



Distributional models and semantic variation across genres

Mariana Montes



RU Quantitative Lexicology and Variational Linguistics

Type level vectors

Church

Semcor, 5L5R

	methodist	steeple	protestant	Sunday	separation	Christ	...	cloud
church	5.59	4.59	3.52	3.31	3.26	2.45	...	0.0

Token level vectors

Semcor, 5L5R

	member	steeple	go	state	separation	school
token1	0	0	0	1	1	0
token2	0	0	0	1	0	0
token3	0	0	0	0	0	1
token4	1	0	0	0	0	0
token5	0	0	1	0	0	0

Second order vectors

Semcor, 5L5R

	protestant	tower	cloud	faculty	sovereign	beyond	join	mission
church	3.52	3.38	0.0	0.0	0.0	2.46	2.57	3.11
member	0.0	0.0	0.0	4.42	0.0	0.0	0.0	1.2
steeple	0.0	6.05	4.93	0.0	0.0	0.0	0.0	0.0
go	1.59	0.0	0.99	0.0	0.0	2.81	1.12	1.29
state	0.0	1.60	0.0	0.0	3.59	0.0	1.08	0.78

Type level vectors

Church

Semcor, 5L5R

	methodist	steeple	protestant	Sunday	separation	Christ	...	cloud
church	5.59	4.59	3.52	3.31	3.26	2.45	...	0.0

Let's visualize it



Observations

- ☁ No agreement between the NMDS (cloud) and the hierarchical clustering (dendrogram)
- ☁ Relatively clear building dimension
- ☁ Different structure in different registers
 - ☁ Preference for a sense
 - ☁ Messier/clearer clouds

Thank you

mariana.montes@kuleuven.be
marianamontes.shinyapps.io/3DClouds/