

Die grammatikale en leksikale eienskappe van geselekteerde Afrikaanse gedigte

Lezandra Grundlingh

Lezandra Grundlingh, Departement Afrikaans en Algemene Literatuurwetenskap,
Universiteit van Suid-Afrika

Opsomming

Elke teksgenre het bepaalde eienskappe. Hierdie eienskappe sluit strukturele aspekte asook taalgebruik in. Taalgebruik word nie net met 'n spesifieke genre geassosieer nie, maar soms ook met 'n bepaalde gender. Die huidige studie ondersoek die grammatikale en leksikale eienskappe in enkele Afrikaanse gedigte wat van 2002 tot 2010 gepubliseer is, asook die moontlike taalgebruik- en genderverskille in hierdie korpora. Deur die ontleding van die korpora vanuit 'n korpuslinguistiekperspektief kan bepaal word hoe frekwent spesifieke leksikale en grammatikale woorde in hierdie gedigte is in vergelyking met ander genres. Dit is ook moontlik om vas te stel of bepaalde woorde meer gereeld deur manlike of vroulike digters gebruik word. Verder kan bepaal word hoe hierdie woorde in konteks gebruik word. Die korpusprogram Antconc is gebruik om die data in die twee korpora van manlike en vroulike digters te ontleed. Sleutelwoordontledings en konkordansiesoektogte is uitgevoer om die moontlike grammatikale en leksikale eienskappe van die korpora as 'n geheel, maar ook in die poësie van die manlike en vroulike digters vas te stel. Die resultate dui aan dat bepaalde taalkundige eienskappe met die Afrikaanse gedigte in hierdie korpora en moontlik met manlike of vroulike digters geassosieer kan word.

Dit is belangrik om kennis te neem dat die data in hierdie studie nie verteenwoordigend is van alle Afrikaanse poësie nie. Dit beteken dat die resultate slegs op die korpora in hierdie studie van toepassing is. Desnietemin kan die resultate steeds insigte bied in nie alleen die potensiele taalkundige eienskappe van die Afrikaanse poësie in sy geheel nie, maar ook in die digkuns van manlike en vroulike digters.

Trefwoorde: Afrikaanse poësie; genderverskille; korpuslinguistiek; manlike digters; taalgebruiksverskille; vroulike digters

Abstract

The grammatical and lexical characteristics of selected Afrikaans poems

The current study explores language use in selected Afrikaans poems from a corpus linguistics perspective. The aim is to determine whether the poems can be distinguished from other general Afrikaans texts, other Afrikaans literature texts and from one another (poems written by men versus poems written by women). By analysing the corpora from a corpus linguistics perspective it is possible to determine whether the poets make use of specific grammatical or lexical characteristics that are unique to the corpora as a whole, but also to the male and female poets. Thus, language use and gender are investigated within the genre of Afrikaans poetry.

This study is based on previous research concerning gender and differences in language use as well as gender and genre. Differences in language use with reference to gender and genre have been researched from a variety of perspectives (Mulac 1975; Mulac and Lundell 1994; Barrette 2004 and Herring and Paolillo 2006). However, only a few studies focus on the Afrikaans language and specifically Afrikaans poetry. Poetry is an interesting genre to research from a linguistic perspective since it conceptualises language differently from other text genres. In poetry, language is probably used in the most creative way and words are often chosen very carefully. This study illustrates that certain characteristics can be identified that not only separate the two main corpora from other genres, but also separate the poetry of male and female poets. This suggests that there are both grammatical and lexical characteristics that could be attributed to male and female poets.

The current study focuses on two corpora of selected Afrikaans poems written from 2002 to 2010. Both corpora consist of 10 datasets (volumes of poetry) by male and female poets. These two corpora (the main corpora) are compared with a corpus of general Afrikaans texts, a corpus of Afrikaans literature texts, another corpus of Afrikaans poetry and also to each other. The corpus analysis program Antconc is used to analyse the different comparisons using keyword lists, concordance lines and concordance plots.

The keyword list analysis compares a target corpus with a reference corpus to determine whether there are words in the target corpus that are significantly key compared with the reference corpus. Significant keyness is indicated through a log-likelihood ratio where a statistical value of 3,84 or higher indicates a p-value of $p < 0,05$. Any words listed as significantly key were analysed to determine their context and whether they are distributed throughout the corpora. The context of a word is determined by generating concordance lines, while the distribution is determined through the concordance plot tool. This tool shows the number of datasets in which a word occurs. For the purposes of this study, a word has to occur in more than 50% of the datasets in order to be considered distributed throughout the corpora. Only words that are distributed throughout the corpora are considered as possibly characteristic of specific corpora. By using the concordance lines to determine the context of the words it is possible to determine whether the words are grammatical or lexical and to divide the lexical words into certain categories. These categories are considered the possible lexical categories of the two main corpora in each comparison. The grammatical characteristics are considered separately.

The comparison of the main corpora and the corpus of general Afrikaans texts indicates a significant use of certain pronouns in the main corpora. Conjunctions are more frequently used in the main corpora as compared with the general Afrikaans texts. 39 lexical categories are distinguished when the main corpora and the general Afrikaans texts are compared. These include words relating to the body and body parts, the senses, cognition, longing, etc.

When the main corpora and the corpus of other Afrikaans literature texts are compared a similar pattern emerges concerning the grammatical characteristics. However, the log-likelihood ratios are lower, which suggests that there is a smaller difference between the main corpora and the other literature texts. This is to be expected, since the corpus of Afrikaans literature texts consists of dramas, novels and prose. The comparison between the main corpora and the corpus of literature texts indicates a few words that can be added to the lexical categories since they are evenly distributed, but they fit into already existing categories mentioned above. This second comparison produces 34 possible lexical categories compared with the 39 categories in the previous comparison. No new categories are identified, but some of the categories in the previous comparison are no longer represented. There are also minor changes to the words included in certain categories. This finding suggests that certain grammatical and lexical characteristics of the main corpora remain constant, even when they are compared with different corpora of various genres.

A third comparison between the main corpora and another corpus of Afrikaans poetry supports the statements above and proves that there are certain characteristics that are shared within a single genre. The keyness is lowest when the main corpora and the additional poetry corpus are compared, suggesting that the poets use similar words in both instances. However, the grammatical characteristics of the main corpora are more or less the same as in the previous comparisons. The lexical characteristics are slightly different with small changes to the categories. This suggests that the poets in the main corpora use specific words in specific ways that allows them to be distinguished from a variety of other corpora, including another poetry corpus.

To determine whether gender plays a role in language use differences, the main corpora are compared with each other. The comparisons indicated that male poets in this corpus tend to use more pronouns, especially pronouns referring to others. It also appears that the male poets use the negative *nie* more often. In terms of the lexical characteristics seven categories are identified in the male/female comparison and 13 in the female/male comparison. Although some categories correspond, there are differences. The poems by male poets contain the categories action, religious concepts and death, which are not as prominent in the poetry of female poets. The female poets can be distinguished through the inclusion of the following categories: cognition, creative processes, food, house or dwelling, nature, colours, fire and light.

Based on the results, the conclusion is drawn that the differences in language use identified are determined by both genre and gender. Although the characteristics that are identified in this study do not reflect the characteristics of all Afrikaans poetry as a whole or all Afrikaans poetry by male or female poets, it does offer insights into possible characteristics of Afrikaans poetry by male and female poets.

Keywords: Afrikaans poetry; corpus linguistics; female poets; gender differences; language use; male poets

1. Inleiding: Genre en korpuslinguistiek

Hierdie studie is hoofsaaklik 'n korpusgebaseerde ontleding van geselekteerde gedigte, en word daarom onder die korpuslinguistiek geklassifiseer. Alhoewel die data uit poësietekste bestaan, word die studie vanuit 'n taalkundige perspektief benader. Kognitiewe benaderings tot taal en taalgebruik word ook oorweeg.

Steen (2002:187) voer aan dat enige kognitiewe benadering tot taal of styl by die taalgebruiker moet begin, aangesien die skrywer of spreker die bron van enige kognitiewe prosesse en uitsette is. Hierbenewens voer Steen (188) aan dat taalgebruikers taal altyd in konkrete situasies prosesseer en oor gedetailleerde kognitiewe modelle en verwagtinge rakende hierdie kontekste beskik. Volgens Steen (188) speel genre 'n kardinale rol in die ontleding van hierdie kognitiewe aspekte:

I assume that language users possess a whole range of cognitive constructs that are individual representations of culturally recognised, distinct genres and subgenres of discourse. Such cognitive genre representations are exploited by language users when they produce or comprehend discourse in all kinds of discourse situations. Cognitive genre representations facilitate such discourse aspects as co-operation with other language users, the selection and organisation of the content of messages, and the selection of the appropriate register.

Dit is duidelik dat die kognitiewe voorstellings van genre 'n rol speel by individuele taalgebruikers se verwerking van taal (produksie en begrip). In die huidige studie word al die data as poësie gekategoriseer. De Lange (2000:17) beklemtoon dat gedigte uit woorde gemaak is en nie uit idees nie, en beskryf die skryfproses in poësie as “kompleks”. Die aard van poësie maak dit moeilik om die data vanuit 'n suiwer taalkundige perspektief te bestudeer. Woorde word soms in spesifieke kontekste gebruik wat die oorspronklike betekenis van die woord beïnvloed. 'n Korpusgebaseerde ontleding van poësie kan egter hierdie probleem oplos aangesien korpusprogramme die ontleder toelaat om woorde deur middel van konkordansiesoektogte in konteks te bestudeer. In hierdie studie word juis dit beoog: om poësie vanuit 'n taalkundige perspektief te ontleed deur op korpusgebaseerde ontledings te steun.

Korpora is die kern van korpusontledings en bestaan uit 'n versameling tekste wat vir 'n bepaalde doel saamgestel is. Een van die hoofkriteria van 'n korpus is dat dit verteenwoordigend moet wees (Cheng 2012:3). Die korpora in die huidige studie is egter nie verteenwoordigend van die Afrikaanse poësie in sy geheel en van die poësie van alle manlike of vroulike Afrikaanse digters nie. Soos in deel 3 bespreek word, vorm slegs 22 gedigte die korpora vir hierdie studie. Alhoewel hierdie korpora klein is, kan die data-ontleding daarvan steeds tot betekenisvolle resultate lei. Vaughan en Clancy (2013:57) voer aan dat studies in sosiolinguietiek dui daarop dat klein korpora, wat nie verteenwoordigend is van die groter gemeenskap nie, steeds voldoende is om sekere eienskappe van taalvariasie aan te dui. Handford (2010: 256) redeneer ook dat kleiner, gespesialiseerde korpora voldoende is wanneer die studie 'n bepaalde doel het, soos om gespesialiseerde genres te ondersoek.

Korpuslinguistiek kan vanuit twee perspektiewe beskou word. Aan die een kant word dit beskou as 'n metodologie wat gebruik word om navorsing in die leksikografie, die taalkunde, diskoers en pragmatiek te versterk (McEnery, Xiao en Tono 2006), en is met ander woorde 'n

metodologie wat ingespan word om bestaande teorieë te toets. Aan die ander kant beskou sommige navorsers dit as 'n dissipline in eie reg (Tognini-Bonelli 2001; Biber 2009) en word dit aangewend om teorieë te ontwikkel (Cheng 2012:6).

In hierdie studie word korpuslinguistiek gebruik om die teorie te toets dat die gedigte in hierdie korpora spesifieke leksikale en grammatikale eienskappe bevat wat dit van ander korpora onderskei en dat daar tussen die grammatikale en leksikale eienskappe van manlike en vroulike digters onderskei kan word.

Korpuslinguistiek en enkele studies oor taalgebruik en gender word in die volgende afdeling verder bespreek.

2. Literatuuroorsig

2.1 Studies oor taalgebruik en gender

Barrette (2004) verwys in 'n aanlyn artikel oor die skryfstyl van mans en vrouens na die webtuiste Gender Genie. Hierdie webtuiste (wat egter nie meer aktief is nie), ontwikkel deur Koppel, Argamon en Shimoni (2002), is gebruik om fiksie, niefiksie en blogs deur manlike en vroulike skrywers van mekaar te onderskei. Dit beteken die webtuiste dui die skrywerskap van 'n teks as “manlik” of “vroulik” aan op grond van die eienskappe wat geïdentifiseer en die genre wat gekies is. Die kriteria wat die webtuiste gebruik, is volgens Barrette (2004) grammatikale verskille eerder as algemene inhoud, maar ook die onderwerp van 'n bepaalde teks is van belang. Koppel e.a. (2002) en Barrette (2004) noem dat daar op grond van vorige navorsing verskeie grammatikale verskille tussen die tekste van manlike en vroulike skrywers onderskei kan word:

This list included a large number of determiners (*a, the, that, these*) and quantifiers (*one, two, more, some*) as male indicators. Moreover [...] a determiner which typically occurs either as the first word in a noun phrase or as the head of a noun phrase, [...] a determiner which typically begins a noun phrase but cannot appear as its head, and [...] cardinal numbers are all strong male indicators. Conversely, the pronouns (*I, you, she, her, their, myself, yourself, herself*) are all strong female indicators. (Koppel e.a. 2002:325)

Barrette (2004), asook Herring en Paolillo (2006:441), redeneer egter dat die verskille in die skryfstyl van mans en vrouens nie so maklik is om te identifiseer nie. Alhoewel sekere eienskappe gewoonlik in tekste van manlike en ander in tekste van vroulike skrywers voorkom, speel ook ander faktore, soos die karaktereieenskappe van die skrywer, die onderwerp van die teks en die bepaalde skryfegnieke wat gebruik word, 'n rol. Hierdie faktore kan veroorsaak dat tekste nie “suiwer” manlike of vroulike eienskappe vertoon nie.

Hierdie komplekse verhouding tussen gender en taal is reeds in 1994 deur Mulac en Lundell (1994) ondersoek. Hulle gebruik die term *gender-linked language effect* om hulle navorsing op te som. Hulle korpus het uit 40 kort essays bestaan – 20 essays deur manlike en 20 essays deur vroulike skrywers. Agt individue wat beskryf word as “untrained transcript raters” (Mulac en Lundell 1994:301) is gevra om die essays te lees en spesifieke kenmerke aan te dui

wat in elk van die essays teenwoordig is. Die kenmerke is gebaseer op die persoonlikheidsdimensies wat Mulac (1975:185, 187) identifiseer:

Socio-intellectual Status – high social status/low social status, white collar/blue collar, literate/illiterate, and rich/poor; *Aesthetic Quality* – pleasant/unpleasant, beautiful/ugly, sweet/sour, and nice/awful; *Dynamism* – strong/weak, active/passive, loud/soft, and aggressive/unaggressive.

Die resultate toon dat vroulike skrywers meer sosio-intellektuele dimensies en estetiese kwaliteite by hul skryfwerk insluit (werkwoorde wat onsekerheid aandui, verwysings na emosies, en langer sinne), terwyl daar in die tekste van manlike skrywers meer dinamiese taalgebruik voorkom (verwysings na plek en hoeveelheid en byvoeglike naamwoorde wat op 'n veroordelende perspektief dui) (Mulac en Lundell 1994:303, 307). Alhoewel hulle verskille in die taalgebruik van manlike en vroulike skrywers kon identifiseer, waarsku Mulac en Lundell (1994:307) dat daar oorvleueling tussen manlike en vroulike skrywers is en dat die skryfwerk nie absolute, duidelik afgebakende onderskeidings bevat nie.

Volgens Newman, Groom, Handelman en Pennebaker (2008:212) het die belangstelling in taalgebruiksverskille tussen manlike en vroulik skrywers ontstaan omdat taal 'n verskynsel is wat aandui hoe mans en vrouens hulle sosiale omgewings benader en interpreteer. Navorsing oor gesproke taal het byvoorbeeld aangedui dat mans taal gebruik om inligting oor te dra terwyl vrouens taal eerder vir sosiale interaksie gebruik. Newman e.a. (2008:213) plaas veral klem op die feit dat genre (in 'n gesproke en geskrewe konteks) 'n groter rol in taalgebruiksverskille speel as gender. Volgens hulle (213) het navorsing konstant aangedui dat vrouens oor die algemeen meer bywoorde, voegwoorde en modale werkwoorde gebruik terwyl mans meer vloekwoorde, langer woorde en lidwoorde gebruik. Mans sluit ook meer plekverwysings in. Hulle waarsku egter dat dit moeilik is om gevolgtrekkings te maak wanneer taalgebruik ontleed word, aangesien groot hoeveelhede data nodig is voordat resultate werklik as “eienskappe” van manlike of vroulike taalgebruik kan kwalifiseer (214).

Vir hulle studie het Newman e.a. (217) twee metodologiese ontwikkelinge gebruik om vas te stel of daar verskille is tussen die taal van manlike en vroulike taalgebruikers. Eerstens maak hulle gebruik van die teksontledingsagteware LIWC (Linguistic Inquiry and Word Count), wat die gebruiker toelaat om groot korpora te ontleed. Elke woord in die korpora word met 'n woordelys van 2 000 woorde vergelyk en in 74 taalkundige kategorieë verdeel. Sommige kategorieë is suiwer grammatikaal (soos die kategorie “articles”), terwyl ander kategorieë, soos “positive emotion words”, geskep is deur onafhanklike beoordelaars wat besluit het watter woorde in elke kategorie pas. Die tweede metodologiese ontwikkeling behels die versamel van 'n tekskorpus met meer as 500 000 tekste wat onder andere uit prosatekste, gedigte en lirieke bestaan.

Die resultate toon dat vrouens meer voornaamwoorde en werkwoorde gebruik. Hulle gebruik ook meer “sosiale” woorde en verwysings na psigologiese prosesse. Vrouens maak ook meer gereeld van ontkenning gebruik, en verwys meer na konsepte wat met die huis te doen het. Mans, daarenteen, gebruik meer lidwoorde, voorsetsels, telwoorde en vloekwoorde (Newman e.a. 223). Benewens die genre van die teks speel die ouderdom van die skrywer ook 'n rol in taalgebruik. Volgens Newman e.a. (225) en Pennebaker, Mehl en Niederhoffer (2003:556) gebruik ouer persone meer woorde wat positiewe emosies aandui, minder eerstepersoonverwysings, en meer woorde wat met die toekomstige tyd geassosieer word.

Argamon, Koppel, Fine en Shimoni (2003:324) waarsku in hul studie dat navorsing oor gendergebaseerde verskille in taal onder baie kritiek deurloop omdat die studies gewoonlik nie van 'n betroubare metodologie gebruik maak nie en navorsers soms aanneem dat daar verskille tussen manlike en vroulike taalgebruik móét wees en dan spesifiek hierdie verskille soek. Om hierdie probleme te vermy, het Argamon e.a. (2003) van 'n groot, hoëgehaltekorpus gebruik gemaak (604 tekste uit die British National Corpus) wat deur rekenaarsagteware ontleed is. Die doel van die studie was om vas te stel of die gender van die skrywer van verskillende tekste vasgestel kan word deur die eienskappe van bepaalde tekste te identifiseer.

Argamon e.a. (325) het bevind dat manlike skrywers meer lidwoorde (bepaald en onbepaald) en kwantifiseerders gebruik. Hulle gebruik ook meer lidwoorde aan die begin van sinne en maak meer gebruik van hooftelwoorde. Vroue, daarenteen, gebruik meer voornaamwoorde. Argamon e.a. (326) voer aan dat die manier waarop vroue skryf, die leser meer betrokke laat voel, terwyl manlike skrywers meer toeligtend skryf:

The pronouns of women's writing, as all pronouns, present things in a relational way: "I know that you know what I am referring to, therefore I will present the information as if we both know it." The specifiers found more frequently in men's writings send the message of: "here are some details about the things being mentioned".

Argamon e.a. (327) noem verder dat voornaamwoorde in beide fiksie en niefiksie meer deur vroulike skrywers gebruik word en dat die skrywers veral die voornaamwoorde *I*, *you* en *she* gebruik. Hulle bevind egter dat manlike skrywers die meervoudsvorme *we*, *us*, *they* en *them* meer as die vroulike skrywers in fiksie gebruik en dat manlike skrywers ook die voornaamwoorde *him* en *he* meer gebruik in beide fiksie en niefiksie (330). Ten opsigte van die manlike skrywers beklemtoon Argamon e.a. (334) die feit dat hulle geneig is om "te spesifiseer" deur onder andere meer lidwoorde te gebruik.

Gender en genre speel volgens Argamon e.a. (336) vir seker 'n rol in taalgebruiksverskille. Nie net gebruik manlike en vroulike skrywers taal verskillend nie, maar hulle gebruik ook taal op spesifieke maniere, afhangende van die genre (fiksie teenoor niefiksie).

Die huidige studie sluit aan by vorige studies oor taalgebruiksverskille en gender, maar beoog 'n alternatiewe benadering. In stede daarvan om net aan te dui watter woorde meer frekwent deur 'n bepaalde gender gebruik word, word ook gepoog om hierdie woorde in kategorieë te verdeel en moontlik nuwe insigte in die taalgebruik van manlike en vroulike digters te bied.

Weer eens word beklemtoon dat hierdie studie nie poog om algemene uitsprake te lewer oor die taalgebruik van mans of vrouens / manlike of vroulike digters oor die algemeen nie. Enige gevolgtrekkings wat gemaak word, is slegs op hierdie korpora van toepassing, maar kan as moontlike beginpunt dien vir toekomstige navorsing in taalgebruiksverskille en gender asook korpuslinguistiek en die letterkunde.

Enkele aspekte van korpusgebaseerde ontledings in die letterkunde word hierna bespreek, waarna die data en metodologie bespreek word.

2.2 Korpuslinguistiek en die letterkunde

Biber (2010:241) meen dat twee begrippe die kern van korpuslinguistiek vorm: *register* en *genre*. Hierdie twee begrippe verteenwoordig twee benaderings in data-ontleding. *Genre* plaas die fokus op die taalkundige eienskappe wat gebruik word om tekste in hul geheel te struktureer. Om hierdie rede word genrestudies op volledige tekste uit 'n bepaalde genre gebaseer (soos akademiese artikels, briewe, nuusberigte, ens.).

Wanneer 'n korpusgebaseerde ontleding egter op die *register* fokus, is die omvattende taalkundige eienskappe van belang en gaan dit eerder om ontledings van taalkundige eienskappe wat met die inhoud van die tekste te make het. Biber (242) beskryf dit soos volg:

The register perspective characterises the typical linguistic features of text varieties, and connects those features functionally to the situational context of the variety. Because the focus is on words and grammatical features that are frequent and pervasive, the analysis can be based on a sample of text excerpts rather than complete texts.

Binne die korpuslinguistiek verwys *genre* met ander woorde veral na die strukturele eienskappe van 'n teks, terwyl *register* na die taalgebruik daarin verwys. Die huidige studie ontleed die data vanuit 'n registerperspektief eerder as 'n genreperspektief (alhoewel die korpora steeds genrespesifiek is) deur te fokus op die taalkundige eienskappe van tekste binne 'n bepaalde genre (die Afrikaanse poësie).

McIntyre en Walker (2010:516, 522) ondersoek spesifiek die gebruik van korpuslinguistiek as instrument vir die ontleding van literêre tekste. Hulle studie fokus op twee gespesialiseerde korpora: 'n versameling gedigte deur William Blake (*Songs of innocence* en *Songs of experience*) en 'n 200 000-woord-korpus uit draaiboeke in die aksie/riller-genre.

Die eerste ontleding, wat fokus op die poësie van William Blake, ondersoek hoe die twee versamelings gedigte van mekaar verskil. Daar word aandag gegee aan spesifiek die leksikale en semantiese verskille tussen die twee versamelings. McIntyre en Walker (517) maak gebruik van die Wmatrix-sagteware vir hul ontleding, aangesien hierdie program 'n outomatiese semantiese annoteerder insluit. In hulle ontleding word die log-waarskynlikheidsstatistiek ("log-likelihood ratio") gebruik om die statistiese beduidendheid van die woorde in 'n spesifieke korpus te bepaal. Dieselfde statistiese metode word ook in die huidige studie gebruik, en word later in meer besonderhede bespreek.

McIntyre en Walker (517–9) het onder andere bevind dat daar 'n verskil is tussen die semantiese domeine en die sleutelwoorde van die twee versamelings gedigte. Deur middel van konkordansiesoektoegte kon hulle (520) vasstel of die sleutelwoorde verteenwoordigend is van die meeste gedigte in die korpora of net in spesifieke gedigte voorkom. Hieruit kon hulle gevolgtrekkings maak oor die gebruik van bepaalde woorde in spesifieke korpora en hoe dit by die semantiese domeine aansluit. Hulle waarsku egter dat korpusontledings alleen nie tot gevolgtrekkings lei nie:

It should be apparent that in some cases our corpus analysis validates some of the subjective critical responses to *Songs*. Of course, key comparisons can only be a starting point. In order to fully understand the lists produced by a computer tool, we

must return to the text. Quantitative analysis guides qualitative analysis, which might guide further quantitative analysis.

Vir hul tweede korpusgebaseerde ondersoek het McIntyre en Walker (522) die volgende navorsingsvrae probeer beantwoord: (1) Is daar 'n verskil in die hoeveelheid manlike en vroulike taalgebruik in aksiefilms? (2) Weerspieël die onderwerpe waaroor die manlike en vroulike karakters gesels, hul gender? (3) Is fisieke krag en dapperheid belangrike semantiese domeine in aksiefilms? (4) Weerspieël die taalgebruik van manlike karakters verset teen gesag?

McIntyre en Walker (523) het die manlike en vroulike taalgebruik geïdentifiseer deur van die aanwysings in die draaiboeke gebruik te maak. Daarna is die Multilingual Corpus Toolkit gebruik om die tekste wat manlike en vroulike taalgebruik verteenwoordig, as aparte lêers te stoor. WordSmith Tools 4.0 is ingespan om die woordfrekwensies asook die “type/token”-ratio's te bepaal, en Wmatrix om die sleutelwoorde, semantiese domeine en n-gramme aan te toon.

McIntyre en Walker (524–7) kon vasstel dat daar 'n voorkeur is vir manlike dialoog in die korpus en dat die taalgebruik van die karakters skynbaar hulle gender weerspieël. “In power” is as 'n belangrike semantiese domein in die dialoog van manlike karakters geïdentifiseer, maar McIntyre en Walker kon nie vasstel of die manlike karakters se taalgebruik verset teen gesag weerspieël nie.

McIntyre en Walker (529) noem dat hul ontleding slegs “potential distinguishing elements” tussen manlike en vroulike taalgebruik in aksiefilms identifiseer, en dat dit tot verdere hipoteses oor taalgebruik in draaiboeke en die filmgenre kan lei. Insgelyks kan ook die huidige studie beskou word as 'n verkennende studie oor die gebruik van korpuslinguistiek in die Afrikaanse poësie.

Die korpora wat in hierdie studie gebruik is, word eers bespreek en daarna die metodologie en die resultate.

3. Die data

Soos aangedui, is die fokus van die studie tweeledig. Eerstens word vasgestel of hierdie korpus van Afrikaanse gedigte (van manlike en vroulike digters) leksikale en grammatikale eienskappe het wat dit van ander korpora onderskei. Tweedens word bepaal of die leksikale en grammatikale eienskappe van manlike en vroulike digters in hierdie studie dié twee korpora van mekaar kan onderskei.

So 'n ontleding vereis dat verskeie korpora met mekaar vergelyk word. Die uiteindelijke ontleding bestaan daarom uit die vergelyking van ses korpora. Die korpora wat vergelyk word, is:

1. 'n Korpus van algemene Afrikaanse tekste (KAT) (Roux, Louw en Niesler 2004).

2. 'n Korpus van Afrikaanse letterkundige tekste (KAL) (Snyman, Van Huyssteen en Daelemans 2012).
3. 'n Korpus van Afrikaanse poësie (KAP) (Snyman, Van Huyssteen en Daelemans 2012).
4. Die Proteakorpus van Afrikaanse poësie (PAP).
5. 'n Korpus van manlike Afrikaanse digters (KMAD).
6. 'n Korpus van vroulike Afrikaanse digters (KVAD).

Die PAP, KMAD en KVAD is die hoofkorpora in hierdie studie. Die PAP is 'n kombinasie van die KMAD en die KVAD. Al die gedigte in die PAP is uit die PUK/Proteakorpus geneem wat deur die Noordwes-Universiteit se Sentrum vir Tekstegnologie (CTexT) versamel en verwerk is. Die Proteakorpus bestaan uit die tekste van 41 digbundels deur verskeie Afrikaanse digters (mans en vroue). Vir die doeleindes van hierdie studie is net 10 digbundels deur manlike skrywers en 10 digbundels deur vroulike skrywers gekies. Die besluit is geneem om digbundels wat Nederlands insluit, uit te laat, asook digbundels waarvan die gedigte vir kinders geskryf is. Die digbundels wat ingesluit is, weerspieël met ander woorde net Afrikaanse gedigte wat op volwasse lesers gemik is. Die woordtelling vir albei korpora, asook die digbundels wat ingesluit is, verskyn in tabel 1. Die digbundels in albei korpora is van 2002 tot 2010 gepubliseer.

Tabel 1. Die woordtelling en struktuur van die KMAD en KVAD

KMAD	
Digter	Digbundel
1. Marius Crous	Aan 'n beentjie sit en kluif
2. M.M. Walters	Braille-briewe
3. Cas Vos	Die afdruk van ons hande
4. Daniel Hugo	Die twaalfde letter
5. Johannes van Jerusalem	Eros ontbind
6. Dolf van Niekerk	Lang reis na Ithaka
7. Melt Myburgh	Oewerbestaan
8. Barend J. Toerien	Om te onthou
9. Johann Lodewyk Marais	Plaaslike kennis
10. Pirow Bekker	Stillerlewe
Totale woordtelling	76 333
KVAD	
Digter	Digbundel
1. Zandra Bezuidenhout	Aardling
2. Jelleke Wierenga	Bloot mens
3. Rosa Smit	Die begeleiding van duiwe
4. Mari Grobler	Eggo's in bewolkte water
5. Ronel Nel	En die Here het foto's geneem oor Vanderbijlpark
6. Sarina Dönges	In die tyd van uile
7. Marlise Joubert	Passies en passasies

8. Martjie Bosman	Toevallige tekens
9. Lina Spies	Tydelse gety
10. Lucie Möller	Watermerke
Totale woordtelling	72 063

Omdat die korpora klein is, is besluit om hulle nie af te rond sodat elke korpus uit presies dieselfde aantal woorde bestaan nie.

Toegang tot die KAT, KAL en KAP is verkry deur middel van die RMA-katalogus wat in samewerking met CText bestuur word. Die KAT bestaan uit verskeie Afrikaanse tekste uit verskeie genres (koerante, webtuistes, ens.) en is nie geannoteer nie. Die KAL bestaan uit verskeie letterkundige tekste, soos dramas, prosatekste en gedigte, en is ook nie geannoteer nie. Vir die doeleindes van hierdie studie is die KAL aangepas sodat dit net uit drama- en prosatekste bestaan. Die KAP verteenwoordig slegs die poësie-afdeling van die KAL.

Tabel 2 dui die woordtellings van hierdie drie korpora aan.

Tabel 2. Die woordtellings van die KAT, KAL en KAP

Korpus	Aantal tekste	Totale woordtelling
KAT	463	1 258 104
KAL	125	320 954
KAP	411	55 755

Die bespreking van die metodologie dui aan hoe die data verwerk en ontleed is.

4. Metodologie

Die korpora in hierdie studie word ontleed deur middel van hoofsaaklik 'n sleutelwoordontleding. Afgesien hiervan word die gebruik van die woorde in konteks met konkordansiesoektogte bepaal. Dit is belangrik aangesien sommige woorde byvoorbeeld as 'n selfstandige naamwoord én werkwoord gebruik kan word. Die korpora is geannoteer in 'n poging om die woordsoorte te identifiseer, maar die annoteerder (bespreek onder 4.2) is nie akkuraat genoeg om betroubare resultate op te lewer nie. Dit het tot gevolg gehad dat die woorde individueel tydens konkordansiesoektogte geklassifiseer is.

Vir die ontleding is die gratis korpusprogram Antconc gebruik. Volgens Anthony (2007:1) het Antconc as 'n eenvoudige konkordansieprogram begin en mettertyd in 'n nuttige teksontledingsprogram ontwikkel. Antconc kan gebruik word om verskeie ontledings op taalkorpora te doen. Die funksies van Antconc sluit onder andere in 'n "wordlist tool", "concordance tool", "keyword list tool" en 'n "clusters/n-gram tool". Vir die doeleindes van hierdie studie is 'n sleutelwoordlys van elke korpus opgestel om vas te stel watter woorde moontlik die twee korpora van mekaar onderskei. Die sleutelwoordontleding bepaal watter woorde buitengewoon frekwent in een korpus gebruik word in teenstelling met 'n ander korpus. Konkordansiesoektogte is op alle woorde met 'n sleutelwaarde bo 3,84 uitgevoer om

die woordsort daarvan te bepaal en vas te stel in watter konteks die woorde gebruik word. Dit is ook gebruik om vas te stel of die woorde verteenwoordigend van 'n bepaalde korpus is.

Omdat die PAP uit verskillende gedigte van manlike en vroulike skrywers bestaan, is dit belangrik dat woorde wat deur die sleutelwoordontleding geïdentifiseer word, wel verteenwoordigend is van die korpus in sy geheel. Dit beteken dat die woord nie net statisties beduidend moet wees nie, maar ook in meer as 50% van die subkorpora moet voorkom. Elke skrywer se versameling gedigte tel as 'n aparte, kleiner subkorpus, wat beteken dat die PAP uit 20 subkorpora bestaan. Om verteenwoordigend te wees van die korpus moet 'n woord in minstens 11 van die subkorpora voorkom. As 'n woord in net 10 voorkom, kan dit beteken dat die woord moontlik slegs in die gedigte deur manlike skrywers of slegs in dié deur vroulike skrywers voorkom. Dieselfde reël geld by die KMAD en KVAD. Hier bestaan die korpora elk uit 10 versamelings gedigte, wat beteken dat 'n woord wat statistiese beduidendheid toon, in minstens ses van die subkorpora moet voorkom.

Twee aparte hoofontledings is gedoen. Die eerste was van die PAP in sy geheel en die tweede van die KMAD en KVAD as aparte korpora. Die vyf vergelykings van die korpora wat hierdie twee ontledings ondersteun, kan soos volg voorgestel word:

1. Die PAP met die KAT.
2. Die PAP met die KAL.
3. Die PAP met die KAP.
4. Die KMAD met die KVAD.
5. Die KVAD met die KMAD.

Omdat hierdie studie die korpora vanuit 'n registerperspektief benader, is die begrippe *grammatikale eienskap* en *leksikale eienskap* van belang. Die grammatikale eienskappe van 'n korpus hou verband met die gebruik van grammatikale woorde in daardie korpus. Grammatikale woorde is woorde wat gebruik word om 'n teks te struktureer. Dit sluit in lidwoorde, voegwoorde, voorsetsels en voornaamwoorde. Die gebruik van hierdie woordsoorte is gewoonlik onder die bewussynsvlak van sprekers of skrywers (Kotzé 2007:390). Leksikale eienskappe van 'n korpus het te doen met die inhoud van die korpus. Die onderwerp van 'n teks of bespreking bepaal die gebruik van die leksikale woorde. Voorbeelde van leksikale woorde is selfstandige naamwoorde, byvoeglike naamwoorde, bywoorde en werkwoorde.

Deur die gebruik van sleutelwoordontledings en konkordansiesoektogte is vasgestel watter grammatikale en leksikale eienskappe moontlik op verskille tussen die bepaalde korpora dui. Hierdie eienskappe is dan in kategorieë verdeel, soos in 4.1 verduidelik word.

4.1 Sleutelwoordontleding

Sleutelwoordontleding word in hierdie studie gebruik omdat dit die statistiese beduidendheid van spesifieke woorde in 'n korpus aandui. Die sleutelwoorde vir 'n bepaalde korpus word bepaal deur 'n doelkorpus met 'n verwysingskorpus te vergelyk. So 'n vergelyking dui aan

hoe statisties beduidend spesifieke woorde in een korpus (die doelkorpus) is in vergelyking met 'n ander korpus (die verwysingskorpus). Hierdie statistiese beduidendheid dui met ander woorde die sleutelwaarde (“keyness”) aan. In Antconc word die sleutelwaarde aangedui deur middel van die log-waarskynlikheid-statistiek. Die woorde in die doelkorpus word gelys vanaf die hoogste beduidende verskil. In statistiek is die nulhipotese altyd dat daar geen verskil tussen die bepaalde veranderlikes is nie – enige verskil word toegeskryf aan toeval. Die p-waarde (veral bekend in Chi²-toetse) word gebruik om te bepaal hoe beduidend verskille tussen twee veranderlikes is en of die nulhipotese verwerp kan word of nie. ('n Nulhipotese word nooit aanvaar nie – 'n navorser kan 'n nulhipotese slegs verwerp of dit nie verwerp nie.) Die log-waarskynlikheid-lesings in Antconc stem ooreen met spesifieke p-waardes. In die sosiale wetenskappe word die vlak van waarskynlikheid dat die resultaat slegs toevallig voorgekom het, op $p < 0,05$ gestel. Enige lesing in Antconc onder 3,84 beteken $p > 0,05$, en dit beteken dat die navorser nie die nulhipotese kan verwerp nie. In sodanige geval beteken dit dat die spesifieke woord nie 'n statisties beduidende verskil tussen die KMAD en KVAD aandui nie, en dat enige verskil wat wel waargeneem word, bloot toeval is.

Net sleutelwoorde wat 'n statisties beduidende verskil van 3,84 of hoër tussen die KMAD en KVAD aandui en verteenwoordigend is van die korpus of subkorpus, is ontleed. Die leksikale woorde is volgens konseptuele kategorieë gegroepeer om aan te dui dat die gedigte in die bepaalde korpora verskillende leksikale eienskappe het. Dieselfde woord kan in meer as een kategorie geplaas word afhangende van die konteks waarin dit gebruik word, maar nie alle woorde kan binne 'n kategorie geplaas word nie. Die grammatikale eienskappe is apart oorweeg en bespreek aangesien dit nie op dieselfde manier as die leksikale woorde gekategoriseer kan word nie. Hierdie eienskappe word in deel 5 verder bespreek.

4.2 Konkordansiesoektogte

Die woordsoort van 'n bepaalde woord (leksikaal of grammatikaal) speel 'n belangrike rol wanneer hierdie woord op spesifieke maniere gegroepeer word. McIntyre en Walker (2010:517) beklemtoon die feit dat die sleutelwaarde van woorde net die beginpunt vir verdere ondersoek is en dat die navorser moet bepaal of sleutelwoorde werklik beduidend is “by exploring in more detail each key item”.

Die woord *verf* kan byvoorbeeld as selfstandige naamwoord of werkwoord gebruik word en *weer* kan 'n selfstandige naamwoord of bywoord of werkwoord wees. In die geval van *verf* sou mens die woord dalk in een konseptuele kategorie kon plaas al word die woord as twee woordsoorte in verskillende kontekste gebruik, aangesien albei woordsoorte met dieselfde konsep te make het. In die geval van *weer* is die drie betekenisse egter glad nie verwant nie, en sou die konteks van die woord bepaal moet word om te besluit of die woord byvoorbeeld saam met *wind* of saam met *opnuut* gegroepeer word. Alhoewel *weer* nie as werkwoord in enige van die bestudeerde korpora gebruik word nie, sou dit moontlik saam met ander werkwoorde (soos *afwend* en *teëhou*) gegroepeer kon word wat 'n soortgelyke konseptuele kategorie voorstel.

Daar is twee maniere om die konteks van woorde en woordsoorte te bepaal. Een is om konkordansiesoektogte uit te voer, en 'n ander is om tekste te annoteer. Die annoteer van tekste kan met die hand gedoen word, maar vir groot hoeveelhede teks moet rekenaarsagteware gebruik word. Vir Afrikaanse korpora is daar annoteerders beskikbaar, maar hulle is nie noodwendig akkuraat vir korpora wat uit gedigte bestaan nie. Die rede

hiervoor is dat die annoteerders slegs korpora wat aan spesifieke vereistes voldoen, kan verwerk. In geval van die twee Afrikaanse POS- (Parts of Speech-) annoteerders wat in hierdie studie gebruik is,¹ moet die data óf as een woord per reël ingelees word óf as volledige sinne met die korrekte interpunksie.

Die probleem met poësie is nie net die struktuur van die tekste nie, maar ook die gebruik van puntuasie, aangesien verskeie digters geen puntuasie of hoofletters gebruik nie. Alhoewel die NCHLT POS-annoteerder 'n algemene puntetelling van 95,71% behaal het (Eiselen en Puttkammer 2014:3702), was die data in hierdie studie steeds problematies. Puttkammer (persoonlike kommunikasie 2018) het bevestig dat die data in hierdie korpus tot probleme en foutiewe annotasies kan lei. Om hierdie rede is besluit om 'n konkordansiesoektog uit te voer op elke woord wat 'n statistiese beduidendheid van 3,84 of hoër toon en verteenwoordigend is van 'n korpus of subkorpus. Om te bepaal of 'n woord verteenwoordigend is, is die konkordansieskets- ("concordance plot") funksie in Antconc gebruik. Hierdie funksie dui aan in hoeveel van die korpora of subkorpora 'n woord (of soekterm) voorkom. Die frekwensie van die woord word aangedui asook die aantal subkorpora waarin die spesifieke soekterm teenwoordig is. In die geval van die soekterm *mond* in die KMAD word aangedui dat die woord in 10 uit die 11 subkorpora voorkom, en daarom verteenwoordigend is van die korpus as geheel.

'n Konkordansiesoektog dui die konteks van 'n woord aan deur die spesifieke soekterm en 'n bepaalde aantal woorde voor of na die soekterm aan te dui. In Antconc word die soekterm in blou aangedui. Figuur 1 is 'n voorbeeld van 'n konkordansiesoektog vir die selfstandige naamwoord *mond*.

Hit	KWIC	File
1	sy onbenutte brein borrel by sy mond en ore uit sy penis gedeel	COPP.Aan 'n
2	odwendig af oor die gesig na sy mond saf van te veel liefde affor	COPP.Aan 'n
3	van are voed ek jou met my mond skut jou teen die dood me	COPP.Aan 'n
4	maar afgaan en jou voël in sy mond neem kan maar jou gat le	COPP.Aan 'n
5	ngevleg op jou divan 'n hand 'n mond lywe 'n tong duisendvouc	COPP.Aan 'n
6	moet toekyk hoe batman robin mond toe neem oorgretig slagg	COPP.Aan 'n
7	'n mangesig die skadu om die mond die engelhare fyn ingepot	COPP.Aan 'n
8	am met sy drinkerspele was sy mond 'n onvulbare (O) Om jou	COPP.Aan 'n
9	uning gebruik die tuinhok In my mond was dit so soet soos heun	COPP.Aan 'n
10	u maag bitter maak, maar in jou mond sal dit so soet soos heunir	COPP.Aan 'n
11	en begeertes leerstellig lê in die mond van iemand anders? Ek w	COPP.Braille
12	fs buitelands gegradeerde wat mond oopmaak en waarsku: Dis	COPP.Braille
13	in veiligheid. 'n Slot is voor elke mond geplaas, dempers op die p	COPP.Braille
14	ande. (Ek hou voorlopig nou my mond om só die stiltes te bewaa	COPP.Die Af

Search Term Words Case Regex Search Window Size 50

mond

Kwic Sort Level 1 1R Level 2 2R Level 3 3R

Figuur 1. 'n Voorbeeld van die resultate van 'n konkordansiesoektog

Die resultate van die sleutelwoordontledings en konkordansiesoektogte word in die volgende deel bespreek.

5. Resultate en bespreking

Die resultate vir die ontledings word apart bespreek. Die vergelykings met die PAP as doelkorpus word voor die K MAD en K VAD as aparte doelkorpora bespreek.

5.1 Sleutelwoordontleding: die PAP

Soos aangedui, word die PAP in drie vergelykings as doelkorpus gebruik. Die rede hiervoor is om vas te stel hoe die PAP van algemene Afrikaanse tekste (KAT), algemene letterkunde (KAL) en ander poësie (KAP) verskil.

5.1.1 Die PAP en die KAT

Die sleutelwoordontleding van hierdie vergelyking dui op 'n maksimumsleutelwaarde van 3114,193. 'n Sleutelwaarde van 15,13 ($p < 0,0001$) of hoër dui reeds aan dat daar slegs 'n 0,01%-kans is dat enige verskille toevallig is, en dat daar met ander woorde 'n beduidende verskil tussen twee korpora is. In hierdie geval is so 'n resultaat te verwagte. Poësie as 'n genre verskil tog noemenswaardig van ander genres, nie net wat struktuur betref nie, maar ook taalgebruik. In die poësie word onder andere aandag gegee aan ritme, klank, rym en strofe-indelings (De Lange 2000:19), iets wat nie in ander genres, soos nuusberigte of akademiese artikels, voorkom nie.

5.1.1.1 Die PAP en die KAT: grammatikale eienskappe

Die voornaamwoord “jou” het die hoogste sleutelwaarde (3114,193), gevolg deur “jy” (1314,007) en “my” (1095,592). Die voornaamwoorde “sy” (754,796) en “haar” (669,931) lê ook hoog op die lys. Ander voornaamwoorde en onbepaalde voornaamwoorde wat 'n beduidende verskil aandui, is “jul” (235,169), “julle” (214,707), “hom” (165,284), “hy” (109,672), “alles” (97,546), “hul” (72,142), “niemand” (47,359), “wie” (32,092) “mekaar” (28,386), “onse” (24,232) en “syne” (18,380).

Die resultate dui nie net daarop dat die skrywers in die PAP oor die algemeen van meer voornaamwoorde gebruik maak nie, maar ook dat die korpus spesifieke voornaamwoorde met 'n hoë frekwensie gebruik. “Jou”, “jy” en “my” word veral frekvent gebruik, wat moontlik aandui dat die skrywers nie net 'n sterk fokus op ander plaas nie (veral die tweede persoon), maar ook die besit van iets of iemand beklemtoon: “jou”, “my” asook “sy” en “haar” sowel as “onse” en “syne”.

Voegwoorde word ook algemeen in PAP gebruik. Die voegwoord “soos” is veral populêr en vertoon 'n sleutelwaarde van 280,015. Die hoë frekwensie van hierdie voegwoord in die PAP dui daarop dat die skrywers in die PAP graag vergelykings tref. Dit is 'n moontlike eienskap van die poësie in die algemeen (vergelyk resultate onder 5.1.2.1).

5.1.1.2 Die PAP en die KAT: die leksikale eienskappe

Om die leksikale eienskappe van ’n bepaalde korpus aan te dui, is besluit om die woorde wat ’n sleutelwaarde van 3,84 of hoër het en verteenwoordigend van die korpus is, in kategorieë te verdeel. Hierdie kategorieë word nie noodwendig slegs op konkrete konsepte gebaseer nie, en kan ook op abstrakte konsepte gebaseer word. Tabel 3 bevat die leksikale eienskappe van die PAP wanneer dit met die KAT vergelyk word. Die gebruik van hakies by kategorie 2 dui aan watter woorde met een sintuig geassosieer kan word. Dit beteken byvoorbeeld dat “sien”, “kyk” en “staar” al drie met sig te make het.

Tabel 3. Leksikale eienskappe van die PAP/KAT-vergelyking

Kategorie	Leksikale woorde
Kategorie 1: Liggaam en liggaamsdele	Selfstandige naamwoorde: oë, lyf, hande, mond, bloed, hart, vingers, hand, hare, voete, tong, arms, kop, vel, gesig, rug, lippe, oog, nek, bors, ken, skoot, voël, borste, bene, keel, lywe, voet, ore, vlees, neus, gesigte, vinger, oor, liggaam
Kategorie 2: Sintuie	Selfstandige naamwoorde: (geur, reuk), (smaak), (klank) Werkwoorde: (sien, kyk, staar) (ruik) (lek, drink, eet, smaak), (streel, voel, raak) (hoor, klink, luister) Byvoeglike naamwoorde: sagte Bywoord: sag
Kategorie 3: Kognisie	Selfstandige naamwoorde: herinnering, gedagtes, geheue, droom, drome Werkwoorde: wonder, vergeet, onthou, weet, ken, vermoed, droom
Kategorie 4: Verlange	Werkwoorde: mis, verlang
Kategorie 5: Vrees	Selfstandige naamwoorde: vrees Byvoeglike naamwoord: bang
Kategorie 6: Blydschap	Selfstandige naamwoorde: lag, glimlag, geluk Werkwoorde: lag, glimlag
Kategorie 7: Hartseer	Selfstandige naamwoorde: pyn, transe Werkwoorde: huil, pyn
Kategorie 8: Liefde/liefkosing	Selfstandige naamwoorde: liefde, soen, lief, geliefde Werkwoorde: druk, soen, vry Byvoeglike naamwoorde: lief Bywoorde: lief
Kategorie 9: Handeling	Werkwoorde: soek, breek, roep, vat, hang, sny, dra, keer, rook, slaan, gryp, vang, wys, ry, stuur, speel, tel, druk, was, lê, val, loop, draai, dans, buig, roer, stoot, trap, vlieg, staan, kruip, pluk, stap, sit, vou, swaai, sak, glip, spring, skeur, kap, skuif, ruk, klim, steek, gooi, pak, bind, blaas, stort
Kategorie 10: Lewe	Selfstandige naamwoorde: lewe Werkwoorde: leef, lewe Byvoeglike naamwoorde: gebore
Kategorie 11: Die dood	Selfstandige naamwoorde: dood, graf, lyk Werkwoorde: sterf Bywoorde: dood

Kategorie 12: Religieuse konsepte	Selfstandige naamwoorde: hemel, kerk, kruis, engele, engel, gees, genade Eiename: God, Here
Kategorie 13: Man-wees	Selfstandige naamwoorde: pa, seun, man
Kategorie 14: Vrou-wees	Selfstandige naamwoorde: vrou, ma, moeder, vroue
Kategorie 15: Familie	Selfstandige naamwoorde: ma, kind, pa, seun, moeder
Kategorie 16: Kommunikasie	Selfstandige naamwoorde: stem, stemme, woorde Werkwoorde: sê, gesê, fluister, groet
Kategorie 17: Woning/tuiste/geboue	Selfstandige naamwoorde: tuin, werf, mure, muur, kamer, vertrek, venster, deure, deur, trap Bywoord: tuis
Kategorie 18: Voedsel	Selfstandige naamwoorde: honger, wyn, brood, melk, vrug, vrugte Werkwoorde: drink, eet
Kategorie 19: Slaap	Selfstandige naamwoorde: slaap, droom, rus, drome Werkwoorde: slaap, droom, rus Byvoeglike naamwoorde: wakker, moeg Bywoorde: moeg
Kategorie 20: Rus	Selfstandige naamwoorde: rus Werkwoorde: rus Byvoeglike naamwoorde: moeg Bywoorde: moeg
Kategorie 21: Kleur	Selfstandige naamwoorde: kleur, wit, geel, groen, bruin, swart Byvoeglike naamwoorde: blou, wit, rooi, geel, groen, grys, bruin, swart, pers, goue Bywoorde: blou, wit, rooi, swart
Kategorie 22: Digkuns	Selfstandige naamwoorde: digter, gedig Werkwoorde: dig
Kategorie 23: Lig en donker	Selfstandige naamwoorde: skadu, ligte, donker Werkwoord: gloei Byvoeglike naamwoorde: lig, helder
Kategorie 24: Nag	Selfstandige naamwoorde: nag, sterre, aand Byvoeglike naamwoorde: donker Bywoorde: snags
Kategorie 25: Beweging	Werkwoorde: lê, val, loop, draai, bly, buig, roer, stoot, trap, vlieg, staan, stap, sit, swaai, spring, skeur, kap, kom, skuif, ruk, klim, gooi, stort, sny Byvoeglike naamwoorde: stil Bywoorde: stil, vas
Kategorie 26: Vloeistof en vloeibaarheid	Selfstandige naamwoorde: see, water, rivier, stroom, waters, sweet Werkwoorde: stroom, vloei, sweet, drup, dryf Byvoeglike naamwoorde: nat, klam
Kategorie 27: Reis	Selfstandige naamwoorde: weg, spore, reis, trek, pad, spoor Werkwoorde: trek, vertrek, reis
Kategorie 28: Afstand	Byvoeglike naamwoorde: nader, naby, ver
Kategorie 29: Posisie	Bywoorde: langs, agter, na, verby

Kategorie 30: Gewig	Byvoeglike naamwoorde: swaar
Kategorie 31: Temperatuur	Selfstandige naamwoorde: koue Byvoeglike naamwoorde: koel, warm Bywoord: koud
Kategorie 32: Die aarde	Selfstandige naamwoorde: aarde, wêreld
Kategorie 33: Die hemelruim	Selfstandige naamwoorde: sterre, maan, hemel, wolke, lug, son
Kategorie 34: Die natuur	Selfstandige naamwoorde: blare, bome, klip, stof, sand, berg, berge, boom, gras, blomme, lug, rivier, blom, saad, stroom, plant, heuwels, vrug, bos, mos, grond, veld, vrugte, son, reën, wind, wolke, mis Werkwoorde: blom, pluk, plant Byvoeglike naamwoorde: wilde
Kategorie 35: Die weer en weersomstandighede	Selfstandige naamwoorde: son, reën, wind, wolke, weer, koue, mis Werkwoorde: reën
Kategorie 36: Diere	Selfstandige naamwoorde: voëls, vlerke, voël, hond Byvoeglike naamwoord: wilde
Kategorie 37: Berge	Selfstandige naamwoorde: berg, berge, hang
Kategorie 38: Vuur	Selfstandige naamwoorde: vuur, vlam, brand, rook Werkwoorde: vlam, brand
Kategorie 39: Tyd	Selfstandige naamwoorde: môre, skemer, dag, uur, dae, jare, tyd, seisoene, winter, somer Werkwoorde: duur Byvoeglike naamwoorde: lang, lank, daaglik Bywoorde: saans, soggens, bedags, snags, altyd, later, nou, skielik, nooit, dan

Uit die tabel is dit duidelik dat daar 39 kategorieë (of leksikale eienskappe) in die PAP geïdentifiseer kan word. Wanneer die PAP met die KAT vergelyk word, blyk dit dat die skrywers in die PAP sekere leksikale woorde met 'n beduidend statistiese verskil gebruik, byvoorbeeld leksikale woorde wat na die liggaam en liggaamsdele verwys, leksikale woorde wat beweging aandui, leksikale woorde wat na die natuur verwys, ens. In sommige gevalle word die leksikale kategorie deur grammatikale eienskappe ondersteun, soos die kategorie “Posisie” wat ook deur voorsetsels in die PAP aangedui word.

Die hipotese is dat die verskille tussen die PAP en die KAT groter sal wees as tussen die PAP en die KAL of die PAP en die KAP. Dieselfde kategorieë sal ook in al drie hierdie vergelykings voorkom, met enkele verskille. Die resultate van die twee laasgenoemde vergelykings word volgende bespreek.

5.1.2 Die PAP en die KAL

Soos vermoed, vertoon die PAP 'n kleiner maksimumsleutelwaarde wanneer die korpus met die KAL vergelyk word (551,530). Dit kan moontlik toegeskryf word aan die feit dat die KAL ook uit letterkundige tekste bestaan en daarom aan 'n verwante genre behoort. Desnietemin is daar steeds beduidende verskille tussen die PAP en die KAL.

5.1.2.1 Die PAP en die KAL: grammatikale eienskappe

Weer eens speel veral die **voornaamwoorde** 'n belangrike rol om die PAP van die KAL te onderskei, maar daar is in die PAP/KAL-vergelyking minder voornaamwoorde wat as sleutelwoorde geïdentifiseer word as in die PAP/KAT-vergelyking. Dit dui daarop dat die skeppers van die tekste in KAL meer voornaamwoorde gebruik as die KAT.

In hierdie geval is “my” die voornaamwoord met die hoogste sleutelwaarde (551,530), gevolg deur “jou” (508,100) en “ek” (164,272). “Ek” verskyn nie met 'n beduidend statistiese waarde in die PAP/KAT-vergelyking nie. Dit kan aandui dat daar in die PAP/KAL-vergelyking 'n groter klem op die skrywer self is. Die voornaamwoorde “jul” (86,346), “jy” (24,784) en “alles” (9,887) kom met laer sleutelwaardes in die PAP/KAL-vergelyking voor. Dit dui daarop dat hierdie voornaamwoorde in die KAL 'n hoër frekwensie het as in die KAT.

Die voornaamwoorde “sy”, “julle”, “hom”, “hy” en “mekaar” word nie as beduidende sleutelwoorde gelys nie. Dit versterk die vermoede dat die skeppers van die tekste in KAL van meer voornaamwoorde gebruik maak as in die KAT. Die voornaamwoord “ons” (7,615) verskyn op die sleutelwoordlys van die PAP/KAL-vergelyking, maar “onse” nie. Dit wil dus voorkom of die skrywers van die PAP in albei vergelykings graag voornaamwoorde gebruik wat inklusiwiteit aandui, maar dat die voornaamwoord “onse” ook gereeld deur die skrywers van die KAL gebruik word.

Weer eens is “soos” die **voegwoord** met die hoogste sleutelwaarde (145,391). Dit bevestig die teorie dat digters in die PAP, en moontlik digters in die algemeen, gereeld van hierdie voegwoord gebruik maak.

5.1.2.2 Die PAP en die KAL: die leksikale eienskappe

Daar is slegs enkele leksikale woorde wat in die PAP/KAT- en die PAP/KAL-vergelykings verskil. Hierdie woorde is:

Selfstandige naamwoorde: taal, huise, paaie, boek, tafel, bestaan

Eienaam: Afrika

Werkwoorde: kleef, waai, lees, sluit, woon, herinner, bestaan

Bywoord: buite.

Die meeste van hierdie woorde hou verband met die leksikale eienskappe wat reeds in 5.1.1.2 geïdentifiseer is, en kan by die bestaande kategorieë in tabel 4 ingedeel word. Verder is daar ook sekere leksikale woorde wat nie as sleutelwoorde in die PAP/KAL-vergelyking gelys word nie en daarom word die kategorieë effens aanpas. Die volledige kategorieë van die PAP/KAL-vergelyking word in tabel 4 uiteengesit.

Tabel 4. Leksikale eienskappe van die PAP/KAL-vergelyking

Kategorie	Leksikale woorde
Kategorie 1: Liggaam en liggaamsdele	Selfstandige naamwoorde: oë, lyf, hande, mond, bloed, vingers, hand, hare, voete, tong, arms, vel, rug, lippe, oog, nek, bors, voël, borste, lywe, voete, ore, vlees, gesigte, oor, liggaam
Kategorie 2: Sintuie	Selfstandige naamwoorde: (geur), (smaak), (klank) Werkwoorde: (staar), (lek, drink, smaak) (streel, raak) Byvoeglike naamwoorde: sagte Bywoord: sag
Kategorie 3: Kognisie	Selfstandige naamwoorde: herinnering, gedagtes, geheue, droom, drome Werkwoorde: wonder, vergeet, onthou, ken, vermoed, droom, herinner
Kategorie 4: Hartseer	Selfstandige naamwoorde: pyn, trane Werkwoorde: pyn
Kategorie 5: Liefde/liefkosing	Selfstandige naamwoorde: liefde, soen, lief, geliefde Werkwoorde: soen, vry Byvoeglike naamwoorde: lief Bywoorde: lief
Kategorie 6: Handeling	Werkwoorde: soek, breek, hang, sny, dra, rook, speel, was, lê, val, dans, buig, roer, stoot, trap, vlieg, kruip, pluk, vou, swaai, glip, kap, skuif, stort, lees waai, sluit
Kategorie 7: Lewe	Selfstandige naamwoorde: lewe Werkwoorde: leef, lewe Byvoeglike naamwoorde: gebore
Kategorie 8: Die dood	Selfstandige naamwoorde: dood Werkwoorde: sterf Bywoorde: dood
Kategorie 9: Religieuse konsepte	Selfstandige naamwoorde: hemel, kerk, kruis, engele, engel, gees, genade Eienaam: God Byvoeglike naamwoord: heilige
Kategorie 10: Vrou-wees	Selfstandige naamwoorde: vrou, moeder, vroue
Kategorie 11: Familie	Selfstandige naamwoorde: ma, kind, moeder
Kategorie 12: Kommunikasie	Selfstandige naamwoorde: stemme, woorde, taal, boek Werkwoorde: sê, gesê
Kategorie 13: Woning/tuiste/geboue	Selfstandige naamwoorde: tuin, werf, mure, muur, venster, deure, deur, trap, vensters, huise Werkwoorde: woon Bywoord: tuis
Kategorie 14: Voedsel	Selfstandige naamwoorde: wyn, brood, melk, vrug, vrugte Werkwoorde: drink
Kategorie 15: Slaap	Selfstandige naamwoorde: slaap, droom, rus, drome Werkwoorde: slaap, droom, rus
Kategorie 16: Rus	Selfstandige naamwoorde: rus Werkwoorde: rus

Kategorie 17: Kleur	Selfstandige naamwoorde: kleur, wit, geel, groen, bruin, swart Byvoeglike naamwoorde: blou, wit, rooi, geel, groen, grys, bruin, swart, pers, goue Bywoorde: blou, wit, rooi, swart
Kategorie 18: Digkuns	Selfstandige naamwoorde: digter, gedig Werkwoorde: dig
Kategorie 19: Lig en donker	Selfstandige naamwoorde: skadu, ligte, donker Werkwoord: gloei Byvoeglike naamwoorde: lig, helder Bywoord: ligte
Kategorie 20: Nag	Selfstandige naamwoorde: nag, sterre Byvoeglike naamwoorde: donker Bywoorde: snags
Kategorie 21: Beweging	Werkwoorde: lê, val, bly, buig, roer, stoot, trap, vlieg, swaai, kap, skuif, stort, sny
Kategorie 22: Vloeistof en vloeibaarheid	Selfstandige naamwoorde: see, water, rivier, stroom, waters, sweet Werkwoorde: stroom, vloei, sweet, drup, dryf Byvoeglike naamwoorde: klam
Kategorie 23: Reis	Selfstandige naamwoorde: spore, reis, paaie Werkwoorde: trek, reis
Kategorie 24: Posisie	Bywoorde: na, verby
Kategorie 25: Gewig	Byvoeglike naamwoorde: swaar
Kategorie 26: Temperatuur	Selfstandige naamwoorde: koue Byvoeglike naamwoorde: koel, warm Bywoord: koud
Kategorie 27: Die aarde	Selfstandige naamwoorde: aarde, wêreld
Kategorie 28: Die hemelruim	Selfstandige naamwoorde: sterre, maan, hemel, wolke, lug, son
Kategorie 29: Die natuur	Selfstandige naamwoorde: blare, bome, klip, stof, sand, berg, berge, boom, gras, blomme, lug, rivier, blom, saad, stroom, plant, heuwels, vrug, grond, vrugte, son, reën, wind, wolke Werkwoorde: blom, pluk, plant Byvoeglike naamwoorde: wilde
Kategorie 30: Die weer en weersomstandighede	Selfstandige naamwoorde: son, reën, wind, wolke, weer, koue Werkwoorde: reën
Kategorie 31: Diere	Selfstandige naamwoorde: voëls, vlerke, voël Byvoeglike naamwoord: wilde
Kategorie 32: Berge	Selfstandige naamwoorde: berg, berge, hang
Kategorie 33: Vuur	Selfstandige naamwoorde: vlam, brand, rook Werkwoorde: vlam, brand
Kategorie 34: Tyd	Selfstandige naamwoorde: môre, skemer, uur, jare, tyd, seisoene, winter, somer Byvoeglike naamwoorde: lank, daaglik Bywoorde: saans, soggens, bedags, snags, altyd

Tabel 4 bevat 34 kategorieë teenoor die 39 kategorieë wat in tabel 3 geïdentifiseer is, maar die resultate ondersteun die redenasie dat sekere kategorieë konstant sal bly selfs wanneer die PAP met 'n ander korpus vergelyk word.

5.1.3 Die PAP en die KAP

Soos voorspel, toon die PAP/KAP-vergelyking die kleinste sleutelwaardes. Dit is te verwagte aangesien albei korpora poësietekste bevat en daarom eienskappe gemeen sal hê wat eie is aan die poësiegenre. Die hoogste sleutelwaarde hier is 161,402 (in vergelyking met 3114,193 in die PAP/KAT-vergelyking en 551,530 in die PAP/KAL-vergelyking).

5.1.3.1 Die PAP en die KAP: die grammatikale eienskappe

Die opvallendste eienskap is die feit dat die hoogste sleutelwaardes nie deur voornaamwoorde gedomineer word nie. “Hul” (117,010), “sy” (68,022) en “hy” (61,547) het die hoogste sleutelwaardes ten opsigte van voornaamwoorde, gevolg deur “haar” (32,673), “jul” (27,923), “julle” (26,827), “hom” (25,704), “wie” (19,705) en “niemand” (12,187). “Mekaar”, “ons”, “alles”, “iets”, “syne”, “onse” en “jy” word nie as sleutelwoorde wat 'n beduidende verskil aandui, gelys nie. Die feit dat daar weer eens verskeie voornaamwoorde is wat nie gelys word nie (in vergelyking met die PAP/KAT-lys), dui daarop dat die skrywers van die PAP en die KAP soortgelyke frekwensies van sekere voornaamwoorde gebruik.

Die voornaamwoord “hulle” word in die PAP/KAP-vergelyking met 'n sleutelwaarde van 18,655 gelys, wat aandui dat alhoewel die skrywers in die KAP ook van hierdie voornaamwoord gebruik maak, die frekwente gebruik daarvan in die PAP steeds 'n moontlike grammatikale eienskap is.

Die bepaalde lidwoord “die” (118,434) word vir die eerste keer in die PAP/KAP-vergelyking met 'n sleutelwaarde gelys wat 'n beduidende verskil aandui. Die onbepaalde lidwoord “'n” kom nie in die sleutelwoordlys voor nie. Die gevolgtrekking kan dus gemaak word dat die gebruik van spesifieke verwysings 'n moontlike kenmerk van die gedigte in die PAP is.

Die **voegwoord** “soos” word nie as sleutelwoord gelys nie, en dit bevestig die vermoede dat dit nie net 'n eienskap van poësie in die PAP is nie, maar ook van poësie in die algemeen. Die gebruik van hierdie voegwoord in die twee korpora stem met ander woorde in so 'n mate ooreen dat dit nie 'n eienskap is wat die PAP van die KAP kan onderskei nie. Die voegwoord “en” is wel 'n voegwoord wat die PAP en KAP van mekaar kan onderskei en word met 'n sleutelwaarde van 28,995 gelys.

5.1.3.2 Die PAP en die KAP: die leksikale eienskappe

Omdat die PAP en die KAP nie soveel van mekaar verskil soos die PAP/KAT- of PAP/KAL-vergelykings nie, is daar veel minder leksikale eienskappe wat die PAP en die KAP van mekaar onderskei. Tabel 5 bevat die leksikale eienskappe van die PAP/KAP-vergelyking.

Tabel 5. Leksikale eienskappe van die PAP/KAP-vergelyking

Kategorie	Leksikale woorde
Kategorie 1: Liggaam en liggaamsdele	Selfstandige naamwoorde: oë, vingers, arms, keel
Kategorie 2: Sintuie	Selfstandige naamwoorde: (geur) Werkwoorde: (lek, drink, eet) (streele) Byvoeglike naamwoorde: sagte
Kategorie 3: Handeling	Werkwoorde: was, lê, stoot
Kategorie 4: Religieuse konsepte	Selfstandige naamwoorde: hemel, kerk, engele, engel, gees, genade, gode Eienaam: Here
Kategorie 5: Familie	Selfstandige naamwoorde: ma, pa, seun
Kategorie 6: Verbale kommunikasie	Werkwoorde: sê, gesê
Kategorie 7: Voedsel	Selfstandige naamwoorde: vrug, vrugte Werkwoorde: drink, eet
Kategorie 8: Slaap	Selfstandige naamwoorde: droom Werkwoorde: droom Byvoeglike naamwoorde: moeg Bywoorde: moeg
Kategorie 9: Kleur	Selfstandige naamwoorde: groen, bruin, swart Byvoeglike naamwoorde: groen, bruin, swart Bywoorde: swart
Kategorie 10: Donkerte	Selfstandige naamwoorde: nag Byvoeglike naamwoorde: donker
Kategorie 11: Beweging	Werkwoorde: lê, stoot
Kategorie 12: Vloeistof	Selfstandige naamwoorde: rivier Byvoeglike naamwoorde: nat
Kategorie 13: Die aarde	Selfstandige naamwoorde: aarde, wêreld
Kategorie 14: Die natuur	Selfstandige naamwoorde: bome, berg, berge, gras, rivier, vrug, grond, vrugte, reën, wind Byvoeglike naamwoorde: wilde
Kategorie 15: Die weer en weersomstandighede	Selfstandige naamwoorde: reën, wind Werkwoorde: reën
Kategorie 16: Diere	Selfstandige naamwoorde: voëls, hond Byvoeglike naamwoord: wilde
Kategorie 17: Berge	Selfstandige naamwoorde: berg, berge
Kategorie 18: Vlam	Selfstandige naamwoorde: vlam Werkwoorde: vlam

Wanneer die PAP met die KAP vergelyk word, kan 18 leksikale eienskappe geïdentifiseer word. Sommige van hierdie eienskappe stem ooreen met die eienskappe wat in die PAP/KAT- en PAP/KAL-vergelykings voorkom, terwyl sommige kategorieë effens aangepas is. Hierdie 18 kategorieë bevestig dat die PAP spesifieke leksikale eienskappe het wat dit van 'n verskeidenheid ander korpora onderskei. Op grond van hierdie resultate kan daar afgelei word dat die skrywers in die PAP onder andere graag na die liggaam en liggaamsdele asook

sekere sintuie verwys. Hulle sluit gereelde verwysings na sekere handeling in en verwys graag na religieuse konsepte. Verwysings na “kleur” en “die natuur” kan ook met hierdie skrywers geassosieer word.

Die gedigte deur die skrywers in die PAP vertoon met ander woorde spesifieke grammatikale en leksikale eienskappe wat hierdie korpus van ander korpora onderskei. Die teorie dat hierdie eienskappe moontlik met gender verbind kan word, word in die volgende deel bespreek.

5.2 Sleutelwoordontleding: Die KMAD/KVAD- en die KVAD/KMAD-vergelykings

Om vas te stel of daar sekere grammatikale of leksikale eienskappe is wat spesifiek met die poësie van manlike of vroulike digters geassosieer kan word, moet die poësie van manlike en vroulike digters met mekaar vergelyk word. Die enigste korpus beskikbaar vir so ’n vergelyking waarvan die gedigte met ’n bepaalde gender verbind word, is die Proteakorpus. Soos in deel 3 aangedui, is die PAP verdeel in twee subkorpora met die tekste van 10 digbundels elk deur manlike en vroulike digters. Om te bepaal watter woorde ’n statistiese beduidendheid in die manlike of vroulike korpus aandui, moet hierdie twee subkorpora met mekaar vergelyk word.

Die algemene metode in ’n sleutelwoordontleding is om ’n kleiner, gespesialiseerde korpus met ’n groter, meer algemene korpus, te vergelyk (Hunston 2002:68). Dit is die metode wat vir die vier voorgaande sleutelwoordontledings gevolg is. Aangesien daar egter net twee subkorpora beskikbaar is om te vergelyk, en aangesien die doel is om te bepaal of die taalgebruiksverskille aan gender gekoppel kan word, word die KMAD en KVAD met mekaar vergelyk.

5.2.1 Die KMAD/KVAD- en die KVAD/KMAD-vergelykings: grammatikale eienskappe

Die eerste twee woorde in die sleutelwoordlys wat die KMAD van die KVAD onderskei, is voornaamwoorde. “Hul” (88,615) en “jul” (48,969) kom met die twee hoogste sleutelwaardes voor, gevolg deur “niemand” (17,644). “Hom” (10,415), “julle” (8,686), “hulle” (6,837) en “hy” (6,758) verskyn ook op die lys. Voornaamwoorde wat na die skrywer self verwys (“ek” of “my”) word egter nie gelys nie. Dit suggereer dat die manlike digters in hierdie subkorpus met ’n hoë frekwensie na ander individue verwys in hul gedigte.

By die KVAD/KMAD-vergelyking word die voornaamwoord “u” met ’n sleutelwaarde van 57,537 gelys. “Als” (12,763) word ook gelys. Die geringe getal voornaamwoorde in die KVAD/KMAD-vergelyking kan daarop dui dat die vroulike skrywers in hierdie subkorpus oor die algemeen van minder voornaamwoorde gebruik maak as die manlike skrywers.

Die **bywoord van ontkenning** “nie” (alhoewel nie ’n grammatikale eienskap nie) word as sleutelwoord gelys wat die KMAD van die KVAD onderskei. Hierdie woord word nie in die sleutelwoordlys van die KVAD gelys nie, en dit dui daarop dat die skrywers in die KMAD meer gereeld van ontkenning gebruik maak.

Die voegwoord “soos” (15,624) word as sleutelwoord by die KVAD/KMAD-vergelyking gelys, maar nie by die KMAD/KVAD-vergelyking nie. Dit is moontlik om die afleiding te maak dat hierdie voegwoord daarom meer gereeld deur vroulike skrywers in die KVAD

gebruik word, en dat hulle moontlik meer vergelykings in hul poësie tref as manlike skrywers.

Die grammatikale eienskappe wat in hierdie studie geïdentifiseer is, stem nie noodwendig ooreen met die resultate van vorige studies in taalgebruiksverskille en gender nie. Die rede hiervoor is moontlik dat hierdie korpora meer spesifiek is (dit bestaan net uit poësie) en ook veel kleiner is as die korpora wat in ander studies gebruik is. Desnietemin bied die resultate insigte in die eienskappe van manlike en vroulike skrywers in hierdie subkorpora, en kan dit steeds tot insigte in potensiele genderverwante taalgebruik in Afrikaanse poësie lei.

Die leksikale eienskappe van die KMAD en KVAD word volgende bespreek.

5.2.2 Die KMAD/KVAD- en die KVAD/KMAD-vergelykings: leksikale eienskappe

Soos in die voorgaande ontledings, is die leksikale eienskappe van die KMAD en KVAD in kategorieë verdeel. Die kategorieë vir die KMAD word in tabel 6 aangedui.

Tabel 6. Leksikale eienskappe van die KMAD/KVAD-vergelyking

Kategorie	Leksikale woorde
Kategorie 1: Liggaam en liggaamsdele	Selfstandige naamwoorde: voete, lippe, hart, arms, dye
Kategorie 2: Sintuie	Selfstandige naamwoorde: smaak Werkwoorde: (smaak), (loer), (strel) Byvoeglike naamwoorde: (suur), (lou, koue), (ruik)
Kategorie 3: Kommunikasie	Selfstandige naamwoorde: boodskap Werkwoorde: praat, vra, sê
Kategorie 4: Handeling	Werkwoorde: slaan, steek, vat, skop, praat, bring, staan, sit, vra, beur, gryp, klou
Kategorie 5: Religieuse konsepte	Selfstandige naamwoorde: gode, kerk, engele, genade Byvoeglike naamwoorde: heilige
Kategorie 6: Die dood	Selfstandige naamwoorde: dood Werkwoorde: sterf
Kategorie 7: Tyd	Bywoorde: vandag, nooit, nou

Sewe kategorieë is geïdentifiseer wat die moontlike leksikale eienskappe van die KMAD insluit. Hierdie kategorieë stem ooreen met die kategorieë wat in die voorgaande ontledings geïdentifiseer is, maar aangesien die KMAD deel uitmaak van die PAP, is so 'n resultaat akkuraat. Voor enige gevolgtrekkings gemaak kan word ten opsigte van 'n onderskeid tussen die KMAD en KVAD, moet die leksikale eienskappe van die KVAD ook oorweeg word. Hierdie eienskappe word in tabel 7 aangedui.

Tabel 7. Leksikale eienskappe van die KVAD/KMAD vergelyking

Kategorie	Leksikale woorde
Kategorie 1: Liggaam en liggaamsdele	Selfstandige naamwoorde: skouers, wang
Kategorie 2: Sintuie	Byvoeglike naamwoorde: koel

	Bywoorde: sag
Kategorie 3: Kommunikasie	Selfstandige naamwoorde: woorde, stem Werkwoorde: skryf, luister, vertel
Kategorie 4: Kognisie	Selfstandige naamwoorde: drome, droom Werkwoorde: droom, gedink, onthou
Kategorie 5: Kreatiewe prosesse / uitsette	Selfstandige naamwoorde: gedig, musiek, kwas Werkwoorde: skryf, vertel
Kategorie 6: Voedsel	Selfstandige naamwoorde: wyn, brood, melk Werkwoorde: drink
Kategorie 7: Woning/tuiste/geboue	Selfstandige naamwoorde: huis, deur, dak, tuin, venster, vertrek, vensters
Kategorie 8: Natuur	Selfstandige naamwoorde: bome, son, duine, wolke, rose, sterre, mis, reën, stroom
Kategorie 9: Vuur	Selfstandige naamwoorde: kerse, vuur Byvoeglike naamwoorde: brandende
Kategorie 10: Lig	Selfstandige naamwoorde: lig, kerse, flits, vuur
Kategorie 11: Kleur	Selfstandige naamwoorde: blou Byvoeglike naamwoorde: blou, bruin, silwer
Kategorie 12: Tyd	Selfstandige naamwoorde: klok, oggend, seisoen, winter Bywoorde: soggens, later, wanneer, daaglik
Kategorie 13: Reis	Selfstandige naamwoorde: reis, pad, spoor Werkwoorde: reis, vertrek

Weer eens stem die meeste kategorieë van KMAD ooreen met dié van die PAP-vergelykings. Daar is ook ooreenstemming tussen sommige kategorieë in die KVAD en KMAD. Nietemin kan gevolgtrekkings gemaak word ten opsigte van die vermoede dat taalgebruiksverskille in poësie ook aan gender gekoppel kan word.

Beide die KMAD en KVAD bevat die volgende kategorieë: *liggaam en liggaamsdele*, *sintuie*, *kommunikasie* en *tyd*. Die inhoud van hierdie kategorieë verskil egter van mekaar wanneer hul vergelyk word. Dit dui daarop dat alhoewel daar in die twee subkorpora dieselfde kategorieë geïdentifiseer kan word, die manlike en vroulike skrywers van verskillende leksikale woorde gebruik maak. Manlike skrywers het byvoorbeeld 'n meer uitgebreide kategorie wat *liggaam en liggaamsdele* asook *sintuie* betref, terwyl die vroulike skrywers weer 'n meer uitgebreide kategorie vir *tyd* het.

Daar is egter kategorieë wat nêr by manlike en nêr by vroulike skrywers aangetref word. Die manlike skrywers se poësie word gekenmerk deur leksikale woorde wat handeling en religieuse konsepte aandui, asook leksikale woorde wat met die dood te make het.

Die vroulike skrywers se poësie word gekenmerk deur leksikale woorde wat na kognisie verwys, kreatiewe prosesse beskryf, en met voedsel en die konsep van “tuiste” verband hou. Vroulike skrywers gebruik ook meer gereeld leksikale woorde wat na die natuur en na kleur en konsepte soos “vuur” en “lig” verwys.

Hierdie resultate ondersteun die oorspronklike teorie dat daar nie net tussen korpora onderskei kan word nie, maar ook tussen die eienskappe van manlike en vroulike poësie. Die resultate van hierdie studie word in deel 6 opsommend bespreek.

6. Gevolgtrekking

Die doel van die studie was eerstens om te bepaal of 'n korpusgebaseerde ontleding daarin kan slaag om die grammatikale en leksikale eienskappe van 'n poësiekorpus te identifiseer en om die bepaalde korpus van ander korpora te onderskei. Tweedens is vasgestel of daar tussen die poësie van manlike en vroulike skrywers in bepaalde subkorpora onderskei kan word. Die resultate van die huidige studie is positief. Nie net is dit moontlik om een genre van 'n ander te onderskei nie, maar dit is ook moontlik om die grammatikale en leksikale eienskappe van een spesifieke korpus te identifiseer. Verder is dit ook moontlik om die grammatikale en leksikale eienskappe van manlike en vroulike skrywers in die subkorpora te identifiseer.

Die korpora is klein en daarom kan daar nie oor Afrikaanse digters veralgemeen word nie, maar dit blyk dat genderverskille en genre, ten minste in hierdie korpora, wel taalgebruik beïnvloed. Die resultate stem egter nie noodwendig ooreen met die resultate uit vorige studies oor gender- en taalgebruiksverskille nie. Dit is nie 'n uitsonderlike geval nie. Newman e.a. (2008:213) let in hul navorsing op dat daar “contradictory reports” uit vorige navorsing is wat hulle toeskryf aan “the different contexts in which the language samples were generated”.

Daar is nie in hierdie studie bespiegel oor die rede vir hierdie taalgebruiksverskille nie. Pennebaker e.a. (2003:548) is egter van mening dat die woorde wat individue gebruik, insigte bied in hul denkprosesse. Woorde kan ook inligting verstrek oor individue se sosiale status, ouderdom en motiewe asook die omgewing of situasie waarin hulle hul bevind. Verdere navorsing oor taalgebruiksverskille en gender asook taalgebruik en kultuur sal onderneem moet word om te bepaal waarom manlike en vroulike Afrikaanse digters sekere leksikale woorde gebruik.

In die huidige studie kan die gevolgtrekking gemaak word dat taalgebruiksverskille in hierdie beperkte korpora gender- en genreverwant is. Alhoewel die eienskappe wat geïdentifiseer is, nie as algemene eienskappe van die Afrikaanse poësie (met betrekking tot die gedigte van manlike en vroulike digters) beskou kan word nie, openbaar die resultate eienskappe wat in die Afrikaanse poësie moontlik as “manlik” of “vroulik” beskou kan word. Verdere navorsing is egter nodig om vas te stel in watter 'n mate die grammatikale en leksikale eienskappe wat in hierdie studie geïdentifiseer is, met eienskappe in die Afrikaanse poësie in die geheel, maar ook die poësie van manlike en vroulike skrywers in die algemeen ooreenstem, en of 'n korpusgebaseerde ontleding gebruik kan word om soortgelyke ontledings op groter skaal aan te pak.

Bibliografie

- Abrams, M.H. 2005. *A glossary of literary terms*. Boston: Thomson Wadsworth.
- Anthony, A. 2007. Read me file for Antconc 3.2.1. <http://www.laurenceanthony.net/software/antconc/releases/AntConc321/help.pdf> (6 Oktober 2017 geraadpleeg).
- Argamon, S., M. Koppel, J. Fine en A.R. Shimoni. 2003. Gender, genre, and writing style in formal written texts. *Text*, 23(3):321–46.
- Barrette, E. 2004. Do women and men really write differently? *The Internet Review of Science Fiction*. <http://www.irosf.com/q/zine/article/10049> (5 Oktober 2017 geraadpleeg).
- Biber, D. 2009. A corpus-driven approach to formulaic language in English: multi-word patterns in speech and writing. *International Journal of Corpus Linguistics*, 14(3):275–311.
- . 2010. What can a corpus tell us about registers and genres? In O’Keeffe en McCarthy (reds.) 2010.
- Cheng, W. 2012. *Exploring corpus linguistics: language in action*. Londen: Routledge.
- CTeXt® (Sentrum vir Tekstegnologie, Noordwes-Universiteit) Korpus, <http://www.nwu.ac.za/ctext>
- De Lange, J. 2000. Hoe om (soos iemand anders) te skryf: oor vakmanskap en die skryf van gedigte. *Literator*, 21(2):17–23.
- Eiselen, E.R. en M.J. Puttkammer. 2014. Developing text resources for ten South African languages. Proceedings of the 9th International Conference on Language Resources and Evaluation, Reykjavik, Ysland. ble. 3698–3703.
- Handford, M. 2010. What can a corpus tell us about specialist genres? In O’Keeffe en McCarthy (reds.) 2010.
- Herring, S.C. en J.C. Paolillo. 2006. Gender and genre variation in weblogs. *Journal of Sociolinguistics*, 10(4):439–59. https://www.researchgate.net/publication/227681436_Gender_and_Genre_Variation_in_Weblogs (5 Oktober 2017 geraadpleeg).
- Hunston, S. 2002. *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Koppel, M., S. Argamon en A.R. Shimoni. 2002. Automatically categorizing written texts by author gender. *Literary and Linguistic Computing*, 7(4):401–12.
- Kotzé, E.F. 2007. Die vangnet van die woord: forensies-linguistiese getuienis in ’n lastersaak. *Southern African Linguistics and Applied Language Studies*, 25(3):385–99.
- McEnery, T., R. Xiao en Y. Tono. 2006. *Corpus-based language studies: an advanced resource book*. Londen: Routledge.
-

McIntyre, D. en B. Walker. 2010. How can corpora be used to explore the language of poetry and drama? In O’Keeffe en McCarthy (reds.) 2010.

Mulac, A. 1975. Evaluation of the speech dialect attitudinal scale. *Speech Monographs*, 42(3):184–9.

Mulac, A. en T.L. Lundell. 1994. Effects of gender-linked language in adults’ written discourse: multivariate tests of language effects. *Language and Communication*, 14(3):299–309.

Newman, M.L., C.J. Groom, L.D. Handelman en J.W. Pennebaker. 2008. Gender differences in language use: an analysis of 14,000 text samples. *Discourse Processes*, 45(3):211–36.

O’Keeffe, A. en M. McCarthy (reds.). 2010. *The Routledge handbook of corpus linguistics*. Londen: Routledge.

Pennebaker, J.W., M.R. Mehl en K.G. Niederhoffer. 2003. Psychological aspects of natural language use: our words, our selves. *Annual Review of Psychology*, 54:547–77.

Romero-Trillo, J. (red.). 2013. *Yearbook of Corpus Linguistics and Pragmatics 2013: New domains and methodologies*. Dordrecht: Springer.

Roux, J.C., P.H. Louw en T.R. Niesler. 2004. The African Speech Technology Project: an assessment. Proceedings of LREC. Lissabon, Portugal, ble. 93–6.

Semino, E. en J. Culpeper (reds.). 2002. *Cognitive stylistics: Language and cognition in text analysis*. Amsterdam en Philadelphia: John Benjamins Publishing.

Snyman, D., G.B. van Huyssteen en W. Daelemans. 2012. Cross-lingual genre classification for closely related languages. Proceedings of the 22nd Annual Symposium of the Pattern Recognition Association of South Africa, ble.132–7.

Steen, G. 2002. Metaphor in Bob Dylan’s “Hurricane”. In Semino and Culpeper (reds.) 2002.

Tognini-Bonelli, E. 2001. *Corpus linguistics at work*. Philadelphia: John Benjamins.

Vaughan, E. en B. Clancy. 2013. Small corpora and pragmatics. In Romero-Trillo (red.) 2013.

Eindnota

¹ Die NCHLT Part of Speech Tagger (<https://rma.nwu.ac.za/index.php/nchlt-part-of-speech-taggers.html>) en die NCHLT Tagger <https://rma.nwu.ac.za/index.php/nchlt-tagger.html>.