

france_belgium
italy_france
russia
italy
switzerland
denmark
europe
japan
asia
india
africa
south_america
mexico
central_america
east_indies
ts_settlements
china_japan
netherlands
paraguay
new_zea
aus_alia
new_zealand
australia
south_africa
lon_economist
bri_tain
vian_countries
elgium_holland
united_states
states
entine_republic
nited_kingdom
argentine
argentina
dominions
britain
eas_dominions
ther_countries
countries
ean_countries
british_isles
united_slates
can_ada
america

Branches in the Canadian Northwest :

WANDON, Man.
A. Maybee, Manager
ALGARY—Alta.

MELFORT—Sask.
E. R. Jarvis, Manager.
MOOSE JAW—Sask.

Lost in (Vector) Space

Using machine learning to investigate discourses
of spatiality in a 33 million word corpus

British Library, 29 April 2019

Anouk Lang, University of Edinburgh

@a_e_lang | anouk.lang@ed.ac.uk | aelang.github.io

VERMILION—Alta.
H. M. Stewart, M'g'r.
MEDICINE HAT—Alta.
E. L. Crawford, Manager.

VERMILION—Alta.
A. C. Brown, Acting Manager
YELLOWGRASS—Sask.
E. J. Turner, Acting Manager



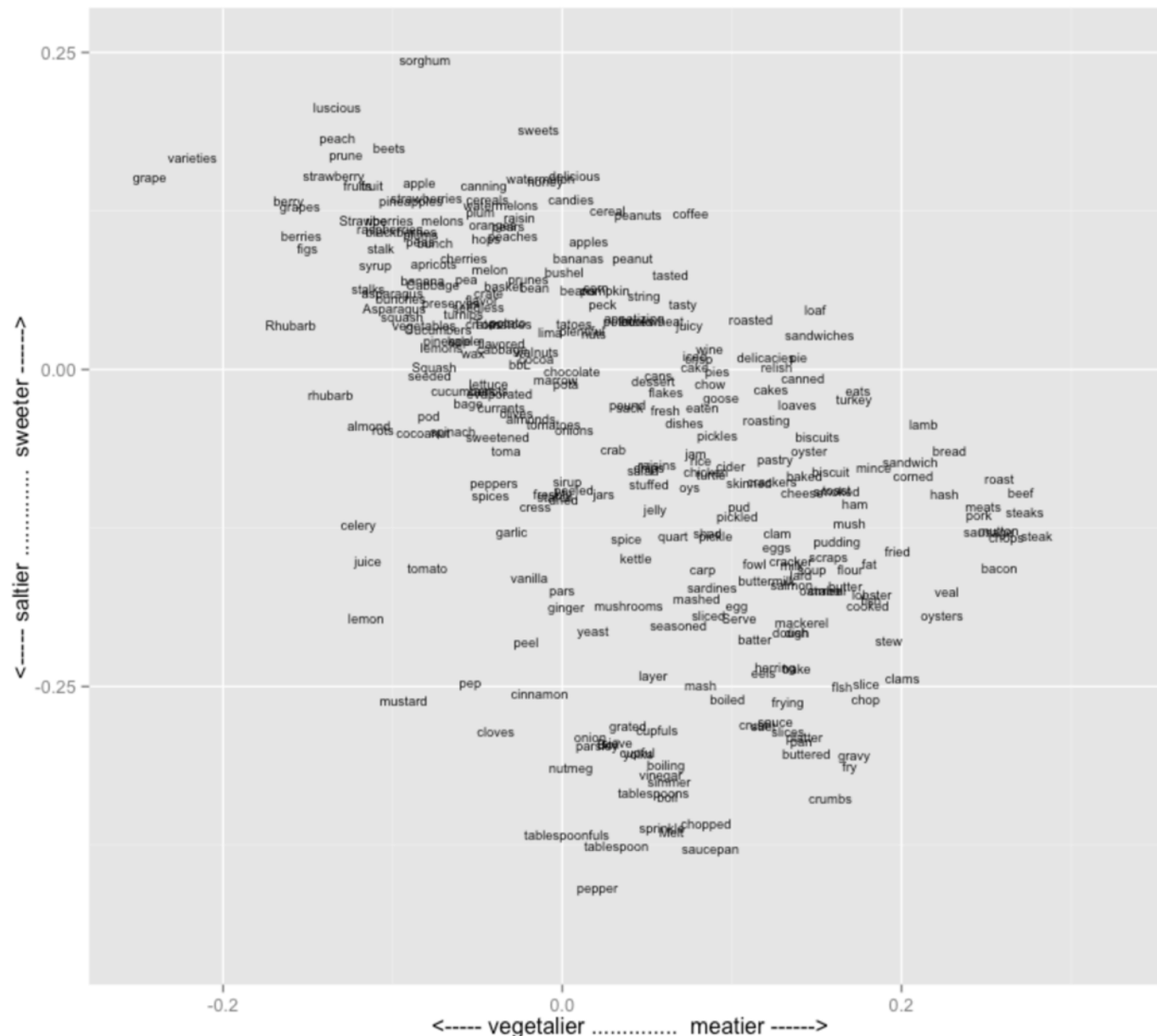
Produced by a GAN (generative adversarial network)
[StyleGAN](#) (Dec 2018) - [Karras](#) et al. and Nvidia
[Original GAN](#) (2014) - [Goodfellow](#) et al.
Don't panic. Learn about [how it works](#).
Help me figure out what was learned by this AI [here](#).
[Check out](#) text generation by another AI
[Click for another person](#) [Link to image](#)

Generating facial images with GANs: thispersondoesnotexist.com



blogs.nvidia.com/blog/2019/03/18/gaugan-photorealistic-landscapes-nvidia-research/

Anouk Lang | University of Edinburgh | @a_e_lang | anouk.lang@ed.ac.uk | aelang.github.io



Discourses of “saltiness”, “sweetness”, “vegetableness” and “meatiness” from a corpus of 6 million newspaper pages digitized by the NEH/Library of Congress project *Chronicling America*.

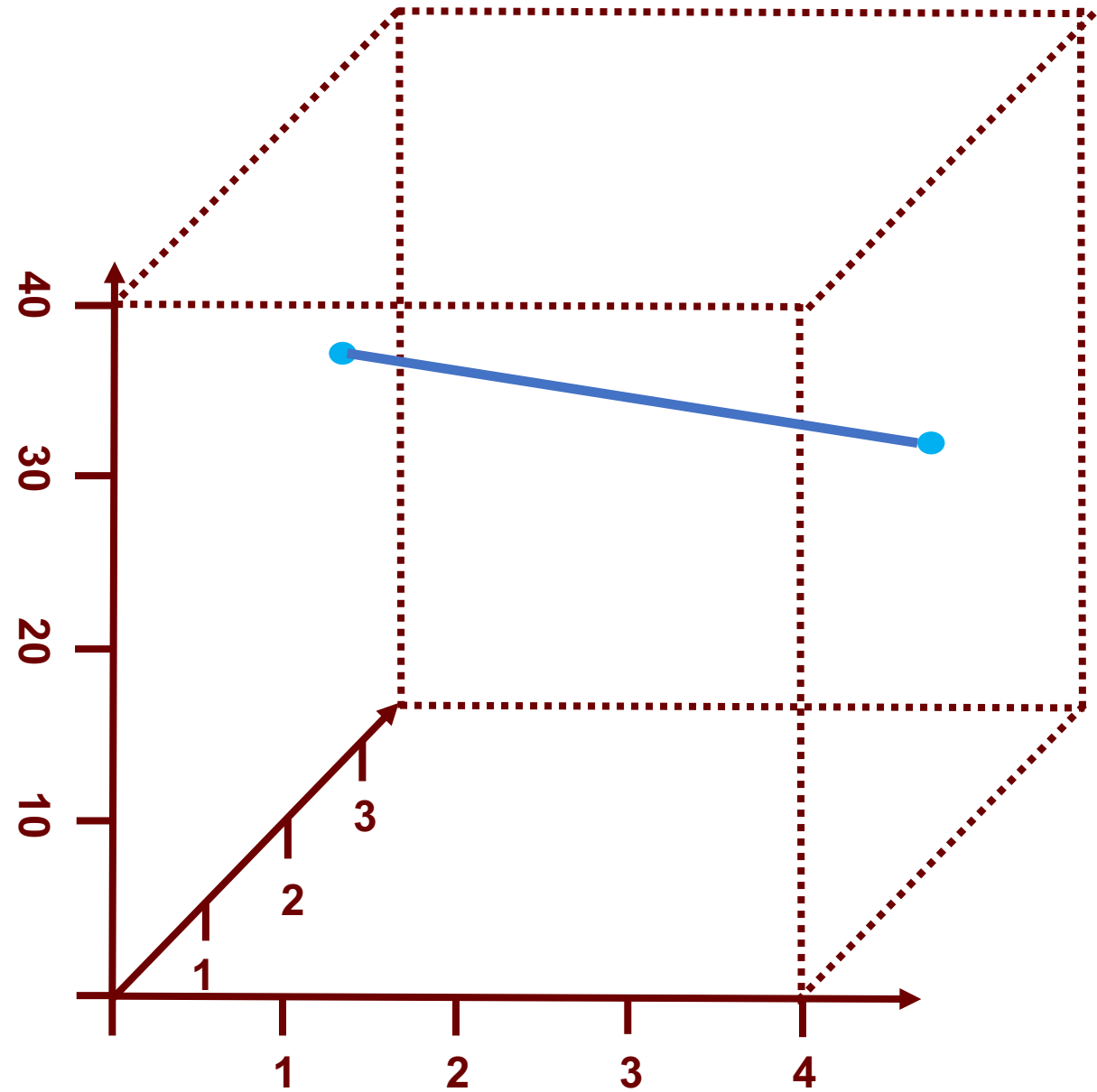
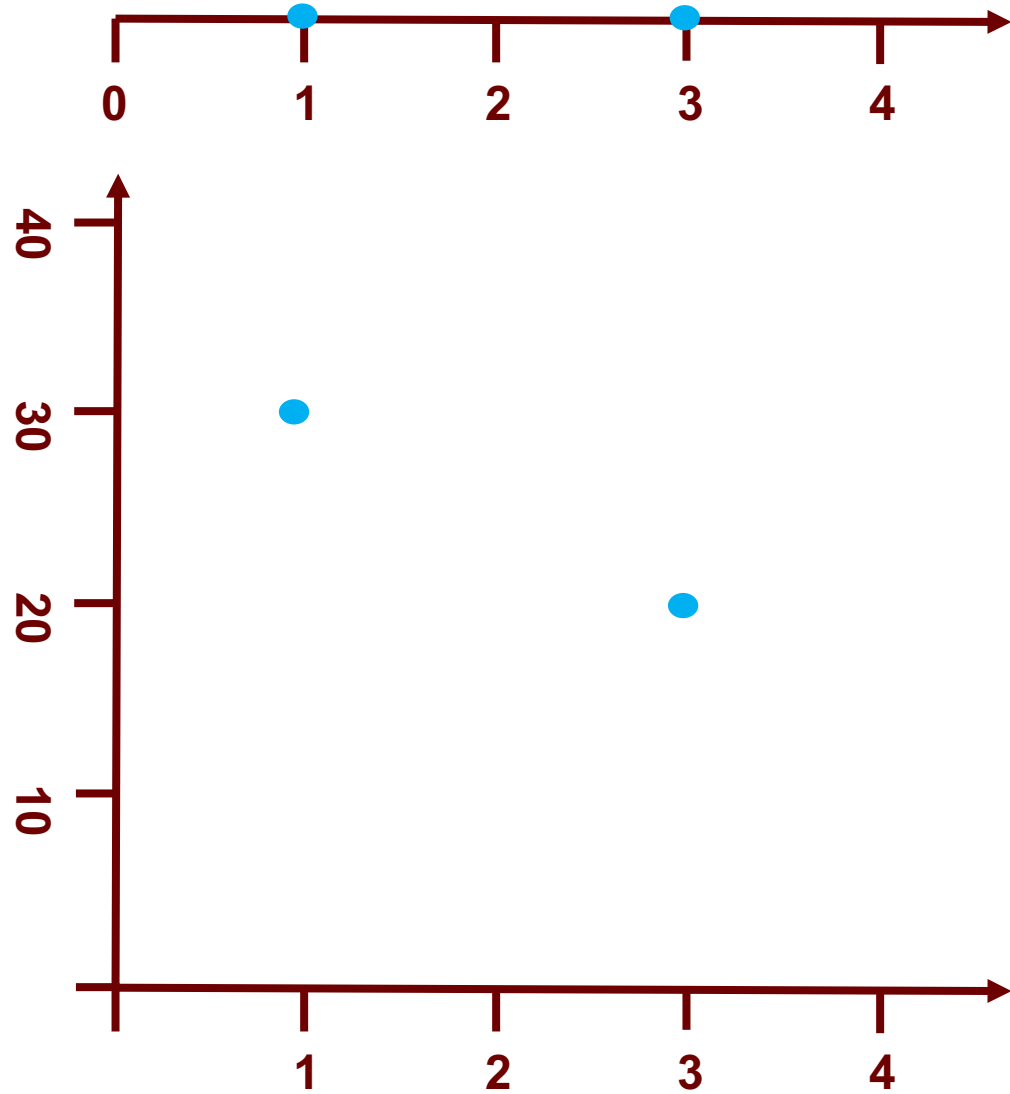
Ben Schmidt, “Vector Space Models for the Digital Humanities.” *Ben’s Bookworm Blog*, 25 Oct. 2015, <http://bookworm.benschmidt.org/posts/2015-10-25-Word-Embeddings.html>.

In modern cultural theory, especially in the post-structuralism associated with the French historian Michel Foucault, the term has been used to denote any coherent body of statements that produces a self-confirming account of reality by defining an object of attention and generating concepts with which to analyse it (e.g. medical discourse, legal discourse, aesthetic discourse). **The specific discourse in which a statement is made will govern the kinds of connections that can be made between ideas**, and will involve certain assumptions about the kind of person(s) addressed. By extension, as a free-standing noun ('discourse' as such), the term denotes language in actual use within its social and ideological context and in institutionalized representations of the world called discursive practices. In general, the increased use of this term in modern cultural theory arises from dissatisfaction with the rather fixed and abstract term 'language' ... by contrast, **'discourse' better indicates the specific contexts and relationships involved in historically produced uses of language.** [emphases added]

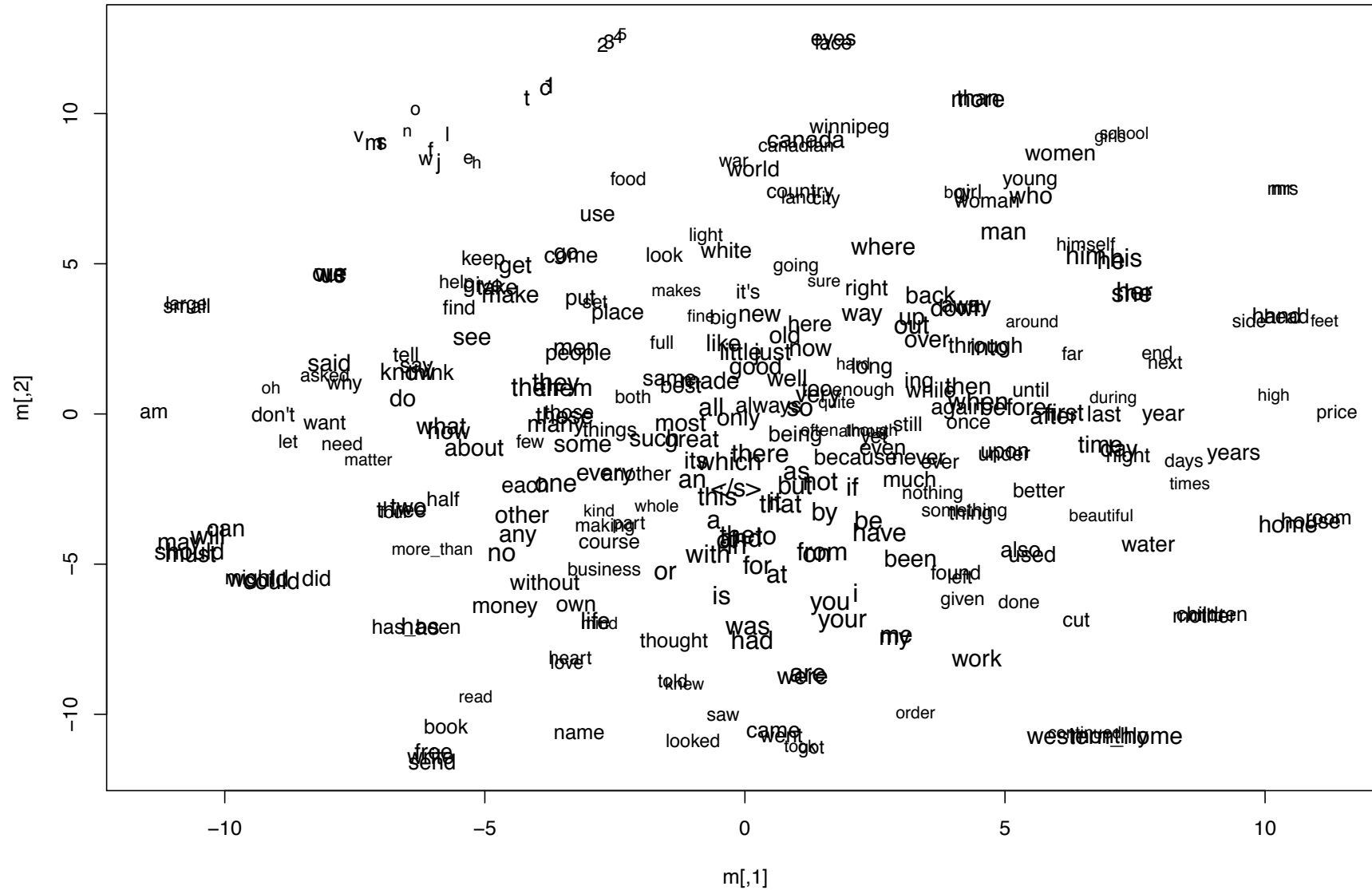
Chris Baldick, ed., *The Oxford Dictionary of Literary Terms*, Oxford UP, 2008.

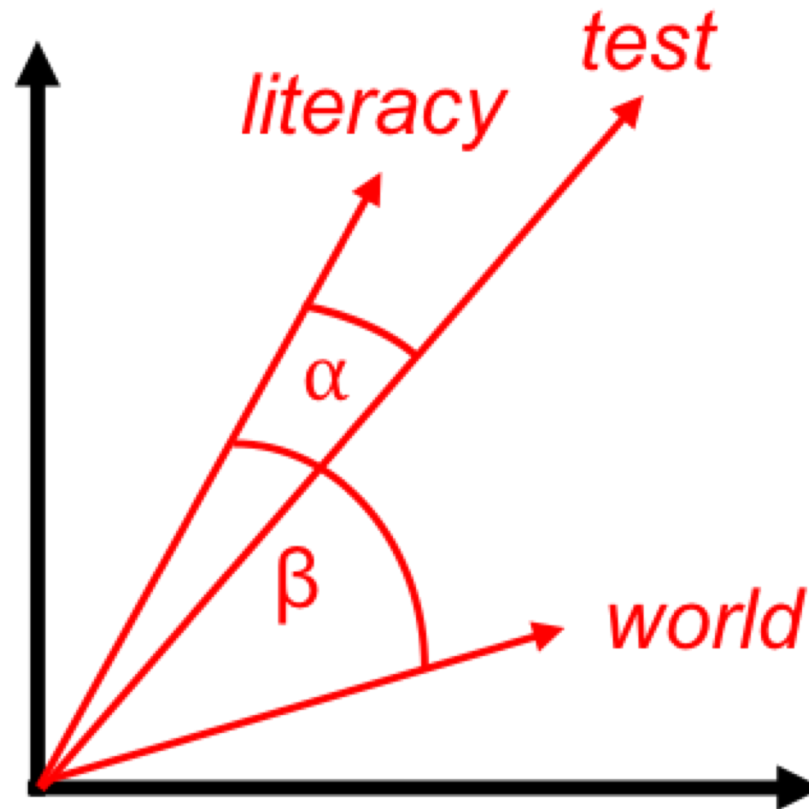
Sentences from <i>WHM</i> containing <i>literacy</i>		Co-occurrence counts		Vector
<i>new provisions of the immigration laws require the passing of a literacy test about which there is still some discussion</i>		the 11		11
		of 9		9
		a 5		5
		in 4		4
		is 3		3
<i>the world has been transformed from illiteracy to literacy</i>		nation 3		3
		to 3		3
<i>those who refused to believe in the literacy of these figurative parts of the bible were hounded as infidels</i>		as 2		3
		believe 2		2
<i>canada stands high among the nations in literacy</i>	→	from 2	→	2
		still 2		2
<i>the following questions are taken from a literacy test published in the memoirs of the national academy of sciences</i>		test 2		2
		that 2		2
		these 2		2
		which 2		2
<i>literacy is a measure of national stability</i>		about 1		2
		academy 1		1
<i>in africa where literacy is less common the public scribes still do a flourishing trade</i>		africa 1		1
		bible 1		1
		canada 1		1
	
<i>peter wished to tell no one that he could read and write believing that his literacy gave him a kind of power over these men which would impress them so long as they did not know its origin</i>		world 1		1
		would 1		1
		write 1		1

aelang.github.io/word-embeddings



A two dimensional reduction of the vector space model using t-SNE



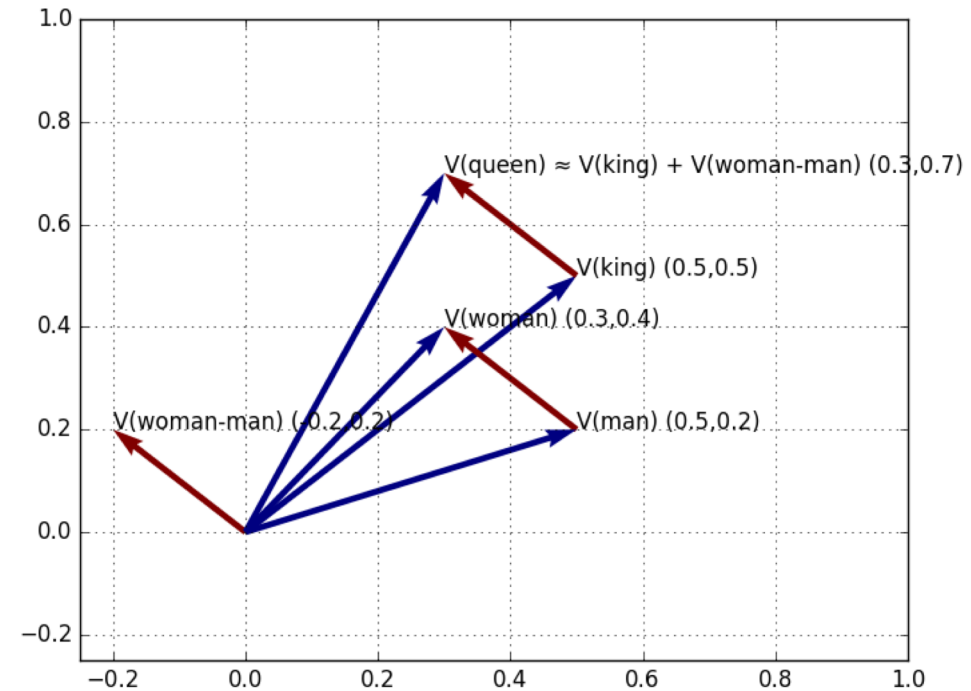


aelang.github.io/word-embeddings

Analogical reasoning

king is to *queen*

as *man* is to ...?



Heuser, Ryan. "Word Vectors in the Eighteenth Century, Episode 2: Methods." *Virtue and the Virtual*, 1 June 2016, <http://ryanheuser.org/word-vectors-2/>.



THE WESTERN HOME MONTHLY.



Vol. I.—No. 5.

WINNIPEG, CANADA, MAY, 1901.

PRICE { 50c. per year.
5c. per copy.



On the "D K" Ranch.

A LOVE STORY.

Written for the Western Home Monthly by
Jack Beeton, Winnipeg.

Dawn was just breaking over the western prairies. Far away to the eastward, beyond leagues of billowy meadow grass, a vast dome of dull gold was rising silently and swiftly. From a belt

out here among the tawny-backed hills and the fresh, fragrant valleys, he had said good-bye to them all a few years ago and had gone to study medicine in the city. Since that time he had ne'er

With one hand grasping the pommel, she hung perilously, her left foot entangled in the stirrup. "The broncho plunged from left to right.

"Hold on for your life," shouted Leslie, darting forward.

In another instant he had the bridle in a firm grip and the horse brought to a standstill. Then with the other arm he drew the girl from her dangerous position. He caught one frightened glance from a pair of dark eyes.

"My ankle," she moaned, "it must be broken." And then she fainted.

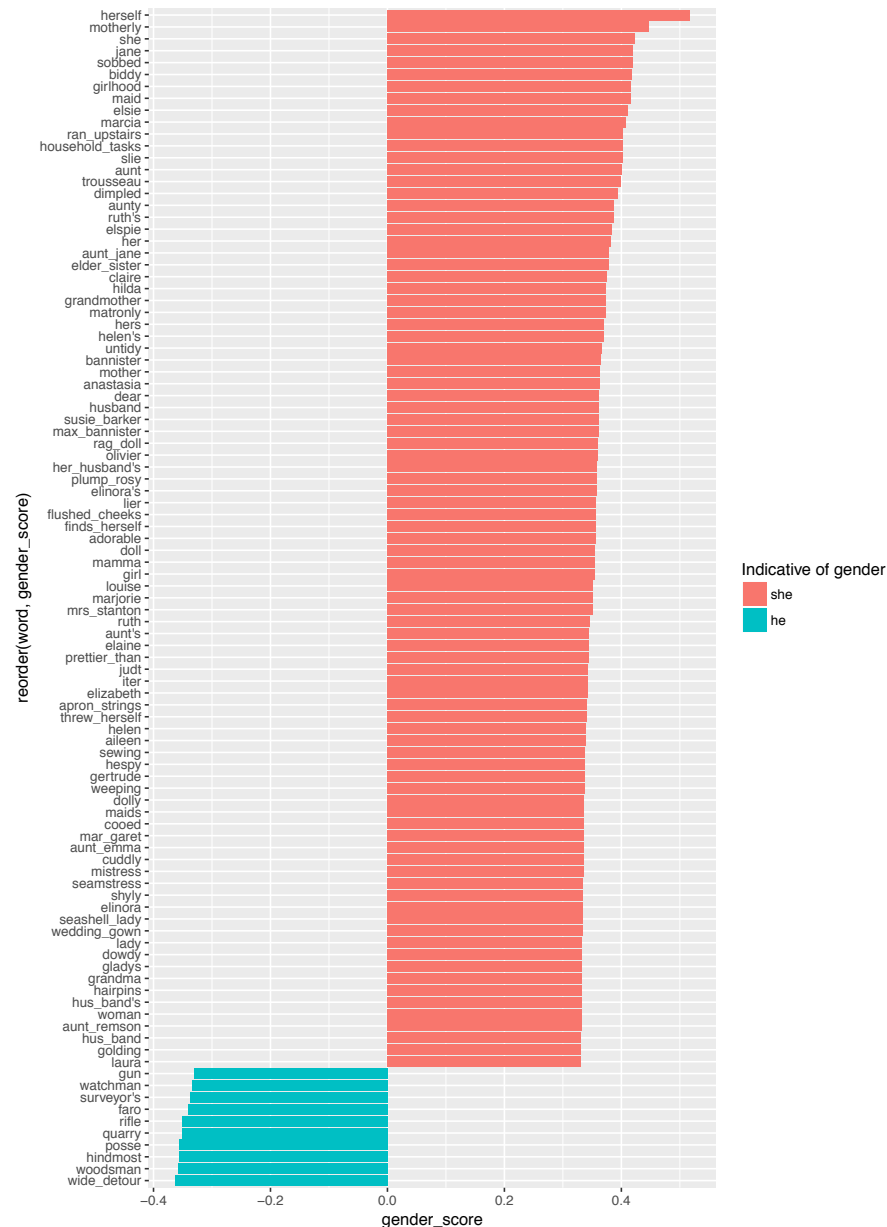
They were close to the edge of the ravine, and he knew every foot of it. Taking her in his arms as though she were a child, he ran down the steep incline towards where, far below at the bottom, a little stream rippled over the stones. Hazel bushes showered them

"Your pony threw you; don't you remember?" he replied, gently, "and I carried you down to the water."

"Oh, I remember," she said, passing her hand across her brow, as if to clear away its confusion. "How foolish of me to be riding so early in the morning, and alone. But, then, I had done it so often before, and nothing ever happened. It is the sunrise," she explained, turning to him half apologetically. "It is so beautiful I can't resist it. I believe I am a perfect slave to beauty in any form," she went on, half laughing, and then as her eyes rested on his face she crimsoned hotly and turned away covered with confusion at the slip she had made. "What a bold compliment he will think I am paying him," she thought to herself angrily, as she bent over the hurt foot. But nothing could be farther from Leslie's mind. Faults he had,

- published in Winnipeg, Manitoba, Canada from 1899-1932
- digitised version contains 24,170 pages and 33,099,536 words, encompassing 348 monthly issues (2 from 1901 and 1903, 10 from 1904, every issue between 1905 and 1932 with the exception of January 1916, September 1919, and March 1922, plus a special illustrated issue *The 1914 War* which supplements the twelve issues from 1915)

Hannah McGregor and Nicholas van Orden, "Remediation and the Development of Modernist Forms in *The Western Home Monthly*," *Reading Modernism with Machines*, ed. Shawna Ross and James O'Sullivan (London: Palgrave, 2016) 135–63.



Words in the *Western Home Monthly* showing the strongest skew along the gender binary.

Words designated by orange bars in the upper part of the plot provide stereotypical descriptions of women, alongside several terms which mark transgressions from approved gender norms (*untidy*, *dowdy*). Other nouns reinforce these associations (*household tasks*, *trousseau*, *rag_doll*, *flushed_cheeks*, *doll*, *apron_strings*, *wedding gown*, *hairpins*) and give a picture of a very clear semantic field of family, home and marriage.

Rank	Word	Similarity to V(<i>winnipeg</i>)	Word	Similarity to V(<i>canada</i>)	Word	Similarity to V(<i>regina, winnipeg, calgary, edmonton, vancouver, montreal, toronto</i>)
1	<i>winnipeg</i>	1.0000	<i>canada</i>	1.0000	<i>vancouver</i>	0.8572
2	<i>factory_599</i>	0.7101	<i>western</i>	0.6494	<i>calgary</i>	0.8478
3	<i>322_eighth</i>	0.7073	<i>canadian</i>	0.6397	<i>regina</i>	0.8445
4	<i>214_second</i>	0.7045	<i>can_ada</i>	0.6207	<i>winnipeg</i>	0.8179
5	<i>325_main</i>	0.6983	<i>dominion</i>	0.6019	<i>kegina</i>	0.8167
6	<i>box_1604</i>	0.6873	<i>waltham_products</i>	0.5862	<i>edmonton</i>	0.8153
7	<i>edmonton_saskatoon</i>	0.6857	<i>c_turnbull</i>	0.5698	<i>edmonton_saskatoon</i>	0.8084
8	<i>1859_scarth</i>	0.6853	<i>rivers_p.q</i>	0.5672	<i>saskatoon_edmonton</i>	0.8075
9	<i>hargrave_street</i>	0.6838	<i>canad</i>	0.5593	<i>winnipeg_regina</i>	0.8020
10	<i>brandon_distributors</i>	0.6824	<i>limite</i>	0.5585	<i>halifax_moncton</i>	0.8015
11	<i>regina_calgary</i>	0.6818	<i>imperial_trading</i>	0.5514	<i>edmonton_vancouver</i>	0.7963
12	<i>john_paton</i>	0.6811	<i>canadian_workmen</i>	0.5475	<i>factory_599</i>	0.7933
13	<i>erin_st</i>	0.6797	<i>established_1848</i>	0.5456	<i>calgary_edmonton</i>	0.7923
14	<i>10153_jasper</i>	0.6781	<i>e.b.eddy</i>	0.5418	<i>214_second</i>	0.7895
15	<i>saskatoon_vancouver</i>	0.6765	<i>halifax_saint</i>	0.5399	<i>325_main</i>	0.7881
16	<i>princess_street</i>	0.6729	<i>toronto</i>	0.5314	<i>saskatoon_lethbridge</i>	0.7798
17	<i>edmonton_vancouver</i>	0.6725	<i>colonization_association</i>	0.5305	<i>10153_jasper</i>	0.7798
18	<i>saskatoon_lethbridge</i>	0.6716	<i>gravenhurst</i>	0.5298	<i>calgary_regina</i>	0.7735
19	<i>mcarthur_bldg</i>	0.6708	<i>limited_halifax</i>	0.5288	<i>322_eighth</i>	0.7731
20	<i>kegina</i>	0.6706	<i>li_mite</i>	0.5229	<i>john_paton</i>	0.7632

Top 20 words closest to the vectors a) *winnipeg*, b) *canada*, and c) *regina*, *winnipeg*, *calgary*, *edmonton*, *vancouver*, *montreal* and *toronto*.

Advertisement for the Canadian Bank of
Commerce from the *Western Home
Monthly*, vol. VII no. 3 (March 1906), p.7.

THE CANADIAN BANK OF COMMERCE

CAPITAL - \$10,000,000 REST - \$4,500,000
HEAD OFFICE, TORONTO
B. E. WALKER, General Manager ALEX. LAIRD, Asst. Gen'l Manager

ONE HUNDRED AND THIRTY BRANCHES IN CANADA,
THE UNITED STATES AND ENGLAND.

Branches in the Canadian Northwest :

BRANDON, Man. A. Maybee, Manager	MELFORT—Sask. E. R. Jarvis, Manager.
CALGARY—Alta. C. W. Rowley, Manager.	MOOSE JAW—Sask. E. M. Saunders, Manager.
CARMAN—Man. E. C. Complin, Manager.	MOOSOMIN—Sask. D. I. Forbes, Manager.
CLARESHOLM—Alta. A. B. Irvine, Manager.	NANTON—Alta. N. F. Ferris, Manager.
DAUPHIN—Man. D. H. Downie, Manager.	NEEPAWA—Man. G. M. Gibbs, Manager.
EDMONTON—Alta. T. M. Turnbull, Manager.	NORTH BATTLEFORD, Sask. A. S. Houston, Manager.
ELGIN—Man. H. B. Haines, Manager.	PINCHER CREEK, Alta. W. G. Lynch, Manager.
ELKHORN—Man. R. H. Brotherhood, Manager.	PONOKA—Alta. E. A. Fox, Manager.
GILBERT PLAINS—Man. G. C. T. Pemberton, Manager.	PORTAGE LA PRAIRIE—Man. A. L. Hamilton, Manager.
GRANDVIEW—Man. G. C. T. Pemberton, Manager.	PRINCE ALBERT—Sask. C. D. Nevill, Manager.
HIGH RIVER—Alta. P. H. Jory, Manager.	RED DEER—Alta. A. Scott, Manager.
INNISFAIR—Alta. H. L. Edmonds, Manager.	REGINA—Sask. H. F. Mytton, Manager.
KINISTINO—Sask. In Charge temporarily of	SASKATOON—Sask. W. P. Kirkpatrick, Manager.
E. R. Jarvis, Manager at Melfort	SWAN RIVER—Man. F. J. Macoun, Manager.
LETHBRIDGE—Alta. C. G. K. Nourse, Manager	TREHERNE—Man. J. S. Munro, Manager.
LLOYDMINSTER—Sask. S. M. Daly, Manager	VEGREVILLE, Alta. W. P. Perkins Manager
MACLEOD—Alta. H. M. Stewart, M'gr.	VERMILION—Alta. A. C. Brown, Acting Manager
MEDICINE HAT—Alta. F. L. Crawford, Manager.	YELLOWGRASS—Sask. F. J. Turner, Acting Manager

WINNIPEG, MAN.—John Aird, Manager.

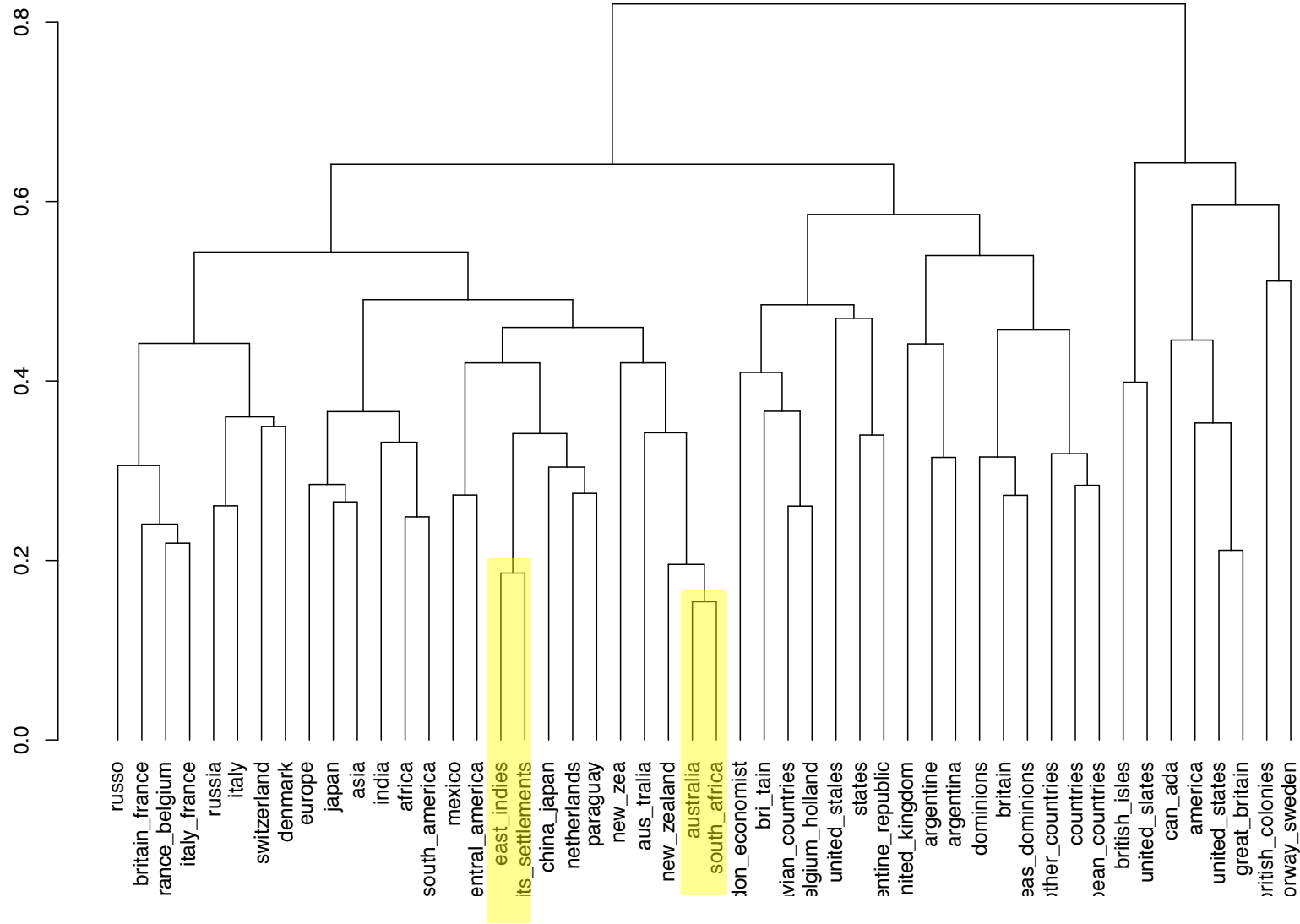
FARMERS' BANKING
Every facility for farmers' banking. Notes discounted, sales notes collected. Deposits
may be made and withdrawn by mail. Every attention paid to out-of-town accounts.
Savings Bank Department at every Branch.
Deposits of \$1.00 and upwards received and interest allowed at current rates.
GENERAL BANKING BUSINESS TRANSACTED.

WHEN WRITING ADVERTISERS PLEASE MENTION THE WESTERN HOME MONTHLY

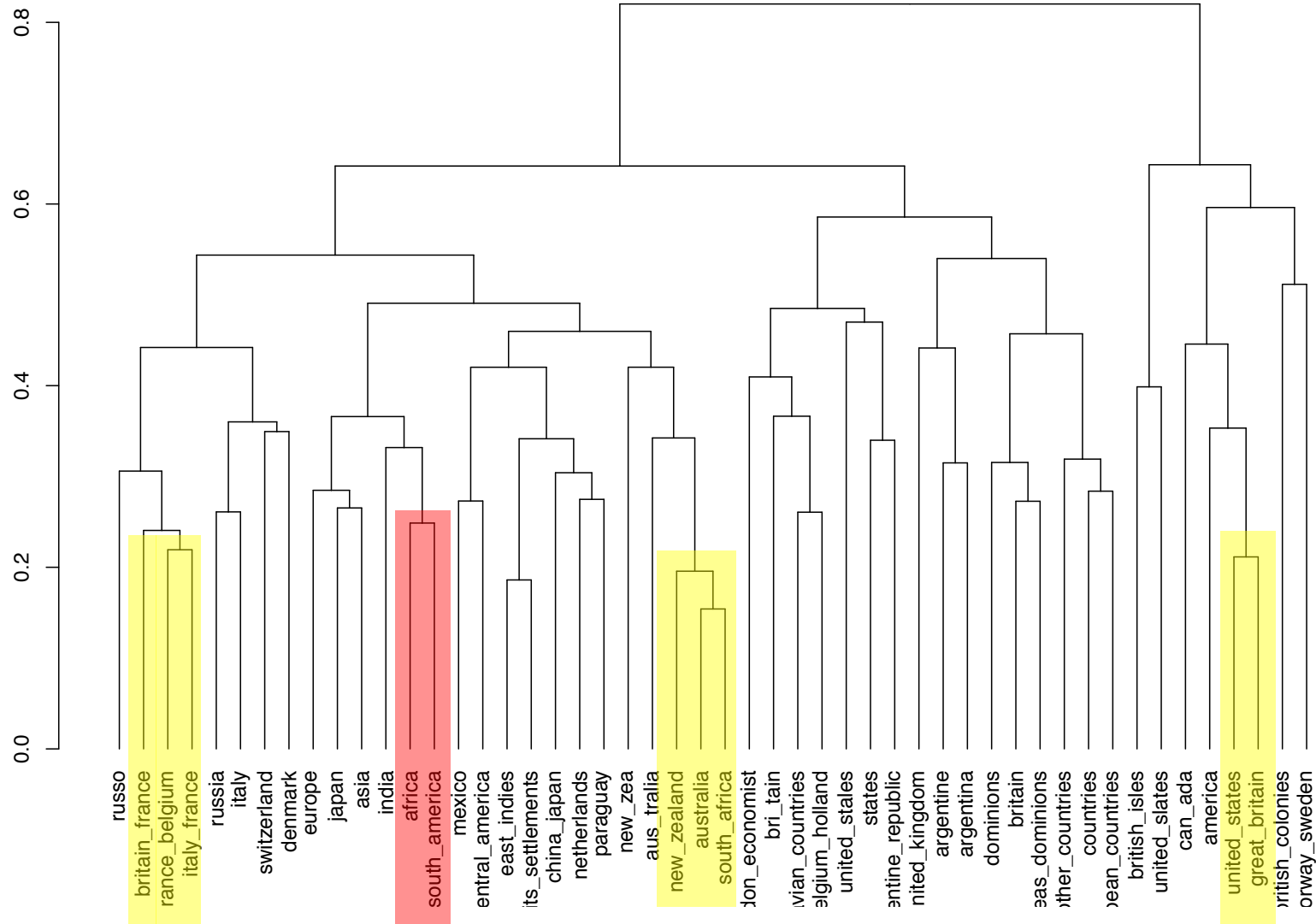
Rank	Word	Similarity to V(<i>united_states, canada, great_britain, united_kingdom, america, british_isles, europe, argentine_republic, australia, mexico, south_africa</i>)
1	<i>great_britain</i>	0.8741361
2	<i>united_states</i>	0.8608055
3	<i>australia</i>	0.8223473
4	<i>south_africa</i>	0.8172986
5	<i>united_kingdom</i>	0.7706215
6	<i>mexico</i>	0.7574097
7	<i>europe</i>	0.7516925
8	<i>belgium_holland</i>	0.7468027
9	<i>america</i>	0.7270632
10	<i>new_zealand</i>	0.7227778
11	<i>argentine_republic</i>	0.7201036
12	<i>india</i>	0.7198081
13	<i>britain</i>	0.7123884
14	<i>countries</i>	0.7104236
15	<i>south_america</i>	0.7091077
16	<i>argentina</i>	0.7042995
17	<i>aus_tralia</i>	0.6959394
18	<i>other_countries</i>	0.6922053
19	<i>states</i>	0.6915724
20	<i>japan</i>	0.6883018

Top 20 words closest to the vectors *united_states, canada, great_britain, united_kingdom, america, british_isles, europe, argentine_republic, australia, mexico, south_africa*.

Cluster dendrogram of the fifty words closest to a placename vector in the Western Home Monthly



Cluster dendrogram of the fifty words closest to a placename vector in the Western Home Monthly

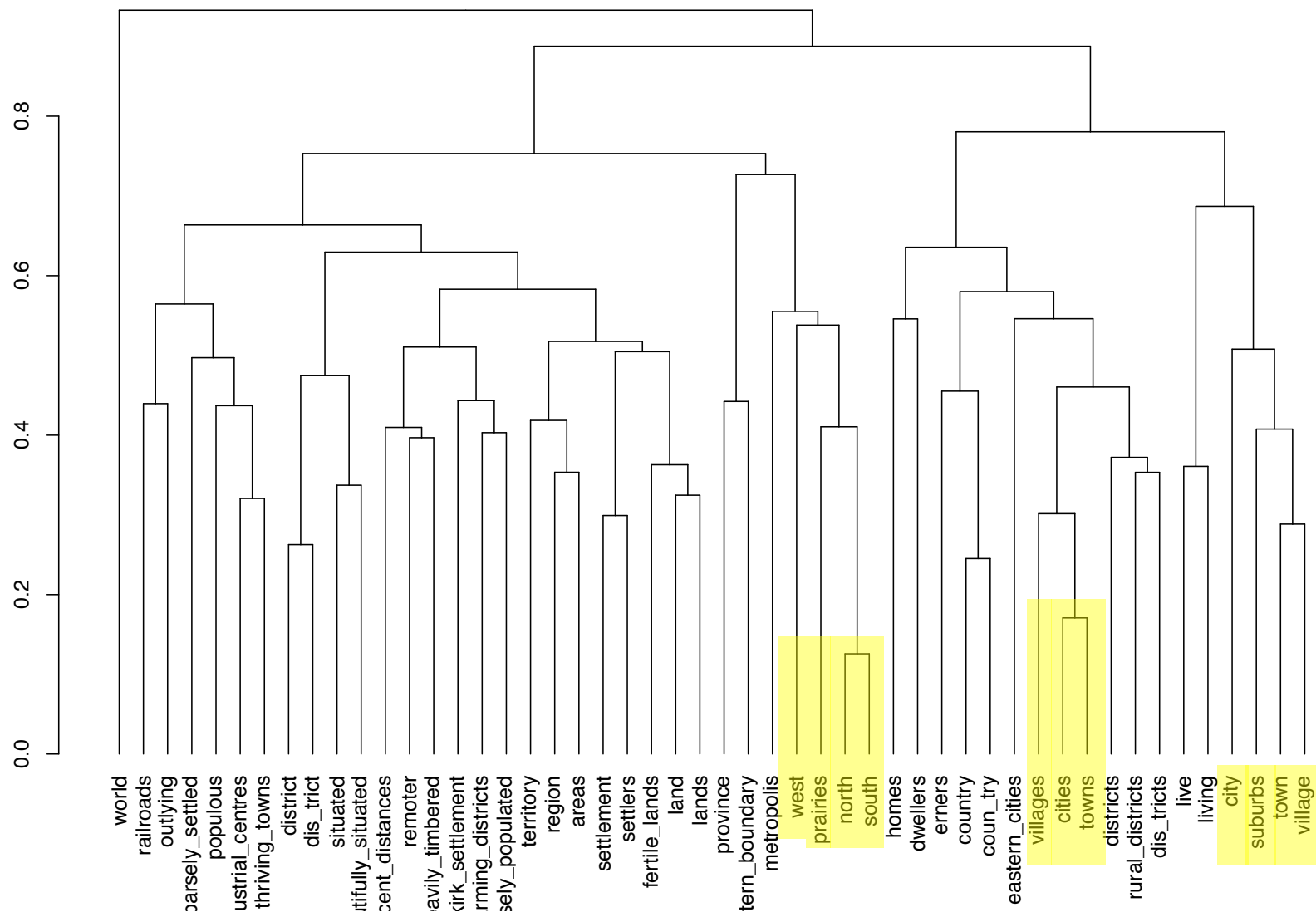


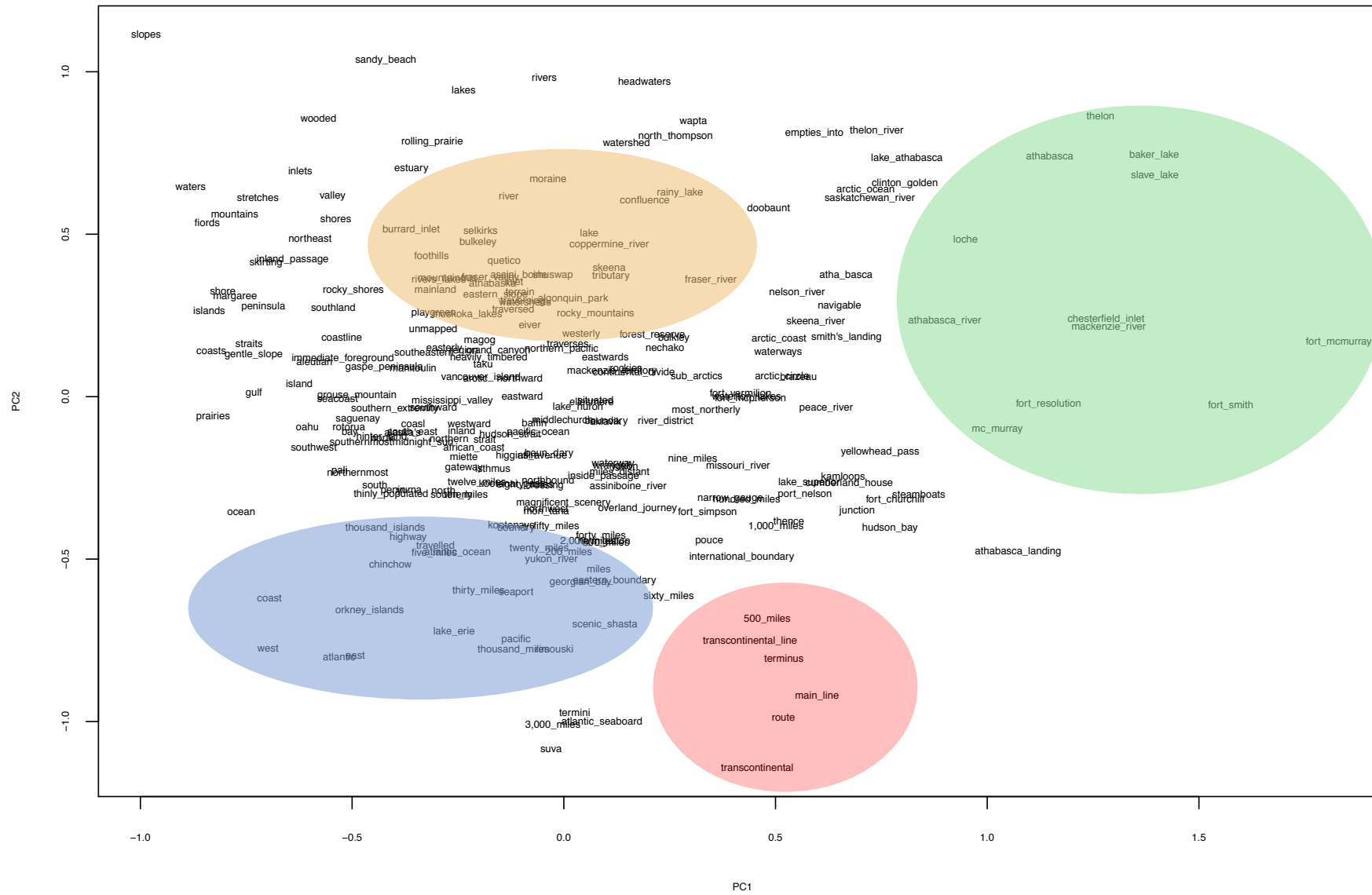
Rank	Word	Similarity to V(<i>land, country, city</i>)
1	<i>country</i>	0.8784156
2	<i>land</i>	0.7943766
3	<i>city</i>	0.7873476
4	<i>coun_try</i>	0.6621502
5	<i>town</i>	0.659282
6	<i>cities</i>	0.6108889
7	<i>districts</i>	0.6034276
8	<i>towns</i>	0.5921221
9	<i>district</i>	0.5771707
10	<i>thriving_towns</i>	0.5736436
11	<i>metropolis</i>	0.5718933
12	<i>dis_tricts</i>	0.5608269
13	<i>prairies</i>	0.5598904
14	<i>west</i>	0.5571857
15	<i>villages</i>	0.5546312
16	<i>fertile_lands</i>	0.5538877
17	<i>rural_districts</i>	0.5455351
18	<i>situated</i>	0.5439044
19	<i>outlying</i>	0.543774
20	<i>eastern_cities</i>	0.5418465

21	<i>farming_districts</i>	0.5410619
22	<i>settlement</i>	0.5392069
23	<i>remoter</i>	0.5364469
24	<i>erners</i>	0.5347513
25	<i>north</i>	0.5335488
26	<i>lands</i>	0.531032
27	<i>south</i>	0.5279061
28	<i>magnificent_distances</i>	0.5272237
29	<i>dis_trict</i>	0.5239724
30	<i>areas</i>	0.5227444
31	<i>densely_populated</i>	0.5226238
32	<i>industrial_centres</i>	0.5192896
33	<i>territory</i>	0.5191257
34	<i>village</i>	0.5179038
35	<i>settlers</i>	0.5164698
36	<i>live</i>	0.5155752
37	<i>suburbs</i>	0.5147471
38	<i>world</i>	0.5128543
39	<i>beautifully_situated</i>	0.5121822
40	<i>homes</i>	0.509831

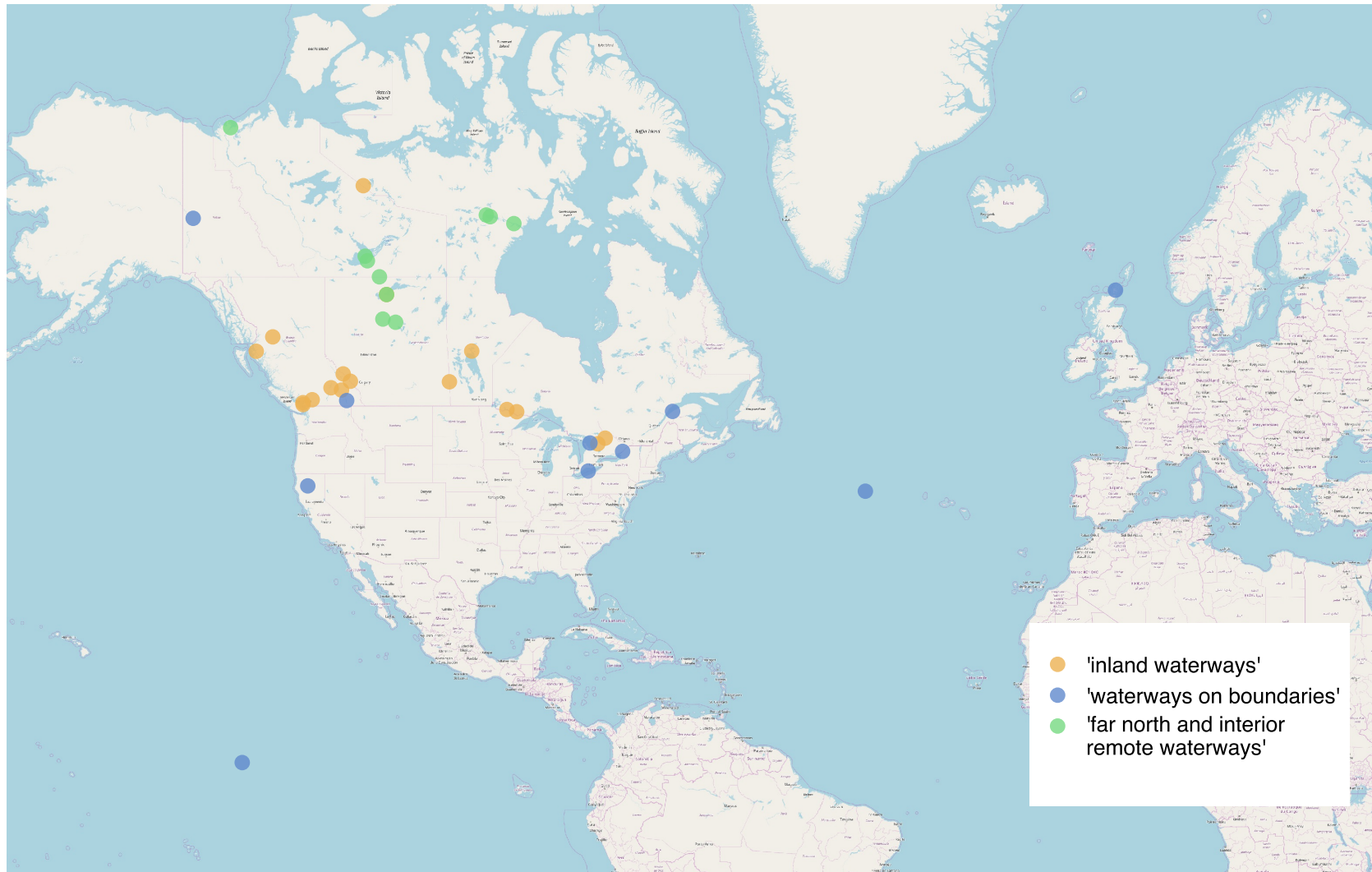
Top 40 words closest to the vectors *land*, *country* and *city*

Cluster dendrogram of the fifty words closest to a 'placename ontology' vector in the Western Home Monthly





Plot showing 250 words closest to the vectors *west*, *river*, *miles*, *north*, *lake*, *east*, *south*, *island*, *coast*, *railway* (list obtained through k-means clustering), reduced to two dimensions using PCA



Map showing distribution of selected placenames from PCA plot of k-means cluster 5. Basemap data © OpenStreetMap contributors and made available under the Open Database Licence: see www.opendatacommons.org/licenses/odbl.

[concepts form] the abstract substrate of language, connecting words in ideational structures that enable thought without necessarily rising to the level of consciousness or explicit expression. Concepts are affiliations among words, sometimes teased out logically but just as often left unspoken, so they demand a different kind of analysis.

Michael Gavin, "Intellectual History and the Computational Turn," *The Eighteenth Century* 58.2 (June 2017): 250.

Acknowledgements

Ben Schmidt, "Rejecting the Gender Binary: A Vector-Space Operation," *Ben's Bookworm Blog*, 30 Oct. 2015, <http://bookworm.benschmidt.org/posts/2015-10-30-rejecting-the-gender-binary.html>.

Ben Schmidt, "Vector Space Models for the Digital Humanities," *Ben's Bookworm Blog*, 25 Oct. 2015, <http://bookworm.benschmidt.org/posts/2015-10-25-Word-Embeddings.html>.

Software

Ben Schmidt and Jian Li, WordVectors, R package, version 2.0.
<http://github.com/bmschmidt/wordVectors>.

Lincoln Mullen and Dimitriy Selivanov, Tokenizers, R package, version 0.1.4. <https://cran.r-project.org/web/packages/tokenizers/>.