



Onzichtbare taal

Stilometrie en literatuurwetenschap

Karina van Dalen-Oskam

30 september 2016



UNIVERSITEIT VAN AMSTERDAM



N e e r l a n d i s t i e k . n l

Over stijl

Ton Anbeek en Arie Verhagen

PUBLICATIEDATUM 14 JUNI 2001

ARTIKELNUMMER 01.01

Stijl

- iets persoonlijks – ‘het DNA van een auteur’
- iets wat analyseerbaar is
- als afwijking van een norm, een standaard, binnen veranderende conventies
- zinvol te onderzoeken ook zonder een allesomvattende taalgebruiksbeschrijving
- contrasteren/confronteren van verschillende teksten
- ‘het DNA van de tekst’ als kenmerkend voor de auteur
- Van der Heijden tegenover Voskuil

Literatuurwetenschap en *stylometry*

- Auteursherkenning
- Genreherkenning
- Personagestijl
- (Een nieuwe definitie van stijl)
- Bestsellers
- Literaire romans
- Van der Heijden tegenover Voskuil



Tracy Chevalier Picks
'Tenant of Wildfell
Hall' for Book Club



What Artists
Take to Coachella



How Dolls
Helped an Author
Through Her Debut
Novel



Martha Graham
Dance Company and
Miami City Ballet
Reviews

SPEAKEASY

The Science That Uncovered J.K. Rowling's Literary Hocus-Pocus

By BEN ZIMMER

Jul 16, 2013 8:33 am ET

0 COMMENTS



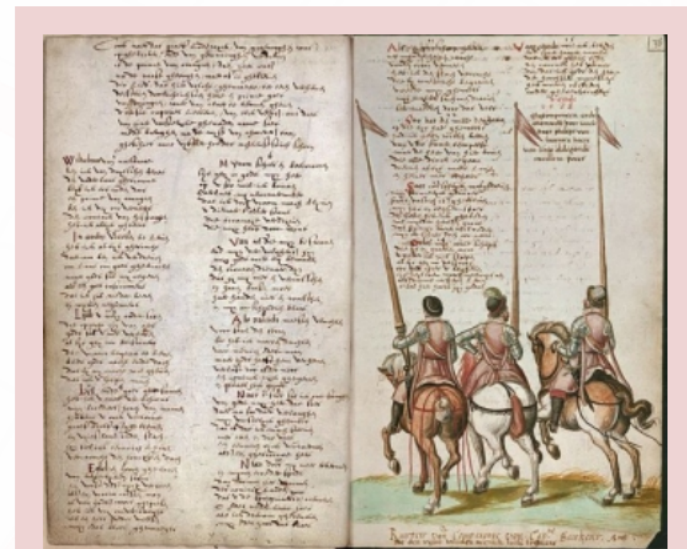
DAVID CHESKIN / PRESS POOL J.K. Rowling

The literary world is still abuzz over [the revelation](#) by London's Sunday Times that J.K. Rowling of "Harry Potter" fame secretly wrote the well-received crime novel "The Cuckoo's Calling" under the pen name Robert Galbraith. In chasing the scoop, the reporters called upon two experts in the field of authorship attribution to determine if "Galbraith" was really Rowling. The experts ran the texts through software programs

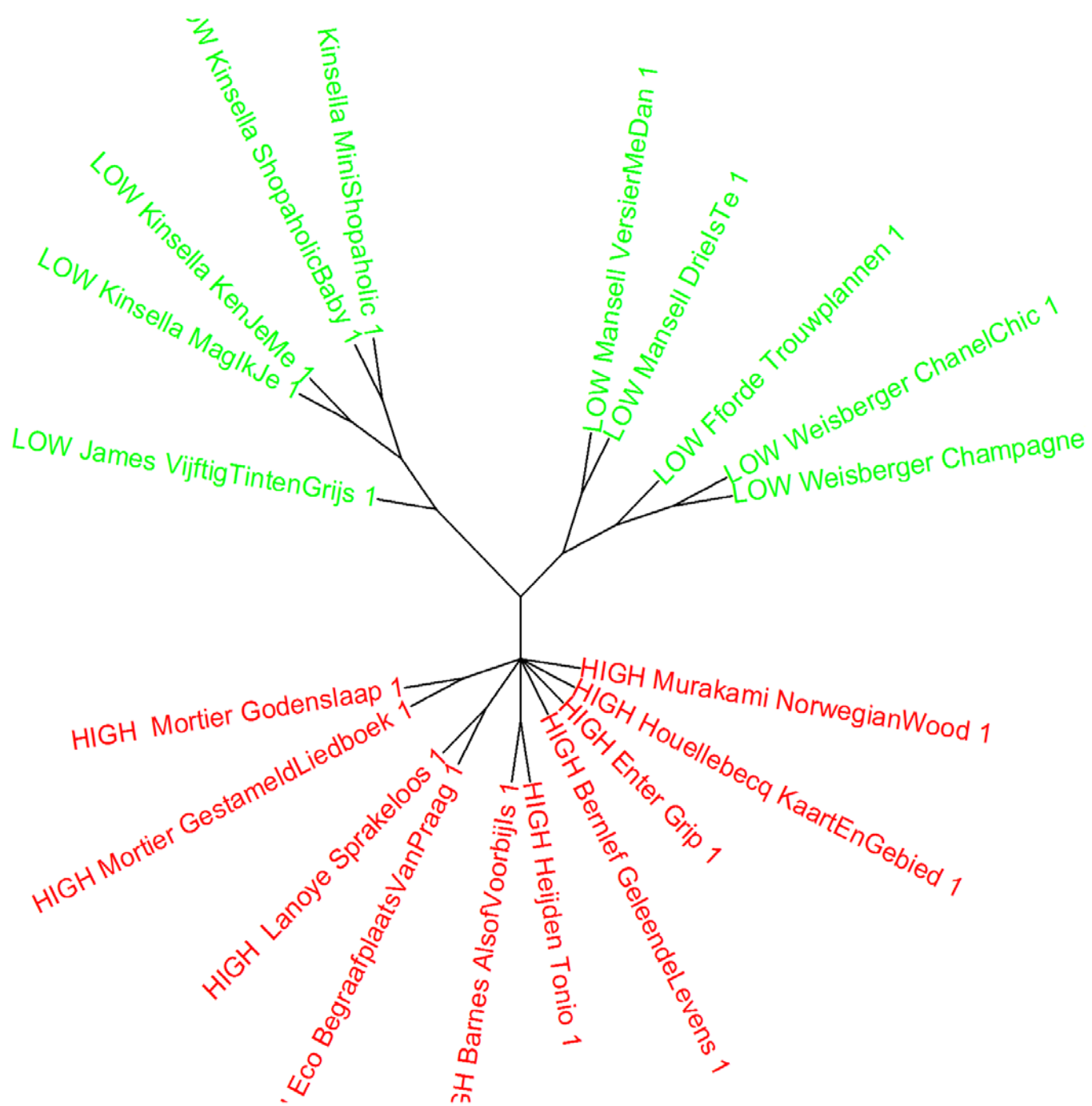
By 2
chan
to

Onderzoekers werpen met de computer nieuw licht op auteurschap Wilhelmus

Het Wilhelmus wordt ook wel het oudste volkslied ter wereld genoemd, maar we weten niet wie het schreef. Met nieuwe computertechnieken zijn onderzoekers uit Antwerpen, Amsterdam en Utrecht een mogelijke auteur op het spoor gekomen. Het is Petrus Datheen, een zestiende-eeuwse dichter waar eeuwenlang niemand meer aan dacht, met een slechte pers. Hij is in de vaderlandse geschiedenis vooral bekend geworden als een van de hoofdpersonen uit de 18-eeuwse psalmoproer. Het waren Datheens slecht-berijmde psalmen die de gereformeerde kerk toen door betere wilde vervangen. Heeft deze Datheen goede kaarten om het auteurschap van de nationale hymne van Nederland te claimen?



Datum 10 mei 2016



The SECRET LIFE of PRONOUNS

[contact](#)

a book by James W. Pennebaker

What our Words Say About us

The smallest, most commonly used, most forgettable words serve as windows into our thoughts, emotions, and behaviors. The ways people use pronouns, articles, and other everyday words are linked to their personality, honesty, social skills, and intentions.



Pronouns (*such as I, you, they*), articles (*a, an, the*), prepositions (*to, of, for*), auxiliary verbs (*is, am, have*), and a handful of other common word categories are called function words. On their own, function words have very little meaning. In English, there are fewer than 500 function words yet they account for more than half of the words we speak, hear, and read every day. By analyzing their use, we begin to learn how speakers are connecting with their audiences, their friends, their

conversational topics, and themselves.

The Secret Life of Pronouns is based on a large-scale research project that links natural language use to real world social and psychological processes

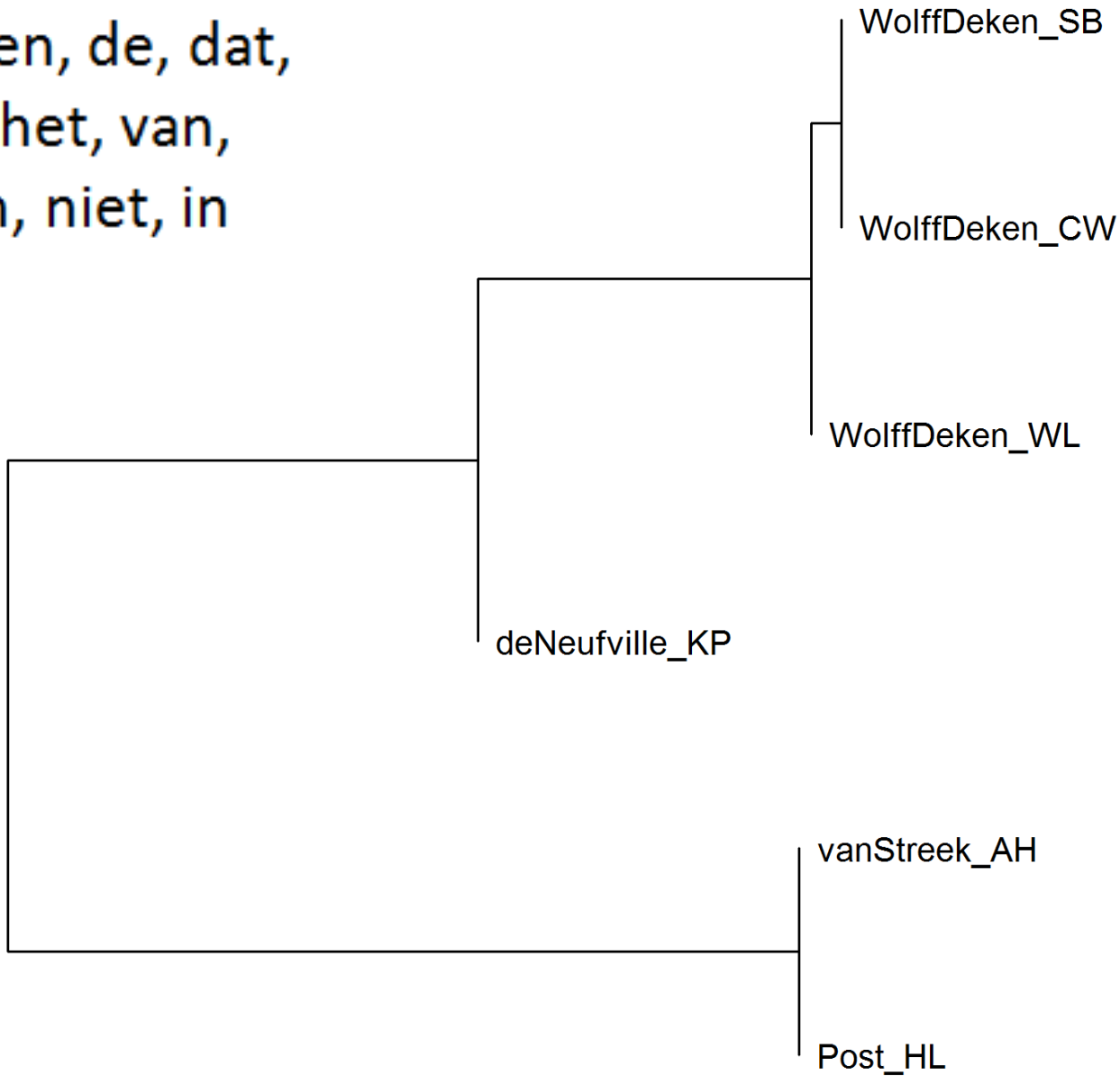
THE SECRET
LIFE OF
PRONOUNS
WHAT OUR
WORDS
SAY ABOUT US

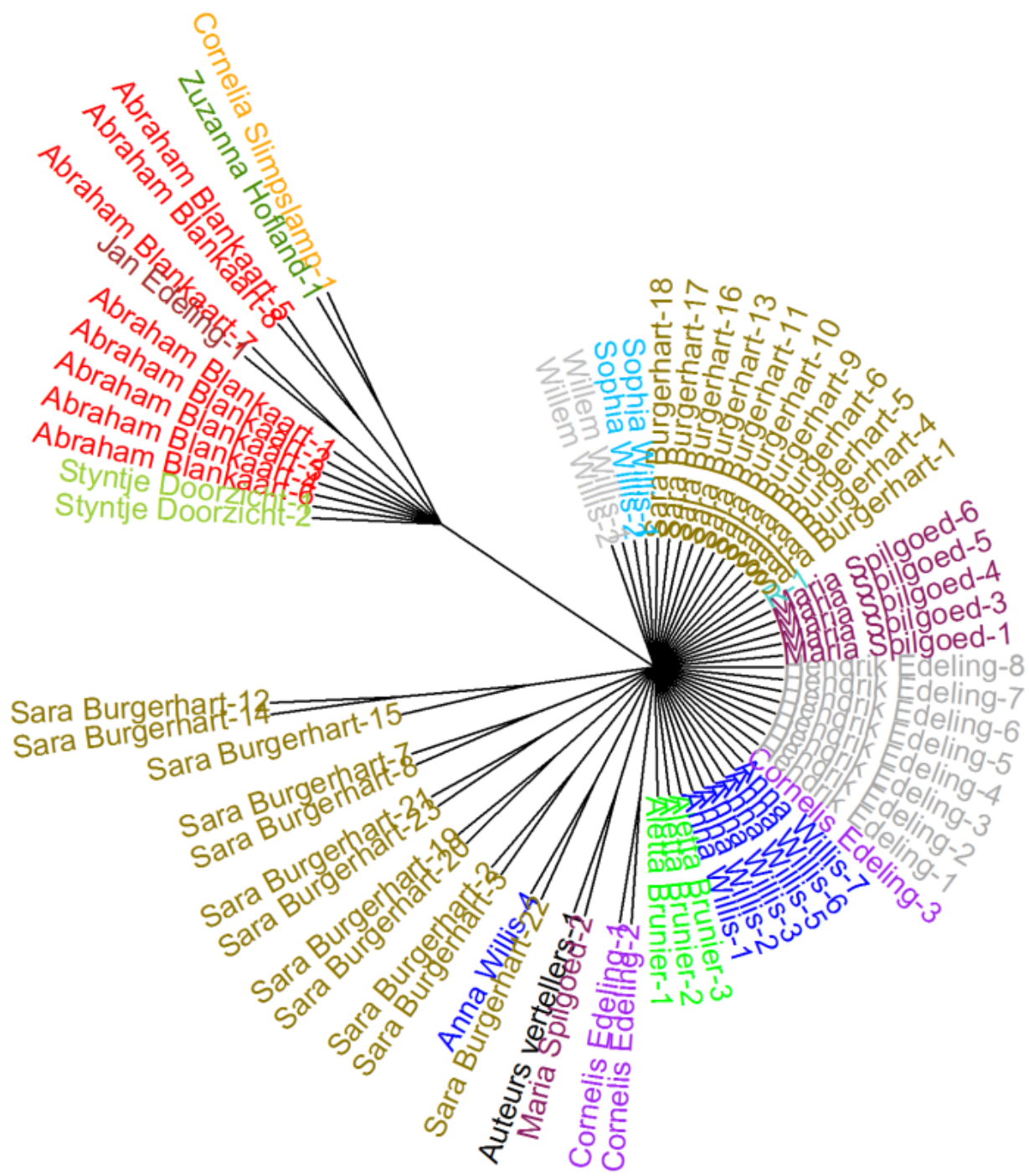
Top 10 woorden

ik, en, de, dat,

te, het, van,

een, niet, in







Abraham Blankaart vs Sara Burgerhart

Abraham Blankaart gebruikt relatief veel vaker:

Blankaart je braaf hoor jy brave **god** zie zoon wil ja
saartje **ouwe** uw denk jan **kind** toch noch **oud** zeg
lui altyd ken eer kom **kinderen** waart vriend
amsterdam **kerk** hele mevrouw doen goed jaar
wou gek

En veel minder vaak

zeer letje vriendin ons zag ô burgerhart geheel naatje
der lotje deeze des hartog dit buigzaam eene
myner sara r myne alle waarlyk reeds zelf
mooglyk zeker zonder eenige waarde

Een nieuwe definitie van stijl

- **Style is a property of texts constituted by an ensemble of formal features which can be observed quantitatively or qualitatively.**

(Stijl is een eigenschap van teksten die wordt uitgedrukt door een combinatie van formele kenmerken die zowel kwantitatief als kwalitatief geanalyseerd kunnen worden.)

J. Berenike Herrmann, Karina van Dalen-Oskam, Christof Schöch, Revisiting Style, a Key Concept in Literary Studies. *Journal of Literary Theory* 2015; 9(1): 25–52

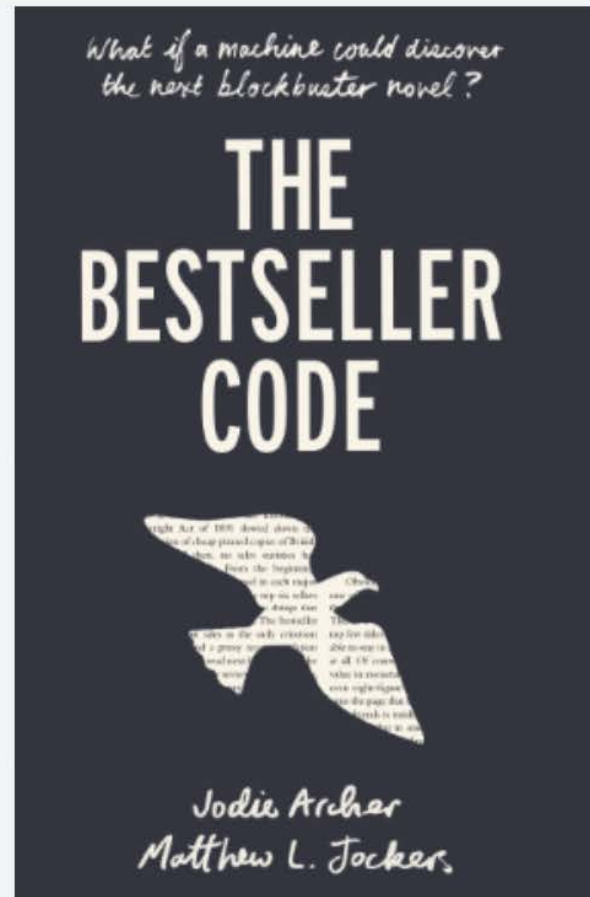
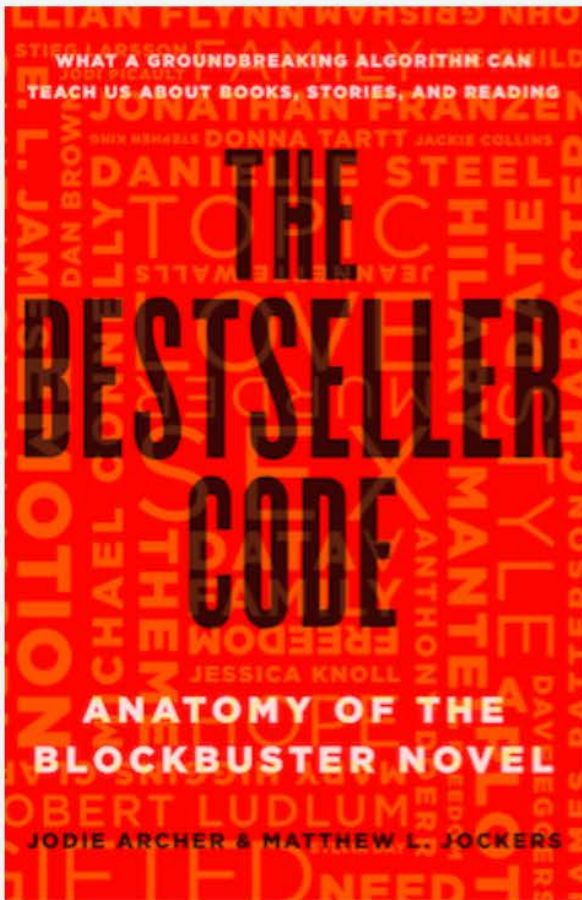


Jodie Archer and Matt Jockers do books.

Jodie has been in books forever. She has been a commissioning editor in London and New York. She has been research lead for a major online book store. She has been a lecturer in contemporary fiction and writing. All those book world roles have led to some expertise in answering two questions: *what works?* And *why?* Jodie got her PhD from Stanford for her work on the bestseller. The work of her first degrees at Cambridge was about how and why we read. She has been published under her own and different names. The press have called her critical work refreshing and revolutionary.

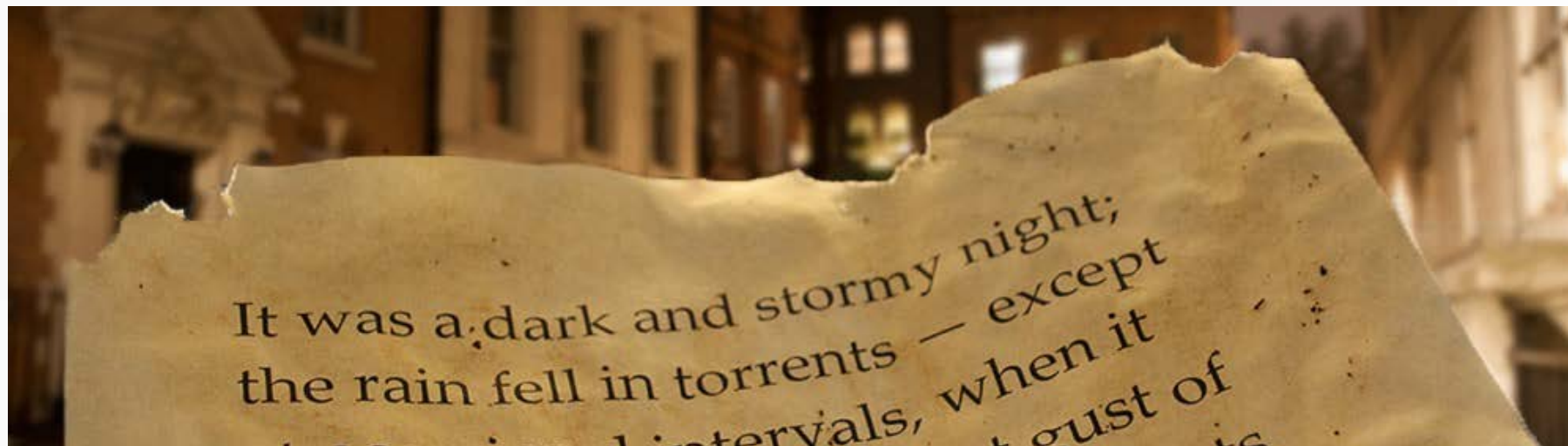


Matt is a book pro from a different perspective. He is considered a leading voice in text mining and the digital humanities--his work is covered world-wide with phrases like "paradigm-shifting" and "massively important." His two books in the field (*Macroanalysis* and *Text Analysis with R*) were groundbreaking. His work has been all about creating algorithms designed to detect the right words in the right order. He is the Susan J. Rosowski Professor of English at the University of Nebraska where he teaches and directs the Nebraska Literary Lab. He lectures all over the world about the intersections of technology and fiction.



Published September 13 (UK) & September 20 (US), 2016

The Riddle of Literary Quality



WAT VIND JIJ
VAN MIJ?



[Home](#) [Enquête](#) [Meer informatie](#)

HET NATIONALE LEZERSONDERZOEK

VUL DE
ENQUÊTE IN



Wat is literaire kwaliteit? Wat maakt een roman goed of slecht? Wanneer is een roman literair, of juist niet? En zijn deze dingen objectief vast te stellen?

Misschien wel. Wellicht zijn er elementen in een boek die bepalen dat het als goed of minder goed wordt gezien. Wetenschappers aan het Huygens Instituut voor Nederlandse Geschiedenis en de Universiteit van Amsterdam gaan dit onderzoeken.

Hiervoor is de mening van de lezer nodig. U dus. Welke boeken vindt u goed (of juist niet goed) en welke vindt u literair (of juist niet) en waarom?

Doe daarom mee aan het Nationale Lezersonderzoek en vul de enquête in. Dat kost ongeveer 5 minuten.

Veel plezier!

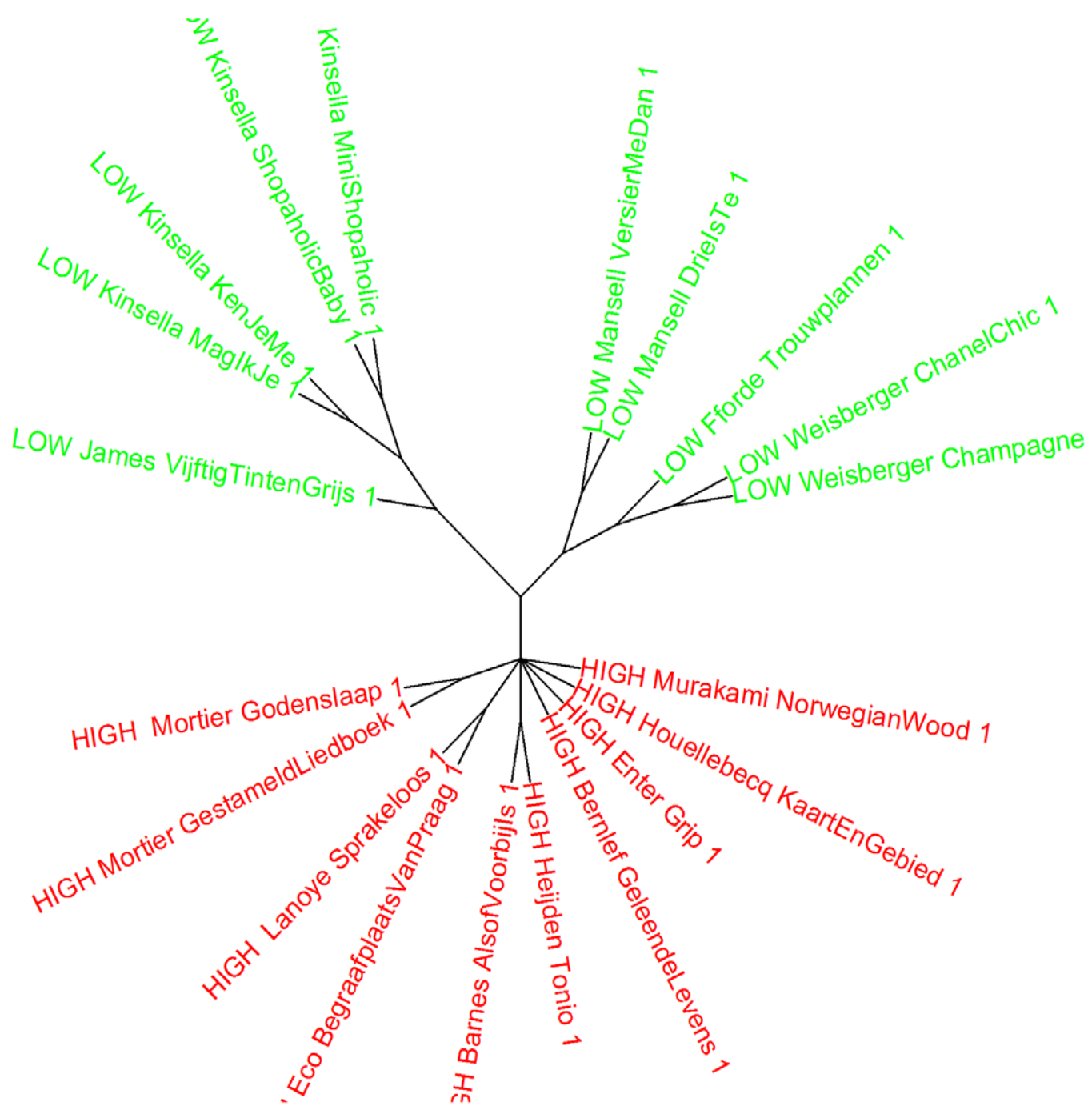
[Delen op Facebook](#)
[Delen op Twitter](#)

Top-10 **minst** 'literair'

		N	Mean
1	E.L. James Vijftig tinten grijs	320	2.1
2	Sophie Kinsella Shopaholic & Baby	313	2.4
3	Sophie Kinsella Mag ik je nummer even?	218	2.4
4	Lauren Weisberger Chanel Chic	212	2.5
5	Sophie Kinsella Mini Shopaholic	190	2.5
6	Sophie Kinsella Ken je me nog?	319	2.5
7	Jill Mansell Versier me dan	192	2.5
8	Lauren Weisberger Champagne in Ch Marmont	115	2.5
9	Katie Fforde Trouwplannen	150	2.6
10	Jill Mansell Drie is te veel	259	2.6

Top-10 **meest** 'literair'

		N	Mean
1	Julian Barnes Alsof het voorbij is	816	6.6
2	Erwin Mortier Godenslaap	810	6.6
3	Erwin Mortier Gestameld liedboek	575	6.5
4	Michel Houellebecq De kaart en het gebied	708	6.4
5	Tom Lanoye Sprakeloos	727	6.4
6	Haruki Murakami Norwegian Wood	848	6.3
7	A.F.Th. van der Heijden Tonio	769	6.3
8	Stephan Enter Grip	752	6.3
9	J. Bernlef Geleende levens	509	6.2
10	Umberto Eco Begraafplaats van Praag	573	6.2



Over stijl

- A.F.Th. Van der Heijden, *Het hof van barmhartigheid* (1996) – De tandeloze tijd 3, eerste deel
- J.J. Voskuil, *Vuile handen* (1996) – Het Bureau 2

Het Nationale Lezersonderzoek

- A.F.Th. Van der Heijden, *Tonio* (2011)
- J.J. Voskuil, *De buurman* (2012)

Top-10 **meest** 'literair'

		N	Mean
1	Julian Barnes Alsof het voorbij is	816	6.6
2	Erwin Mortier Godenslaap	810	6.6
3	Erwin Mortier Gestameld liedboek	575	6.5
4	Michel Houellebecq De kaart en het gebied	708	6.4
5	Tom Lanoye Sprakeloos	727	6.4
6	Haruki Murakami Norwegian Wood	848	6.3
7	A.F.Th. van der Heijden Tonio	769	6.3
8	Stephan Enter Grip	752	6.3
9	J. Bernlef Geleende levens	509	6.2
10	Umberto Eco Begraafplaats van Praag	573	6.2
.....			
28	J.J. Voskuil De buurman	605	6.2

Lezers die *De buurman* 'literair' vonden (score 6)

- Diepgang, tijdloos, prachtige zinnen.
- Goed geschreven, herkenbare karakters
- Het is vooral uit welwillendheid om het andere werk dat ik gelezen heb. Het complementeert de boel. Voskuils stijl is volstrekt eigen, zonder te willen behagen, zonder tegemoet te willen komen aan de snelle vluchtige consumerende lezer. Het is traag en vraagt om aandacht. Mooi Nederlands sterke psychologie, veel onuitgesproken gedachten. Het is pointillisme .

Lezers die *De buurman* ‘eerder niet literair’ vonden (score 3)

- Ik vind het niet literair omdat ik de stijl zo eenzijdig en lang niet altijd goed geschreven vind. Dit boek laat zo uitvoerig slechts 1 gezichtspunt vrij negatief zien (van Maarten Koning) dat je hoe langer hoe meer nieuwsgierig wordt naar de visie van degenen hij beschrijft.
- dagboek

Lezers die *Tonio* 'literair' vonden (score 6)

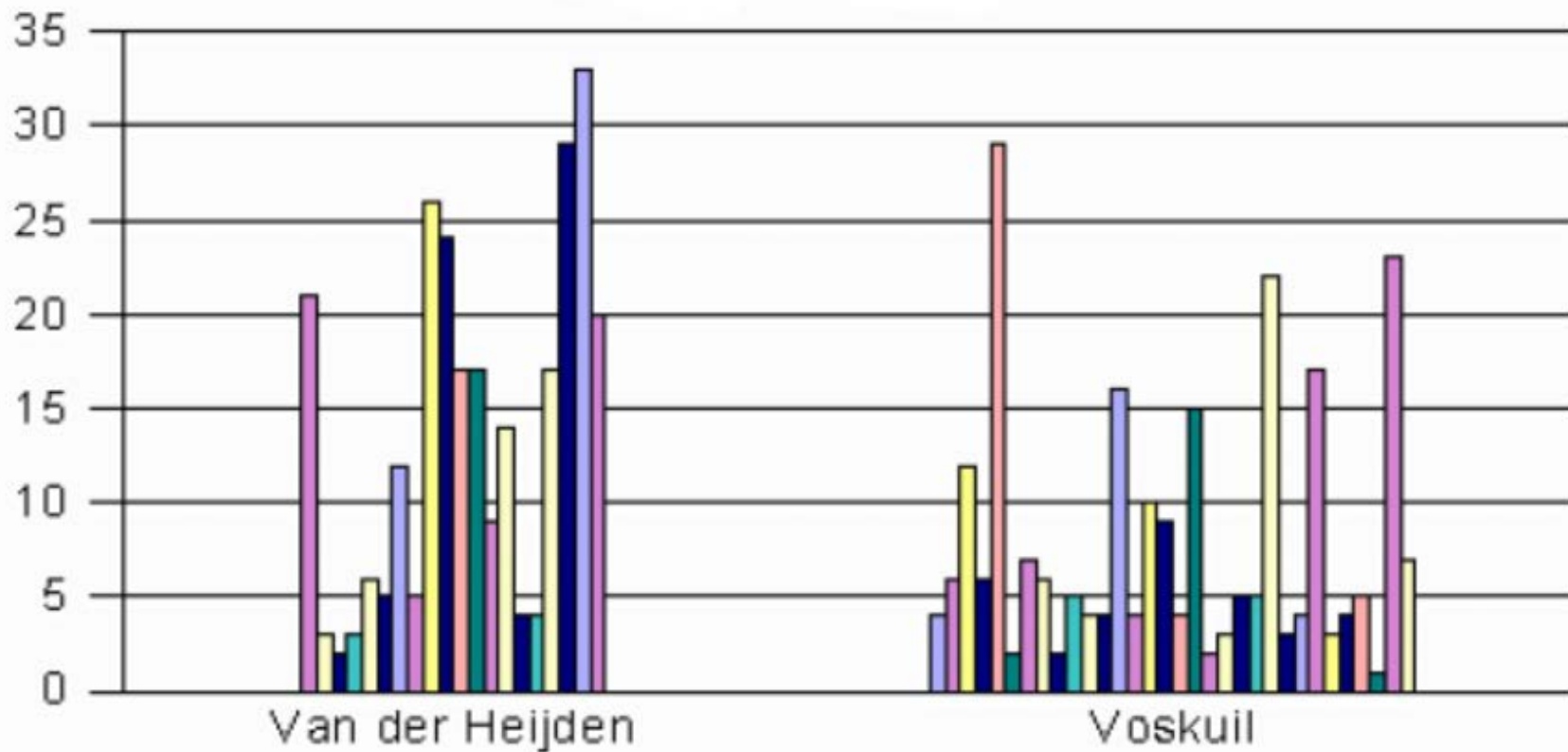
- 'Emotioneel verrassend, uitgedrukt in niet veel voorkomende taal een zekere zin melodieus.'
- De stijl is goed
- Diepgang
- Emoties goed verwoord
- Erg goed geschreven. Thema van het boek goed uitgewerkt.
- Het boek boeit de lezer. Is goed, in een mooie stijl geschreven.
- Het stilistisch goed en maakt gebruik van literaire middelen waardoor het uitstijgt boven de vertelde feiten.

Lezer die *Tonio* ‘absoluut niet literair’ vond (score 1)

- Wanneer ik het vergelijk met andere auteurs dienovereenkomstig dit aangrijpende thema schrijven (het verlies van een kind) dan vond ik dat van der Heijden alleen maar medelijden opwekt en verder niets. Dat is bv bij David grosmann (uit de tijd vallen) heel anders. Ik kwam ook geen mooie zinnen, woordspelingen en gedachtes tegen, alleen een gevestigde veelschrijver..

Zinslengte (1996, sample)

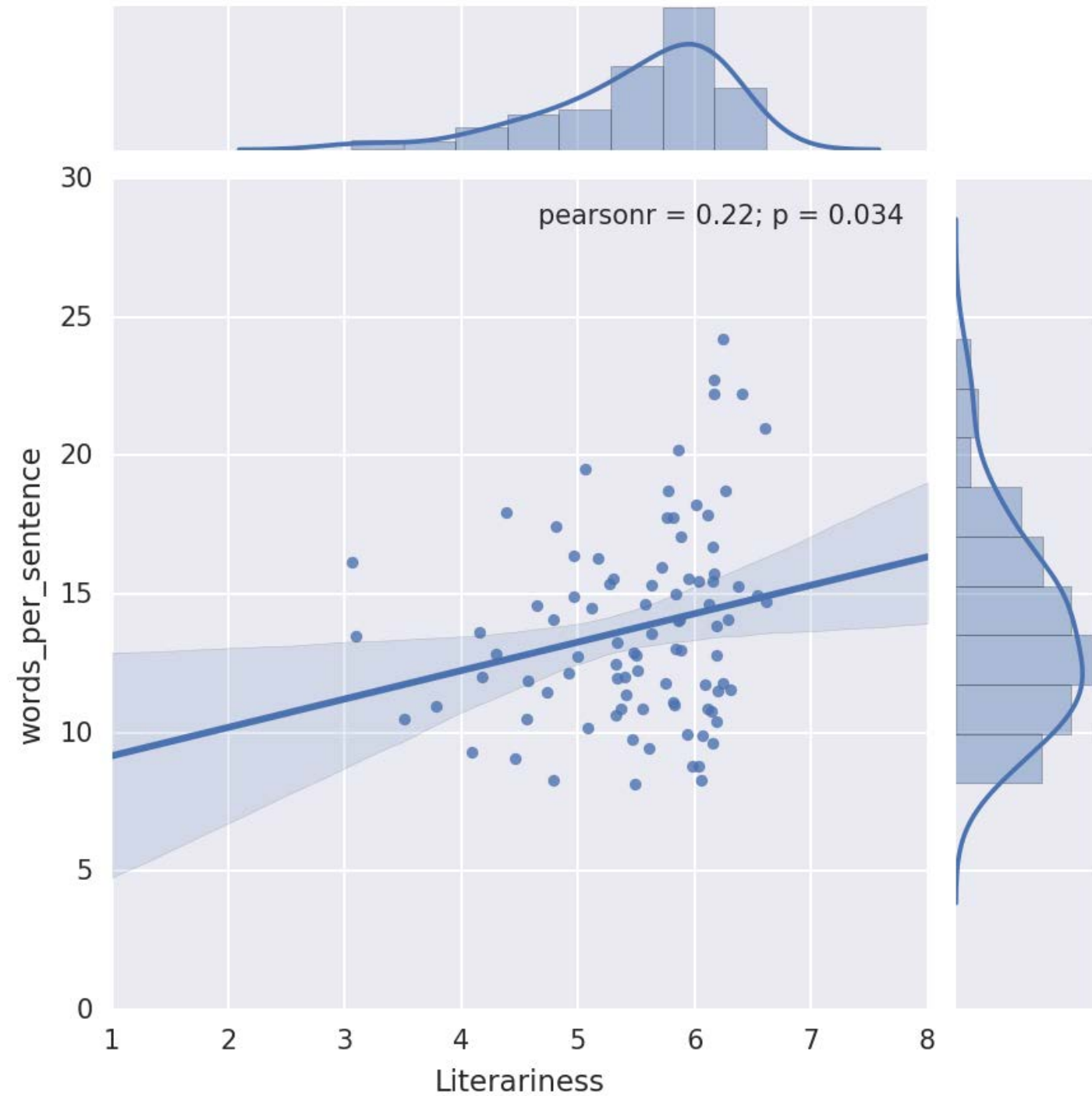
Zinslengte



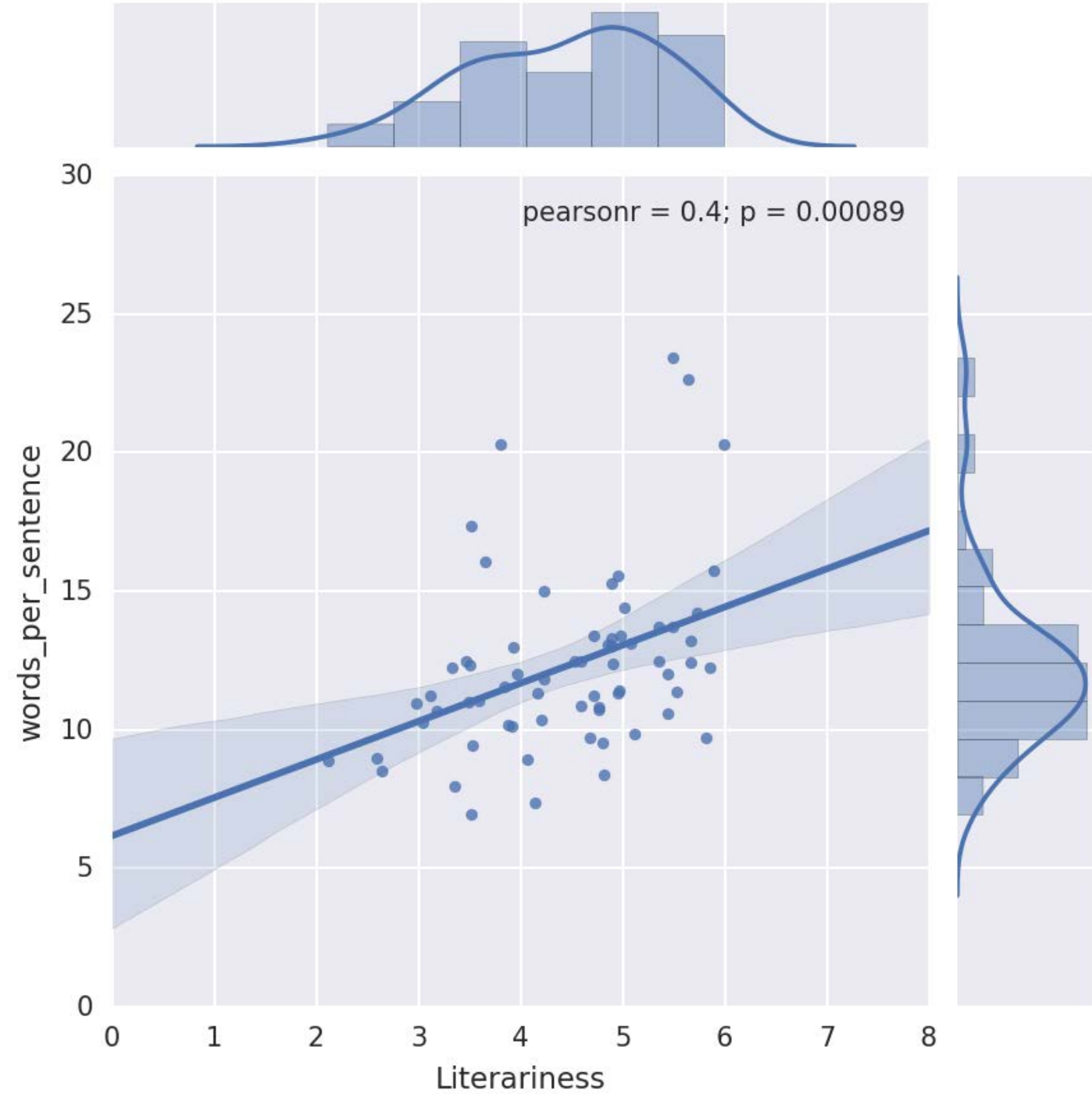
Zinslengte volledig / 2011 en 2012

- Van der Heijden, *Tonio*:
gemiddeld 13.3 woorden per zin
- Voskuil, *De buurman*:
gemiddeld 13.1 woorden per zin

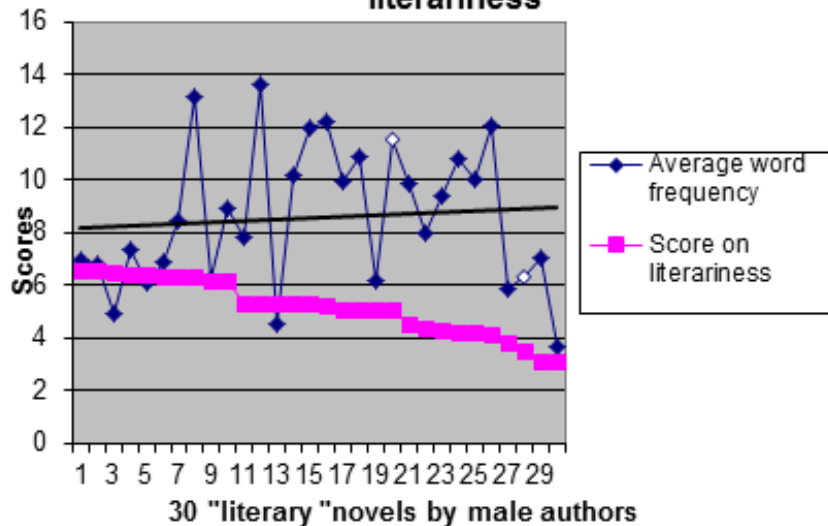
male



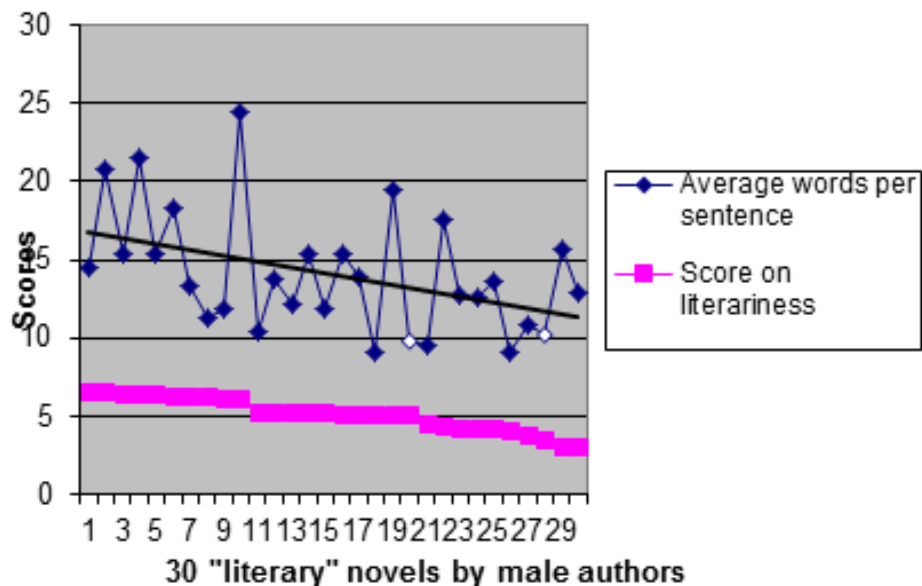
female



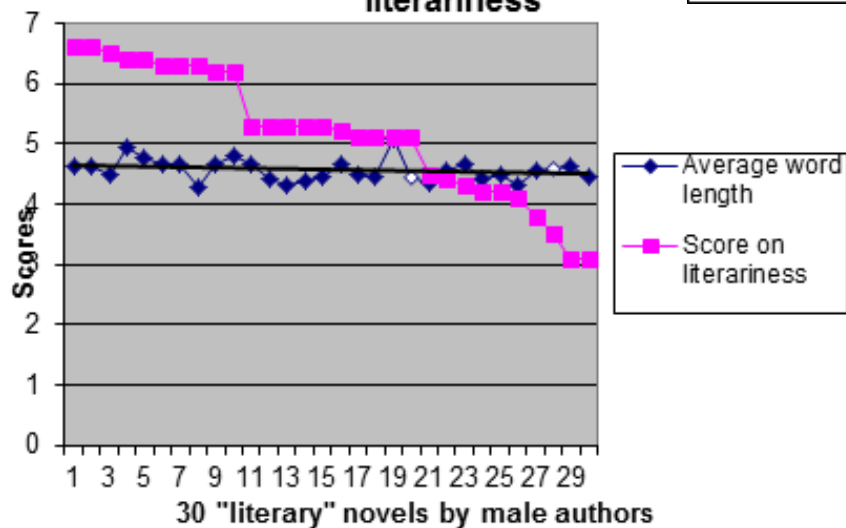
Average word frequency and score for literariness



Average words per sentence and score for literariness



Average word length and score for literariness



Confrontaties en conventies

- Percepties van stijl
- Conventies van literariteit
- Breder blik mede dankzij technische mogelijkheden
- Taal is voor literatuurwetenschappers weer zichtbaar geworden



Links naar websites en artikelen

<http://literaryquality.huygens.knaw.nl/>

<http://www.hetnationalelezersonderzoek.nl/>

http://www.oratiereeks.nl/upload/pdf/PDF-1290Weboratie_Van_Dalen-Oskam.pdf

<http://lcc.oxfordjournals.org/content/29/3/443.full.pdf?keytype=ref&ijkey=aRCuD3n825cEkmP>

<http://www.aclweb.org/anthology/W13-1410>

