

Faktore wat die weglating van die Afrikaanse onderskikker *dat* bepaal

*Factors that determine the omission of the Afrikaans complementiser *dat* “that”*

BERTUS VAN ROOY

Vaaldriehoekcampus, Noordwes-Universiteit
E-pos: bertus.vanrooy@nwu.ac.za

HAIDEE KRUGER

Macquarie University
Sydney, Australia
Vaaldriehoekcampus, Noordwes-Universiteit
E-pos: haidee.kruger@mq.edu.au



Bertus van Rooy



Haidee Kruger

BERTUS VAN ROOY is professor in Engelse taalkunde binne die fokusarea *Verstaan en Verwerking van Taal in Komplekse Omgewings* (UPSET) en die Skool vir Tale aan die Vaaldriehoekcampus van die Noordwes-Universiteit. Hy spesialiseer op navorsing oor variëteite van Engels vanuit 'n world Englishes en korpus-linguistiese benadering. Verder werk hy, saam met Haidee Kruger, aan die ontwikkeling van 'n navorsingsprogram genaamd “Constrained Language” waarin die effek van beperkings op taalverwerking en taalgebruik oor verskillende tipes taalvorme heen ondersoek word. Hy is 'n medeskrywer van gedeeltes oor die werkwoordstuk binne die Afrikaanse grammatikaprojek van VivA.

BERTUS VAN ROOY is professor in English linguistics in the research focus area *Understanding and Processing Language in Complex Settings* (UPSET) and the School of Languages of the Vaal Triangle Campus of the North-West University. He specialises in research on varieties of English from the perspective of world Englishes and corpus linguistics. He is also engaged, with Haidee Kruger, in the development of a research programme called “Constrained Language” in which the effect of constraints on language processing and language use is investigated across a range of different types of language usage. He is co-author of parts of the verb phrase within the Afrikaans grammar project of VivA.

HAIDEE KRUGER is 'n navorsingsgenoot in die Departement Taalkunde by Macquarie Universiteit. Sy is ook buitengewone professor binne die fokusarea *Verstaan en Verwerking van Taal in Komplekse Omgewings* (UPSET) aan die Vaaldriehoekcampus van die Noordwes Universiteit, waar sy saam met Bertus van Rooy binne die “Constrained Language” navorsingsprogram werk. Haar onlangse navorsing fokus veral op vorme van taalgebruik wat beïnvloed word deur taalkontak, insluitende vertaling en tweedetaalvariëteite; sowel as die rol van teksredaksie in die konvensionaliseringprosesse vir geskrewe taal, en maak gebruik van korpus-linguistiese sowel as psigolinguistiese metodes. Sy is 'n medewerker aan die Afrikaanse grammatikaprojek van VivA.

HAIDEE KRUGER is a research fellow in the Department of Linguistics at Macquarie University. She also holds a position as extraordinary professor in the research focus area *Understanding and Processing Language in Complex Settings* (UPSET), at the Vaal Triangle Campus of North-West University, where she works in the “Constrained Language” research programme with Bertus van Rooy. Her recent research focuses on forms of language use influenced by language contact, including translation and second-language varieties; as well as the role of text editing in the conventionalisation processes of written language, using corpus-linguistic as well as psycholinguistic methods. She is also a collaborator on the Afrikaans grammar project of VivA.

ABSTRACT

Factors that determine the omission of the Afrikaans complementiser dat “that”

Verb complement clauses in written Afrikaans have two formal variants. One form has an overt complementiser dat “that”, followed by dependent word order (dat+SXV), while the other form has no complementiser and independent word order (Ø+SVX). Previous research on Afrikaans has not yielded conclusive findings about the factors that influence the choice between the variants, although the factors conditioning the alternation have been studied extensively in English and German. Based on that research, a number of potential conditioning factors are identified, which relate to the syntactic complexity of the main clause (its subject, modification of tense and modality features of the main verb, passivisation, negation or additional modifiers between the main clause and the complement clause); the semantics, lexical choice and frequency of the verb of the main clause; and the formality of the register.

This article adopts corpus linguistics as method to determine which factors best predict the choice between the variants. The data are drawn from the Taalkommissiekorpus (“Corpus of the Language Commission”), which is a 57 million word corpus of contemporary written Afrikaans. A sample of 10 084 instances of the declarative complement clause were extracted from the corpus by using 104 different verb lemmas (see Appendix A) as extraction terms. These were classified manually for their complementation pattern, and all the potential features identified as potential conditioning factors in the literature review were annotated for every valid instance in the sample. The classified data were then subjected to decision tree modelling, to yield a classification tree that identifies the most important factors related to the choice between the two variants of the declarative complement construction in Afrikaans.

The results indicate that the verb of the main clause is the most important factor. A few very general verbs with high frequency, such as sê “say”, weet “know” and dink “think” are strongly associated with the variant Ø+SVX. Verbs that are more specific in their meaning and less frequent tend to take the variant dat+SXV. In addition, register is an important conditioning factor for verbs that allow combination with both variants. More formal registers, such as academic writing, study guides and published popular books are more likely to make use of the variant dat+SXV, while the less formal registers of newspapers, magazines and fiction make more frequent use of the variant without a complementiser (Ø+SVX).

The findings are interpreted as support for the view that thematic prominence is the underlying force behind the variability. When the main clause is very general and low in informativity, the most prominent information is contained in the complement clause, which is therefore presented without an overt complementiser and with main clause word order (Ø+SVX). However, when the main clause is thematically prominent, carries a higher information load, and demands the attention of the reader, the complement clause is overtly marked as subordinate by means of an overt complementiser and the use of dependent word order (dat+SXV).

KEYWORDS: Afrikaans, syntax, word order, variation, complement clause, subordinator, complementiser, dependent word order, register, verb lemma, informativity, corpus linguistics

TREFWOORDE: Afrikaans, sintaksis, woordorde, variasie, stelbysin, onderskikker, afhanklike woordorde, register, werkwoordlemma, informatiwiteit, korpuslinguistiek

OPSOMMING

Die stelbysin in geskrewe Afrikaans vertoon twee variante. Die onderskikker *dat* word saam met afhanklike woordorde (SXV) gebruik in die een vorm, terwyl die ander vorm sonder die

onderskikker voorkom en hoofsinwoordorde (SVX) gebruik. In hierdie artikel word die korpuslinguistiek as metode gebruik om ondersoek in te stel na die faktore wat verband hou met die keuse tussen die twee variante. ’n Analise van die Taalkommissiekorpus dui daarop dat die werkwoord van die hoofsin veral deurslaggewend is vir die keuse. Werkwoorde met ’n algemene betekenis wat met hoë frekwensie voorkom word eerder saam met die stelbysin sonder ’n onderskikker gebruik, terwyl werkwoorde met ’n meer spesifieke betekenis en dikwels laer frekwensie kombineer met die stelbysin wat ’n eksplisiete onderskikker bevat. Verder speel die register van tekste ’n rol, deurdat meer formele registers eerder die variant met *dat* verkies. Die bevindinge dui daarop dat die vorm sonder *dat* geassosieer word met hoër tematische prominensie vir die stelbysin, terwyl die vorm met *dat* die prominensie verskuif na die hoofsin.

1. INLEIDING

Die stelbysin in Afrikaans, wat prototipies as naamwoordsin gebruik word (Ponelis 1979:450-453), het drie verskillende konstruksievorme.¹ Daar is twee standaardvorme, wat onderskei word deur die aan- of afwesigheid van ’n overte onderskikker *dat*, soos geïllustreer deur voorbeeld (1) en (2). Die derde vorm met die onderskikker *dat*, soos geïllustreer deur voorbeeld (3), word as niestandaard beskou (Biberauer 2002; Feinauer 1989, 1990; Ponelis 1979:440-441, 1993:340).²

- (1) Kenners bespiegel nou **dat** [hy moeg geraak het daarvoor dat van sy slagoffers nie gevind is nie]. (TK)
- (2) “Ek dink [hy is dood].” (TK)
- (3) Hy’t gevoel **dat** [die Voortrekkers moet nog so baie dinge leer]. (PKGA)

Die stelbysin wat ingelei word deur *dat* met afhanklike woordorde is die ouer vorm in Afrikaans, terwyl sinne waarin *dat* weggelaat word eers teen die negentiende eeu gereeld in Afrikaanse tekste aangetref word (Ponelis 1993:308). Die sinne verskil nie net ten opsigte van die teenwoordigheid of afwesigheid van die onderskikker nie, maar ook wat woordorde betref. Indien ’n stelbysin deur *dat* ingelei word, volg die afhanklike woordorde met die werkwoord aan die einde, soos geïllustreer deur die plasing van *gevind is* in voorbeeld (1). Hierdie variant word afgekort tot *dat*+*SXV* in die res van die artikel. Indien *dat* weggelaat word in die stelbysin, volg die onafhanklike woordorde, wat normaalweg in hoofsinne aangetref word, soos geïllustreer deur die plasing van *is* in voorbeeld (2). Hierdie variant word afgekort tot \emptyset +*SVX* in die res van die artikel. Daar is dus duidelik sprake van taalverandering in Afrikaans, deurdat \emptyset +*SVX* geleidelik in kompetisie met *dat*+*SXV* tree, wat lei tot sinchroniese variasie.

¹ ’n Uiteenlopende versameling terminologie word gebruik om na die konstruksies te verwys wat in hierdie artikel bespreek word. Sover dit Engelstalige literatuur betref, ingesluit literatuur wat na Afrikaans, Nederlands of Duits verwys, word die stelbysin ook ’n komplement sin (“complement clause”) of argument sin (“argument clause”) genoem. Die onderskikker *dat* staan bekend as komplementiseerder (“complementiser”). Hoofsin- en bysinvolgorde is ekwivalent aan onafhanklike en afhanklike woordorde. Ter wille van aansluiting by die tradisie van grammatikale beskrywing in Afrikaans word die terminologie van Ponelis (1979) hoofsaaklik in hierdie artikel gebruik.

² Die onderskikker *dat* word in vetdruk weergegee in alle voorbeelde in hierdie artikel, en die stelbysin met vierkantige hakies omsluit. Waar dit spesifiek ter sprake kom, word die hoofsin onderstreep. Voorbeelde in hierdie artikel word geneem uit die Taalkommissiekorpus (afgekort tot TK) wat in meer besonderhede in afdeling 3 van die artikel bekendgestel word, of in enkele gevalle waar daar nie voldoende illustrasiemateriaal is nie, uit die Poneeliskorpus van Gesproke Afrikaans (PKGA).

Die derde variant van die stelbysin verteenwoordig 'n vermenging van die eerste twee, deurdat die onderskikker *dat* aanwesig is, maar tog deur onafhanklike woordorde gevolg word, soos geïllustreer deur die plasing van *moet* in voorbeeld (3). Die meeste bestaande navorsing oor stelbysinne in Afrikaans konsentreer op die voorkoms van die niestandaardvariant *dat*+SVX, met spesifieke aandag aan hoe algemeen dit voorkom en hoe dit ontstaan het (Donaldson 1988:279; Feinauer 1989:33-35; Ponelis 1993:335; Steyn 1976:48). Die *dat*+SVX konstruksie kom met matige frekwensie in die gesproke taal voor, in ongeveer 20% tot 50% van alle gevalle (Biberauer 2002; Feinauer 1989; Stell 2011), maar in die geskrewe taal is die frekwensie van die konstruksie weglaatbaar klein (Biberauer 2002; Feinauer 1989). In hierdie artikel word geskrewe data ontleed, en daarom word daar nie verder aandag gegee aan hierdie niestandaardvariant nie.

Bosch (1998, 1999) bied 'n uitgebreide analise van die *dat*+SXV konstruksie, met klem op die gebruik in onderwerp- en voorwerpposisie, en die werkwoorde wat tipies voorwerpsinne neem. Ponelis (1993) bevind dat die Ø+SVX konstruksie eers deur die loop van die negentiende eeu gevestig geraak het, terwyl Malherbe (1920:117) vroeg in die twintigste eeu rapporteer: “Baie gebruiklik is die onderdrukking van *dat* as inleidende woord vir die voorwerpsin: *Hy sê vir my, ek moet maar gaan. Dink jy, ons sal dit regkry*”. Die herkoms van die variante konstruksie sonder die onderskikker word ten dele aan 'n soortgelyke, maar veel minder frekwente, proses in Nederlands toegeskryf, maar daar word veral klem geplaas op die rol van Engels in die toename in frekwensie van die konstruksie (Donaldson 1988: 279; Feinauer 1990:119; Ponelis 1979:242).

Biberauer (2002) bevind dat die weglating van *dat* in 46% van alle moontlike gevalle in haar gesproke korpus aangetref word, maar die konstruksie is minder frekwent in haar geskrewe data, waar dit in 35% van gevalle voorkom. Stell (2011:175) rapporteer dat die onderskikker *dat* tussen 70% en 100% van die tyd ná die hoëfrekwensiewerkwoorde in sy gesproke korpus weggelaat word.

Daar is egter min bestaande navorsing oor die faktore wat die keuse om *dat* weg te laat beïnvloed, en dus bestaan daar 'n leemte wat die verklaring van die waargenome variasiepatrone betref. Sommige skrywers bied relevante opmerkings sonder dat hulle noodwendig sistematies vir die faktore kontroleer. Malherbe (1966:13) identifiseer die moontlikheid dat die ontwikkeling van die Ø+SVX konstruksie versterk word deur die groter kognitiewe gemak waarmee die hoofsinwoordorde geprosesseer word. Terselfdertyd bestempel hy die keuse tussen *dat*+SXV en Ø+SVX as arbitrêr en onverbandhoudend met enige semantiese of strukturele faktore (Malherbe 1966:14-15). Feinauer (1990:117) argumenteer dat die weglating van die onderskikker groter klem op die voorwerpsin plaas, en dus tipies voorkom ná meer neutrale werkwoorde. Ook Malherbe (1966:13) identifiseer die moontlike rol van werkwoorde wat basiese menslike gevoelens en waarnemings uitdruk as verbandhoudend met weglating. Braeckeveldt (2013) se studie bevat die mees sistematiese inligting oor die keuse tussen die twee konstruksies tot op hede. Sy argumenteer dat werkwoorde wat informeel, morfologies simpleks en semanties algemeen is, tipies meer aangetrokke is tot die Ø+SVX konstruksie (sien ook Coleman, Feinauer en Braeckeveldt 2016, hierdie uitgawe). Teenoor hierdie enkele opmerkings oor Afrikaans bestaan daar 'n omvangryke literatuur oor die faktore wat 'n rol speel by die ooreenstemmende variasie in Engels tussen stelbysinne met of sonder die onderskikker *that*,³ asook die weglating van *dass* in Duits (Auer 1998; Weinert 2012). Daar word verwag dat die situasie in Afrikaans meer kompleks is as in Engels (maar wel soortgelyk aan Duits), aangesien dit by Afrikaans nie net handel oor die teenwoordigheid en afwesigheid van die onderskikker nie, maar ook oor die verskil tussen afhanklike en onafhanklike woordorde.

³ Biber (1999); Boye & Harder (2007); Dor (2005); Kearns (2007a),(2007b); McGregor (2013); Rohdenburg (1996); Tagliamonte & Smith (2005); Thompson (2002); Thompson & Mulac (1991); Torres Cacoullous & Walker (2009); Yaguchi (2001).

Die doel van hierdie artikel is om faktore te identifiseer wat 'n rol speel by die voorkoms van die *dat*+SXV en Ø+SVX variante van die stelbyisin in Afrikaans, en om aan die hand van hierdie faktore vas te stel wat die beste verklaring vir die variasie is. Ten einde hierdie doel te bereik, word faktore wat in die literatuur oor Engels en Duits geïdentifiseer is, geïntegreer met die enkele opmerkings oor Afrikaans, sodoende potensiele bydraende faktore te identifiseer. Die korpuslinguistiese metode word in hierdie artikel gebruik, en vervolgens word die metode uiteengesit wat betref die korpusdata, die wyse van dataonttrekking, en die manier hoe die data geklassifiseer is. Ter ondersteuning van die variasie-analise is dit nodig om die data statisties te modelleer, en daarom word besluitnemingsbome ('n tipe regressiemetode) as deel van die metode uiteengesit. In die laaste gedeeltes van die artikel word die bevindinge gerapporteer en geïnterpreteer om tot 'n slotsom te kom oor die omvang van variasie in die vorm van stelbyisinnne in Afrikaans, en 'n verklaring te bied vir die faktore wat hiertoe aanleiding gee.

2. FAKTORE WAT VERBAND HOU MET DIE VARIASIE TUSSEN STELBYSINKONSTRUKSIES

Sintaktiese, semantiese en leksikale faktore is in vorige navorsing (primêr oor Engels en Duits) geïdentifiseer as moontlike verklarings vir die variasie tussen stelbyisinnne met en sonder die onderskikker, terwyl die variasie ook met verskille in register verband hou. Volgens Kearns (2007a:296) is 'n vertrekpunt vir baie van die navorsing die aanname dat die keuse tussen die twee konstruksies nie betekenisdraend is nie, en dat dit daarom 'n besondere geskikte verskynsel is vir die ondersoek van hoe verskillende talige en buitetalige faktore 'n rol speel by die keuse tussen variante. Daar is egter navorsers wat beweer dat die keuse tog op pragmatiese en diskoersvlakke betekenisdraend word (Biber 1999; Boye & Harder 2007; McGregor 2013; Yaguchi 2001). In hierdie artikel word die data-analise onderneem sonder om 'n spesifieke semantiese verskil tussen die konstruksies aan te neem, omdat dit uit die analise moet blyk of faktore wat 'n funksionele of semantiese verskil ondersteun, die sterkste verklarende faktore is.

Sintaktiese eienskappe wat saamhang met die keuse tussen die twee variante hou verband met die kompleksiteit van die elemente in die hoofsin, soos aangedui vir Duits deur Auer (1998) en vir Engels deur Rohdenburg (1996).⁴ Indien die hoofsin uit minder komplekse elemente bestaan, verhoog die waarskynlikheid dat die onderskikker weggelaat sal word. Dit geld vir 'n aantal spesifieke veranderlikes wat deur vorige navorsers geïdentifiseer is. 'n Voornaamwoord as sintaktiese onderwerp van die hoofsin verhoog die kans vir weglating van die onderskikker (Tagliamonte & Smith 2005; Thompson & Mulac 1991:242-243), en by uitstek as die onderwerp van die hoofsin die enkelvoudige eerstepersoonsvoornaamwoord is (Torres Cacoullous & Walker 2009:26). Indien die hoofsin egter 'n passiefkonstruksie bevat, verhoog die kans dat die onderskikker oort sal realiseer (Biber et al. 1999:682). Addisionele elemente wat met die werkwoord van die hoofsin geassosieer is, verhoog ook die kans op die overte realisering van die onderskikker. Hierdie elemente sluit in modale hulpwerkwoorde (Thompson & Mulac 1991:246-247; Torres Cacoullous & Walker 2009:27), negativering (Auer 1998:291) en indirekte

⁴ Op teoretiese gronde onderskryf ons nie sonder voorbehoud die gebruik van die begrip "hoofsin" in hierdie konteks nie, en verkies ons die benadering van Boye en Harder (2007) en Thompson (2002) om eerder hieraan te dink as die predikaat wat 'n komplement sin inlei ("complement-taking predicate"). Aangesien hierdie teoretiese punt nie omvangryke implikasies vir die analise in hierdie artikel het nie, word egter volstaan by die geïntegreerde Afrikaanse term "hoofsin", verwysende na veral die onderwerp en werkwoord van die sin waarby die stelbyisin as voorwerp ingelyf is.

voorwerpe of adverbiale uitdrukkings wat tussen die hoofwerkwoord en die begin van die stelbysin voorkom (Tagliamonte & Smith 2005:304; Thompson & Mulac 1991:245; Torres Cacoullous & Walker 2009:27). Voorbeeld (4) illustreer 'n hoofsin (onderstreep) met relatief meer kompleksiteit: dit is 'n passiefkonstruksie en daar volg 'n bywoordelike bysin tussen die hoofsin en die stelbysin. In hierdie tipe voorbeelde word ver wag dat die onderskikker *dat* overt voorkom.

- (4) Die proses om 'n liggaam te balsem neem sowat ses weke, dus word daar geskat, as 'n mens terugwerk, [*dat* hy in Desember of Januarie dood is – die middel van die winter en jagseisoen in die Noordelike Halfrond.] (TK)

In die geval van 'n hoofsin met lae kompleksiteit, soos in voorbeeld (5), is die verwagting dat die stelbysin eerder sonder die onderskikker sal voorkom in die konstruksievorm Ø+SVX.

- (5) Ek hoor [sy naam is Frederik.] (TK)

Een verklaring vir waarom sintaktiese kompleksiteit in die hoofsin 'n rol speel in die keuse tussen die konstruksie met of sonder die overte onderskikker, is dat hierdie kompleksiteit 'n groter prosesseringslading op die spreker plaas, wat die noodsaaklikheid verhoog om die aanvang van die stelbysin eksplisiet met 'n onderskikker af te baken (Rohdenburg 1996:160-161). 'n Analise wat 'n alternatiewe perspektief op die kompleksiteitsanalise bied, word voorgehou deur navorsers wat die sintakties eenvoudige hoofsin beskou as 'n vaste roetine wat eintlik in adverbiale funksie gebruik word en nie as afsonderlike hoofsin geprosesseer word nie (Boye & Harder 2007; Thompson & Mulac 1991; Thompson 2002). Sulke hoofsinne het dus lae tematiese prominensie (Auer 1998:291-292; Feinauer 1990:117; Kearns 2007b:503). Die weglating van die onderskikker funksioneer in hierdie geval om die verhoging in tematiese prominensie van die bysin (en die gepaardgaande verlaging van die prominensie van die hoofsin) te merk. 'n Sintese wat deur Tagliamonte en Smith (2005:307) voorgestel word, is dat sintaktiese kompleksiteit in die agtergrond funksioneer, maar ander faktore meer prominent kan wees en die keuse tussen die twee konstruksies afdwing sonder dat kompleksiteit self 'n rol speel.

Versekeie semantiese en leksikale eienskappe van die werkwoord van die hoofsin is ook al as moontlike verklaring vir die keuse tussen die overte en afwesige onderskikker ondersoek. Werkwoorde wat kommunikatiewe of mentale betekenis oordra word veral met die weglating van die onderskikker verbind (Auer 1998; Dor 2005; Malherbe 1966; Torres Cacoullous & Walker 2009; Yaguchi 2001). Navorsers probeer aan die hand van die leksikale assosiasies van onderskikkerweglating 'n dieper verklaring vind. Die laer tematiese prominensie van die hoofsin vergestalt in werkwoorde wat baie algemeen in hulle betekenis is (Braeckeveldt 2013:45), teenoor meer spesifieke betekenis wat meer aandag op die hoofsin vestig. Dor (2005:348-349) stel voor dat die stelbysin met groter assertiwiteit (oftewel graad van stelligheid) as informasie aangebied word indien dit nie 'n overte onderskikker het nie, en net sekere werkwoordkategorieë is semanties versoenbaar met hierdie tipe verskuiwing van assertiwiteit van die hoofsin na die stelbysin. 'n Mentale werkwoord soos *weet* in voorbeeld (6) illustreer hierdie soort gebruik. Die informasie wat na *geweet* volg, word weer in die daaropvolgende sin verder uitgebrei, en die hoofsin *ek het geweet* is minder sentraal tot die voortsetting van die diskoers.

- (6) “Ek het twee jaar gelede ('n Grand Slam-sege) ervaar. Ek het geweet [ek is nie klaar nie, daar steek nog baie in my.] Dit was net die begin,” het Sharapova gesê. (TK)

Aan die ander kant is daar werkwoorde wat veel meer klem op die proses of wyse van kommunikeer plaas, soos *e-pos* of *neerskryf*, of werkwoorde wat om ander redes aandag vir die hoofsin opeis, waar die overte onderskikker die voorkeursposisie is. Die werkwoord *rondvertel* in voorbeeld (7)

inkorporeer 'n adverbiale betekenis by die basiese kommunikatiewe betekenis, en sal daarom eerder 'n bysin met overte onderskikker inlei, as gevolg van die groter semantiese spesifiekheid wat outomaties tot groter tematiese prominensie vir die hoofsin aanleiding gee.

- (7) Esau Gedenkskool se leerlinge vertel gou-gou oral rond [dat met elke blik mis vir Kallie se tuin, draers meer as net 'n oulap kan verdien.] (TK)

Die frekwensie van die werkwoord van die hoofsin is ook as moontlike verklarende faktor vir die voorkoms van onderskikkerweglating geïdentifiseer (Biber et al. 1999:681; Kearns 2007a:476; Thompson & Mulac 1991:244). Die meer frekwente werkwoorde laat eerder weglating toe in Engels, terwyl die minder frekwente werkwoorde eerder die onderskikker eksplisiet sal gebruik.

Ten slotte word register ook as 'n faktor in die keuse tussen die overte en weggelate onderskikker geïdentifiseer. In gesproke Engels (Biber 1999; McGregor 2013) en gesproke Duits (Auer 1998; Weinert 2012) is weglating van die onderskikker algemeen, maar weglating kom minder gereeld in geskrewe registers voor, waar die mees formele register, akademiese skryfwerk, gekenmerk word deur die laagste voorkoms van weglating. Hoe meer formeel die register, hoe groter is die waarskynlikheid dat die onderskikker overt realiseer. Rohdenburg (1996:160) voer ook aan dat die werkwoorde wat tipies in die meer formele registers voorkom, oor die algemeen 'n laer frekwensie het, en dus is daar moontlik 'n onderlinge verband tussen frekwensie van die hoofsinwerkwoord en die effek van formaliteit op die weglating van die onderskikker.

Uit die uiteensetting hierbo blyk dit dat daar verskeie sienings is oor die interpretasie van die weglating van onderskikkings. Die een argument is dat meer komplekse hoofsinne 'n groter prosesseringslading meebring, en gevolglik is die gebruik van die onderskikker 'n meganisme om die kompleksiteit te hanteer deur die grens tussen hoofsin en stelbysin eksplisiet te maak (bv. Rohdenburg 1996). Daarteenoor is daar 'n groep navorsers wat argumenteer dat die variant sonder die onderskikker eerder gebruik word in gevalle waar die hoofsin laer tematiese prominensie het in vergelyking met die stelbysin, en dat die weglating van die onderskikker dus juis die groter vormlike onafhanklikheid van die bysin vergestalt (Auer 1998; Boye & Harder 2007; Feinauer 1990; Kearns 2007b; Thompson 2002). Die twee verklarings is nie noodwendig met mekaar in botsing nie, en kan mekaar dikwels ondersteun. In die operasionalisering van veranderlikes in hierdie studie word wel gepoog om tot die mate moontlik tussen die twee benaderings te onderskei, om hulle relatiewe verklarende krag te beoordeel.

3. NAVORSINGSMETODE

Die analise van die faktore wat met die weglating van die onderskikker *dat* verband hou, word met behulp van die korpuslinguistiek en statistiese modellering onderneem. 'n Groot korpus van geskrewe Afrikaans, die Taalkommissiekorpus (Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns 2011), word as basis gebruik, waaruit voorbeelde van die twee variante van die stelbysin op 'n sistematiese wyse onttrek word. Hierna word die voorbeelde geklassifiseer in terme van die faktore wat moontlik verband hou met die keuse. Hierdie proses lewer 'n geannoteerde datastel op, wat vervolgens statisties ontleed word om te bepaal óf daar bepaalde faktore is wat verband hou met die onderskeid, en indien wel, watter faktore 'n rol speel en hoe hulle met mekaar in interaksie tree. Daaruit word 'n interpretasie van die redes vir die verskil tussen die konstruksies afgelei.

Die Taalkommissiekorpus is 'n groot versameling tekste wat geskrewe Afrikaans verteenwoordig. In totaal bevat die korpus tans ongeveer 57 miljoen woorde, en sluit die volgende registers in: fiksie (kortverhale en romans), koerante (dagblaaie en Sondagkoerante), tydskrifte (algemeen/

gesinstydskrifte en vrouetydskrifte), populêre boeke (subkategorieë: algemeen, finansies, godsdienstig, kuns en kultuur, leefstyl, sport en ontspanning, tegnologie, vroue), administratiewe tekste (notules van vergaderings), akademiese tekste (vakwetenskaplike artikels, proefskrifte en verhandelings), en universiteitstudiegids.

Met behulp van die rekenaarprogram *Wordsmith Tools* (Scott 2014) is 'n omvattende lys onttrek van alle kontekste waarin die woord *dat* voorkom. Hieruit is 'n lys werkwoorde geïdentifiseer wat direk links van *dat* voorkom en wat dus dui op kontekste waar daar moontlik stelbysinne gevind kan word. 'n Totaal van 104 werkwoorde is gekies as soekterme vir die analise. Die werkwoorde is verteenwoordigend van die volle omvang van werkwoordfrekwensie. Vervolgens is alle voorbeelde in die korpus waar elkeen van hierdie werkwoorde voorkom, afsonderlik onttrek uit die data. Die basisvorm (bv. *sê*), verlede deelwoordvorm (bv. *gesê*), en in die geval van skeibare werkwoorde ook die geskeide vorm (bv. *sê aan*, naas *aansê* en *aangesê*), is onttrek. Waar daar meer as 550 gevalle van 'n bepaalde vorm in die korpus was, is die data verminder tot 550 voorbeelde per vorm met behulp van ewekansige steekproefneming, buiten vir enkele hoogs frekwente werkwoorde (*sê*, *dink*, *weet*, *aanbeveel*) waar daar tog groter steekproewe behou is. Hierdie prosedure het meer as 50 000 moontlike datapunte opgelewer, wat vervolgens deur die twee navorsers met die hand geklassifiseer is in een van drie kategorieë: *dat*+SXV, Ø+SVX, of irrelevant (waaronder veral gevalle waar die betrokke werkwoord sonder 'n voorwerp gebruik word; waar die werkwoord 'n naamwoordstuk as voorwerp neem; of waar die bysin nie 'n stelbysin is nie, maar byvoorbeeld 'n vraagbysin). Op hierdie wyse is die data verminder tot 10 084 voorbeelde, wat die primêre datastel vir die artikel vorm.

Die sintaktiese vorm van die stelbysin (*dat*+SXV of Ø+SVX) is die afhanklike veranderlike in hierdie studie. Die onafhanklike veranderlikes wat ingespan word om 'n model te ontwikkel vir die voorspelling van die keuse tussen die twee waardes van die onafhanklike veranderlike, is vervolgens met behulp van *Wordsmith Tools* en *Microsoft Excel* (2010) geannoteer deur die twee navorsers, wat mekaar oor en weer gekontroleer het vir akkuraatheid en alle gevalle van botsende analyses opgelos het deur 'n geformaliseerde sintaktiese analise van die tersaaklike voorbeelde te onderneem. Hierdie veranderlikes, met hulle waardes, word in tabel 1 saamgevat.

TABEL 1: Onafhanklike veranderlikes

Veranderlike	Waardes
Onderwerp van hoofsin	Voornaamwoorde: eerste persoon enkelvoud, tweede persoon enkelvoud, derde persoon enkelvoud, eerste persoon meervoud, tweede persoon meervoud, derde persoon meervoud, vraend Naamwoordstukke: enkelvoud, meervoud Geen onderwerp (by infinitiewe)
Werkwoord van hoofsin	Elke lemma is afsonderlik gekodeer (sien bylaag A vir die volledige lys werkwoorde)
Breë semantiese kategorie van werkwoord ⁵	Kommunikasieproses, mentale proses, wyse van kommunikasieproses, waarborg, ander

⁵ Hoofsaaklik gebaseer op die werk van Dor (2005).

TABEL 1: Onafhanklike veranderlikes (*vervolg*)

Veranderlike	Waardes
Frekwensie van die werkwoord	Vanuit die steekproef van die data is geprojekteer na die frekwensie van die konstruksie per werkwoord vir die korpus in geheel, en daardie waarde is as 'n kontinue veranderlike in die analise gebruik, d.w.s. nie die absolute frekwensie van die werkwoord in die korpus nie, maar die frekwensie van die grammatikale konstruksie van die werkwoord in kombinasie met 'n stelbysin
Tyd- en modaliteits-vorm van hoofsin	Eenvoudige teenwoordige tyd, eenvoudige verlede tyd, modale hulpwerkwoord, imperatief, infinitief, skakelwerkwoord, infinitief + modaal, infinitief + skakelwerkwoord, imperatief + skakelwerkwoord, modale hulpwerkwoord + skakelwerkwoord, meerdere modale hulpwerkwoorde, meerdere modale hulpwerkwoorde + skakelwerkwoord, modale hulpwerkwoord + imperatief, imperatief + skakelwerkwoord
Polariteit	Positief, negatief
Sintaktiese afstand	Aantal woorde tussen die laaste element van die onderwerp en werkwoord van die hoofsin, en die eerste woord van die bysin
Register	Fiksie, koerante, tydskrifte, populêre boeke, administratiewe tekste, akademiese tekste, studiegids

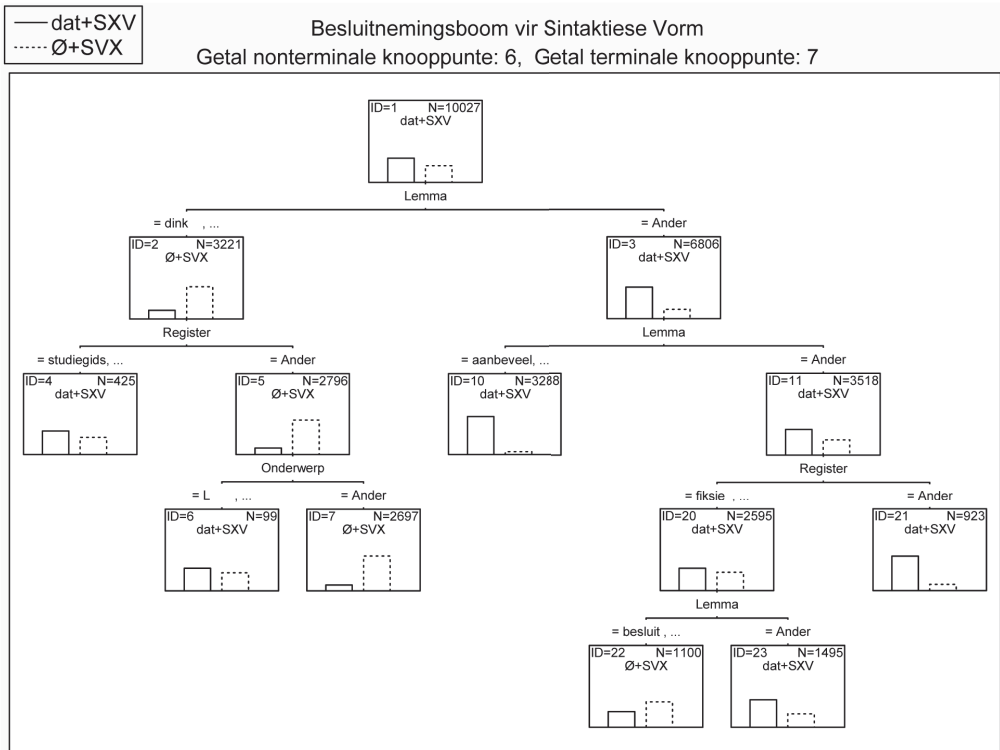
Nadat die klassifikasie van die data voltooi is, het ons 'n statistiese model met behulp van die sagtewareprogram *Statistica* (Statsoft 2013a) onttrek. Die modelleringstechniek staan bekend as 'n besluitnemingsboom.⁶ Dit bestaan uit 'n reeks agtereenvolgende besluite wat hiërargies voorgestel word, en wat ten doel het om die datastel te verdeel in kleiner groepe observasiepunte wat hoofsaaklik of uitsluitlik een van die twee waardes van die afhanklike veranderlike het. Die onderliggende algoritme is 'n meervoudige regressie-analise van die data. Die algoritme kan 'n groot aantal besluitnemingsbome genereer, en gebruik dan 'n χ^2 -toets om die optimale besluitnemingsboom te bepaal. Die kriteria wat tydens die evaluering van modelle in ag geneem word, sluit in die beste moontlike verdeling van die data in afsonderlike groepe en die kleinste aantal verdelingspunte.⁷

⁶ Die spesifieke tipe besluitnemingsboom in hierdie studie gebruik is 'n CART ("Classification and Regression Tree"). Vir meer besonderhede oor besluitnemingsbome kan Breiman et al. (1984) geraadpleeg word. Die detail van hoe die CART-model deur *Statistica* geïmplementeer word, word uiteengesit deur *Electronic Statistics Textbook* (StatSoft 2013b) op www.statsoft.com/Textbook/Classification-and-Regression-Trees.

⁷ 'n Soortgelyke bevinding word verkry deur 'n logistiese meervoudige regressie-analise, maar sodanige resultate word nie in hierdie artikel gerapporteer nie, aangesien dit nie nuwe inligting bydra om die onderskeid tussen die twee variante van die stelbysin te verstaan nie.

4. BEVINDINGE

Die statistiese model vir die onderskeid tussen die twee variante van die stelbysin in Afrikaans word in figuur 1 weergegee. Dit is eerstens belangrik om op te merk dat dit wel vir die algoritme moontlik is om 'n model te vind, wat nie 'n gegewe is nie – indien daar nie op grond van die beskikbare onafhanklike veranderlikes 'n statisties beduidende onderskeid tussen die twee waardes van die afhanklike veranderlike getref kan word nie, lewer die statistiese analise geen model op nie. Hiervolgens kan ons alreeds die bevinding stel dat daar 'n sistematiese onderskeid bestaan tussen die *dat*+SXV en \emptyset +SVX variante in Afrikaans.



Figuur 1: Besluitnemingsboom vir die variante van die stelbysin in Afrikaans

Die besluitnemingsboom dui aan watter veranderlikes die grootste bydraes lewer om die onderskeid te tref. Die heel eerste verdeling van bo af beskou is die verdeling tussen 'n groep werkwoorde (die linkerkantste tak van die boom) wat 'n sterk voorkeur (79%) vertoon vir die variant \emptyset +SVX, en alle ander werkwoorde.⁸ Hierdie lemmas wat 'n sterk voorkeur vir \emptyset +SVX vertoon word nie volledig in figuur 1 getoon nie, maar is die volgende:

⁸ Persentasies word bereken deur die aantal observasies in die voorkeurklas te deel deur die totale aantal observasies binne daardie vertakkingspunt. (Die waarde wat telkens vir N gespesifiseer word, is die totale aantal observasies wat by daardie vertakkingspunt in 'n afsonderlike subgroep geplaas word.) Die persentasie gee 'n aanduiding van hoe waarskynlik dit is dat die voorkeurvariant aangetref sal word, gegewe die spesifikasies van die subklasse tot op daardie punt in die besluitnemingsboom.

(8) afkondig, belowe, dink, gestel,⁹ hoor, sê, skat, skryf, voel, wed, weet, wens, wis

Die lys werkwoorde in (8) sluit die mees frekwente werkwoorde in die hele datastel in, naamlik *sê*, *dink* en *weet*, naas 'n aantal verdere werkwoorde wat grootliks op mentale prosesse dui en enkele kommunikasiewerkwoorde. Gegewe die prominensie van dié groep werkwoorde in die korpus, kan ons tot die gevolgtrekking kom dat hulle die mees prototipiese groep vorm. Tydens die verwerking van die data is daar by wyse van steekproewe slegs 550 voorbeelde van elke werkwoordvorm ontleed, wat die gevolg het dat hierdie groep in terme van rou frekwensie in die data onderverteenvoerdig word. Indien ons die data in die steekproef gebruik om te projekteer na die verdeling van data in die volle korpus, blyk dit dat die werkwoorde *sê*, *dink* en *weet* meer as die helfte van alle data verteenwoordig (51%).

Die statistiese model identifiseer 'n subgroep van die observasies in die eerste linkerkantste vertakking van die besluitnemingsboom, waar die voorkoms van die ander konstruksie, *dat+SXV*, wel meer algemeen is. Hierdie verdeling word gedoen op grond van register. Indien die stelbysin saam met die stel werkwoorde in voorbeeld (8) in die meer formele registers voorkom, naamlik akademiese skryfwerk, studiegidse en populêre boeke, verhoog die waarskynlikheid dat die vorm met die onderskikker, *dat+SXV*, gebruik word na 68%. In al die ander registers (fiksie, koerante en tydskrifte vernaamlik) word die vorm sonder die onderskikker, \emptyset +*SVX*, gebruik in 84% van alle gevalle. Dit is in hierdie verband belangrik om daarop te let dat 'n normatiewe bron soos *Skryf Afrikaans van A tot Z* (Müller & Pistor 2011:708), wat as stylgids dien vir die koerant- en tydskrifpublikasies van Media24, die aanbeveling maak dat indirekte rede wat deur die werkwoord *sê* ingelei word, verkieslik nié die onderskikker *dat* moet gebruik nie. Die effek hiervan is om die prominensie van die hoofsin te verlaag, en groter prominensie aan die inhoud van die bysin te verleen. Tydskrifte en koerante dra dus by om die gebruik van die variant \emptyset +*SVX* prominent te hou in die blootstelling wat taalgebruikers aan die taal het.

Die laaste verdeling aan die linkerkantste tak van die besluitnemingsboom is die verdeling tussen 'n kleiner aantal vorme (steeds met die werkwoorde in (8)) wat die passief- of infinitiefvorm in die hoofsin bevat, en dus geen agentiewe sintaktiese onderwerp het nie, saam met sinne waarvan die onderwerp 'n vraende voornaamwoord is. In hierdie gevalle bestaan daar dus in 'n sekere sin 'n inligtingsgaping binne die hoofsin, en die effek daarvan is om die waarskynlikheid dat die variant *dat+SXV* gekies word te verhoog na 56%. Die inligtingsgaping in die hoofsin vereis klaarblyklik meer aandag en verleen so meer prominensie aan die hoofsin, met die gevolg dat die afhanklike vorm *dat+SXV* in die bysin gebruik word. Indien die onderwerp van die hoofsin egter ekplisiet bekend gemaak word, is die waarskynlikheid dat die variant \emptyset +*SVX* gekies word 86%.

Die data wat by die eerste vertakking in die besluitnemingsboom aan die regterkant saamgegroepeer word, is gevalle waar die variant *dat+SXV* in die meerderheid van gevalle voorkom (77%). Hierdie subgroep verteenwoordig observasies waar die hoofsin 'n werkwoord bevat wat nié in lys (8) hierbo weergegee is nie, en verteenwoordig dus die oorblywende 91 werkwoorde wat ingesluit is by die studie. Die verdere vertakkings aan die regterkant het ten doel om kleiner groepe in hierdie subgroep te vind waar die verdeling nog fyner gedoen kan word. Die eerste vertakking identifiseer 'n groep werkwoorde in die linkerkantste vertakking wat 'n nog hoër voorkeur van 93% het vir die gebruik van die *dat+SXV* konstruksie. In die geval van hierdie

⁹ Die lemma *gestel* word apart van *stel* hanteer, omdat dit hoofsaaklik in imperatiefkonstruksies gebruik word, bv. "Gestel X is die geval..." Dit toon ooreenkomste met die passiefvorm van *stel*, asook voorsetsels of onderskikkende voegwoorde, maar behou bepaalde verbale eienskappe en funksioneer baie soortgelyk aan 'n uitdrukking soos "sê nou maar". Derhalwe word dit as afsonderlike lemma hanteer in hierdie ondersoek.

werkwoorde is daar dus sprake van 'n byna kategorieese voorkeur vir die konstruksie met die onderskikker. Die lys werkwoorde word in (9) weergegee:

- (9) aanbeveel, aansê, aanstip, aframmel, afspreek, agiteer, antisipeer, argumenteer, baklei, begryp, beken, bekendmaak, bemerk, bepleit, bespiegel, beveel, bewys, bluf, deklameer, deursein, dikteer, gewaar, gis, herbeklemtoon, herbevestig, insinueer, konkludeer, konstateer, mompel, neul, ontdek, ontgaan, ontken, paai, postuleer, prewel, profeteer, proklameer, propageer, protesteer, raaklees, rondvertel, sinspeel, skel, skimp, smeek, snap, spekleer, stateer, stel, stipuleer, teëkap, teoretiseer, terugkap, terugskryf, toesnou, uitbasuin, vergeet, verifieer, verklaar, vermaan, verordineer, verswyg, verwittig, vind, volg, vra, wink

Die lys sluit werkwoorde in wat met baie lae frekwensie in die data voorkom (tipies minder as een keer per miljoen woorde), soos *teëkap* of *insinueer*, en werkwoorde wat 'n baie duidelike funksie het in die strukturering van tekstuele verhoudings in 'n argument, soos *argumenteer*, *teoretiseer* en *herbeklemtoon*.

Laefrekwensiewerkwoorde is geneig om semanties baie spesifiek te wees, en omdat hulle lae frekwensies het, is hulle minder bekend vir die taalgebruiker. Albei hierdie redes lei daartoe dat sodanige werkwoorde baie meer klem op hulleself vestig, en daardeur verhoog die prominensie van die hoofsin in sy geheel. Dit het weer die gevolg dat die stelbysin in prominensie verlaag, en daarom word die afhanklike woorde saam met 'n eksplisiete merker van ondergeskiktheidstatus gebruik. Dit sluit ook aan by die verklaring wat in afdeling 2 hierbo gegee is vir die interaksie tussen die keuse van die variant met die onderskikker en faktore soos lae frekwensie en formele taalgebruik. In die besluitnemingsboom is die gewig van die spesifieke lemmas egter meer deurslaggewend om die onderskeid te tref as die ander faktore.

Die oorblywende gevalle in die besluitnemingsboom word afgebaken op grond daarvan dat hulle voorkom saam met hoofsinne wat nie die werkwoorde bevat wat in lys (8) en (9) ingesluit is nie. In kombinasie met hierdie werkwoorde is die vorm *dat+SXV* steeds dominant (63%), maar daar word tog subgroepe gevind waarin die vorm \emptyset +SVX met 'n groter waarskynlikheid voorkom. Die eerste verdere verdeling is in terme van register, waar die minder formele registers (fiksie, koerante en tydskrifte) oorhoofs 'n hoër waarskynlikheid van *dat*-weglating of dan die gebruik van die konstruksie \emptyset +SVX saam met die betrokke werkwoorde vertoon, terwyl die meer formele registers (akademiese skryfwerk, populêre boeke, studiegids en administratiewe tekste) met 'n waarskynlikheid van 85% die vorm *dat+SXV* saam met die betrokke werkwoorde verkies.

Die subgroep observasies in die minder formele registers word ten slotte verdeel in terme van die voorkeure van individuele werkwoorde. In gevalle waar die werkwoord een van die werkwoorde in lys (10) hieronder is, word die vorm \emptyset +SVX in 62% van die gevalle verkies. Hierdie werkwoorde is oorwegend mentale werkwoorde wat 'n meer spesifieke vorm van die betekenis van *dink* en *weet* verteenwoordig. Soos wat die geval met *dink* en *weet* is, is die voorkeurvariant wat met hierdie werkwoorde kombineer \emptyset +SVX. Daar word tog 'n hoër persentasie (38%) observasies in hierdie groep gevind waar stelbysinne met die vorm *dat+SXV* voorkom. Ons skryf dit toe aan die meer spesifieke betekenis van die werkwoorde en die wyse waarop hulle gebruik word, wat daartoe lei dat in sommige gevalle wel meer aandag op die hoofsin gevestig word.

- (10) beloof, besluit, erken, fantaseer, glo, kerm, onthou, sien, uitvind, vermoed

Vir alle oorblywende werkwoorde in die minder formele registers is die variant *dat+SXV* die dominante keuse met 'n voorkeur van 67%. Dit impliseer dat die werkwoorde oorwegend aandag op die hoofsin vestig, maar dit is moontlik in 'n derde van die gevalle wat hierdie werkwoorde

gebruik word, om hulle tog te kombineer met die Ø+SVX variant, wat suggereer dat die stelbysin in hierdie gevalle groter prominensie verkry.

Die analise van die data dui daarop dat 'n klein aantal van die faktore wat ondersoek is, bydra tot 'n verdeling van die data in subgroepe waarin die variante *dat+SXV* of Ø+SVX onderskeidelik dominant is. Die faktore wat veral met die konstruksie sonder die onderskikker, Ø+SVX, saamhang, is bepaalde lemmas, insluitend die hoëfrekwensiewerkwoorde *sê*, *dink* en *weet*, en die minder formele registers van fiksie, tydskrifte en koerante. Daarteenoor word die konstruksie met die ekplisiete onderskikker en afhanklike woordorde, *dat+SXV*, gebruik vir laefrekwensiewerkwoorde met spesifieker betekenisse, werkwoorde wat die verloop van 'n argument ekplisiet maak, en die meer formele registers van akademiese tekste, studiegidse en gepubliseerde populêre boeke. In 'n klein aantal gevalle is 'n inligtingsgaping oor die identiteit van die onderwerp van die hoofsin ook verantwoordelik daarvoor dat 'n groter persentasie van die *dat+SXV* konstruksie gebruik word.

Die ander faktore wat uit vorige navorsing geblyk het moontlik van belang kan wees en hoofsaaklik 'n meer direkte verband met die kompleksiteit van die hoofsin vertoon, speel 'n statisties onbeduidende rol in die analise van Afrikaans, waaronder die sintaktiese onderwerp, veral of dit voornaamwoordelik is of uit 'n naamwoordstuk bestaan; negativering; addisionele elemente tussen die hoofsin en die bysin; en die tyd- en modaliteitseienskappe van die hoofsin.

5. GEVOLGTREKKING

Die doel van hierdie ondersoek was om vas te stel of daar 'n sistematiese onderskeid getref kan word tussen die faktore wat verband hou met die keuse tussen twee variante van die stelbysin in geskrewe Afrikaans, *dat+SXV* en Ø+SVX. Die vernaamste bevindinge is dat die onderskeid wel sistematies is, en dat dit hoofsaaklik verband hou met die spesifieke werkwoord in die hoofsin en die graad van formaliteit van die register.

Die interpretasie van die bevindinge dui daarop dat die relatiewe prominensie van die hoofsin en die stelbysin die belangrikste funksie van die onderskeid tussen die twee konstruksies is. Waar die stelbysin meer prominent is, soos in voorbeeld (11), word hoofsinvolgorde in die bysin gebruik en word geen eksplisiete onderskikker gebruik nie.

- (11) Hy sê die Springbokke se spel is soms ietwat vervelig maar tog doeltreffend. “Hulle kan met enige span in die wêreld afreken as hulle voor die oorhand het. Dit is lank reeds hul resep vir welslae. Hulle het ook genoeg vaart op vleuel.” *De Marigny sê* [Samoa het die Bokke in hul eerste wedstryd probeer boelie, maar die Bokke het hulle van hul eie medisyne gegee en genoeg spelers van gehalte gehad om punte aan te teken.] (TK)

In hierdie voorbeeld word 'n kenner se mening oor die Springbokke se rugby gerapporteer, maar die kenner is nie tematies prominent in die uittreksel nie. Die teks handel oor die Springbokke se spel, en die inhoud van die stelbysin is daarom in terme van die ontwikkeling van die teks tematies sentraal. Indien die onderstreepte hoofsin weggelaat sou word, sou dit weinig verskil maak aan die vloei van die teks, terwyl die weglating van die stelbysin die teks wesenlik sou verskraal.¹⁰

Wanneer die hoofsin egter substantief meer bydra tot die ontplooiing van die betekenis van die teks, is daar 'n voorkeur vir die konstruksie *dat+SXV*, waardeur die stelbysin se prominensie verlaag word en meer nadruk op die hoofsin geplaas word, soos in voorbeeld (12).

¹⁰ Hierdie argument is gebaseer op Verhagen (2005:95-96), hoewel hy homself tot die funksie van die ooreenstemmende konstruksie in Nederlands beperk en nie vormlike variasie behandel nie.

- (12) Want die vlekkelose presiesheid waarmee die land se kolfmaestro gister in sy beurt van 186 lopies op die Wanderers in Johannesburg te werk gegaan het, is niks minder as krieket-adel nie. Niemand sal harder as Kallis mor [dat die magiese 200 hom vanoggend steeds ontwyk nie], maar vir die grootste deel van sy marathon-beurt was dit tone-krul-lekker om te sien hoe hy die Black Caps sonder weerga martel. (TK)

In hierdie voorbeeld val die klem op Jacques Kallis, die krieketspeler, wat ongelukkig is omdat hy (op daardie stadium) steeds nie 'n 200-tal in 'n kriekettoets kon behaal nie. Die joernalis beklemtoon dus veel meer die kolfvernuf van Jacques Kallis, die sintaktiese onderwerp van die onderstreepte hoofsin, en rapporteer nie in die eerste plek dat hy 'n mylpaal misgeloop het nie. Kallis self is dus tematies sentraal in die uittreksel, en gevolglik word die *dat*+SXV konstruksie gekies, waardeur die prominensie van die inhoud van die stelbysin verlaag word.

Die bevindinge van hierdie studie ondersteun Thompson (2002) en Verhagen (2005) se interpretasie van die funksie van die komplementkonstruksie, naamlik dat dit prototipies gebruik word om die inligting wat deur die stelbysin aangebied word, prominent te maak. Die Afrikaanse konstruksie sonder 'n eksplisiete onderskikker (en met hoofsinvolgorde) is meer geskik vir dié funksie, soos Auer (1998) vir Duits en Boye en Harder (2007), Kearns (2007b) en Thompson (2002) vir Engels argumenteer, en ook deur Feinauer (1990) vir Afrikaans gesuggereer word. Die siening dat die eksplisiete onderskikker in die *dat*+SXV konstruksie gebruik word omdat die hoofsin of die verhouding tussen die hoofsin en stelbysin te kompleks geword het (soos deur Rohdenburg 2006 vir Engels geargumenteer word), word nie deur die bevindinge van hierdie studie ondersteun nie, aangesien veranderlikes wat met hierdie verklaring verband hou, nie 'n statisties-beduidende bydrae lewer om tussen die konstruksievorme van die stelbysin te onderskei nie.

BIBLIOGRAFIE

- Auer, P. 1998. Zwischen Parataxe und Hypotaxe: “Abhängige Hauptsätze” im gesprochenen und geschriebenen Deutsch. *Zeitschrift für Germanistische Linguistik*, 26(3):284-307.
- Biber, D. 1999. A register perspective on grammar and discourse: Variability in the form and use of English complement clauses. *Discourse Studies*, 1(2):131-150.
- Biber, D., Johansson, S., Leech, G., Conrad, S. & Finegan, E. 1999. *Longman grammar of spoken and written English*. Harlow: Longman.
- Biberauer, T. 2002. Verb second in Afrikaans: Is this a unitary phenomenon? *Stellenbosch Papers in Linguistics*, 34:19-69.
- Boye, K. & Harder, P. 2007. Complement-taking predicates: Usage and linguistic structure. *Studies in Language*, 31(3):569-606.
- Bosch, B. 1998. Die onderskikker *dat*: 'n Korpus-gebaseerde bespreking (Deel 1). *Suid-Afrikaanse Tydskrif vir Taalkunde*, 16(4):120-126.
- Bosch, B. 1999. Die onderskikker *dat*: 'n Korpus-gebaseerde bespreking (Deel 2). *Suid-Afrikaanse Tydskrif vir Taalkunde*, 17(1):2-15.
- Braeckeveldt, C. 2013. *De skoon bysin in Afrikaanse krantetaal: Voorkomen en parameters*. Ongepubliseerde MA verhandeling. Gent: Universiteit Gent.
- Breiman, L., Friedman, J., Stone, C.J., & Olshen, R.A. 1984. *Classification and regression trees*. New York: Chapman & Hall.
- Colleman, T., Feinauer, I. & Braeckeveldt, C. 2016. Over lexicale voorkeure in de alternantie tussen de “skoon bysin” en de “*dat*-bysin”: Een distinctieve collexeemanalyse. *Tydskrif vir Geesteswetenskappe*, 56(1):117-133.
- Donaldson, B.C. 1988. *The influence of English on Afrikaans*. Pretoria: Serva.
- Dor, D. 2005. Toward a semantic account of *that*-deletion in English. *Linguistics*, 43(2):345-382.
- Feinauer, A.E. 1989. Plasing in Afrikaanse afhanklike sinne. *Suid-Afrikaanse Tydskrif vir Taalkunde*, 7(1):30-37.

- Feinauer, A.E. 1990. Skoon afhanklike sinne in Afrikaanse spreektaal. *Suid-Afrikaanse Tydskrif vir Taalkunde*, 8(3):116-120.
- Kearns, K. 2007a. Regional variation in the syntactic distribution of null finite complementizer. *Language Variation and Change*, 19:295-336.
- Kearns, K. 2007b. Epistemic verbs and zero complementizer. *English Language and Linguistics*, 11(3):475-505.
- Malherbe, D.F. 1920. *Afrikaanse Taalboek*. (3e uitgawe) Bloemfontein: Nasionale Pers.
- Malherbe, G.H. 1966. *Die gebruik van "dat" as verbindingswoord in Afrikaans*. Ongepubliseerde MA verhandeling. Stellenbosch: Stellenbosch Universiteit.
- McGregor, W.B. 2013. Optionality in grammar and language use. *Linguistics*, 51(6):1147-1204.
- Müller, D. & Pistor, S. 2011. *Skryf Afrikaans van A tot Z*. (2e uitgawe) Kaapstad: Pharos.
- Ponelis, F.A. 1979. *Afrikaanse sintaksis*. Pretoria: JL van Schaik.
- Ponelis, F.A. 1993. *The development of Afrikaans*. Frankfurt am Main: Peter Lang.
- Rohdenburg, G. 1996. Cognitive complexity and increased grammatical explicitness in English. *Cognitive Linguistics*, 7(2):149-182.
- Scott, M. 2014. *Wordsmith Tools*. Liverpool: Lexical Analysis Software.
- StatSoft. 2013a. Statistica (data analysis software system), version 12. <http://www.statsoft.com> [30 July 2015]
- StatSoft. 2013b. *Electronic Statistics Textbook*. Tulsa, OK: StatSoft. www.statsoft.com/textbook [30 July 2015]
- Stell, G. 2011. *Ethnicity and language variation: Grammar and code-switching in the Afrikaans speech community*. Frankfurt am Main: Peter Lang.
- Steyn, J.C. 1976. 'n Paar sintaktiese onvastighede in Afrikaans. *Taalfasette*, 20(3):25-53.
- Taalkommissie van die Suid-Afrikaanse Akademie vir Wetenskap en Kuns. 2011. Taalkommissiekorpus 1.1. Noordwes-Universiteit: CText.
- Tagliamonte, S. & Smith, J. 2005. No momentary fancy! The zero "complementizer" in English dialects. *English Language and Linguistics*, 9(2):289-309.
- Thompson, S.A. 2002. "Object complements and conversation": Towards a realistic account. *Studies in Language*, 26(1):125-164.
- Thompson, S.A. & Mulac, A. 1991. The discourse conditions for the use of the complementizer *that* in conversational English. *Journal of Pragmatics*, 15(3):237-251.
- Torres Cacoullos, R. & Walker, J.A. 2009. On the persistence of grammar in discourse formulas: A variationist study of *that*. *Linguistics*, 47(1):1-43.
- Verhagen, A. 2005. *Constructions of intersubjectivity: Discourse, syntax and cognition*. Oxford: Oxford University Press.
- Weinert, R. 2012. Complement clauses in spoken German and English: Syntax, deixis and discourse-pragmatics. *Folia Linguistica*, 46(1):233-265.
- Yaguchi, M. 2001. The function of the non-deictic *that* in English. *Journal of Pragmatics*, 33:1125-1155.

BYLAAG A

Die volgende werkwoorde is in hierdie studie gebruik om data te onttrek:

aanbeveel, aansê, aanstip, afkondig, aframmel, afspreek, agiteer, antisipeer, argumenteer, baklei, beaam, begryp, beken, bekendmaak, beloof, belowe, bemerk, bepleit, besluit, bespiegel, beveel, bewys, bieg, bluf, byvoeg, deklameer, deursein, dikteer, dink, erken, fantaseer, gewaar, gis, glo, herbeklemtoon, herbevestig, hoor, insinueer, kerm, konkludeer, konstateer, meedeel, mompel, mor, neul, ontdek, ontgaan, onthou, onthul, ontken, paai, postuleer, prewel, profeteer, proklameer, propageer, protesteer, raai, raaklees, rondvertel, sê, sien, sinspeel, skat, skel, skerts, skimp, skryf, smee, snap, soebat, spekuleer, stateer, stel, stipuleer, teëkap, teoretiseer, terugkap, terugskryf, toesnou, uitbasuin, uitblaker, uitvind, verduidelik, vergeet, verifieer, verklaar, vermaan, vermoed, verordineer, verstaan, verswyg, vertel, verwittig, vind, voel, volg, vra, waarsku, wed, weet, wens, wink, wis