

## 'n Metodologie vir die beskrywing van konstruksionaliseringsnetwerke: Konstruksies met [in] as gevallestudie

---

Gerhard B. van Huyssteen 

Sentrum vir Tekstegnologie (CTexT), Noordwes-Universiteit, Potchefstroom, Suid-Afrika  
E-pos: [gerhard.vanhuysteen@nwu.ac.za](mailto:gerhard.vanhuysteen@nwu.ac.za)

Adri Breed 

Understanding and Processing Language in Complex Settings (UPSET), Noordwes-Universiteit,  
Vanderbijlpark/Potchefstroom, Suid-Afrika  
E-pos: [adri.breed@nwu.ac.za](mailto:adri.breed@nwu.ac.za)

Anneke Butler 

Understanding and Processing Language in Complex Settings (UPSET), Noordwes-Universiteit,  
Vanderbijlpark/Potchefstroom, Suid-Afrika  
E-pos: [anneke.butler@nwu.ac.za](mailto:anneke.butler@nwu.ac.za)

Lande Botha 

Understanding and Processing Language in Complex Settings (UPSET), Noordwes-Universiteit,  
Vanderbijlpark/Potchefstroom, Suid-Afrika  
E-pos: [lande.botha@nwu.ac.za](mailto:lande.botha@nwu.ac.za)

Maristi Partridge 

Understanding and Processing Language in Complex Settings (UPSET), Noordwes-Universiteit,  
Vanderbijlpark/Potchefstroom, Suid-Afrika  
E-pos: [maristi.partridge@nwu.ac.za](mailto:maristi.partridge@nwu.ac.za)

Suléne Pilon 

Departement Afrikaans, Universiteit van Pretoria, Suid-Afrika  
E-pos: [sulene.pilon@up.ac.za](mailto:sulene.pilon@up.ac.za)

### 1. Inleiding

Beskouings oor konstruksionalisering (“constructionalisation”) en konstruksieverandering (“constructional change”) is aan die voorpunt van konstruksiegrammatikabenaderings tot taalverandering (kyk o.a. Barðdal et al. 2015; Colleman 2015, 2016; Enghels & Garachana Camarero 2021; Hilpert 2013, 2021; Sommerer & Smirnova 2020; Traugott & Trousdale 2013).

‘Konstruksionalisering’ kan gedefinieer word as die proses waardeur ’n nuwe simboliese verbinding tussen ’n bepaalde vorm en ’n bepaalde betekenis ontstaan (“establish”) en in ’n spraakgemeenskap gevestig (“entrench”) word. Binne ’n netwerkbeskouing van die grammatika as ’n gestruktureerde inventaris van konstruksies (vorm-betekenispare), verteenwoordig konstruksionalisering dus die daarstelling van nuwe nodes (konstruksies) in die netwerk (die konstruktikon). Konstruksieverandering is bloot netwerkmodulasies (aanpassings, verskuiwings, ensovoorts) wat voor en ná konstruksionalisering deur die gebruik van konstruksies te weeg gebring word (Traugott 2021).

Ten einde oor konstruksionalisering en konstruksieverandering in ’n bepaalde deel van die konstruktikon te kan praat, is dit nodig dat ’n mens ’n diachroniese én sinchroniese blik op daardie netwerk van konstruksies moet hê. Anders gestel: As ’n mens oor taalverandering wil praat, moet ’n mens weet hoe die konstruksienetwerk gelyk het, hoe dit tans daaruit sien, en hoe dit moontlik in die toekoms sou kon lyk.

Die oorhoofse doel van hierdie bydrae is om ’n metodologie (d.i. ’n stel metodes, beginsels, reëls, ens. vir ’n bepaalde doel binne ’n bepaalde teorie/dissipline) voor te stel vir die beskrywing en verklaring van konstruksionaliseringsnetwerke, d.i. netwerke met ’n tyddimensie waarmee die ontwikkeling van konstruksies aangedui word (Van Huyssteen & Eiselen 2021). Om hierdie raamwerk te illustreer, kyk ons na die konstruksionaliseringsnetwerk van *in* as preposisionele leksikale item (voortaan **P-item** in navolging van Fontaine (2017)<sup>1</sup>). Ons vooronderstel dat **P-konstruksies** (d.i. konstruksies met sulke P-items) ’n deel van die konstruktikon is wat geredelik aan dinamiese verandering in die netwerk onderworpe is en dus vrugbare grond vir navorsing oor taalverandering is (kyk o.a. Fagard & Mardale 2012; Rhee 2018). Onder P-items word ingesluit:

- i. voorsetsels (soos **in die kar**)
- ii. agtersetsels (soos **die wêreld in**)
- iii. alkantsetsels (soos **in baie staatsgeboue in**)
- iv. adjektiewe (soos *Glans op die gesig en lippe is weer in* (Huisgenoot, 25 Aug. 2016))
- v. bywoorde (soos *met spelfoute in*)
- vi. werkwoordpartikels (soos *samel fondse in*)<sup>2</sup>

In §2 gee ons ’n bondige programmatiese uiteensetting van ons voorgestelde metodologiese raamwerk, waarna ons in §3 die eerste stap in die raamwerk illustreer, naamlik om die protokonstruksie en protoscène van *in* te beskryf. In §4 bied ons ’n polisemiese netwerk van *in* aan op grond van data uit woordeboeke (d.i. as stap 2) en in §5 op grond van korpusdata (stap 3). Stap 4 van die metodologie word in §6 gedemonstreer aan die hand van enkele konstruksies met taboewoorde, met enkele voorlopige bevindinge in §7. Aangesien ons nie stappe 5 tot 7 van die metodologie hier demonstreer nie, sluit ons in §8 af met idees vir verdere navorsing.

<sup>1</sup> Fontaine (2017) stel: “... let us assume there is a class, called simply P, that includes items such as *in, on, out, up, behind, to, under, over,* and many more. We can then refer to these items as ‘P-items’. ‘Item’ as a term is preferred to ‘word’ to avoid the problems associated with ‘word’, especially the prevalence of the orthographic word as the primary reference ...”

<sup>2</sup> Tensy anders vermeld, kom alle voorbeelde in hierdie bydrae uit die NWU-Kommentaarkorpus 1.5 wat uit kommentare bestaan wat gebruikers onderaan verskeie webbladsye aangebring het. Hierdie weergawe (1.5) het in Oktober 2021, toe die soektogte gedoen is, uit 24 717 080 woorde bestaan; die korpus is beskikbaar as deel van die Virtuele Instituut vir Afrikaans (VivA) se eksklusiewe korpusversameling (VivA-KPO), toeganklik by <https://viva-afrikaans.org/>.

Weens die omvang van die netwerke, word slegs gedeeltes van die volledige netwerke<sup>3</sup> in die figure in hierdie artikel aangebied.

## **2. Metodologie**

Breedweg gebaseer op die werk wat Tyler en Evans (2001) gedoen het om die polisemiese netwerk van Engelse *over* te beskryf, stel ons onderstaande konkrete, praktiese stappe in die identifisering van konstruksienetwerke voor.

1. Identifiseer die protoscène, d.i. die skematiese mentale voorstelling van die konstruksie se primêre betekenis (“abstracted mental representation of the primary sense” (Tyler & Evans 2001:735)), soos wat dit op sinchroniese en/of diachroniese gronde geïdentifiseer kan word. Die vertrekpunt is dus betekenis wat op algemene kognitiewe vermoëns (soos subjektivering, konstruering, profilering, kategorisering, ens.) gebaseer is (kyk Langacker 1987). Die protoscène verteenwoordig daarom ook die konseptualiseringspool (betekenis) van die protokonstruksie, wat gedefinieer kan word as die oorspronklike/oudste vorm-betekenispaar.
2. Konstrueer ’n literatuurgebaseerde konstruksienetwerk. Vir leksikale konstruksies is omvattende en standaardwoordeboeke ’n goeie vertrekpunt; vir ander grammatikale konstruksies is vorige taalkundige beskrywings ’n wegspringplek.
3. Konstrueer ’n korpusgebaseerde konstruksienetwerk, onafhanklik van die literatuurgebaseerde netwerk. Die bedoeling is dat ’n mens uit korpusdata ’n konstruksienetwerk laat ontluik, met so min as moontlik kontaminasie deur vorige beskouings. Dit is daarom verkieslik dat stap 2 en 3 onafhanklik deur twee of meer navorsers, navorserspanne of annoteerders uitgevoer word.
4. Amalgameer die twee netwerke met mekaar deur hulle te vergelyk. Dit is veral belangrik om te let op konstruksies wat in die een maar nie in die ander netwerk voorkom nie; in sulke gevalle kan daar na data gesoek word om te bevestig of ’n bepaalde node ingesluit moet word al dan nie. Dit kan ook dui op verouderde konstruksies en/of ontluikende konstruksies.
5. Rangskik die netwerk op ’n temporele as, op basis van historiese bronne (byvoorbeeld historiese korpora, historiese literatuur, ouer woordeboeke of etimologiese woordeboeke). Let ook weer hier op verouderde en/of ontluikende konstruksies.
6. Dui uitbreidings- en verlengingsverhoudings, en ooreenkomste tussen nodes aan. Dui ook op prototipiese nodes, nodes wat sterk gevestig is, ensovoorts. In hierdie stap moet die rol van aanlyn betekeniskonstruering (gebaseer op ensiklopediese kennis), pragmatiese versterking en die konseptuele belangrikheid van grammatikale konstruksies (byvoorbeeld morfologiese en sintaktiese konstruksies) verreken word.
7. Maak gevolgtrekkings en voorspellings ten opsigte van konstruksionaliserings (d.i. nuwe nodes in die netwerk) en faktore wat bydra tot konstruksieverandering (byvoorbeeld taalkontak).

Uiteraard kan hierdie programmatiese raamwerk nog in veel meer detail uitgewerk en aangepas word, maar ons beskou bostaande as genoegsame praktiese stappe om konkrete uitkomst te kan bereik. Ons demonstreer vervolgens stappe 1 tot 4 bondig aan die hand van die setsel *in*; stappe 5 tot 7 bied geleentheid vir verdere navorsing.

---

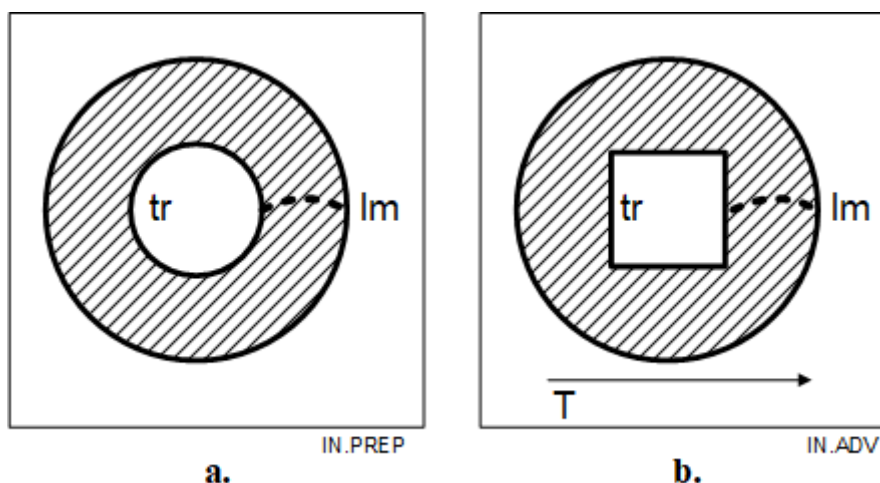
<sup>3</sup> Die volledige netwerke, tesame met al die data wat in hierdie artikel bespreek word, is beskikbaar by <https://gerhard.pro/publications/vanhuyssteenetal2023a/>.

### 3. Die protoscène van *in*

Philippa et al. (2003-2009) toon aan dat *in*.PREP reeds aan die einde van die 8ste eeu in Oudnederlands opgeteken is: *gelōbistu in got alamehtigen fadaer* ‘glo jy in God die almagtige Vader?’. In die 10de eeu tref ons ook reeds “in”, “aan”, “op”, “na” en “te” as polisemiese betekenisonderskeidings met betrekking tot beide ruimte en tyd (as abstraktum) aan, asook vir die gebruik van *in* (met wisselvorme *inna* en *inne*) as bywoord met die betekenis “binne”. Opmerklik is dat *in*.ADV in die 10de eeu al as deel van samekoppelings (“separable complex verbs”) voorkom, byvoorbeeld *ingān* ‘ingaan’ (en in die 11de eeu: *indragon* ‘na binne dra’).

In die protogeskiedenis word *in*.PREP en *in*.ADV onderskeidelik na Proto-germaanse *\*in-* en *\*inna* teruggelei, en beide na Proto-Indo-Europese *\*h<sub>1</sub>en* ‘in’. Die moontlikheid bestaan egter ook dat dit verband kan hou met, of teruggelei of beïnvloed kan wees deur Proto-Indo-Europese *\*h<sub>1</sub>on* ‘in, na’ en sy afgeleide *\*h<sub>1</sub>enter* ‘tussen, binne’ (Philippa et al. 2003-2009).

Ons sien dus van die vroegste tye af dat *in* gekonseptualiseer word as ’n verhouding tussen een entiteit – die trajektor (tr in Figuur 1) – wat bevat of omvat word deur ’n ander entiteit – die landmerk (lm) – binne ’n bepaalde soekdomein (“search domain”, wat in Figuur 1 deur arsering aangedui word). Hoewel dieselfde INHOUD/HOUER-beeldskema die setsel en die bywoord onderlê, word die verhouding in die protoscène van *in*.PREP as ’n simplekse atemporele verhouding verbeeld (Figuur 1a), terwyl dit in *in*.ADV as ’n komplekse atemporele verhouding (met ’n ongeprofileerde temporele dimensie en ’n verhouding as trajektor) verbeeld word (Figuur 1b) (vergelyk onder andere Langacker (1987), hoofstuk 6 en 7, maar veral pp. 242-243). Vir doeleindes van hierdie bydrae is dit belangrik om te verstaan dat *in*.PREP en *in*.ADV se konseptualisering nie onderling uitsluitend is nie, maar eerder ’n kategorie met oorvleuelende grense verteenwoordig. Onses insiens bied ’n Langackeriaanse beskrywingsraamwerk by uitstek die nodige konstrakte en terme om enige protoscène omvattend en gedetailleerd te beskryf.



Figuur 1. Protoscène van (a) *in*.PREP; en (b) *in*.ADV

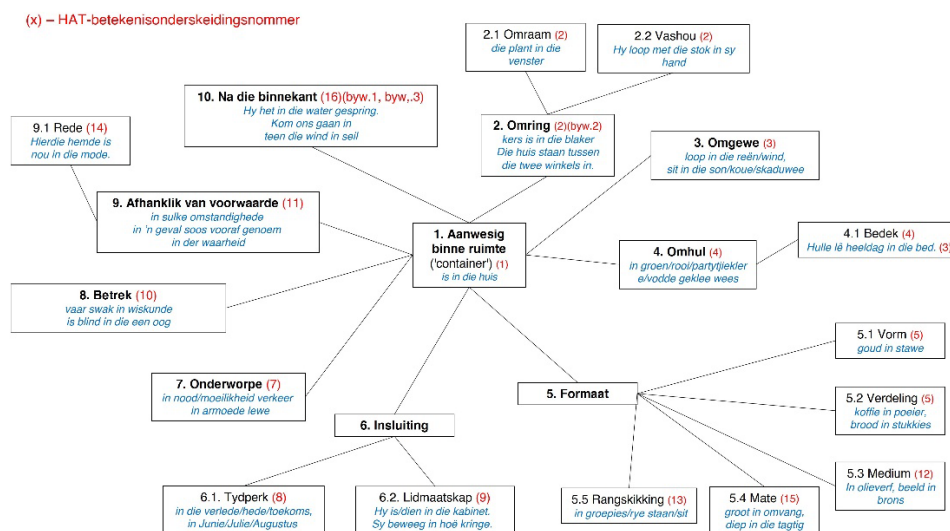
### 4. *In* in woordeboeke

Stap 2 in die voorgestelde metodologie bepaal dat ’n mens konstruksienetwerke op grond van bestaande literatuur saamstel. Aangesien ons hier op leksikale aspekte van *in* fokus, is ’n goeie

vertrekpunt die inligting wat in standaard- en omvattende woordeboeke vervat is. Die veronderstelling van dié metode is dat leksikograwe van gerenommeerde woordeboeke oor uitstekende vermoëns beskik om betekenisonderskeidings uit groot hoeveelhede taaldata te ontwar. Boonop verlaat jy jou uiteindelik nie net blindweg op hulle interpretasie van hulle data nie, maar word hulle interpretasies later in die metodologie ook geverifieer, gekontroleer en verfynd aan die hand van jou eie data (kyk §5).

Ons gebruik die *Handwoordeboek van die Afrikaanse Taal* (HAT, 2015) as 'n standaardwoordeboek en die *Woordeboek van die Afrikaanse Taal* (WAT, 1961) as 'n omvattende woordeboek; die veronderstelling is dat 'n standaardwoordeboek 'n bondige, oorsigtelike beeld bied, terwyl 'n omvattende woordeboek 'n volledige, gedetailleerde beeld gee. HAT (2015) onderskei by die lemma *in* tussen vier woordsoorte: PREP met 16 poliseme, ADV met drie, ADJ slegs een en V twee. WAT (1961) erken dieselfde vier woordsoorte: PREP met sewe primêre betekenisonderskeidings en 18 sekondêres, ADV met vyf primêres en vyf sekondêres, ADJ met twee, en V ook slegs twee (dus 39 in totaal). Aangesien ons hier slegs met PREP en ADV werk, is daar onderskeidelik 19 (HAT) en 35 (WAT) betekenisonderskeidings om mee te werk.

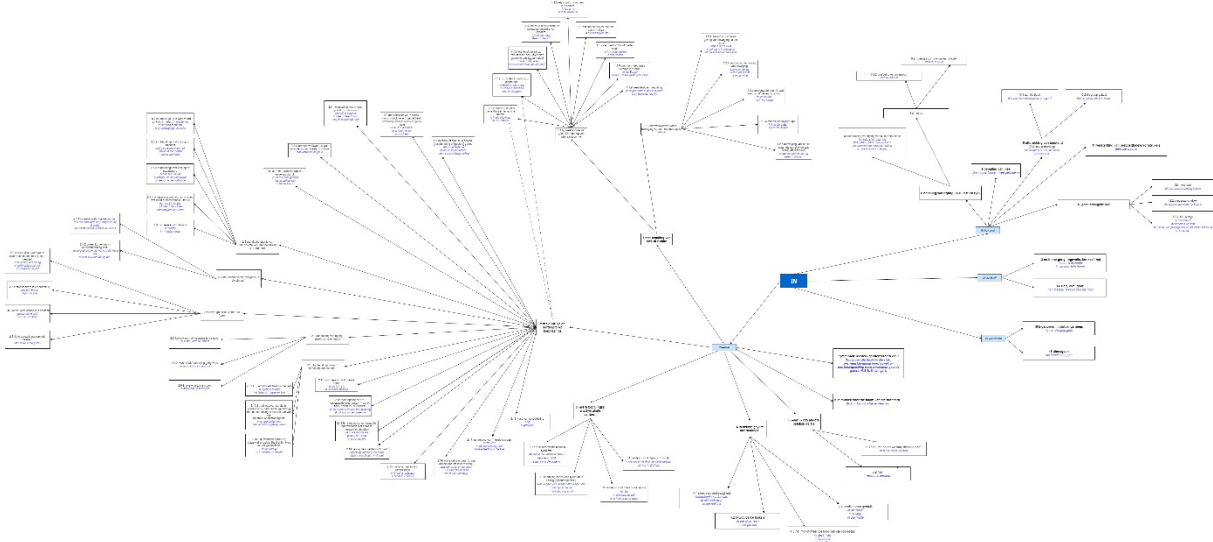
Twee van die navorsers (hierna die woordeboekspan) het met die HAT-data begin deur eers die betekenisonderskeidings op 'n wyse te herformuleer wat meer taalkundig teoreties van aard (eerder as eindgebruikersvriendelik) is, met die oogmerk dat 'n mens uiteindelik verskillende woordeboeke en verskillende datastelle eenvormig met mekaar kan vergelyk. Elk van die betekenisonderskeidings (tesame met 'n voorbeeld uit die woordeboek) is toe in blokkies in grafiesevisualiseringsprogrammatuur (soos Microsoft Visio of Draw.io) geplaas, sodat dit as nodes in 'n visuele netwerk kan dien. Daarna is die nodes gekategoriseer deur betekenis saam te groepeer; sommige te verenig in een node, maar ander te skei; tipes verhoudings tussen die nodes aan te dui; ens. Vir die data uit HAT (2015) het dit tot 'n netwerk met 'n sentrale nodus (aangedui met [1.] in Figuur 2), agt sekondêre nodes (soos [2. *omring*], [3. *omgewe*], ens.) en elf tersiêre nodes (soos [2.1 *omraam*] en [2.2 *vashou*] by [2. *omring*]) gelei.



Figuur 2. Polisemiese netwerk van *in* in HAT<sup>4</sup>

<sup>4</sup> 'n Volgroottegewing van die hierdie skematiese voorstelling is aanlyn beskikbaar by <https://gerhard.pro/publications/vanhuyssteenetal2023a/>.

Hierna is dieselfde proses vir die WAT-data (1961) deur die woordeboekspan gevolg, met die resultaat soos te sien in Figuur 3. Dié veel omvattender netwerk het 16 primêre, 38 sekundêre en 33 tersiêre nodes. Wat veral gerieflike was, is dat die WAT (1961) ook die hoofonderskeidings gemaak het wat ooreenstem met die protoscène (§3).



Figuur 3. Polisemiese netwerk van *in* in WAT<sup>5</sup>

Afhangend van wat 'n mens se doel is, sou 'n mens die verskillende woordeboeke se netwerke met mekaar kon amalgameer (soos wat Van Huyssteen [2021, 2022] gemaak het) om ooreenkomste en verskille tussen verskillende woordeboeke uit te lig, om slegs 'n enkele netwerk oor te hou, ensovoorts. Vir ons doeleindes het ons eerder gefokus op 'n addisionele netwerk wat op gebruiksdata in korpora gebaseer is.

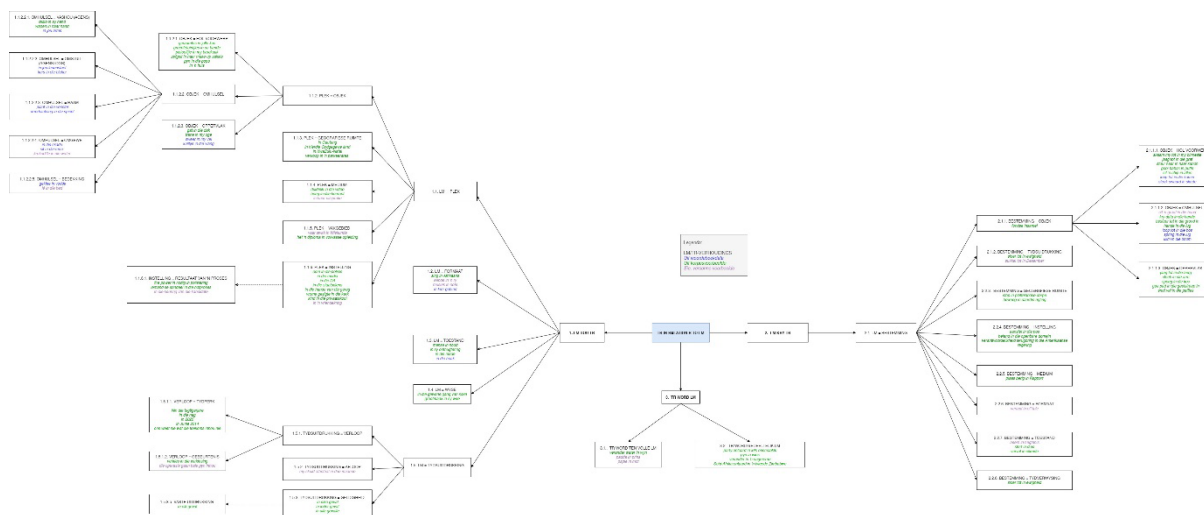
## 5. *In* in korpusdata

Onafhanklik van die woordeboekspan maar met inagneming van hulle formalisme, het 'n tweede span (die korpusspan) 'n polisemiese netwerk uit korpusdata saamgestel. Vir doeleindes van hierdie ondersoek het ons slegs data uit die NWU-Kommentaarkorpus 1.5 (kyk voetnoot 2) gebruik. Alle trefslae van *in* met 'n konteks van vyf woorde links en vyf regs van *in* is onttrek, wat 'n totaal van 339 610 trefslae opgelewer het. Hieruit is 'n lukrake gerieflikheidsteekproef van 500 trefslae onttrek en tussen die lede van die korpusspan verdeel. Elke geval is geannoteer met betrekking tot vorm, funksie en betekenis, byvoorbeeld wat die woordsoort van die woorde voor en ná *in* is, watter tipe entiteite *in* in verhouding tot mekaar stel, wat die funksie van die setselstuk (met *in* as kern) in die sin is, ensovoorts.

In Figuur 4 word die netwerk aangebied wat uit hierdie data saamgestel is. Vanuit die sentrale node [TR IN VERHOUDING R TOT LM] ontspring drie primêre nodes:

1. [LM HOU TR]
2. [LM KRY TR]
3. [TR VERANDER IN LM]

<sup>5</sup> 'n Volgrootweergawe van die hierdie skematiese voorstelling is aanlyn beskikbaar by <https://gerhard.pro/publications/vanhuyssteenetal2023a/>.



Figuur 4. Polisemiese netwerk van *in* in korpusdata<sup>6</sup>

Benewens bogenoemde vier nodes, het die netwerk nog 40 ander nodes, wat dit dus aansienlik minder kompleks/gedetailleerd as die WAT-netwerk (Figuur 3) maak. ’n Mens moet egter baie versigtig wees om gevolgtrekkings oor die korrektheid en/of kwaliteit van die woordeboeke of die korpusdata op grond van die grootte van ’n netwerk te maak. In die WAT se geval word baie fyn onderskeide getref wat nie in die steekproefdata teenwoordig was nie, gewoon omdat die WAT se skopus vereis dat alle moontlikhede opgeteken moet word. Nietemin, indien die netwerke vergelyk en/of geamalgameer word, kan nuttige afleidings oor byvoorbeeld prototipiese betekenisonderskeidings gemaak word. Vir ons doeleindes wou ons kyk of daar dalk op grond van ander gebruikdata nuwe nodes in hierdie netwerk sou wees, wat op hulle beurt weer aanduidings van konstruksionele verandering en/of konstruksionalisering sou kon wees. Ons het dus vervolgens na gebruikdata van taboewoorde gekyk, soos in die volgende afdeling bespreek word.

## 6. In in konstruksies met taboewoorde

Om die polisemiese netwerke wat op woordeboekdata (Figuur 2 en Figuur 3) of korpusdata (Figuur 4) gebaseer is, te verifieer, verfyn of aan te vul, kan ’n mens dit met data uit **ander woordeboeke** (of selfs anderstalige woordeboeke – kyk Van Huyssteen 2021), uit ’n **addisionele steekproef** van dieselfde tipe data, en/of teen ’n **volledige** dataset van ’n bepaalde domein/subkategorie doen. Ons het laasgenoemde opsie gekies deur ’n volledige dataset van konstruksies waar *in* met bepaalde vloekwoorde voorkom, te gebruik.

Waarom spesifiek sulke konstruksies met taboewoorde? ’n Sentrale vooronderstelling wat ons maak, is dat perifere taalinnovasies ’n belangrike faktor is wat taalverandering stu (Blythe & Croft 2021). Vloek, skel en ander vorme van taboetaalgebruik is goeie voorbeelde van sulke kreatiewe innovasies wat taalverandering kan fasiliteer; vergelyk in dié verband Burridge en Benczes (2019) se hoofstuk “Taboo as a driver of language change” in *The Oxford Handbook of Taboo Words and Language*. Op grond hiervan stel ons dan breedweg die vloek-is-’n-omen-van-konstruksionaliseringshipotese (VOK-hipotese):

<sup>6</sup> ’n Volgroottegeweegawe van die hierdie skematiese voorstelling is aanlyn beskikbaar by <https://gerhard.pro/publications/vanhuyssteenetal2023a/>.

Veranderinge in 'n konstruksienetwerk word (onder andere) gekenmerk deur die teenwoordigheid van taboekonstruksies in daardie deel van die netwerk.

Metodologies stel ons dus:

As 'n mens na konstruksieverandering en konstruksionaliserings in 'n netwerk soek, kan konstruksies met taboewoorde as een van die vertrekpunte geneem word om 'n soektog van stapel te stuur.

'n Derde span navorsers (die vloekspan) het soektogte gedoen met tien prototipiese vloekwoorde wat as naamwoorde gebruik kan word, te wete *fok, poes, bliksem, doos, kak, donder, kont, moer* en *naai*.<sup>7</sup> Daar is in die korpus na alle gevalle gesoek waar *in* binne die bestek van drie woorde voor of na die bepaalde vloekwoord voorkom. Die data is daarna handmatig gekontroleer om gevalle te verwyder waar die vloekwoord nie 'n argument van *in* is nie (soos *Ek hoop hulle rape die kak uit hom uit in die tjoekie!*, waar *kak* 'n argument van *uit* is, maar nie van *in* nie).

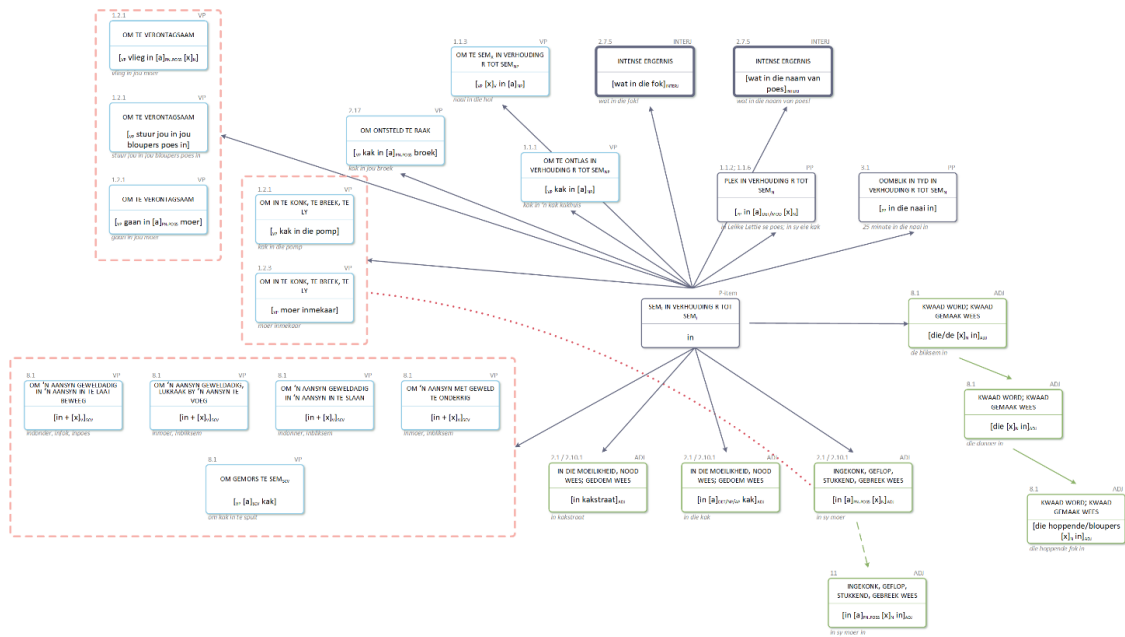
In die finale datastel van 432 gevalle (wat saam met die res van die data beskikbaar is) is *in* in verreweg die meeste gevalle (n=213) 'n voorsetsel, gevolg deur *in* as alkantsetsel (n=105), agtersetsel (n=77) en as deel van 'n skeibare komplekse werkwoord (n=37). Op grond van frekwensie kan vier prototipiese konstruksies geïdentifiseer word (waarvan [b] 'n versterkte vorm van [a] is):

- (1) [in [a]<sub>PN.POSS</sub> [X]<sub>N</sub>]<sub>ADJ</sub> ⇔ [ingekonk, geflop, stukkend, gebreek wees]  
*Op die ou end is hulle almal in hulle moer.*
- (2) [in [a]<sub>PN.POSS</sub> [X]<sub>N</sub> in]<sub>ADJ</sub> ⇔ [ingekonk, geflop, stukkend, gebreek wees]  
*En Argentinië is nou ook in hulle moer in.*
- (3) [in [a]<sub>DET/NP/AP</sub> kak]<sub>ADJ</sub> ⇔ [in die moeilikheid, nood wees; gedoem wees]  
*Ons volk is in die kak is 'n understatement.*
- (4) [die/de [x]<sub>N</sub> in]<sub>ADJ</sub> ⇔ [kwaad word; kwaad gemaak wees]  
*Dit maak my somer die donder in!*

Die vloekspan het ook 'n konstruksienetwerk van die volledige vloekdatastel opgestel – kyk Figuur 5. Dié netwerk is anders as die ander in die sin dat dit nie 'n polisemiese netwerk van *in* is nie, maar wel 'n konstruksienetwerk met konstruksie-idiome, spesifieke instansiërings, konstruksie-spesifieke betekenis, ensovoorts. Dié netwerk het dan uiteindelik saam met al die ander netwerke gedien om 'n geamalgameerde netwerk op te stel, soos in die volgende afdeling bespreek word.

<sup>7</sup> Die keuse van dié tien woorde is baseer op Van der Merwe (2022).

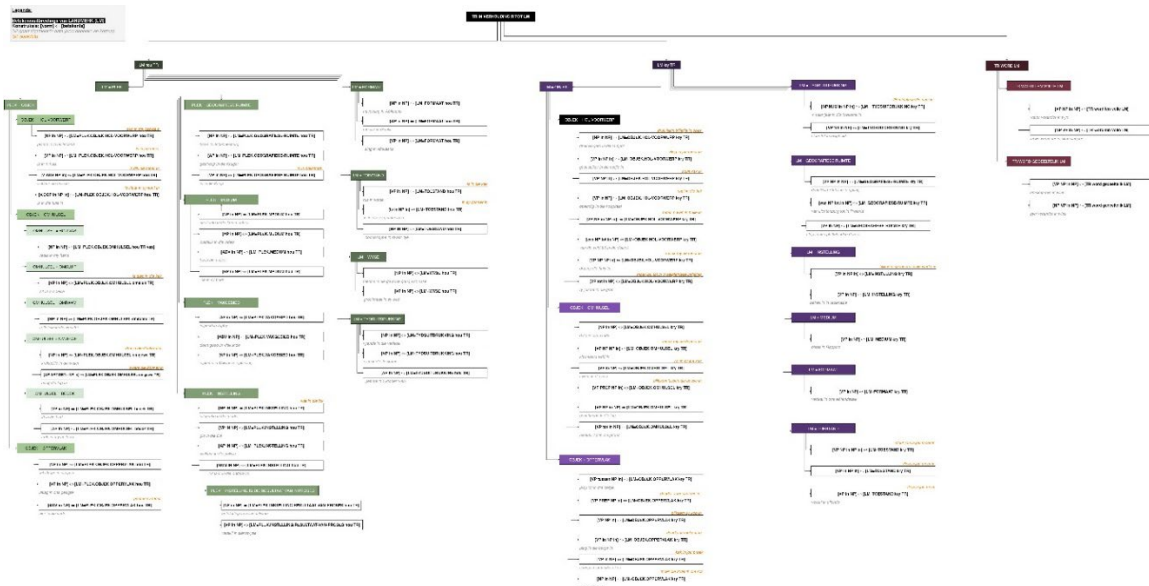




Figuur 5. Netwerk van *in*-konstruksies met vloekwoorde<sup>8</sup>

## 7. Enkele bevindinge

Met vier konstruksienetwerke wat op vier datastelle gebaseer is in die hand, kon die navorsingspan saam-saam 'n geamalgameerde netwerk opstel; die resultaat is in Figuur 6 te sien. As 'n mens nou al hierdie netwerke met mekaar en met die geamalgameerde netwerk vergelyk, kan 'n mens reeds enkele interessante waarnemings doen, selfs sonder om al die stappe in die metodologie uit te voer.



Figuur 6. Geamalgameerde polisemiese netwerk van *in*<sup>9</sup>

<sup>8</sup> 'n Volgroottegewag van die hierdie skematiese voorstelling is aanlyn beskikbaar by <https://gerhard.pro/publications/vanhuyssteenetal2023a/>.

<sup>9</sup> 'n Volgroottegewag van die hierdie skematiese voorstelling is aanlyn beskikbaar by <https://gerhard.pro/publications/vanhuyssteenetal2023a/>.

- *In* in woordeboeke
  - Vir omvattende taalkundige beskrywings (soos waarmee ons hier besig is) kan 'n mens vanuit die staanspoor fokus op data uit 'n omvattende woordeboek. Die inligting in 'n handwoordeboek kan wel gebruik word om jou netwerk te verifieer, verfyn, ensovoorts, maar dit is nie nodig om 'n addisionele netwerk vir sulke data saam te stel nie.
  - 'n Omvattende woordeboek met diachroniese inligting (soos die *Oxford English Dictionary* of die *Woordenboek der Nederlandse Taal*) is onontbeerlik vir die opstel van konstruksionaliseringsnetwerke; daarsonder is dit 'n veel intensiewer navorsingstaak.
- *In* in korpusdata
  - Anders as Tyler en Evans (2001) het ons slegs die landmerk se betekenis gebruik om die verskillende polisemiese betekenis te onderskei. Ons poging was onses insiens suksesvol en bied dus 'n moontlike werkswyse vir soortgelyke beskrywings – spesifiek ook vir die kategorisering van P-items.
  - Die gebruik van korpusdata in beskrywings van hierdie aard is onontbeerlik, aangesien dit nie net tot verbetering in die kategoriseringsverhoudings gelei het nie, maar ook nuwe nodes in die netwerk aan die lig gebring het. 'n Mens moet dus daarteen waak om nie eksklusief op woordeboekdata staat te maak nie.
  - Uit frekwensies blyk dit dat *in* dikwels in frase (d.i. vaste uitdrukkings, of nievrye frases) gebruik word, byvoorbeeld:
    - in vaste verbinding met 'n spesifieke werkwoord, soos *belangstel in; ontaard in; glo in*; en so meer;
    - in vaste verbinding met 'n spesifieke werkwoord en 'n spesifieke setselkomplement, soos *gaan terug in tyd; stort in duie; neem in oënskou*; en so meer;
    - in idiome, soos *om iemand 'n gat in die kop te praat; om jou in 'n ander bloedgroep in te skrik*; en so meer.
  - Voorsetselstukke met *in* kom baie meer dikwels voor as adjunkte (soos *speel in die Super 12; laat in die nag*, of *doen goed in wiskunde*), of komplemente (kyk volgende kolpunt), as wat dit as nabepalers in naamwoordstukke (soos *die man in die kamer*) voorkom.
  - Voorsetselstukke met *in* kom verrassend dikwels as komplemente voor, waaronder as:
    - kopulapredikate (soos *is in die moeilikheid*);
    - objekkomplemente (soos *gooi melk in die kan*); en
    - komplemente van oorganklike adjektiewe (soos onskuldig **in die ongeluk** wees, of welkom **in die polisie** wees).
- *In* in taboewoordkonstruksies
  - Die voorsetselkonstruksie [in [a]<sub>PN.POSS</sub> [x]<sub>N</sub>]<sub>ADJ</sub> (kyk (7) hier bo) het 'n emfatiese, versterkte weergawe in die alkantsetselkonstruksie [in [a]<sub>PN.POSS</sub> [x]<sub>N</sub> in]<sub>ADJ</sub> (kyk (8) hier bo).
  - In teenstelling met die woordeboek- en korpusnetwerk, beklee die agtersetselkonstruksie [die/de [x]<sub>N</sub> in]<sub>ADJ</sub> 'n baie prominenter rol in die vloeknetwerk. Die redes hiervoor, en met spesifieke verwysing na die oorsprong van hierdie konstruksie, is vrugbare grond vir verdere navorsing.
  - Die taboewoordkonstruksies het die netwerk op ten minste tien plekke uitgebrei met konstruksies wat nie ook in die algemene datastel te vinde was

nie. Ons metodologiese stelling – dat konstruksies met taboewoorde as een van die vertrekpunte geneem kan word om soektogte na konstruksieverandering en konstruksionaliserings in 'n netwerk te begin – blyk dus voorlopig tog van waarde te wees, maar moet uiteraard ook verder beproef word.

## **8. Volgende stappe**

Ons het in hierdie artikel 'n sewestapmetodologie voorgelê en stappe 1 tot 4 aan die hand van konstruksies met die P-item *in* geïllustreer. In §7 het ons enkele nuttige, praktiese wenke verskaf, asook enkele voorlopige waarnemings oor die data gemaak. Vir opvolgwerk wil ons ander navorsers uitnoui om dieselfde metodologie te volg en verbeteringe aan die metodologie aan te bring. Aan die ander kant sal dit ook interessant wees om ander metodologieë met hierdie sewestapmetodologie te vergelyk; in dié verband moet veral ook gelet word op ander navorsing wat oor talige netwerke gedoen word (byvoorbeeld Mehler et al. 2016).

Die laaste woord oor konstruksies met *in* – om van ander P-items nie te praat nie – is nog lank nie gespreek nie. In Pilon et al. (2023) verken ons *in* verder, met die primêre fokus om die onderskeid tussen *in* as verskillende soorte setsels, werkwoordpartikels, adjektiewe en bywoorde op te helder (sien ook Pilon et al. 2023).

Ons ondersoek ook die moontlikhede om netwerke outomaties eerder as handmatig te kan saamstel en teken. Idealiter wil 'n mens sulke netwerke met geen (of slegs minimale) annotering uit rou korpusdata aflei, en daarom ondersoek ons ook tans van die tegnieke wat byvoorbeeld in genetiknavorsing gebruik word (kyk byvoorbeeld Wolf et al. 2016).

## **Erkenning**

Die projek *Wat de Vloekwoord! Multidissiplinêre navorsing en wetenskapskommunikasie oor vloek* waarbinne hierdie navorsing uitgevoer is, is ten dele moontlik gemaak deur finansiële ondersteuning van die Suid-Afrikaans Akademie vir Wetenskap en Kuns. Die in natura-bydraes van BlueTek Computers, WatKykJy.co.za, Afrikaans.com en Maroela Media word ook hiermee met dank erken. Geen van die opinies in hierdie artikel kan egter aan enige van hierdie instansies en/of hulle medewerkers toegedig word nie.

## **Outeursbydrae**

Al ses outeurs is deel van 'n virtuele konstruksiegrammatikawerkgroep en werk sodoende voortdurend in gelyke dele saam aan die skep van intellektuele eiendom. Vir hierdie artikel het Gerhard B van Huyssteen die meeste van die teks van die artikel geskryf, en hy was ook verantwoordelik vir die vloekdata-annotasie en die opstel van die vloeknetwerk; Adri Breed was verantwoordelik vir die konseptualisering van die tema, die eerste weergawe van die woordeboek- en korpusnetwerke, algemenedata-annotasie en projekbestuur; Anneke Butler was verantwoordelik vir kwaliteitskontrole (netwerke en annotasies) en algemenedata-annotasie. Die ander drie skrywers (in alfabetiese volgorde) was verantwoordelik vir algemenedata-annotasie. Alle outeurs het in gelyke mate deelgeneem aan gesprekke oor die konseptuele raamwerk, navorsingsontwerp en data-analises.

## Belangeverklaring

Die outeurs verklaar dat: (i) hulle vertrou is met die inhoud van die artikel en daarmee saamstem; (ii) die artikel hulle oorspronklike intellektuele eiendom is, sonder enige plagiaat; (iii) die manuskrip nie gelyktydig by verskillende tydskrifte vir keuring voorgelê is nie; en (iv) daar geen belangebotsing – insluitende finansiële belang – is om te verklaar nie. Al die outeurs gee toestemming dat die artikel in *Stellenbosch Papers in Linguistics* gepubliseer mag word. Dit word wel hier verklaar dat daar in oorleg met die Noordwes-Universiteit besluit is dat 'n onafhanklike projekwebblad vir bogenoemde projek opgerig en onderhou word, ten einde die navorsing ten beste te operasionaliseer. Dié webblad, [vloek.co.za](http://vloek.co.za), is ontwikkel en word besit deur Viridevert NPC (2016/411799/08), 'n maatskappy sonder winsbejag wat by die SAID geregistreer is as 'n opvoedkundige openbare weldaadsorganisasie (OWO) met belastingvrystelling ('n sogenaamde “artikel 18(a)-maatskappy”) en met 'n eie bankrekening by FNB. Gerhard B van Huyssteen is een van die direkteure van die maatskappy

## Etiëklaring

Hierdie bydrae verskyn onder die vaandel van 'n projek waarvan die oorhoofse etiëklaring op 21 Mei 2019 by die Noordwes-Universiteit se Language Matters Ethics Committee geregistreer is; die registrasienommer is NWU-00632-19-A7.

## Verwysings

- Barðdal, Jóhanna, Elena Smirnova, Lotte Sommerer & Spike Gildea (eds). 2015. *Diachronic construction grammar*. Constructional approaches to language, Vol. 18. Amsterdam: John Benjamins. <https://doi.org/10.1075/cal.18>.
- Blythe, Richard A & William Croft. 2021. *How individuals change language*. *PLoS One* 16(6):e0252582. <https://doi.org/10.1371/journal.pone.0252582>.
- Burridge, Kate & Reka Agnes Benczes. 2019. Taboo as a driver of language change. In Keith Allan (ed), *The Oxford handbook of taboo words and language*, 180–198. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198808190.013.10>.
- Colleman, Timothy. 2015. Constructionalization and post-constructionalization: The constructional semantics of the Dutch *krijgen*-passive in a diachronic perspective. In Barðdal, Smirnova, Sommerer & Gildea (eds), 213–255. <https://doi.org/10.1075/cal.18.07col>.
- Colleman, Timothy. 2016. A reflection on constructionalization and constructional borrowing, inspired by an emerging Dutch replica of the ‘time’-away construction. *Belgian Journal of Linguistics* 30:91–113. <https://doi.org/10.1075/bjl.30.05col>.
- Engels, Renata & Mar Garachana Camarero. 2021. Grammaticalization, lexicalization, and constructionalization. In Xu Wen & John R. Taylor (eds), *The Routledge handbook of Cognitive Linguistics*, 314–332. New York: Routledge. <https://doi.org/10.4324/9781351034708>.

- Fagard, Benjamin & Alexandru Mardale. 2012. The pace of grammaticalization and the evolution of prepositional systems: Data from Romance. *Folia Linguistica* 46(2):303–340. <https://doi.org/10.1515/flin.2012.011>.
- Fontaine, Lise. 2017. On prepositions and particles: A case for lexical representation in systemic functional linguistics. *WORD* 63(2):115–135. <https://doi.org/10.1080/00437956.2017.1309029>
- HAT: Luther, Jana, Fred Pfeiffer & Rufus H. Gouws. 2015. *Handwoordeboek van die Afrikaanse taal [Standard dictionary of the Afrikaans language]*. 6th ed. Cape Town: Pearson.
- Hilpert, Martin. 2013. *Constructional change in English: Developments in allomorphy, word formation, and syntax*. Cambridge: Cambridge University Press.
- Hilpert, Martin. 2021. *Ten lectures on diachronic Construction Grammar*. Distinguished lectures in Cognitive Linguistics. Leiden: Koninklijke Brill.
- Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar: Theoretical prerequisites*. Vol. 1. Stanford: Stanford University Press.
- Mehler, Alexander, Andy Lücking, Sven Banisch, Philippe Blanchard, & Barbara Job (eds). 2016. *Towards a Theoretical Framework for Analyzing Complex Linguistic Networks*. Berlin: Springer-Verlag.
- Philippa, Marlies, Frans Debrabandere, Arend Quak, Tanneke Schoonheim & Nicoline van der Sijs. 2003–2009. *Etymologisch woordenboek van het Nederlands*. 4 vols. Amsterdam: Amsterdam University Press.
- Pilon, Suléne, Lande Botha, Maristi Partridge, Adri Breed, Anneke Butler & Gerhard B. van Huyssteen. 2023. Postpositions take centre stage: What can we learn about Afrikaans postpositions from descriptions in Dutch and English?. *Proceedings of the 4th International Afrikaans Grammar Workshop*. <https://gerhard.pro/publications/PilonEtAl2023>.
- Rhee, Seongha. 2018. Grammaticalization of adpositions: An introduction. *Language and Linguistics* 81:1–20. <https://doi.org/10.20865/20188101>.
- Sommerer, Lotte & Elena Smirnova (eds). 2020. *Nodes and networks in diachronic Construction Grammar*. Constructional approaches to language, Vol. 27. Amsterdam: John Benjamins. <https://doi.org/10.1075/cal.27>.
- Traugott, Elizabeth Closs. 2021. *Language change*. Online course. Brazil: Abralín EAD. <https://ead.abralin.org/course/view.php?id=20> (datum van besoek 11 Desember 2021).
- Traugott, Elizabeth Closs & Graeme Trousdale. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199679898.001.0001>.

Tyler, Andrea & Vyvyan Evans. 2001. Reconsidering prepositional polysemy networks: The case of *over*. *Language* 77(4):724–765. <https://doi.org/10.1353/lan.2001.0250>.

Van der Merwe, Mart-Mari. 2022. Prototipiese Afrikaanse taboewoorde. Ongepubliseerde MA-verhandeling. Pretoria: Universiteit van Pretoria.

Van Huyssteen, Gerhard B. 2021. “When a word is befok.” Afrikaans Grammar Workshop III, Amsterdam, The Netherlands, 29 September – 1 Oktober. <https://vloek.co.za/leesstof/kongresmateriaal/when-a-word-is-befok-agw-2021> (datum van besoek 11 Desember 2021).

Van Huyssteen, Gerhard B. 2022. *Wat ons van 'fok' weet (en nie weet nie)*. *LitNet Akademies (Geesteswetenskappe)*, 19:428–452. <https://doi.org/10.56273/1995-5928/2022/j19n3b11>.

Van Huyssteen, Gerhard B & Roald Eiselen. 2021. Oor feekse en helleveë. *Tydskrif vir Geesteswetenskappe*, 61(4-1):1–27. <https://doi.org/10.17159/2224-7912/2021/v61n4-1a9>.

WAT: Buro van die WAT. 1961. *Woordeboek van die Afrikaanse Taal – Deel IV: H-I*. Stellenbosch: Buro van die WAT. <https://viva-afrikaans.org/> (datum van besoek 11 Desember 2021).

Wolf, Sebastian, Sara Barbosa, Joachim Bucher & Ralf Takors. 2017. Automatic network generation describes differential gene data in user friendly and expeditiously analyzable network views. *Biomedical Genetics and Genomics* 2(1):1–14. <https://doi.org/10.15761/bgg.1000124>.