

Afrikaans as 'n komplekse netwerk: Die woordko-voorkomsnetwerke van woorde in André P. Brink se *Donkermaan* in Afrikaans, Nederlands en Engels

**Authors:**

Burgert A. Senekal¹
Cornelia Geldenhuys¹

Affiliations:

¹Unit for Language Facilitation and Empowerment, University of the Free State, South Africa

Corresponding author:

Burgert Senekal,
burgertsenekal@yahoo.co.uk

Dates:

Received: 04 Nov. 2015

Accepted: 14 Mar. 2016

Published: 31 Aug. 2016

How to cite this article:

Senekal, B.A. & Geldenhuys, C., 2016, 'Afrikaans as 'n komplekse netwerk: Die woordko-voorkomsnetwerke van woorde in André P. Brink se *Donkermaan* in Afrikaans, Nederlands en Engels', *Suid-Afrikaanse Tydskrif vir Natuurwetenskap en Tegnologie* 35(1), a1368. <http://dx.doi.org/10.4102/satnt.v35i1.1368>

Copyright:

© 2016. The Authors.
Licensee: AOSIS. This work is licensed under the Creative Commons Attribution License.

Taal word reeds sedert De Saussure as 'n sisteem benader, en onlangs is die teorie van komplekse sisteme ook binne die taalkunde toegepas. Komplekse sisteme kan egter ook as komplekse netwerke gemodelleer word, en daar is reeds 'n verskeidenheid studies in die buiteland onderneem wat ondersoek instel na die netwerkstruktuur van taal. Die huidige studie volg in die voetspoor van oorsese studies en ondersoek die netwerkstruktuur van Afrikaans deur 'n woordko-voorkomsnetwerk te ontleed, soos saamgestel uit André P. Brink se roman *Donkermaan*. Skakelverspreidingspatrone en die kleinwêreldfenomeen word ondersoek en vergelyk met die Engelse en Nederlandse vertalings van dié roman. Die huidige studie verteenwoordig die eerste netwerkstudie in Afrikaans, en maak eerstens gebruik van Erdős en Rényi se lukrake netwerkmodel, sowel as Barabási en Albert se skaalvrye netwerkmodel, om aan te toon dat die skakelverspreidingspatrone in 'n woordko-voorkomsnetwerk van Afrikaans beter volgens Barabási en Albert as volgens Erdős en Rényi se netwerkmodel beskryf word. Verder word gebruik gemaak van die metode soos voorgestel in Humphreys en Gurney om kleinwêreldsheid (S) presies te kwantifiseer in die Afrikaanse teks, maar ook in die Engelse en Nederlandse vertalings. Met $522 \leq S \leq 797$ word aangedui dat Afrikaans, Engels en Nederlands al drie duidelike kleinwêreldnetwerke is. Voorstelle word ook vir verdere navorsing gemaak.

Afrikaans as a complex network: The word co-occurrence network in André P. Brink's *Donkermaan* in Afrikaans, Dutch and English. Language has already been approached as a system since De Saussure, and recently the theory of complex systems has been applied within Linguistics as well. Complex systems, however, can also be modelled as complex networks, and a variety of studies investigating the network structure of language have already been undertaken worldwide. The current study follows in the footsteps of overseas studies and investigates the network structure of Afrikaans by analysing a word co-occurrence network compiled from André P. Brink's novel *Donkermaan*. Link distribution patterns and the small-world phenomenon are investigated and then compared to the English and Dutch translations of this novel. The current study represents the first network study of Afrikaans. Firstly, the random network model of Erdős and Rényi and the scale-free network model by Barabási and Albert are used to indicate that the link distribution patterns in a word co-occurrence network of Afrikaans are better described according to the network model of Barabási and Albert than by that of Erdős and Rényi. Furthermore, the method proposed by Humphreys and Gurney to define small-worldness (S) was used to quantify this phenomenon for the Afrikaans, as well as English and Dutch versions of the text. With $522 \leq S \leq 797$, it is indicated that Afrikaans, English and Dutch are all clearly small-world networks. Suggestions are also made for further research.

Inleiding

Taal word reeds sedert De Saussure (1966) as 'n sisteem beskou. Latere teoretici het op sy insigte voortgebou (Odendal 1969), en onlangs het die teorie van komplekse sisteme ook inslag begin vind in die taalkunde (Weideman 2009, 2011). Taal word tans algemeen as 'n komplekse sisteem beskou (Cilliers 1998:3; Kwapien & Drożdż 2012:129–130; Solé *et al.* 2010:3), maar volgens Cong en Liu (2014:599) kom die sisteemteoretiese perspektief op taal gewoonlik bloot op 'n metaforiese gebruik neer en word dit nie veel verder gevoer nie.

Binne die teorie van komplekse netwerke word die interafhanklike samehang tussen woorde in 'n taal egter nie bloot in metaforiese terme bespreek nie, en 'n verskeidenheid studies is reeds

Read online:

Scan this QR code with your smart phone or mobile device to read online.

gedoen oor taal as 'n komplekse netwerk (Cong & Liu 2014). Die meerderheid studies konsentreer op Engels, byvoorbeeld i Cancho en Solé (2001), Dorogovtsev en Mendes (2001), Masucci en Rodgers (2006) en Newman (2006), maar ander tale is ook reeds sodanig bestudeer, insluitende Chinees (Liu & Zhao 2010), Pools (Grabska-Gradzińska *et al.* 2012), Kroaties (Margan, Martinčić-Ipšić & Meštrović 2014), Hebreeus (Kenett *et al.* 2011), en verskeie ander tale (Liu & Cong 2013). Daar is egter nog nie 'n studie van Afrikaans as 'n komplekse netwerk onderneem nie.

Netwerkstudies van die ko-voorkoms van woorde maak gereeld gebruik van literêre tekste as datastelle. Grabska-Gradzińska *et al.* (2012) ondersoek byvoorbeeld onder meer literêre tekste; Masucci en Rodgers (2006) bestudeer George Orwell se *1984*; Newman (2006) ondersoek 'n netwerk van byvoeglike naamwoorde in Charles Dickens se *David Copperfield*; Liu en Cong (2013) bestudeer twaalf vertalings van Nikolai Ostrovsky se roman *Kak zakaljalas' stal'* (*How the steel was tempered*); en Margan, Martinčić-Ipšić en Meštrović (2014) bestudeer tien Kroatiese boeke. Die huidige studie volg in die voetspore van hierdie studies en ondersoek die woordko-voorkomsnetwerk in André P. Brink se *Donkermaan* (Brink 2000a), en vergelyk ook die resultate met dieselfde ko-voorkomsnetwerke van die Engelse (Brink 2000b) en Nederlandse (Brink 2001) vertalings van die roman. Let egter daarop dat Brink self die Engelse vertaling behartig het, maar die Nederlandse vertaling is deur Anneke Goddijn en Rob van der Veer onderneem.

Een van die voorstelle vir verdere navorsing wat Solé *et al.* (2010:7) maak, is: Is daar statistiese verskille tussen die netwerke van verskillende tale? Dít is dan die primêre oogmerk van die huidige artikel, naamlik om ondersoek in te stel na of Afrikaans enige beduidende statistiese verskille met ander tale ten opsigte van sy netwerkstruktuur vertoon.

Verder is die topologiese aspek van kleinwêreldsheid, soos voorgestel en gemodelleer deur Watts en Strogatz (1998), wel ondersoek in bogenoemde internasionale studies, maar dié verskynsel is nog nie presies gekwantifiseer in woordko-voorkomsnetwerke nie. Volgens Solé *et al.* (2010:22) en Liang *et al.* (2009:4901) vertoon alle taalnetwerke kleinwêreldsheid, terwyl 'n studie deur Markoavová en Nather (2006:3) ook die siening ondersteun dat Engels 'n kleinwêreldnetwerk is. Buiten Kennett *et al.* (2011) se kwantifisering van kleinwêreldsheid in Hebreeuse semantiese netwerke, is kleinwêreldsheid egter nog nie vir taalnetwerke gekwantifiseer nie. Bykomend tot die bestudering van die netwerkeienskappe van Afrikaans, word kleinwêreldsheid in die huidige artikel in navolging van Kennett *et al.* (2011) volgens die metode soos voorgestel deur Humphries en Gurney (2008) vir die woordko-voorkomsnetwerk in Afrikaans gekwantifiseer, terwyl kleinwêreldsheid ook met dié in Engels en Nederlands vergelyk word deur gebruik te maak van bogenoemde twee vertalings van Brink se roman. Die huidige studie verteenwoordig gevolglik, onder meer, die eerste kwantifisering van kleinwêreldsheid in 'n woordko-voorkomsnetwerk, en verder het die onderhawige studie betrekking op Afrikaans, Engels en Nederlands.

Die kenmerke van taal as 'n komplekse sisteem

Taal word lank reeds as 'n sisteem beskou. In De Saussure (1966:114) se bekende definisie is taal '*a system of interdependent terms in which the value of each term results solely from the simultaneous presence of the others*'. Taal kan ook as 'n komplekse (teenoor 'n eenvoudige) sisteem beskryf word (Cilliers 1998:3; Albert & Barabási 2002:253; Grabska-Gradzińska *et al.* 2012:2; Cong & Liu 2014:599), en deel heelwat eienskappe met ander komplekse sisteme (Strogatz 2004:256; Cong & Liu 2014:602). Newman (2011:1) omskryf 'n komplekse sisteem soos volg:

... most researchers in the field would probably agree that it is a system composed of many interacting parts, such that the collective behavior of those parts together is more than the sum of their individual behaviors. The collective behaviors are sometimes also called 'emergent' behaviors, and a complex system can thus be said to be a system of interacting parts that displays emergent behavior.

Soos Newman in dié omskrywing te kenne gee, word komplekse sisteme onder meer deur opkoms ('*emergence*') gekenmerk, wat omskryf kan word as '*a spontaneous occurrence of macroscopic order from a sea of randomly interacting elements on a microscopic level*' (Kwapien & Drożdż 2012:118; Mitchell 2006:1196; DeLaurentis 2007:364). Opkoms beteken dat die funksie van die totale sisteem meer is as die funksie van individuele onderdele. Taal word ook gekenmerk deur opkoms: terwyl taal op een vlak uit 'n kombinasie van klanke bestaan, voorspel daardie klanke nie dat 'n persoon gelukkig, hartseer, teleurgesteld, ensovoorts sal wees wanneer hy 'n sekere kombinasie aanhoor nie. Dieselfde geld vir letters: die letters *e*, *o*, *p* en *s* kan op maniere gekombineer word wat onder meer 'n literêre vorm (epos), kommunikasiemiddel (e-pos) en vulgariteit (poes) voorstel – wat tel, is *hoe* die letters saamhang. In sisteme val die klem altyd op 'geordenheid en samehang' (Senekal 1987:24), wat die eienskap van sisteme is wat hul van aggregate onderskei (Viljoen 1986:3); opkoms is egter 'n kenmerk van *komplekse* sowel as *eenvoudige* sisteme.

'n Komplekse sisteem is ook daartoe in staat om aan te pas, en dít is wat 'n komplekse van 'n eenvoudige sisteem onderskei (Holland 1992:19; Amaral & Ottino 2004:159). Aanpasbaarheid kom nie in eenvoudige sisteme voor nie: alhoewel 'n motorvoertuig byvoorbeeld gekompliseer is en sy onderdele saamwerk om die sisteem se totale funksie te verrig (om te ry), kom die sisteem tot 'n stilstand wanneer 'n onderdeel onklaar raak. In komplekse sisteme pas die sisteem egter aan, byvoorbeeld die immuunstelsel wat aanpas by 'n nuwe patogeen, of 'n ekosisteem wat aanpas by die verlies van 'n spesie of by klimaatsverandering. Volgens Kwapien en Drożdż (2012:129) is taal ook aanpasbaar:

Under the influence of interactions between members of a community sharing the language, the existing grammar rules are modified, some words are pushed out by other words, there appear completely new ones whose task is to name new objects and concepts, etc. (Dorogovtsev & Mendes 2001:2603.)

Alle komplekse sisteme is noodwendig oop sisteme, wat beteken dat hulle bestaan deur middel van hulle interaksie

met hul omgewing (wat ander komplekse sisteme insluit), en hierdie vlak van interaksie kan beskryf word as 'n horisontale interaksie tussen komplekse sisteme (Senekal 2013:674–675). Tale verkeer natuurlik in interaksie met ander tale, soos Deens en Saksies in elfde-eeuse Engeland, Frans en Engels in Engeland ná 1066 (McCrum, Cran & MacNeil 1992), of Afrikaans en Engels in die hedendaagse Suid-Afrika. Tale ontwikkel deur kontak en beïnvloed mekaar: net soos die Skandinawiërs woorde soos *skill*, *skin*, *want* en *wrong* tydens die okkupasie van Suidoos-Engeland in die vroeë elfde eeu tot Engels bygedra het (McCrum, Cran & MacNeil 1992:71), het Afrikaans woorde soos *braai*, *kraal* en *sommer* tot Suid-Afrikaanse Engels bygedra, terwyl Engels tegelykertyd 'n groot hoeveelheid woorde tot Afrikaans bygedra het. Tale funksioneer ook 'at the interface between biology and social interactions' (Kwapień & Drożdż 2012:129). Taal (mondeling of skriftelik) is afhanklik van biologiese aanpassings soos Broca se area in die brein en die FOXP2-geen, waarsonder dit nie kan bestaan nie, en verkeer gevolglik ook in 'n wisselwerking met biologiese sisteme. Daarbenewens verkeer taal in interaksie met die sosiale en kulturele sisteme waar dit ingebed is, insluitende tegnologie, ideologie, ensovoorts.

Komplekse sisteme is gewoonlik ook sisteme-van-sisteme, waarin 'n hiërargiese struktuur geïdentifiseer kan word waar sisteme verder uit subsisteme bestaan, en terselfdertyd deel uitmaak van groter supersisteme (Senekal 2013:674; Cong & Liu 2014:603). 'n Mens bestaan byvoorbeeld uit, onder meer, 'n senuweestelsel en immuunstelsel, wat opgebou is uit organe en selle, maar behoort ook tot 'n familie, gemeenskap en die mensdom self. Volgens Kwapień en Drożdż (2012:130) kan so 'n hiërargiese struktuur ook in taal waargeneem word:

Language has a hierarchical structure. At the most basic level, it consists of phonemes (spoken language) and characters or ideograms (written language). Typically, [the] number of either of these elements in a given language is small and usually reaches several tens (for example, in British English there are exactly 26 characters and about 45–50 phonemes in active use; the latter number varies depending on a source). The phonemes and characters group themselves in morphemes, which play an important role of fundamental carriers of meaning. The morphemes are not self-reliant, however. The function of the smallest self-reliant components of language is played by the words consisting of one or more connected morphemes. A higher level of language organization is formed by clauses and sentences which are the most important units of information transfer. In the case of written language, there can be distinguished also other levels of the organization hierarchy (paragraphs, chapters, texts, and so on).

Komplekse sisteme kan as komplekse netwerke gemodelleer word (Kuhnert 2011:9), en die modellering van tale as netwerke is in 'n verskeidenheid studies gedoen (vir 'n oorsig van navorsing oor taal as 'n komplekse netwerk, kyk Cong & Liu 2014). Die topologie van taalnetwerke vertoon baie ooreenstemming met ander tipes komplekse netwerke, insluitend 'n kleinwêreldargitektuur, en oneweredige skakelverspreiding tussen woorde (met ander woorde nie 'n Poisson-verspreiding nie) (Dorogovtsev & Mendes 2001:2603; Strogatz 2004:256; Masucci & Rodgers 2006:2; Solé *et al.* 2010:22). Die huidige artikel fokus op kleinwêreldseheid (S), maar neem ook skakelverspreidingspatrone ($P[k]$) in ag.

Netwerke (ook genoem grafieke) bestaan uit nodusse (n) sowel as hul skakels (m), (vir 'n oorsig oor die netwerkteorie, kyk Newman 2010). Newman (2010) onderskei tussen vier soorte netwerke: biologiese netwerke (bv. voedselwebbe, proteïeninteraksienetwerke en neurale netwerke), tegnologiese netwerke (bv. die internet, kragvoorsienings- en vervoernetwerke), sosiale netwerke, en inligtingsnetwerke, byvoorbeeld die Wêreldwye Web, verwysingsnetwerke in die wetenskap, en taal.

Markošová (2008:663) verdeel studies van taal as 'n komplekse netwerk in twee kategorieë: posisionele studies, byvoorbeeld studies van die ko-voorkomsnetwerke van woorde, en konseptuele studies, byvoorbeeld studies van semantiese netwerke (bv. hiponieme, antonieme en sinonieme). In navolging van i Cancho en Solé (2001), Mendes (2003), Masucci en Rodgers (2006), Antiqueira *et al.* (2007), Markošová (2008), Solé *et al.* (2010) en Grabska-Gradzi ska *et al.* (2012) is die ko-voorkoms van woorde hier bestudeer, wat beteken dat 'n skakel aangedui is vir woorde wat langs mekaar in 'n teks voorkom. Ko-voorkomsnetwerke kan die direkte ko-voorkoms van woorde aandui, maar tweedegraadse voorkoms kan ook in berekening gebring word. In die sin *Piet skop die bal* kan skakels aangedui word as *Piet* → *skop* → *die* → *bal*, maar ook as *Piet* → *die* en *skop* → *bal*. Antiqueira *et al.* (2007:819) dui egter aan dat eerstegraadse ko-voorkoms (met ander woorde direkte ko-voorkoms) genoeg is om betekenisvolle resultate te verkry, en daarom word slegs eerstegraadskakels hier ondersoek.

Anders as in Masucci en Rodgers (2006) laat ons egter leestekens uit, soos i Cancho en Solé (2001:2262) en ander ook doen, aangesien die fokus hier op die interaksies tussen woorde in tekste val. Die woordko-voorkomsnetwerke wat in die onderhawige studie ondersoek word, is met ander woorde 'n netwerk wat slegs leksikale items in ag neem.

Antiqueira *et al.* (2007:813) verwyder stopwoorde (bv. lidwoorde en voorsetsels) uit hul datastel. Stopwoorde is nie in die huidige studie verwyder nie, omdat stopwoorde hier juis vergelyk word met die kort lys Engelse woorde met die hoogste gebruiksfrekwensie, soos verskaf deur i Cancho en Solé (2001:2265), en ook bereken is in Engels en Nederlands met behulp van die vertalings van *Donkermaan*.

Teoretiese konsepte

Ten einde die teoretiese konsepte te illustreer wat in hierdie afdeling bespreek word, is 'n voorbeeld van 'n ko-voorkomsnetwerk tussen woorde uit die volgende drie sinne saamgestel:

- Piet skop die bal.
- Ek gaan haal die fiets.
- Ek haal die vleis uit.

Die gemiddelde getal skakels wat 'n nodus in 'n netwerk het, $\langle k \rangle$, is in 'n ko-voorkomsnetwerk gelykstaande aan daardie woord se gebruiksfrekwensie (Masucci & Rodgers 2006:1).

Dit is reeds sedert Zipf (1935) bekend dat die gebruiksfrekwensie van woorde deur 'n kragwet bepaal word, wat in eenvoudige terme beteken dat daar 'n baie klein getal woorde met 'n hoë gebruiksfrekwensie is, terwyl die oorgrote meerderheid woorde 'n lae gebruiksfrekwensie het (Cong & Liu 2014:601). Verskeie studies van taal as netwerk het Zipf se bevinding ondersteun en die kragwet ook in die skakelverspreiding van woorde gevind (Masucci & Rodgers 2006:2; Markošová 2008:662; Grabska-Gradzińska *et al.* 2012:8). Die skakelverspreidingspatroon in 'n netwerk ($P[k]$) dui eenvoudig op die waarskynlikheid dat enige nodus 'n graad of getal skakels k sal hê, en die kragwet beteken dat skakelverspreidingspatrone deur $P(k) \sim k^{-\gamma}$ beskryf word (Cong & Liu 2014:601). Dit kan verwag word dat dieselfde verskynsel in Afrikaans gevind sal word. Kyk byvoorbeeld na die voorbeeldnetwerk in Figuur 1: daar is een woord met vyf skakels (*die*), een woord met drie skakels (*haal*), drie woorde met twee skakels elk (*ek*, *gaan* en *vleis*), en vyf woorde met net een skakel. Soos die getal skakels dus afneem, neem die getal woorde wat daardie getal skakels het toe. Hierdie oneweredige skakelverspreidingspatroon word vervat in Barabási en Albert (1999) se skaalvrye netwerkmodel, en in die huidige studie word die ko-voorkomsnetwerke van die drie romans ook met hierdie netwerkmodel vergelyk ten einde te bepaal of 'n oneweredige skakelverspreidingspatroon ook dié netwerke kenmerk. Ten einde hierdie vergelyking te kan tref, word ekwivalente van die woordko-voorkomsnetwerke met behulp van beide Barabási en Albert (1999) se skaalvrye netwerkmodel as Erdős en Rényi (1959, 1960) se lukrake netwerkmodel saamgestel, met ander woorde, netwerke met dieselfde getal nodusse (n) en skakels (m) as wat in die werklike netwerk aangetref word. Die korrelasies tussen die skakelverspreidingspatrone van die betrokke woordko-voorkomsnetwerke en die ekwivalente netwerk van enersyds Barabási en Albert (1999) en dié van

Erdős en Rényi (1959, 1960) andersyds, word daarna met behulp van Pearson se korrelasiekoëffisiënt bereken.

'n Ander aspek van woordko-voorkomsnetwerke wat hier in navolging van Masucci en Rodgers (2006:4) en Antiquiera *et al.* (2007:813) in berekening gebring word, is die getal kere wat woorde saam voorkom. Dié aspek word gemeet deur die gewig wat aan 'n skakel toegeken word, en 'n skakel met 'n gewig van twee beteken bloot dat woorde twee keer saam voorkom. Masucci en Rodgers (2006:4) het gevind dat die ko-voorkoms van twee woorde in Engels die hoogste is vir *of the*, gevolg deur *it was* en *in the*.¹ In die voorbeeldnetwerk in Figuur 1 kan 'n dikker skakel tussen *haal* en *die* waargeneem word, wat daarop dui dat dié twee nodusse die meeste in hierdie voorbeeldnetwerk ko-voorkom (twee keer). In die netwerke wat in die huidige artikel bestudeer is, dui die gewig van skakels ook op hoeveel keer woorde saam voorkom.

'n Belangrike kenmerk van netwerke is dat hulle gewoonlik ook kleinwêreldnetwerke is. Kleinwêreldsheid, soos deur Watts en Strogatz (1998) gemodelleer, verwys na netwerke wat 'n kort gemiddelde pad tussen nodusse het (nodusse is gereeld slegs 2 of 3 stappe van mekaar verwyderd, en selde meer as 10), sowel as 'n mate van groepsvorming wat beduidend hoër is as wat aan toeval toegeskryf kan word. Müller-Linow (2008:15), Solé *et al.* (2010:3) en Heidtmann (2013:442) definieer 'n gemiddelde pad (L) eenvoudig as die kleinste getal skakels wat twee nodusse verbind. Kyk byvoorbeeld na die voorbeeldnetwerk in Figuur 1. Die deursnee van die netwerk is 4, wat die pad is wat vanaf byvoorbeeld *ek* na *uit* geneem moet word (aangedui in rooi), of tussen *gaan* en *Piet*, of *ek* en *Piet*. Die gemiddelde pad in dié netwerk is egter 2.205, wat die gemiddeld verteenwoordig in die hele netwerk. Gemiddelde pad word met behulp van Vergelyking 1 bereken (König & Battiston 2009:32):

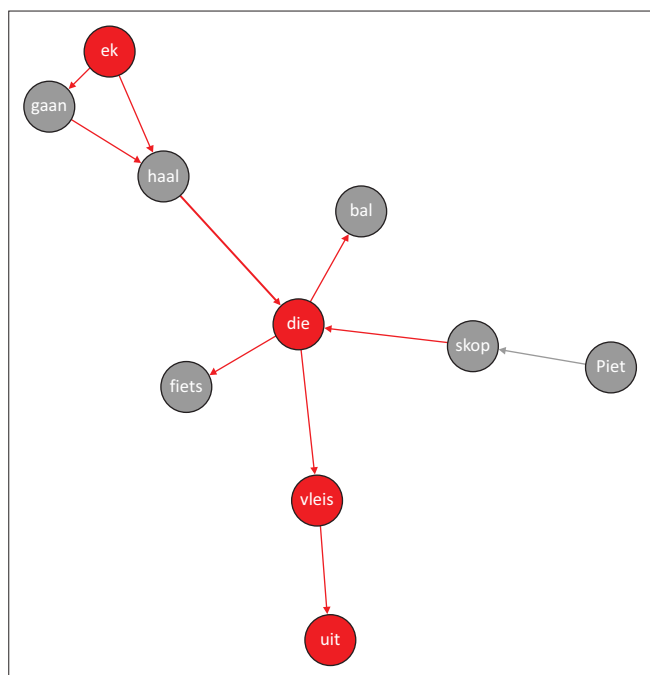
$$L = \frac{1}{\frac{1}{2}n(n-1)} \sum_{i \geq j} d_{ij} \quad [\text{Vergl. 1}]$$

In Vergelyking (1) is d_{ij} die geodetiese afstand (die kortste pad) tussen nodus i en j . In die geval van L is daar soms 'n statisties beduidende verskil tussen tale gevind: L is byvoorbeeld hoër in Russies en Tsjeggies as vir Sweeds en Deens (Cong & Liu 2014:601). Die huidige studie se vergelykende aanpak laat ons dus ook toe om L in drie tale te vergelyk.

Solé *et al.* (2010:3) definieer groepsvorming (C) as 'the probability that two vertices (e.g. words) that are neighbors of a given vertex are neighbors of each other'. In Watts en Strogatz (1998) word die groepsvorming C_i van nodus i deur Vergelyking 2 bereken:

$$C_i = \frac{2E_i}{k_i(k_i - 1)} \quad [\text{Vergl. 2}]$$

1. Masucci en Rodgers (2006:4) se lys van nodusse met die hoogste ko-voorkoms sluit ook leestekens in, byvoorbeeld [, and].



FIGUUR 1: 'n Voorbeeld van 'n woordko-voorkomsnetwerk.

In Vergelyking (2) verwys E_i na die getal skakels tussen die bure van nodus i , en k_i na die aantal direkte bure van nodus i . Die groepsvorming van die hele netwerk is dan die gemiddeld van C_i vir al die nodusse. Taal het gewoonlik 'n lae groepsvormingskoëffisiënt, aangesien woorde gereeld nie onderling met mekaar kan skakel nie. Masucci en Rodgers (2006:3) skryf dat die woord *like* in Engels kan skakel met beide die bepaalde en onbepaalde lidwoord, maar hierdie woorde skakel nooit onderling met mekaar in 'n ko-voorkomsnetwerk nie (dieselfde geld vir Afrikaans, waar 'n en *die* ook nooit langs mekaar in 'n sin voorkom nie). In die voorbeeldnetwerk hierbo kan daar gesien word dat naamwoorde nie gereeld in hierdie netwerk skakel nie; daar is byvoorbeeld nie 'n skakel tussen *fiets* en *bal* nie. 'n Driehoek kom egter tussen *ek*, *gaan* en *haal* voor. C verskil ook tussen tale, en Grabska-Gradzińska *et al.* (2012:11–12) het 'n groter groepsvorming in Engelse as Poolse tekste gevind. Ons studie vergelyk dan ook C in drie tale.

Om 'n netwerk as 'n kleinwêreldnetwerk te klassifiseer soos deur Watts en Strogatz (1998) gemodelleer, moet die netwerk vergelyk word met 'n nulmodel waar n en m gelyk is aan die werklike netwerk wat bestudeer word. Müller-Linow (2008:31) skryf: 'Die Erzeugung von Null-Modellen einer Graphenarchitektur ist für solche Untersuchungen von Bedeutung, bei denen die Signifikanz verschiedener struktureller Merkmale bestimmt werden soll.' Die nulmodel wat gewoonlik vir vergelykingsdoeleindes gebruik word, is die lukrake netwerkmodel van Erdős en Rényi (1960, 1959) (Humphries & Gurney 2008), en gevolglik word daar tydens die vergelyking 'n netwerk volgens Erdős en Rényi se model saamgestel wat ekwivalent is aan die netwerk wat ondersoek word, beide ten opsigte van n en m ('n sogenaamde $G(n,m)$ -model). Die gemiddelde pad van die netwerk of grafiek (L_g) word dan met die gemiddelde pad in die ekwivalente Erdős en Rényi-netwerk (L_{rand}) vergelyk, en die groepsvorming van die netwerk (C_g) met die groepsvorming in die ekwivalente Erdős en Rényi-netwerk (C_{rand}). Kleinwêreldnetwerke is met ander woorde netwerke waar $L_g \approx L_{rand}$ en $C^{\Delta} C_{rand}$ (i Cancho en Solé 2001:2663; Markošová 2008:662; Solé *et al.* 2010:3; Heidtmann 2013:445; Cong & Liu 2014:602). Dié definisie is uiteraard nie presies nie, en daarom stel Humphries en Gurney (2008) die volgende metode voor om kleinwêreldsheid deur middel van die kleinwêreldsindeks (S) te kwantifiseer, soos vervat in Vergelykings 3, 4 en 5:

$$\gamma_g = \frac{C_g}{C_{rand}} \quad [\text{Vergl. 3}]$$

$$\lambda_g = \frac{L_g}{L_{rand}} \quad [\text{Vergl. 4}]$$

$$S = \frac{\gamma_g}{\lambda_g} \quad [\text{Vergl. 5}]$$

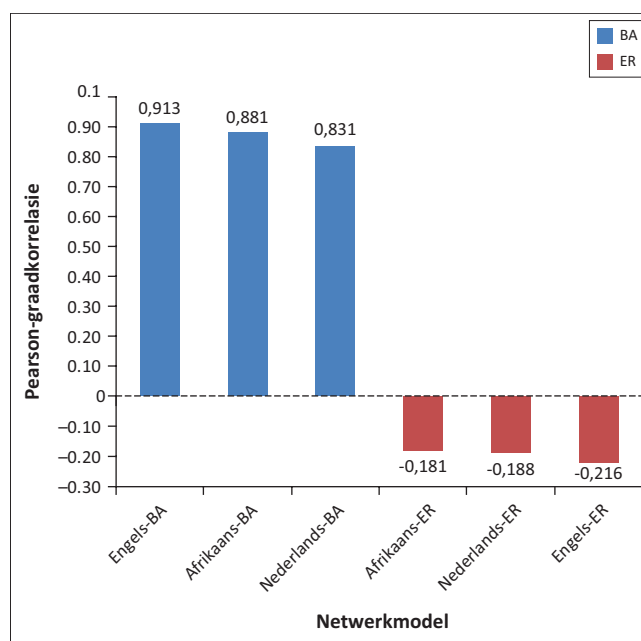
'n Kleinwêreldnetwerk is 'n netwerk waar $S > 1$, en waardes van $1 \leq S \leq 3$ is grensgevalle; dus is gevalle waar $S \geq 3$ duidelike kleinwêreldnetwerke (Humphries & Gurney 2008).

Resultate

Ten einde die woordko-voorkomsnetwerke in die drie weergawes van *Donkermaan* te onderneem, is al drie tekste gedigitaliseer en vir karaktererkenningfoute nagegaan. Aangesien hierdie digitalisering slegs vir ons eie navorsingsdoeleindes onderneem is en die digitale weergawes nie enigsins versprei word nie, is kopiereg nie geskend nie. Die Afrikaanse teks bestaan uit 104 751 woorde, die Engels uit 98 043 woorde, en die Nederlands uit 108 969 woorde.

Die woordko-voorkomsnetwerke in Brink se *Donkermaan* is in die Afrikaanse, Nederlandse en Engelse tekste saamgestel, en eerstens vergelyk met netwerke van dieselfde grootte (ten opsigte van beide n as m) wat volgens Erdős en Rényi (1959, 1960) se lukrake netwerkmodel en Barabási en Albert (1999) se skaalvrye netwerkmodel saamgestel is. Die vergelyking is met ander woorde tussen die netwerke wat in die drie weergawes van *Donkermaan* enersyds, en die twee netwerkmodelle andersyds saamgestel is. Die oogmerk hiervan is om te bepaal of die skakelverspreidingspatroon nader aan 'n Poisson-verspreiding val, soos wat aangetref word in die skakelverspreidingspatroon van Erdős en Rényi-netwerke (Liang *et al.* 2009), of 'n oneweredige patroon volg, waar 'n klein getal nodusse 'n buitengewone hoë getal skakels het, soos aangetref in Barabási en Albert (1999) se skaalvrye netwerkmodel. Die korrelasies is met behulp van Pearson se korrelasiekoëffisiënt tussen die werklike netwerke en dié twee netwerkmodelle bereken. Die resultate kan in Figuur 2 gesien word.

Dit is duidelik dat al drie woordko-voorkomsnetwerke beter voorgestel word deur die skaalvrye netwerkmodel van Barabási en Albert (1999) as die lukrake netwerkmodel van Erdős en Rényi (1959; 1960), wat bloot beteken dat 'n klein getal woorde 'n buitengewoon groot getal skakels het. Soos in Figuur 2 gesien



FIGUUR 2: 'n Vergelyking van netwerkmodelle met woordko-voorkomsnetwerke.

kan word, lewer die skakelverspreidingspatrone r -waardes van $0.831 \leq r \leq 0.913$ volgens die skaalvrye netwerkmodel van Barabási en Albert, teenoor r -waardes van $-0.216 \leq r \leq -0.181$ volgens Erdős en Rényi se netwerkmodel, wat beteken dat daar 'n sterk korrelasie met die BA-model aangedui kan word, teenoor 'n swak korrelasie met die ER-model. Dié bevinding verbaas geensins nie omdat die verskynsel vroeër in 'n verskeidenheid Engelse tekste gevind is. Die verskynsel is egter nog nie met betrekking tot Afrikaans ondersoek nie.

'n Belangrike vraag wat uit bogenoemde vergelyking tussen netwerkmodelle en werklike netwerke ontstaan, is watter woorde so 'n groot getal skakels het. In i Cancho en Solé (2001:2265) is daar aangedui dat *and, the, of, in, a, to, 's, with, by* en *is* die woorde met die grootste getal skakels in Engels is, en in Solé *et al.* (2010:5) word dit beklemtoon dat die woorde met die grootste getal skakels in Engels 'low semantic content' bevat, met ander woorde eerder leksikale items wat 'n grammatikale funksie vervul. Tabel 1 verskaf die woorde met die grootste getal skakels in die Afrikaanse teks van *Donkermaan* (Brink 2000a), sowel as in Engels (Brink 2000b) en Nederlands (Brink 2001).

Dit is opvallend dat die bepaalde lidwoord die grootste getal skakels in al drie tekste het, gevolg deur die neweskikkende voegwoord, en dertens die onbepaalde lidwoord in Nederlands en Afrikaans, terwyl die onbepaalde lidwoord ietwat laer op die lys van Engelse woorde voorkom. Verder domineer voorsetsels en voornaamwoorde al drie dié lyste, wat onder meer daarop dui dat soortgelyke woorde in al drie hierdie tale die meeste voorkom. Let ook daarop dat *nie* in Afrikaans baie meer skakels het as *niet* in Nederlands, terwyl *not* nie onder die twintig woorde in Engels tel wat die meeste skakels het nie (*not* het die 25ste grootste getal skakels in dié netwerk). Die dubbele ontkenning in Afrikaans is natuurlik daarvoor verantwoordelik dat dié woord soveel skakels in Afrikaans het.

Dit is ook interessant om te let op watter woordpare die meeste saam voorkom. Tabel 2 verskaf die woordpare wat die hoogste ko-voorkoms in die drie weergawes van *Donkermaan* het.

Die woordpare wat dus in al drie netwerke die meeste saam voorkom, is *in die*, terwyl *op die*, *van die*, *in my*, *met 'n* en ander ook gereeld in al drie tale saam voorkom.

Die huidige studie konsentreer egter nie op blote tellings van woorde se getal skakels nie, maar eerder op L en C . Tabel 3 verskaf die getal nodusse (n), getal skakels (m), gemiddelde pad (L) en groepsvorming (C) vir die woordko-voorkomsnetwerke wat in vier vorige studies bestudeer is.

Tabel 3 wys dat 'n kort gemiddelde pad 'n kenmerk van woordko-voorkomsnetwerke is, selfs wanneer verskillende tale onder die loep kom (kyk ook verder aan), en in dié studies val L in die spektrum $2.54 \leq L \leq 3.59$. Alhoewel die woordko-voorkomsnetwerke wat deur Masucci en Rodgers (2006), Newman (2006) en Margan, Martinčić-Ipšić en Meštrović (2014) bestudeer is 'n beduidend kleiner groepsvorming het as dié wat deur i Cancho en Solé (2001), bestudeer is, geld $C^A C_{rand}$ steeds in al dié netwerke. Ons het $C_{rand} = 0.003$ bereken in die netwerk soos bestudeer deur Masucci en Rodgers (2006), en $C_{rand} = 0.073$ in die netwerk van Newman (2006), teenoor onderskeidelik $C = 0.19$ en 0.16 . Dié groot verskil in groepsvormingskoeffisiënte (tesame met $L_g \approx L_{rand}$ soos in die oorspronklike studies gerapporteer) suggereer dat hierdie netwerke kleinwêreldnetwerke is, alhoewel S nie in dié studies bereken word nie (Humphries en Gurney 2008 bereken egter $S = 2.13$ in die netwerk van Newman 2006).

Bostaande syfers sowel as S is vervolgens bereken in die netwerke wat in die huidige artikel bestudeer word. Tabel 4 verskaf n , m , L , L_{rand} , C , C_{rand} en S in al drie netwerke.

TABEL 1: Getal skakels van woorde in *Donkermaan*.

Woord (Afrikaans)	Getal skakels	Woord (Engels)	Getal skakels	Woord (Nederlands)	Getal skakels
die	2609	the	2568	de	2126
en	1879	and	1830	en	2099
'n	1691	of	1666	een	1989
het	1632	to	1612	het	1551
van	1304	a	1516	van	1531
nie	1294	in	1103	te	1313
te	1041	her	984	ik	1107
my	999	I	948	in	1043
in	962	was	763	haar	983
haar	893	she	762	ze	978
wat	857	it	718	dat	975
ek	847	that	672	was	960
was	834	my	645	die	921
sy	820	with	638	op	849
om	724	on	578	met	763
is	721	for	530	maar	684
se	667	you	470	niet	673
met	611	me	445	om	641
dit	600	or	436	mijn	622
maar	546	as	431	had	614

TABEL 2: Ko-voorkoms van woorde in *Donkermaan*.

Ko-voorkoms (Afrikaans)	Voorkoms	Ko-voorkoms (Nederlands)	Voorkoms	Ko-voorkoms (Engels)	Voorkoms
in die	677	in de	422	in the	531
ek het	521	van de	349	it was	388
het ek	422	dat ik	308	of the	347
van die	393	op de	245	on the	260
dit was	307	dat ze	205	to the	222
op die	290	met een	205	I was	178
om te	229	het was	200	she was	170
het nie	201	in het	168	there was	160
het sy	200	in mijn	165	in a	140
in my	196	van haar	161	she said	132
aan die	187	ik het	161	from the	127
met die	184	ik heb	152	was the	125
wat ek	182	naar de	149	at the	124
met 'n	159	van het	143	was a	123
na die	153	het is	139	to be	122
ek is	146	van een	133	in my	122
by die	142	ik me	133	with a	117
dat ek	141	als een	130	of her	114
het die	139	op een	129	when I	109
van haar	137	en ik	128	of a	109

TABEL 3: Netwerkdigheid, gemiddelde pad en groepsvorming in 'n getal woordko-voorkomsnetwerke.

Studie	Taal	n	m	L	C
i Cancho en Solé (2001)	Engels	478,773	17,700,000	2.63	0.687
i Cancho en Solé (2001)	Engels	460,902	16,100,000	2.67	0.437
Masucci en Rodgers (2006)	Engels	8,992	117,687		0.19
Margan, Martinčić-Ipšić, en Meštrović (2014)	Kroaties	9,530	22,305	3.59	0.15
Margan, Martinčić-Ipšić, en Meštrović (2014)	Kroaties	40,809	156,857	3.25	0.25
Margan, Martinčić-Ipšić, en Meštrović (2014)	Kroaties	91,647	464,029	3.10	0.32
Newman (2006)	Engels	112	425	2.54	0.16

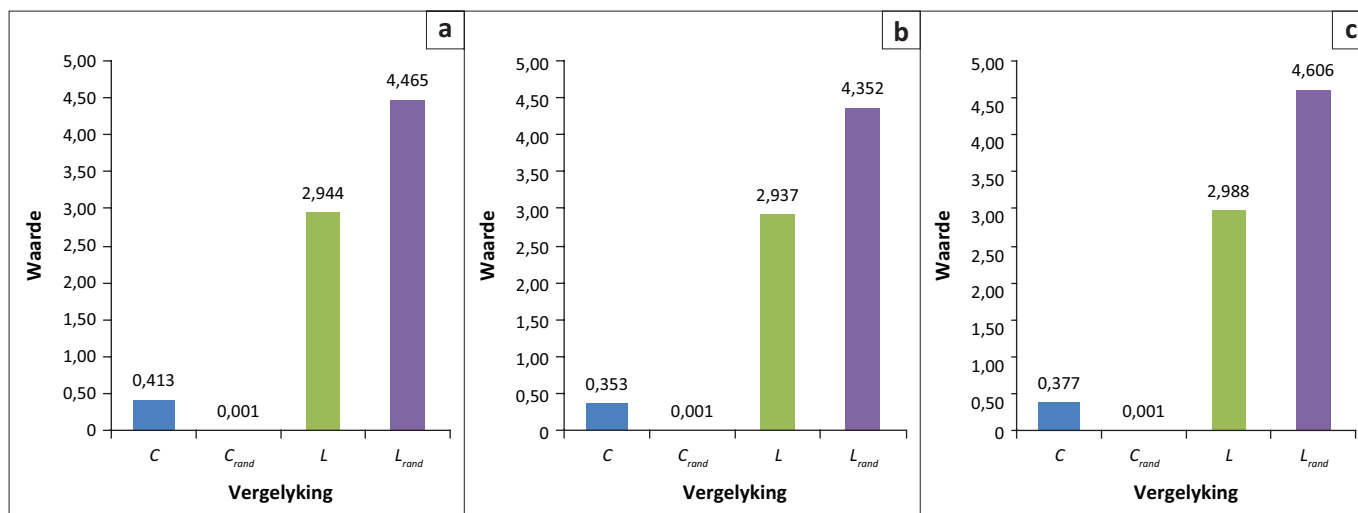
TABEL 4: Resultate van die huidige studie.

Netwerk	n	m	L	L_{rand}	C	C_{rand}	S
<i>Donkermaan</i> Afrikaans	10051	44732	2.94442	4.464754	0.4132	0.000893117	701.569675
<i>Donkermaan</i> Engels	9178	42466	2.93743	4.35159	0.3528	0.001	522.606192
<i>Donkermaan</i> Nederlands	11854	50803	2.9877	4.60648	0.3767	0.000728	797.713362

Dié figure vergelyk goed met vorige studies. Grabska-Gradzińska *et al.* (2012) het 'n gemiddelde pad in die spektrum van $2.7 \leq L \leq 3.8$ in Engelse en Poolse tekste gevind, terwyl Margan, Martinčić-Ipšić en Meštrović (2014) 'n gemiddelde pad in die spektrum van $3.10 \leq L \leq 3.59$ in Kroatiese boeke bereken het, en Liang *et al.* (2009) het $3.29 \leq L \leq 4.95$ in Engelse en Chinese tekste gevind. Soos in die geval met Engels, Pools, Kroaties en Chinees, word Afrikaans en Nederlands dus ook deur 'n kort gemiddelde pad gekenmerk, en die onderhawige studie se bevindinge ondersteun ook vorige bevindinge rakende Engels. 'n Vergelyking van L en L_{rand} word in al drie tekste in Figuur 3 verskaf. Die feit dat L baie soortgelyk in al drie tale is, verbaas nie: al drie tale is Germaanse tale. Boonop is Afrikaans oorwegend op Nederlands gebaseer en in 'n sterk interaksie met Engels. In die geval van L is daar dus nie 'n beduidende verskil tussen die tale waarin hoër L -waardes, byvoorbeeld in Russies en Tsjeggies, as in byvoorbeeld Sweeds en Deens gevind is nie (Cong & Liu 2014:601). Gegewe Cong en Liu se bevinding is dit duidelik dat tale wat tot dieselfde groep behoort, soortgelyke L -waardes het, en dus onderskryf die onderhawige studie ook hulle resultate.

C is ook baie soortgelyk wanneer dié koëffisiënte met vorige studies vergelyk word. Grabska-Gradzińska *et al.* (2012) het C in die spektrum $0.09 \leq C \leq 0.34$ bereken, terwyl Margan, Martinčić-Ipšić en Meštrović (2014) $0.15 \leq C \leq 0.32$ gevind het, en Liang *et al.* (2009) kleiner koëffisiënte van $0.0013 \leq C \leq 0.264$ gevind het (let daarop dat $C^{\Delta} C_{rand}$ selfs vir dié klein waardes van C geld). C is dus ook in die huidige studie baie soortgelyk aan wat in vorige studies gevind is. Let daarop dat $C^{\Delta} C_{rand}$ duidelik in al drie netwerke geld wat in die huidige studie ondersoek is, soos ook in Figuur 3 gesien kan word. Groepsvorming is ook baie soortgelyk tussen dié drie tale, wat weer eens toegeskryf kan word daaraan dat al drie tale Germaanse tale is. Grabska-Gradzińska *et al.* (2012:11–12) se bevinding dat groepsvorming hoër is in Engelse as Poolse tekste is natuurlik 'n studie wat tale vanuit verskillende taalgroepe vergelyk, wat nie in die huidige studie die geval is nie. Dit is egter interessant om daarop te let dat Afrikaans 'n hoër groepsvorming as Nederlands of Engels het, waarskynlik vanweë die dubbele ontkenning.

Die huidige studie verteenwoordig die eerste kwantifisering van S volgens Humphries en Gurney (2008) se metode vir



FIGUUR 3: 'n Vergelyking van L en L_{rand} en C en C_{rand} : (a) Donkermaan Afrikaans, (b) Donkermaan Engels, (c) Donkermaan Nederlands.

woordko-voorkomsnetwerke, en met $522 \leq S \leq 797$ kan duidelik gesien word dat Afrikaans, Engels en Nederlands al drie duidelik kleinwêreldnetwerke is (S^A 3). Woordko-voorkomsnetwerke in Afrikaans, Engels en Nederlands kan dus volgens sowel Watts en Strogatz (1998) se kleinwêreldnetwerkmodel as Barabási en Albert (1999) se skaalvrye netwerkmodel beskryf word, eerder as volgens Erdős en Rényi (1959; 1960) se lukrake netwerkmodel.

Die vraag is wat die waarde daarvan is om kleinwêreldsheid in woordko-voorkomsnetwerke te kwantifiseer. Kleinwêreldsheid het belangrike implikasies vir diffusie oor komplekse netwerke van alle soorte: die hoë mate van groepsvorming sowel as die kort gemiddelde pad wat nodusse verbind, maak dit byvoorbeeld makliker vir inligting en virusse om oor kleinwêreldnetwerke te versprei. Kleinwêreldsheid is ook van belang in woordko-voorkomsnetwerke: Volgens Cong en Liu (2014:602) vergemaklik die kleinwêreldstruktuur van taal die taalgebruiker se navigasie tussen woordkeuses, en Bordag en Bordag (2003) het geïllustreer hoe 'n kleinwêreldargitektuur dit moontlik maak om vinnig die betekenis van dubbelsinnige woorde te bepaal. Ook is kleinwêreldargitektuur 'n kenmerk van 'n mens se taalgebruik na die ouderdom van twee jaar (Corominas-Murtra, Valverde & Solé 2009).

Die bestudering van die gebruik van taal as komplekse netwerk kan aangewend word om taalverwerwing te ondersoek. Volgens Cong en Liu (2014:611) en Solé *et al.* (2010:4) verskil die woorde van kinders met die hoogste gebruiksfrekwensie van dié van volwassenes; kinders wend naamwoorde meer gereeld aan omdat lidwoorde en voorsetsels nog nie so gereeld in hul sinstrukture voorkom nie. Antiqueira *et al.* (2007) stel ook voor dat gebruiksfrekwensies L en C skryfvaardighede kan evalueer ten opsigte van kohesie en koherensie, sowel as tematiese ontwikkeling. Hul toepassing is juis interessant omdat hulle 'n metode voorstel om universiteitstoelating te bepaal (Ke *et al.* 2014).

Slot

Hierdie artikel het die eerste ontleding van 'n Afrikaanse teks as woordko-voorkomsnetwerk onderneem. Daar is aangedui dat dié netwerk ook, soos in ander tale, gekenmerk word deur 'n oneweredige skakelverspreidingspatroon, kort gemiddelde pad en hoër mate van groepsvorming as wat by Erdős en Rényi-netwerke gevind word. Die artikel het ook Humphreys en Gurney (2008) se metode aangewend om kleinwêreldsheid in die Afrikaanse, Engelse en Nederlandse tekste te kwantifiseer, en daar is gevind dat in al drie gevalle hulle duidelik kleinwêreldnetwerke is.

'n Belangrike vraag wat ontstaan, is hoe veralgemeenbaar die resultate van die huidige studie is, aangesien slegs drie boeke en slegs een skrywer (sowel as twee vertalers) bestudeer is. Let egter daarop dat slegs een skrywer ook in Masucci en Rodgers (2006), Newman (2006), asook Liu en Cong (2013) bestudeer is. Gegewe die feit dat vorige studies aangedui het dat taal 'n kleinwêreldnetwerk is, sowel as die hoë waardes van S wat in die huidige studie gevind is, is dit ons insiens onwaarskynlik dat tot 'n ander gevolgtrekking gekom sal word indien meer tekste in berekening gebring word, alhoewel klein verskille wel sal voorkom. Ten einde hierdie vermoede te bevestig, moet meer tekste en meer outeurs egter bestudeer word.

'n Verdere vraag wat uit dié artikel ontstaan, is of daar 'n statisties beduidende verskil tussen die topologiese kenmerke van werklike taalgebruik in Afrikaans en die literêre gebruik van Afrikaans is. Ook wonder 'n mens of daar 'n verskil is tussen die poësie, prosa en drama, en selfs of daar unieke fasette aan sommige skrywers se taalgebruik is. Gegewe die feit dat taal deur kleinwêreldsheid gekenmerk word, is dit onwaarskynlik dat werklike taalgebruik of die bestudering van ander genres of ander skrywers se werke tot ander bevindinge sal lei, maar dit is moontlik dat geringe, dog betekenisvolle verskille gevind kan word. In toekomstige studies kan meer tekste en verskillende datastelle dus

vergelyk word ten einde te bepaal of sodanige statisties beduidende verskille bestaan.

Erkenning

Mededingende belange

Die outeurs verklaar dat hulle geen persoonlike verhouding het wat hulle voordelig of nadelig kon beïnvloed het in die skryf van hierdie artikel nie.

Outeurs se bydraes

B.S. was die projekteier, het die data-analise gedoen en meegeskrif aan die artikel. C.G. het die data ingewin en voorberei en aan die artikel meegeskrif.

Literatuurverwysings

- Albert, R. & Barabási, A.-L., 2002, 'Statistical mechanics of complex networks', *Reviews of Modern Physics* 74, pp. 47–97. <http://dx.doi.org/10.1103/RevModPhys.74.47>
- Amaral, L. & Ottino, J., 2004, 'Complex networks. Augmenting the framework for the study of complex systems', *European Physical Journal* 38, 147–162. <http://dx.doi.org/10.1140/epjb/e2004-00110-5>
- Antiqueira, L., Nunes, M.d.G.V., Oliveira, O. & Da Costa, F., 2007, 'Strong correlations between text quality and complex networks features', *Physica A: Statistical Mechanics and its Applications* 373, 811–820.
- Barabási, A.-L. & Albert, R., 1999, 'Emergence of scaling in random networks', *Science* 286, 509–11. <http://dx.doi.org/10.1126/science.286.5439.509>
- Bordag, S. & Bordag, D., 2003, 'Advances in automatic speech recognition by imitating spreading activation', in V. Matoušek & P. Mautner (eds.), *Text, speech and dialogue*, pp. 158–164, Springer, Heidelberg. http://dx.doi.org/10.1007/978-3-540-39398-6_23
- Brink, A.P., 2000a, *Donkermaan*, Human & Rousseau, Kaapstad.
- Brink, A.P., 2000b, *The rights of desire*, Vintage, London.
- Brink, A.P., 2001, *Donkermaan*, Meulenhoff, Amsterdam.
- Cilliers, P., 1998, *Complexity and postmodernism: Understanding complex systems*, Routledge, London.
- Cong, J. & Liu, H., 2014, 'Approaching human language with complex networks', *Physics of Life Reviews* 11, 598–618. <http://dx.doi.org/10.1016/j.phlev.2014.04.004>
- Corominas-Murtra, B., Valverde, S. & Solé, R., 2009, 'The ontogeny of scale-free syntax networks: Phase transitions in early language acquisition', *Advances in Complex Systems*, 12(3), 371–392. <http://dx.doi.org/10.1142/S0219525909002192>
- De Saussure, F., 1966, *Course of general linguistics*, McGraw-Hill, New York.
- DeLaurentis, D., 2007, 'Role of humans in complexity of a system-of-systems', in V. Duffy (ed.), *Digital human modeling*, Springer (pp. 363–371), Berlin. http://dx.doi.org/10.1007/978-3-540-73321-8_42
- Dorogovtsev, S.N. & Mendes, J.F.F. 2001, Language as an evolving word web, *Proceedings of the Royal Society of London, Series B: Biological Sciences* 268(1485), 2603–2606. <http://dx.doi.org/10.1098/rspb.2001.1824>
- Erdős, P. & Rényi, A., 1959, 'On random graphs', *Publicationes Mathematicae* 6, 290–297.
- Erdős, P. & Rényi, A., 1960, 'On the evolution of random graphs', *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* 5, 17–61.
- Grabska-Gradzińska, I., Kulig, A., Kwapień, J., & Drożdż, S., 2012, 'A complex network analysis of literary and scientific texts', *International Journal of Modern Physics C* 23(7), 1250051. <http://dx.doi.org/10.1142/S0129183112500519>
- Heidtmann, K., 2013, 'Internet-Graphen', *Informatik_Spektrum* 6, 440–448.
- Holland, J.H., 1992, 'Complex adaptive systems', *Daedalus* 121(1), 17–30.
- Humphries, M.D. & Gurney, K., 2008, 'n etwork "small-world-ness": A quantitative method for determining canonical network equivalence', *PLoS one* 3(4), e0002051. <http://dx.doi.org/10.1371/journal.pone.0002051>
- i Cancho, R.F. & Solé, R., 2001, 'The small world of human language', *Proceedings of the Royal Society B, London* 268, 2261–2265.
- Kenett, Y.N., Kenett, D.Y., Ben-Jacob, E. & Faust, M., 2011, 'Global and local features of semantic networks: Evidence from the Hebrew Mental Lexicon', *PLoS ONE* 6(8), e23912. <http://dx.doi.org/10.1371/journal.pone.0023912>
- Ke, X., Zeng, Y., Ma, Q. & Zhu, L., 2014, 'Complex dynamics of text analysis', *Physica A: Statistical Mechanics and its Applications* 415, 307–314.
- Kuhnert, M.-T., 2011, 'Komplexe dynamische Systeme als funktionelle Netzwerke: Möglichkeiten und Grenzen der datengetriebenen Analyse', Doktorarbeit, Rheinischen Friedrich-Wilhelms-Universität, Bonn.
- König, M.D. & Battiston, S., 2009, 'From graph theory to models of economic networks', in A. Naimzada (ed.), *Networks, topology and dynamics*, pp. 23–63, Springer, Heidelberg.
- Kwapień, J. & Drożdż, S., 2012, 'Physical approach to complex systems', *Physics Reports* 515(3), 115–226. <http://dx.doi.org/10.1016/j.physrep.2012.01.007>
- Liang, W., Shi, Y., Tse, C.K., Liu, J., Wang, Y. & Cui, X., 2009, 'Comparison of co-occurrence networks of the Chinese and English languages', *Physica A*, 388, 4901–4909. <http://dx.doi.org/10.1016/j.physa.2009.07.047>
- Liu, H. & Cong, J., 2013, 'Language clustering with word co-occurrence networks based on parallel texts', *Chinese Scientific Bulletin* 58, pp.1139–1144. <http://dx.doi.org/10.1007/s11434-013-5711-8>
- Liu, H. & Zhao, Y.H.W., 2010, 'How do local syntactic structures influence global properties in language networks?', *Glottometrics* 20, 38–58.
- Margan, D., Martinčić-Ipšić, S. & Meštrović, A., 2014, 'Preliminary report on the structure of Croatian linguistic co-occurrence networks', *arXiv* 1405.4433v1, 1–8.
- Markoová, M. & Nather, P., 2006, 'Language as a small world network', *6th International Conference on Hybrid Intelligent Systems*, 13–15 December 2006.
- Markošová, M., 2008, 'n etwork model of human language', *Physica A* 387, 661–666. <http://dx.doi.org/10.1016/j.physa.2007.09.027>
- Masucci, A. & Rodgers, G., 2006, 'n etwork properties of written human language', *Physical Review E* 74(2), 026102. <http://dx.doi.org/10.1103/PhysRevE.74.026102>
- McCrum, R. Cran, W. & MacNeil, R., 1992, *The story of English*, Faber & Faber, London.
- Mendes, J., 2003, 'Effect of accelerated growth on networks dynamics', in R. Pastor-Satorras, M. Rubi & A. Diaz-Guilera (eds.), *Statistical mechanics of complex networks*, pp. 88–113, Springer-Verlag, Berlin. http://dx.doi.org/10.1007/978-3-540-44943-0_6
- Mitchell, M., 2006, 'Complex systems: Network thinking', *Artificial Intelligence* 170(18), 1194–1212. <http://dx.doi.org/10.1016/j.artint.2006.10.002>
- Müller-Linow, M., 2008, 'Analyse der Verbindung von Topologie und Dynamik in abstrakten Graphen und biologischen Netzwerken', Doktorarbeit, Technischen Universität Darmstadt.
- Newman, M.E.J., 2006, 'Finding community structure in networks using the eigenvectors of matrices', *Physical Review E* 74, 036104. <http://dx.doi.org/10.1103/PhysRevE.74.036104>
- Newman, M.E.J., 2010, *Networks*, Oxford University Press, Oxford. <http://dx.doi.org/10.1093/acprof:oso/9780199206650.001.0001>
- Newman, M., 2011, 'Complex systems: A survey', *arXiv* 1112.1440v1, 1–10. <http://dx.doi.org/10.1119/1.3590372>
- Ondal, F., 1969, 'Taal as sisteem', rede gelewer by die aanvaarding van die amp van hoogleraar aan die Randse Afrikaanse Universiteit, pp. 5–29.
- Senekal, B.A., 2013, 'Die gebruik van die netwerkteorie binne 'n sisteemteoretiese benadering tot die Afrikaanse letterkunde: 'n Teorie-oorsig', *Tydskrif vir Geesteswetenskappe* 53(4), 668–682.
- Senekal, J.H., 1987, *Literatuuroppattings: 'Wese' en 'waarhede' van 'n nuwe literêre teorie*, Universiteit van die Oranje-Vrystaat, Bloemfontein.
- Solé, R.V., Corominas-Murtra, B., Valverde, S. & Steels, L., 2010, 'Language networks: Their structure, function, and evolution', *Complexity* 15(6), 20–26. <http://dx.doi.org/10.1002/cplx.20326>
- Strogatz, S., 2004, *Sync. The emerging science of spontaneous order*, Penguin, London.
- Viljoen, H., 1986, 'Die Suid-Afrikaanse romansisteem. 'n Vergelykende studie', Ph.D.-thesis, Potchefstroomse Universiteit vir Christelike Hoër Onderwys.
- Watts, D.J. & Strogatz, S.H., 1998, 'Collective dynamics of "small-world" networks', *Nature* 393(6684), 409–410. <http://dx.doi.org/10.1038/30918>
- Weideman, A., 2009, 'Uncharted territory: A complex systems approach as an emerging paradigm in applied linguistics', *Per Linguam* 25(1), 61–75.
- Weideman, A., 2011, *A framework for the study of linguistics*, Van Schaik, Pretoria.
- Zipf, G.K., 1935, *The psychobiology of language*, Houghton-Mifflin, Boston, MA.