beskikbaar nie, en leksikale items is slegs na die oordeel van TK-lede bygevoeg. So byvoorbeeld verskyn *kernwapenbeperkingskonferensie* vir die eerste keer in AWS⁸, waarskynlik om rekursiwiteit in komposita te illustreer. Dié woord kom egter nêrens anders as in AWS⁸ en AWS⁹ voor nie – nie in ander woordeboeke nie, en ook nie in enige korpusse of Googlesoektogte nie. Die woordelys van die AWS is nog nooit in sy geheel onderwerp aan ’n sistematiese frekwensieondersoek nie. So ’n ondersoek sou esoteriese gevalle soos *kernwapenbeperkingskonferensie* kon uitskakel.

TABEL 1: Stratum van TK-korpus (2011)

Vlak	Kat.	Genre	Domein	TK 1.1	
Romans	Prosa	Romans		5 168 392	
		Kortverhale		684 559	
Nie-fiksie	Akademies	Proefskrif/Verhandeling Artikel	Geesteswetenskappe	240 476	
			Tydskrif vir Geesteswetenskappe	186 634	
			In die Skriflig	1 095 273	
			Koers	588 452	
			Woord en Daad	265 139	
			Literator	737 450	
			Tydskrif vir Natuurwetenskappe en Tegnologie	172 355	
		Studiegids	Verskeie wetenskappe 1	2 628 075	
			Verskeie wetenskappe 2	4 115 694	
			Verskeie wetenskappe 3	5 723 929	
		Verskeie wetenskappe 4	1 957 055		
	Nie-akademies	Koerante	Noord	4 825 742	
			Suid	6 343 360	
			Landswyd	5 186 000	
		Tydskrifte	Algemeen	158 419	
			Vrouetydskrif	195 005	
		Boeke	Algemeen	2 866 041	
			Vroue	221 053	
			Leefstyl	182 405	
			Sport & Ontspanning	195 357	
			Tegnologie	241 261	
			Finansies	200 085	
			Kuns en Kultuur	638 151	
			Ander	1 632 474	
			Godsdienstige tekste	Primêr: Christelik	837 788
			Notules	Amptelik	34 720
					47 321 344

- **Hoeveel keer** moet 'n leksikale item of spellingvorm voorkom alvorens dit opgeneem word? In dié verband sê Lehmann (2006): “A frequency threshold may be defined for a dictionary. Typically, hapax legomena are thrown out.” Daar is tot nog toe geen vaste, formele frekwensiedrempelwaarde vir die AWS bepaal nie, omdat die AWS juis ook voorsiening maak vir hapaxlegomena wat spel- of skryfprobleme sou kon oplewer, of waarvan die status erken wil word (en omdat die korpusse waarmee gewerk word, gewoon te klein is om frekwensie as kriterium te verabsoluteer). In die samestelling van die woordelys word daar dus eerder gewerk met praktiese beduidendheid (d.i. of'n verskil groot genoeg is om van praktiese waarde in toegepaste

omgewings te wees) as statistiese beduidendheid (d.i. wiskundige waarskynlikheid), wat dikwels bepaal word deur die vraag te beantwoord: *Kom hierdie spel-/skryfvorm genoeg kere in betroubare bronne voor om erkenning daaraan te verleen?* (waar beide “genoen kere” en “betroubare bronne” relatief en subjektief is). Binne die TK is nog geen metodologiese beleid in dié verband geformaliseer nie, en dit is sonder twyfel ’n saak wat in die toekoms verder nagevors en ondersoek moet word.

In hierdie artikel (en in Van Huyssteen 2017) stel ek voor dat die frekwensieklas (Perkuhn et al. 2012) van ’n bepaalde ortografiese vorm ’n nuttige maatstaf vir opname of eliminerings kan wees. Die frekwensieklas N (na onder afgerond) word bepaal deur:

$$N = \left\lfloor 0.5 - \log_2 \left(\frac{\text{frekwensie van woord}}{\text{frekwensie van frekwentste woord}} \right) \right\rfloor$$

Aangesien met ortografiese vorme gewerk word, word die frekwensie van die geskrewe vorm (en nie gelemmatiseerde vorme nie) gebruik soos dit in die korpus voorkom. Dusdanig word byvoorbeeld *die* en *Die*, of *soek* en *gesoek* as aparte woord(vorm)e hanteer. Die frekwentste ortografiese vorm (*die* = 3 198 933 in die TK-korpus) se frekwensieklas is 0, terwyl die tweede frekwentste vorm (*van* = 1 393 371) in frekwensieklas 1 is. Die frekwensieklasse vertoon ’n tipiese Zipfiaanse verspreiding, waar ’n hapaxlegomenon soos *dramakritikus* in die TK-korpus in frekwensieklas 22 is. Deur sistematiese ondersoek sou die Taalkommissie ’n drempel vir die frekwensieklas kon vasstel, byvoorbeeld dat slegs woorde in frekwensieklasse 4 tot 18 vir opname oorweeg word.

- **Paradigmavoltooiing** (kyk punt (6) hier bo): ’n Spesifieke eienskap van die AWS is dat by sommige groepe leksikale items afleidings van of samestellings met die basiswoord opgeneem word ten einde die woordvormingsparadigma van die basiswoord te illustreer. Dit geld onder andere taalbenamings (byvoorbeeld *Sanskrit*; *Sanskritis* (*b.nw.*); *Sanskritis*), religieuse strominge (byvoorbeeld *Hindoeïsme*; *Hindoe* of *hindoe*; *Hindoeïsties* of *hindoeïsties*; *Hindoeokaste*) en leksikale items van klassieke herkoms (byvoorbeeld *reglement*; *reglementeer*; *reglementêr*).
- **Illustrasie**: ’n Ander AWS-spesifieke eienskap is dat leksikale items wat aan geen van bostaande kriteria voldoen nie, maar wat ter illustrasie van ’n spesifieke spelreël ingesluit word, soms in die woordelys opgeneem word. ’n Woord soos *kelder* lewer byvoorbeeld nie spellingprobleme op nie en oor die status daarvan in Standaardafrikaans bestaan geen twyfel nie, maar aangesien dit ’n voorbeeld is by reël 11.4 (AWS¹¹), word dit tog ingesluit in die woordelys.
- **Ruimtelike beperking**: Aangesien ’n uitgewer bepaalde beperkinge op die omvang van ’n boek kan lê (byvoorbeeld die aantal bladsye wat nie oorskry mag word nie ten einde die boek steeds teen ’n bekostigbare prys te verkoop), moet die TK “...constantly ... make pragmatic decisions on what to include in a dictionary to conform to the dictates of space available” (Walter 1996:640). In dié verband merk Martin et al. (1983: 81-82) ook op: “The decision what to include in the dictionary still has to be made by the lexicographer himself, however, and this depends in turn upon the nature and size of the dictionary and its intended users.” Sodanige besluite kan egter nie lukraak geneem word nie, maar moet liefies aan die einde van ’n hersieningsproses geskied. Hier kan allerlei outomatiese metodes gebruik word om finale keuses te

maak, soos byvoorbeeld om na die gebalanseerdheid van verskillende letters te kyk (Prinsloo & De Schryver 2003), die sistematiese hantering van opgeneemde spesialisterme, paradigmas, ensovoorts.

Die woordelysgedeelte van die AWS is sedert 2010 gratis aanlyn beskikbaar by www.woordelys.co.za. Hiermee het die AWS die digitale era betree, en ruimtelike beperkings (soos tradisioneel in 'n boekformaat) is nie meer 'n hoofoorweging nie. AWS¹ is ook in sy geheel digitaal beskikbaar op die uitgewer se woordeboekplatform, en dié beperking word des te minder belangrik.

4.2 Mikrostrukturele kriteria

Mikrostrukturele oorwegings is dié wat betrekking het op/as motivering verskaf word vir die keuse en insluiting van kategorieë grammatikale inligting oor elke opgeneemde lemma. Die vraag is dus: *Hoe moet die leksikale item opgeneem word?*, oftewel *Watter ortografiese inligting kan aan die lemmateken onttrek word?* Die volgende subkriteria is onderling afhanklik, maar word tog in volgorde van belangrikheid gerangskik:

- **Taalkundige korrektheid/motivering** (kyk punt (1) hier bo): Myns insiens geniet die heersende konvensies van die taalsisteem en grammatika voorrang bo enige ander kriteria. Die taak van die TK is juis om te verseker dat sodanige konvensies sistematies en konsekwent toegepas word in die ortografie/skryfwyse van die taal – daarom ook dat daar altyd ervare taalkundiges en taalpraktisyns in die TK dien. Daar moet 'n passing gemaak word tussen die spelreëls onderling (d.i. reëls moet samehang toon en mekaar nie weerspreek nie) en tussen die spelreëls en die woordelys (d.i. leksikale items wat opgeneem word, moenie die spelreëls weerspreek/oortree nie). Wanneer dit uit die taalkundig wetenskaplike proses (en/of uit onderstaande kriteria) blyk dat daar 'n diskrepansie bestaan tussen die spelreëls en/of woordelys aan die een kant, en die taalwerklikheid aan die ander kant, behoort die TK te oorweeg of dit nodig is om aanpassings aan die spelreëls en/of woordelys te maak. Op dié manier beïnvloed die TK (as standaardiseringsliggaam) dus die konvensies van die taal (d.i. deur die normatiewe funksie van die TK), eerder as om uitsluitlik 'n beskrywing van taalpatrone te gee (d.i. sy deskriptiewe funksie).

Byvoorbeeld, al kom die spelvorm **onmiddelik* (en selfs **onmiddellik* en **onmidelik*) met 'n hoë frekwensie in korpusse voor, is dit die TK se taak om taalkundig te motiveer dat *onmiddellik* die korrekte spellingvorm is, aangesien dit uit die morfeme *on-*, *middel* en *-lik* bestaan. In dié sin “troef” hierdie kriterium dus as ‘t ware die frekwensiekriterium (kyk hier bo).

Ook hierdie kriterium kan egter nie verabsoluteer word nie. Dit gebeur soms dat die spelling van 'n leksikale item die uitspraak in die volksmond navolg, ongeag die taalkundige korrektheid daarvan. Sulke uitsonderings word dan tradisioneel op 'n ander wyse geskryf of gespel as wat deur reëls of taalkundige beginsels voorspel sou kon word. Vergelyk byvoorbeeld die woord *vadoek* “lap om skottelgoed mee te was/af te droog”. Diachronies gesien moet die spelling *vaatdoek* wees, aangesien *vaat* die Nederlandse woord vir “skottelgoed” is. McLachlan (2016:499) toon aan dat die wisselvorme *vadoek* of *vatdoek* van AWS¹ tot AWS³ erken is; van AWS⁴ tot AWS⁷ word *vatdoek* verander na die taalkundig meer korrekte *vaatdoek* (naas *vadoek* as wisselvorm); vanaf AWS⁸ word slegs *vadoek* erken. In beskikbare korpusse van

hedendaagse Afrikaans kom *vaatdoek* nooit voor nie; *vadoek* kom een keer voor in die TK-korpus (Taalkommissie 2011); *vadoek* kom 119 keer voor. Vir die hedendaagse TK om die taalkundige versene teen die talige prikkels te slaan, sou gewoon futiel wees.

’n Belangrike aspek van taalkundige korrektheid het betrekking op die tradisie en ontwikkelingsgeskiedenis van die Afrikaanse ortografie waarteen “korrektheid” gestel word. Dit is ’n feit dat daar aan die beginjare van Afrikaans se normatiewe kodifiseringsgeskiedenis sterk aangeleun is by die Nederlandse tradisie (vergelyk Van Rensburg (2017) vir ’n oorsig en o.a. Kotzé (2014); McLachlan (2016); en Odendaal (2012) vir kritiese beskouings). By die beoordeling van ’n bepaalde ortografiese patroon moet daar m.i. steeds heden ten dage oorweging geskenk word aan die Germaanse (spesifiek Wes-Germaanse en meer spesifiek Dietse) tradisie, omdat Afrikaans steeds vormlik binne dié taalfamilie getipologiseer kan word. As kriterium kan dit egter eweneens nie verabsoluteer word nie, omdat Afrikaans ook sedert sy vroegste ontstaansgeskiedenis ’n eie tradisie ontwikkel het, en aangesien onder andere internasionalisering en globalisering ook bydra tot die skep van ’n nuwe tradisie (byvoorbeeld ten aansien van transliterasie, verafrikaansing (al dan nie) van eiename, en insigte wat deur nuwe taalkundige navorsing verkry word).

- **Institusionalisering** (kyk punt (3) hier bo): Ten einde te bepaal tot watter mate ’n bepaalde spel- of skryfvorm reeds geïnstusionaliseer (algemeen gebruik en aanvaarbaar) is, en aangesien die kollektiewe insig van die taalkunde- en leksikografiegemeenskap ook besluite van die TK moet/kan rig, kan ander gesaghebbende bronne geraadpleeg word. Dit sluit onder andere omvattende woordeboeke (soos die WAT), standaardwoordeboeke (soos die HAT), vakwoordeboeke, (taalkunde)boeke, grammatikas en ander wetenskaplike publikasies in.

Byvoorbeeld, in AWS¹⁰ word aangedui dat *endomorfse* meervoud óf *endomorfe* óf *endomorwe* is. In die WAT (2016) word slegs *endomorfe* as meervoud aangedui, en die TK het dus vir AWS¹¹ slegs dié vorm erken en verder geïnstusionaliseer.

- **Frekwensie** (kyk punt (2) hier bo): Een van die moeiliker besluite wat dikwels deur die TK geneem moet word, is of enige wisselvorm vir ’n bepaalde spel- of skryfvorm aanvaar word. Benewens die gebruik van bogenoemde kriteria, kan die TK hom ook in dié verband deur frekwensietellings in korpusse laat lei. Soos hier bo genoem, bestaan daar in ondersoeke van hierdie aard dikwels ’n spanning tussen statisties beduidende verskille en prakties beduidende verskille. Tot op hede maak die TK eintlik, tereg, alleen staat op prakties beduidende verskille – die enigste vraag wat beantwoord word, is of ’n bepaalde wisselvorm met redelike frekwensie in korpusse voorkom, en of daar ’n groot genoeg waarskynlikheid is dat die wisselvorm(e) naas ’n ander vorm gebruik sal word. Indien wel, en indien dit ook aan van die ander kriteria voldoen (soos taalkundige korrektheid), word dit as ’n variante lemma in die AWS opgeneem. Soos ook reeds genoem, is verdere grondige ondersoek met betrekking tot die gebruik van frekwensies en statistiek in die samestelling van ’n woordeboek soos die AWS nodig.

In hierdie artikel stel ek voor dat daar – veral met betrekking tot wisselvorm – gewerk word binne die statistiese raamwerk wat deur Wallis (2012a; 2012b; 2013a; 2013b) voorgehou word. Binne ’n **keusemodel** word ortografiese variasie gesien as keuseopsies wat vry is om te varieer, d.i. “a genuine choice exists and all cases could theoretically be of one type or the other” (Wallis et al. 2012:4). Die waarskynlikheid

P dat 'n bepaalde vorm n_i uit die totale aantal werklike keuseopsies n_i gekies sal word, is:

$$P(n_1 | \{n_1, n_2, n_3, \dots, n_i\}) = \frac{F(n_1)}{F(n_1, n_2, n_3, \dots, n_i)}$$

Waar $F(n_i)$ die totale aantal gevalle (ongenormaliseerde frekwensie) van die eerste ortografiese variant is, $F(n_j)$ van die tweede variant, ensovoorts (Wallis et al. 2012:4). Die waarskynlikheid P word uitgedruk as 'n waarde tussen 0 (onwaarskynlik) en 1 (volkome waarskynlik). Aangesien daar met waarskynlikhede gewerk word, kan ook met die gemiddeld van waarskynlikhede oor verskillende korpusse heen gewerk word (kyk Van Huyssteen (2016; 2017) vir 'n toepassing van dié model).

Om 'n aanduiding te kry of die waarskynlikheidsverspreiding enigsins ook statisties beduidend is (indien dit nie op die oog af duidelik is nie), maak Wallis (2013a) gebruik van die enkelsteekproef- z -toets. Kortliks kom dit daarop neer dat, gegewe 'n waargenome waarskynlikheid p en 'n bekende, verwagte populasiewaarskynlikheid P (0,5 in die geval waar twee vorme vergelyk word), die enkelsteekproef- z -toets bepaal of die twee waardes statisties beduidend verskil. Die z -waarde word bereken deur:

$$z = \frac{p - P}{S}$$

waar:

$$\text{Populasie se standaardafwyking } S = \sqrt{\frac{P(1 - P)}{S}}$$

en n die totale aantal gevalle in die steekproef is. Daarna word getoets of die absolute waarde $|z|$ minder is as die kritieke waarde van z by 'n bepaalde foutdrempelwaarde (0,05 in Wallis (2013a) se implementasie). Indien $|z| > 1,95996$ is, is die verskil in frekwensie statisties beduidend en kan die twee vorme nie eintlik as wiskundig/statisties waarskynlike wisselvorme gesien word nie. Met dié gegewens in pag, moet dan besluit word of die verskil ook prakties beduidend is.

Byvoorbeeld, in die korpusse in VivA se korpusportaal (VivA 2016) kom *weer eens* 1 820 keer voor en *weereens* 1 199 keer. Die verskil tussen hierdie twee frekwensies is statisties beduidend ($z=11,30212$; $p<0,00$), wat beteken dat dié twee vorme eintlik nie as wisselvorme opgeneem behoort te word nie. As 'n mens egter die waarskynlikheidsverdeling ($\sim 0,6:0,4$) in ag neem, is dit tog dalk van praktiese waarde om *weer eens* en *weereens* as wisselvorme te behou. In Van Huyssteen (2017) word hierdie metode geïllustreer aan die hand van 'n gevallestudie.

- **Woordeboekstyl en redaksionele beleid:** Die stylgids van 'n woordeboek bepaal nie net wat opgeneem word nie, maar ook hoe dit opgeneem word. Oor 'n tydperk van honderd jaar het die opnamestyl in die AWS in wese nie verander nie: 'n Artikel bestaan uit 'n lemma (en wisselvorme), gevolg deur fleksie-inligting, en soms word addisionele inligting (woordsoort, voorbeeld, betekenisonderskeiding of gebruiksdomein) in die vorm van woordeboeketikette aangebring. Vanaf AWS¹⁰ word sommige artikels ook van reëlverwysings voorsien. Soos in afdeling 1 aangetoon is, het die

TK sedert 2006 begin om die redaksionele beleid in 'n interne, huishoudelike stylgids te formaliseer, waarin byvoorbeeld die volgorde van etikette, die volgorde van spel- en skryfvariante en die hantering van morfologiese inligting uitgestippel word. Die voorskryfte in die stylgids is daarom bepalend vir hoe 'n lemma uiteindelik in die AWS opgeneem word.

Bostaande kriteria geld uiteraard nie alle gevalle nie, en kan in die praktyk nie noudeset en slaafs vir elke moontlike inskrywing gevolg word nie (gewoon omdat veral die woordelys organies oor meer as honderd jaar gegroei het). Die eksplisiering hiervan dra egter hopelik by tot 'n konsekwenter hantering van veral kontensieuse gevalle en gevalle waar keuses gemaak en besluite geneem moet word.

5. SLOT

In TABEL 2 word die opname- en elimineringskriteria prakties opgesom en geoperasionaliseer in 'n aantal vrae. Hoe meer ja-antwoorde op die makrostrukturele vrae verkry word vir 'n voorgestelde leksikale item, hoe groter sou die kans wees dat die leksikale item in die AWS opgeneem kan word. Die antwoorde op die mikrostrukturele vrae verskaf leiding oor hoe die inskrywing uiteindelik in die AWS daar sal uitsien. Soos telkens hier bo geargumenteer is, moet geen van die kriteria egter verabsoluteer word nie – 'n holistiese beskouing en benadering is van kardinale belang, soos geïllustreer word in Van Huyssteen (2017).

TABEL 2: Operasionalisering van makro- en mikrostrukturele kriteria

Makrostrukturele kriteria: <i>Moet die leksikale item vir opname (of skrapping) oorweeg word?</i>		
Spel-/skryfprobleme	1. <i>Lewer die leksikale item potensiële spel- of skryfprobleme op?</i>	Ja/Nee
Spel-/skryfprobleme	2. <i>Is die leksikale item se spelling/skryfwyse/morfologiese struktuur ondeursigtig?</i>	Ja/Nee
Status	3. <i>Moet die leksikale item (se status?) in Standaardafrikaans erken word?</i>	Ja/Nee
Status	4. <i>Is die leksikale item (nog) in hedendaagse Standaardafrikaans in gebruik?</i>	Ja/Nee
Status	5. <i>Kan die leksikale item as 'n unieke benoemer beskou word?</i>	Ja/Nee
Status	6. <i>Is die leksikale item van spesifieke sosiokulturele belang?</i>	Ja/Nee
Status: Spreiding	7. <i>Is die leksikale item wyd genoeg bekend, of sal dit wyd gebruik kan word?</i>	Ja/Nee
Status: Frekwensie	8. <i>Kom die leksikale item met prakties beduidende frekwensie in betroubare, verteenwoordigende, gebalanseerde korpusse voor?</i>	Ja/Nee
AWS-spesifiek	9. <i>Moet die leksikale item opgeneem word ter wille van paradigmatolting?</i>	Ja/Nee
AWS-spesifiek	10. <i>Word die leksikale items elders in die AWS as voorbeeld gebruik?</i>	Ja/Nee
AWS-spesifiek	11. <i>Is dit nodig om die leksikale item op te neem, ruimtelike beperkinge ten spyt?</i>	Ja/Nee

TABEL 2: Operasionalisering van makro- en mikrostrukturele kriteria (*vervolg*)

Mikrostrukturele kriteria: Hoe moet die leksikale item opgeneem word?		
Taalkunde	12. <i>Hoe moet die leksikale item gespel-/skryf word ingevolge taalkundige oorwegings (ook ten aansien van die Germaanse taalsisteem)?</i>	Spel-/skryfvorm(e)
Taalkunde	13. <i>Watter spel- en skryfreëls is ter sake ten aansien van die spelling/ skryfwyse van die leksikale item?</i>	Reëlverwysing(s)
Taalkunde	14. <i>Indien die leksikale item 'n uitsondering is wat tradisioneel anders gespel of geskryf word as wat deur die spel- en skryfreëls gedikteer word, op watter spel- en skryfreëls is dit 'n uitsondering?</i>	Reëlverwysing(s)
Institusionalisering	15. <i>Hoe word die leksikale item in ander gesaghebbende woordeboeke of (taalkunde)boeke opgeneem en bewerk?</i>	Spel-/skryfvorm(e)
Frekwensie	16. <i>Watter wisselvorme moet vir die leksikale item opgeneem word (slegs indien daar geen prakties beduidende verskil tussen verskillende vorme is nie)?</i>	Spel-/skryfvorm(e)
AWS-spesifiek	17. <i>Hoe moet die artikel bewerk word om te voldoen aan die AWS se styl- en redaksionele beleid (insluitende fleksie-inligting en gebruikerleiding)?</i>	Opnamevoorstel

Deur die loop van hierdie artikel is talle temas vir verdere navorsing geïdentifiseer.

Navorsingsvrae soos die volgende kan in die toekoms aandag geniet:

- Navorsing oor eindgebruikers:
 - Wie is die primêre gebruikers van die AWS?
 - Word die AWS inderdaad deur die primêre gebruikers as “dié normatiewe gesagsbron oor Afrikaanse spelling” (Carstens 2011:146) beskou?
 - Wat is die voorkennis van die primêre gebruikers?
 - Wat is die primêre en sekondêre behoeftes van hierdie gebruikers?
 - Wat is die primêre gebruikers se toegangstrategie? Is dit hoofsaaklik die woordelys- of die spelreëlgedeelte wat eerste geraadpleeg word?
- Navorsing oor die woordelysgedeelte van die AWS:
 - Hoe gebalanseerd is die AWS se lemmaverspreiding oor letters heen (Prinsloo & De Schryver 2003) in vergelyking met ander Afrikaanse woordeboeke?
 - Tot watter mate is die woordelys van die AWS verteenwoordigend van hedendaagse formele, geskrewe Standaardafrikaans? Byvoorbeeld: Is daar (variante) lemmas wat slegs in die AWS voorkom en nie in werklike taalgebruiksdata of ander woordeboeke nie?
- Navorsing oor die metodologie van die TK:
 - Hoe kan die TK-korpus uitgebrei en verbeter word?
 - Hoe vergelyk verskeie “geslote” korpora (soos die TK- en Leipzigkorpus) met data wat deur Googlesoektogte verkry word? Kan “web hunting” (Fletcher 2005) as data-insamelingstrategie laat vaar word?
 - Kan objektiewe drempelwaardes vir frekwensieklasse bepaal word om as kriterium vir opname en eliminerings te dien?

- Met betrekking tot variante ortografiese vorme, is dit moontlik om objektiewe drempelwaardes vir waarskynlikheidsverspreiding daar te stel? Kan statistiese beduidendheid prakties gebruik word om die aantal wisselvorme in die AWS te beperk (of uit te brei)?

ERKENNING

Hierdie artikel is baseer op 'n aanvanklike werksdokument wat in 2006 deur my en Tom McLachlan vir die Taalkommissie opgestel is. Erkenning word hiermee nie net gegee aan sy insette in daardie dokument nie, maar veral ook aan sy jarelange invloed op my denke oor die aard, doel en omvang van die AWS.

Sedert 2006, toe die aanvanklike werksdokument opgestel is, het verskeie TK-lede persoonlik of tydens TK-sittings kommentaar gelever op die dokument en bygedra tot die verfyning van die werkwyse van die TK. Erkenning word hiermee gegee aan iedere TK-lid se direkte of indirekte invloed op die vorming van die gedagtegang in hierdie artikel.

Erkenning word ook gegee aan die volgende persone vir perspektiewe gestel in persoonlike gesprekke: Wannie Carstens, Hans du Plessis, Rufus Gouws, Danie Prinsloo en Piet Swanepoel.

Alle potensiële denk- en argumentasiefoute bly egter myne en kan nie voor die deur van enige van hierdie individue of die TK gelê word nie.

BIBLIOGRAFIE

- AWS¹: Le Roux, T.H., Malherbe, D.F. & Smith, J.J., in opdrag van die Suidafrikaanse Akademie vir Taal, Lettere en Kuns. 1917. *Afrikaanse woordelys en spelreëls*. 1e uitgawe. Bloemfontein: Het Volksblad-Drukkerij.
- AWS²: Le Roux, T.H., Malherbe, D.F. & Smith, J.J., in opdrag van die Suidafrikaanse Akademie vir Taal, Lettere en Kuns. 1918. *Afrikaanse woordelys en spelreëls*. 2e uitgawe. Bloemfontein: De Nasionale Pers.
- AWS³: Bosman, D.B., Le Roux, T.H., Malherbe, D.F. & Smith, J.J., in opdrag van die Suidafrikaanse Akademie vir Taal, Lettere en Kuns. 1921. *Afrikaanse woordelys en spelreëls*. 3e uitgawe. Bloemfontein: Die Nasionale Pers.
- AWS⁴: Boshoff, S.P.E., Bosman, D.B., Le Roux, T.H., & Malherbe, D.F., in opdrag van die Suid-Afrikaanse Akademie vir Taal, Lettere en Kuns. 1931. *Afrikaanse woordelys en spelreëls*. 4e uitgawe. Bloemfontein, Kaapstad & Stellenbosch: Nasionale Pers.
- AWS⁵: Boshoff, S.P.E., Bosman, D.B., Hiemstra, L.W., Le Roux, T.H., & Malherbe, D.F., in opdrag van die Suid-Afrikaanse Akademie vir Taal, Lettere en Kuns. 1937. *Afrikaanse woordelys en spelreëls*. 5e uitgawe. Bloemfontein, Kaapstad & Pretoria: Nasionale Pers.
- AWS⁶: Taalkommissie (samest.), in opdrag van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 1953. *Afrikaanse woordelys en spelreëls*. 6e uitgawe. Kaapstad, Bloemfontein & Johannesburg: Nasionale Boekhandel.
- AWS⁷: Taalkommissie (samest.), in opdrag van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 1964. *Afrikaanse woordelys en spelreëls*. 7e uitgawe. Kaapstad & Johannesburg: Tafelberg-Uitgewers.
- AWS⁸: Taalkommissie (samest.), in opdrag van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 1991. *Afrikaanse woordelys en spelreëls*. 8e uitgawe. Kaapstad: Tafelberg.
- AWS⁹: Taalkommissie (samest.), in opdrag van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2002. *Afrikaanse woordelys en spelreëls*. 9e uitgawe. Kaapstad: Pharos.
- AWS¹⁰: Taalkommissie (samest.), in opdrag van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2009. *Afrikaanse woordelys en spelreëls*. 10e uitgawe. Kaapstad: Pharos.
- AWS¹¹: Taalkommissie (samest.), in opdrag van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2017. *Afrikaanse woordelys en spelreëls*. 11e uitgawe. Kaapstad: Pharos.

- Bergenholtz, H. 1990. Lexikographische Instruktionen für ein zweisprachiges Wörterbuch. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*, 43:19-37.
- Carstens, W.A.M. 2011. *Norme vir Afrikaans*. 5e uitgawe. Pretoria: J.L. van Schaik.
- De Schryver, G-M. & Prinsloo, D.J. 2001. Corpus-based activities versus intuition-based compilations by lexicographers: The Sepedi lemma-sign list as a case in point. *Nordic Journal of African Studies*, 10(3):374-398.
- De Villiers, M. 1977. Die plek van die normatiewe in die taalkunde. *Taalfasette*, 22(3):1-9.
- Fletcher, W.H. 2007. Concordancing the web. Promise and problems, tools and techniques. In Hundt, M., Nesselhauf, N. & Biewer, C. (eds). *Corpus Linguistics and the Web*. Amsterdam: Rodopi, pp.1-22.
- Fuertes-Olivera, P.A. 2012. Lexicography and the internet as a (re-)source. *Lexicographica*, 28:49-70.
- Gouws, R.H. 1989. *Leksikografie*. Pretoria: Academica.
- Gouws, R.H. & Prinsloo, D.J. 2005. *Principles and practice of South African lexicography*. Stellenbosch: SUN PRESS.
- Gove, P.B. (ed.). 1961. *Webster's Third New International Dictionary of the English Language*. Springfield: Merriam-Webster.
- HAT⁶: Luther, J., Pfeiffer, F. & Gouws, R.H. (reds.). 2015. *Handwoordeboek van die Afrikaanse taal*. 6e uitgawe. Kaapstad: Pearson. (In VivA Woordeboekportaal. Weergawe 1.2. [Toepassing]). <https://www.microsoft.com/en-gb/store/apps/windows> [30 November 2016]
- Heid, U., Evert, S., Säuberlich, B., Debus-Gregor, E. & Scholze-Stubenrecht, W. 2004. Tools for upgrading printed dictionaries by means of corpus-based lexical acquisition. *Proceedings of the 2004 Language Resources and Evaluation Conference*. Lisbon: LREC, pp. 419-423.
- Instituut voor Nederlandse Lexicologie. 2005. *Het Groene Boekje: Woordenlijst Nederlandse Taal van de Nederlandse Taalunie*. Den Haag: Sdu Uitgevers.
- Kotzé, E.F. 2014. Afrikaans as besitting, en die vraagstuk van herstandaardisering. *Tydskrif vir Geesteswetenskappe*, 54(4):635-655.
- Lehmann, C. 2006. Daten – Korpora – Dokumentation. In Institut für Deutsche Sprache – Jahrbuch 2006. *Sprachkorpora – Datenmengen und Erkenntnisfortschritt*. Berlin: De Gruyter, pp. 9-27.
- Martin, W.J.R., Al, B.P.F. & Van Sterkenburg, P.J.G. 1983. On the processing of a text corpus: From textual data to lexicographical information. In Hartmann, R.R.K. (ed.). *Lexicography: Principles and practice (Applied language studies 5)*. London: Academic Press, pp. 77-87.
- McLachlan, J.D. 2016. Standaardafrikaans, standaardspelling en die AWS. *Tydskrif vir Geesteswetenskappe*, 56(2-1):477-502.
- Nederland. 2005. Spellingwet van 15 September 2005.
- Odendaal, B.G. 2012. Die herstandaardisering van Afrikaans: 'n Praktiese benadering met die AWS as gevallestudie. Ongepubliseerde proefskrif. Stellenbosch: Universiteit Stellenbosch.
- Perkuhn, R., Keibel, H. & Kupietz, M. 2012. *Korpuslinguistik*. Paderborn: Wilhelm Fink Verlag.
- Prinsloo, D.J. & De Schryver, G-M. 2003. Effektiewe vodering met die 'Woordeboek van die Afrikaanse Taal' soos gemeet in terme van 'n multidimensionele liniaal. In Botha, W. (red.). *'n Man wat beur: Huldigingsbundel vir Dirk van Schalkwyk*. Stellenbosch: Buro van die WAT, pp. 106-126.
- Projekt Deutscher Wortschatz. 2016. *Afrikaans corpus*. Leipzig: Universität Leipzig, Abteilung Automatische Sprachverarbeitung. <http://corpora.uni-leipzig.de/en>
- Reynaert, M. 2009. Parallel identification of the spelling variants in corpora. *Proceedings of the Third Workshop on Analytics for Noisy Unstructured Text Data. Barcelona*, pp. 77-84.
- Singh, R.A. 1982. *An Introduction to Lexicography*. Michigan: University of Michigan, Central Institute of Indian Languages. <http://www.ciil-ebooks.net/html/lexico/link5.htm> [2 Desember 2016].
- Suid-Afrika. 1959. Wet op die Suid-Afrikaanse Akademie vir Wetenskap en Kuns 54 van 1959.
- Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2015. Reglement: Taalkommissie. Hersiene weergawe goedgekeur in September. Pretoria: Suid-Afrikaanse Akademie vir Wetenskap en Kuns.
- Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2011. Taalkommissiekorpus 1.1. Potchefstroom: CText, Noordwes-Universiteit.
- Tarp, S. & Fuertes-Olivera, P.A. 2016. Advantages and disadvantages in the use of internet as a corpus: The case of the online dictionaries of Spanish Valladolid-UVa. *Lexikos*, 26:273-295.

- Tomaszczyk, J. 1983. On bilingual dictionaries: The case for bilingual dictionaries for foreign language learners. In: Hartmann, R.R.K. (ed.). *Lexicography: Principles and practice (Applied language studies 5)*. London: Academic Press, pp. 41-51.
- Van Huyssteen, G.B. 2016. Die ortografiese realisering van komposita met en afleidings van multiwoordeiname. *LitNet Akademies*, 13(3):424-456.
- Van Huyssteen, G.B. 2017. Opname- en elimineringskriteria vir die *Afrikaanse Woordelys en Spelreëls*: Die geval *emeritus*. *Tydskrif vir Geesteswetenskappe*, 57(2-1):346-368.
- Van Rensburg, M.C.J. 2017. Ná honderd jaar: die Afrikaans van die eerste taalkommissie. *Tydskrif vir Geesteswetenskappe*, 57(2-1):249-270.
- Van Sterkenburg, P.J.G. 2004. Resensie: Afrikaanse Woordelys en Spelreëls. *Lexikos*, 14:428-431.
- Virtuele Instituut vir Afrikaans (VivA). 2016. Korpusportaal. <http://viva-afrikaans.org> [20 November 2016].
- Wallis, S. 2012a. Measures of association for contingency tables. <http://www.ucl.ac.uk/english-usage/staff/sean/resources/phimeasures.pdf%5Chttps://corplingstats.wordpress.com/2012/04/09/measures-of-association/> [2 Desember 2016].
- Wallis, S. 2012b. Robust and sound? <https://corplingstats.wordpress.com/2012/04/04/robust-and-sound/>. [2 Desember 2016].
- Wallis, S. 2013a. Comparing frequencies within a discrete distribution. <https://corplingstats.wordpress.com/2013/05/20/comparing-frequencies/> [2 Desember 2016].
- Wallis, S. 2013b. Goodness of fit measures for discrete categorical data. <http://www.ucl.ac.uk/english-usage/staff/sean/resources/gofmeasures.pdf>. [2 Desember 2016].
- Wallis, S., Bowie, J. & Aarts, B. 2012. That vexed problem of choice. Some reflections on experimental design and statistics with corpora. In *International Conference on Advances in Mechanics Engineering (ICAME 33)*, Leuven. <http://www.ucl.ac.uk/english-usage/staff/sean/resources/vexedchoice.pdf>. [6 Julie 2016].
- Walter, E. 1996. Parallel development of monolingual and bilingual dictionaries for learners of English. In Gellerstam, M., Järborg, J., Malmgren, S-G., Norén, K., Rogström, L. & Pappmehl, C.R. (eds). *Euralex '96 Proceedings I-II*. Gothenburg: Göteborg University, Department of Swedish, pp. 635-641.
- WAT: Woordeboek van die Afrikaanse Taal. 2016. Stellenbosch: Buro van die WAT. (In VivA Woordeboekportaal. Weergawe 1.2. [Toepassing]). <https://www.microsoft.com/en-gb/store/apps/windows> [30 November 2016].
- Wette *kyk* Suid-Afrika.
- Zgusta, L. 1971. *Manual of lexicography*. The Hague: Mouton.