

Polarization of political speech: a text analysis of Finnish parliamentary speech

Salla Simola
Aalto University

May 3, 2019

Political polarization vs. polarization of political speech

- **Political polarization:** political attitudes diverging towards ideological extremes
- Interest by social scientists e.g. due to associations with **legislative inefficiency** and **political instability**
- (DW-)NOMINATE score based on roll-call/plenary voting used to position political representatives ideologically (Poole and Rosenthal 1984, ...)

Political polarization vs. polarization of political speech

- **Political polarization**: political attitudes diverging towards ideological extremes
- Interest by social scientists e.g. due to associations with **legislative inefficiency** and **political instability**
- (DW-)NOMINATE score based on roll-call/plenary voting used to position political representatives ideologically (Poole and Rosenthal 1984, ...)
- **Polarization in political speech**: language/topics used by representatives from different parties becoming increasingly different
- **Political polarization** and **polarization in political speech** are related but distinct phenomena
 - ▶ Subject to different incentives and constraints

Approach

- **Data:** Records for plenary sessions of the Finnish Parliament, 1907-2017; MP background characteristics
 - ▶ **Preprocessing:** remove punctuation etc., parenthetical additions, Swedish speech, stopwords; stem; tokenize into two-word phrases (bigrams)
 - ▶ **Construct a dictionary** of unique bigrams
 - ▶ 14,620 speeches, 895,086 unique bigrams in 2015

→ Speaker-year level bag-of-bigrams representation
- **Method:**
 - ▶ **Lasso-penalized maximum likelihood** estimation of a Poisson model (approximation to multinomial logit)
 - ▶ Addresses finite-sample bias and facilitates parallelization
- Builds heavily on Gentzkow, Shapiro and Taddy (*Econometrica*, forthcoming), Taddy (2013), and Taddy (2015)

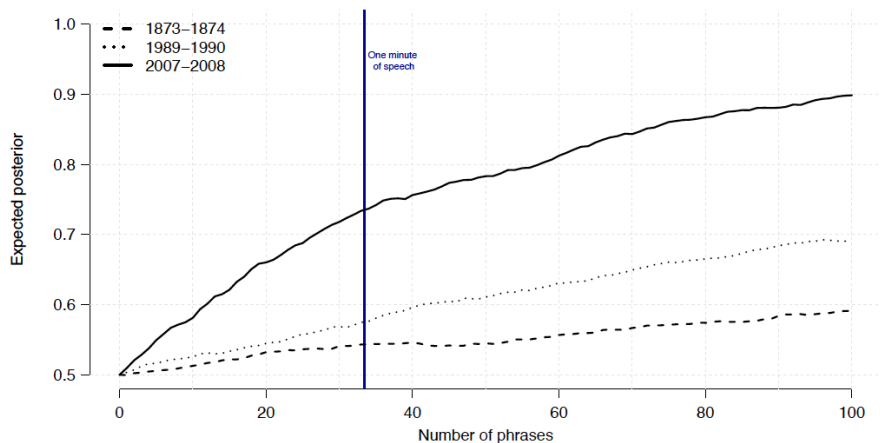
Example: Most spoken bigrams by select Finnish MPs in 2015

Raw counts in parentheses

- **Petteri Orpo**, Minister of Finance, National Coalition party:
'rouva.puhemies' (13) 'harm.taloud' (19) 'sisäis.turvallisuud' (20) 'täl.hetk'
(27) 'arvois.puhemies' (52)
- **Ville Niinistö**, MP, Green Party:
'täl.hetk' (12) 'uusiutuv.energia' (14) 'toise.ast' (19) 'miljoon.euro' (19)
'arvois.puhemies' (133)
- **Jan Vapaavuori**, Minister of Economic Affairs, National Coalition Party:
'täl.hetk' (6) 'viime.käde' (7) 'jalostusarvo.tuotanto' (8)
'korkeam.jalostusarvo' (8) 'arvois.puhemies' (31)

Results from the US: Partisanship of a minute of speech at 73 % in 2007-2008

Figure 3: Informativeness of Speech by Speech Length and Session



Results from the US: Most polarized phrases, US Congress Session 2015-2016

Session 114 (2015-2016)

Republican	#R	#D	Democratic	#R	#D
american peopl	327	205	homeland secur	96	205
al qaeda	50	7	climat chang	23	94
men women	123	83	gun violenc	3	74
side aisl	133	93	african american	11	71
human traffick	60	26	vote right	2	62
colleagu support	123	89	public health	24	83
religi freedom	34	4	depart homeland	48	93
taxpay dollar	47	19	plan parenthood	66	104
mental health	59	32	afford care	40	77
radic islam	22	0	puerto rico	42	79