

Observing political and societal changes in Finnish parliamentary speech data, 1980–2010, with topic modelling

Anna Ristilä & Kimmo Elo

To cite this article: Anna Ristilä & Kimmo Elo (2023): Observing political and societal changes in Finnish parliamentary speech data, 1980–2010, with topic modelling, Parliaments, Estates and Representation, DOI: [10.1080/02606755.2023.2213550](https://doi.org/10.1080/02606755.2023.2213550)

To link to this article: <https://doi.org/10.1080/02606755.2023.2213550>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 22 May 2023.



Submit your article to this journal [↗](#)



Article views: 13





View related articles [↗](#)



View Crossmark data [↗](#)

Observing political and societal changes in Finnish parliamentary speech data, 1980–2010, with topic modelling

Anna Ristilä ^a and Kimmo Elo ^b

^aSchool of Languages and Translation Studies, University of Turku, Turku, Finland; ^bCenter for Parliamentary Studies, University of Turku, Turku, Finland

ABSTRACT

Parliamentary speech reflects many events, changes and developments in society, as well as shaping them by influencing legislation and public interest. Knowing what topics have been dominant in parliamentary discussions can reveal what has been considered important at the time the speech was given. This knowledge can be achieved computationally with topic modelling, which can identify latent topics in large numbers of texts. Currently, the method is still underused in parliamentary studies and has only previously been used once with Finnish parliamentary speeches. This article aims to create and validate a topic model offering a robust overview of Finnish parliamentary speeches from 1980 to 2010, and to demonstrate the validity of the model by examining peaks in topic occurrences and comparing them to the historical and societal context at the times. The topics ‘energy’, ‘employment’ and ‘democracy’ were selected for closer inspection.

KEYWORDS

Plenary debates; Finnish Parliament; topic modelling

Introduction

Topics discussed in parliament reflect both what is considered important in society at the time and how they affect societal discussions. The occurrence of these topics can vary over time and between political parties. Some topics, such as budget debates, government formation and interpellations, emerge periodically and form a kind of parliamentary calendar. Other topics may vary due to historical events and other factors. This article studies changes in topics as a way to observe the development of parliamentary interest during times of those major societal changes that have made Finland increasingly international, for example, the rise and explosive expansion of the Internet, and Finland joining the European Union. The model will be used as a basis for further analysis where compact size, easy interpretability and general robustness are the preferred qualities.

Topic modelling is a method of unsupervised machine learning, a statistical tool for discovering abstract topics in a collection of texts. It takes in a set of texts,

CONTACT Anna Ristilä  ankriss@utu.fi  FI-20014 Turun yliopisto, Finland

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

uses a probabilistic framework to infer relationships between words and assigns words to certain topics, then assigns these topics to the texts. The topics defined by a topic model are abstract, which means that the researcher must give meaningful labels to the topics afterwards, as a model cannot understand the text it has processed. The topics are also often broader and not as fine-grained as topics defined by human readers. The power of topic modelling becomes apparent when the number of texts is so large that it is impossible for a human reader to process everything in a reasonable amount of time. Topic modelling can process millions of words in mere hours.

Studies of parliamentary speeches can reveal the impact they have on our society and help monitor the use of power, consequently there have been many studies from various perspectives. Numerous studies have examined the occurrence and aspects of a single theme or topic, such as illegal drugs, higher education, or the war against ISIS. Topic modelling is a fairly recent development in the context of studying parliamentary speeches. Most international studies have used topic modelling with parliamentary data in English¹ but studies in other languages are also emerging as the digitized materials become available.² Topic modelling has only been used previously on Finnish parliamentary speech by Loukasmäki & Makkonen³ who analyzed how the topics of the government and opposition differ from each other. Other types of quantitative studies on Finnish parliamentary speech have been conducted by Elo,⁴ Kettunen & la Mela,⁵ and Simola.⁶ This scarcity is explained by the fact that Finnish parliamentary speech data has not been previously available in a computer-readable format.

The Finnish parliament, the *Eduskunta*, is a public arena where individual MPs can freely access the floor. Although factions can – depending on the saliency of the issue, limited debating time, or the expertise required for a proper contribution – exercise some level of control, parliamentary factions generally respect their members' right to speak.⁷ A major parliamentary reform in Finland in the 1990s sought to elevate the plenary session to one of the most important forums for national politics and debates

¹J.O. Guldi, 'Parliament's Debates about Infrastructure: An Exercise in Using Dynamic Topic Models to Synthesize Historical Change', *Technology and Culture* 60, (2019), pp. 1–33; B. Curran, K. Higham, E. Ortiz, & D.V. Filho, 'Look Who's Talking: Two-mode Networks as Representations of a Topic Model of New Zealand Parliamentary Speeches', *PLoS One* 13, (2018), p. e0199072; D. Greene & J.P. Cross, 'Exploring the Political Agenda of the European Parliament Using a Dynamic Topic Modeling Approach', *Political Analysis* 25, (2017), pp. 77–94.

²T. Sakamoto & H. Takikawa, 'Cross-National Measurement of Polarization in Political Discourse: Analyzing Floor Debate in the US and the Japanese Legislatures', *Proceeding of 2017 IEEE International Conference on Big Data*, (2017), pp. 3104–110; M. Moilanen & S. Østbye, 'Doublespeak? Sustainability in the Arctic – A Text Mining Analysis of Norwegian Parliamentary Speeches', *Sustainability* 13, (2021), p. 9397; F. Müller-Hansen, M.W. Callaghan, Y.T. Lee, A. Leipprand, C. Flachsland, & J.C. Minx, 'Who Cares About Coal? Analyzing 70 Years of German Parliamentary Debates on Coal with Dynamic Topic Modeling', *Energy Research & Social Science* 72, (2021), p. 101869.

³P. Loukasmäki & K. Makkonen, 'Eduskunnan täysistunnon puheenaiheet 1999–2014 – Miten käsitellä LDA-aihemalleja', *Politiikka* 61, (2019), pp. 127–59.

⁴K. Elo, 'Debates on European Integration in the Finnish Parliament (Eduskunta) 1990–2020', *Proceedings of the Digital Parliamentary Data in Action Workshop on the DHNB2022 Conference*, CEUR Workshop Proceedings 3133, (2022), pp. 129–45.

⁵K. Kettunen & M. la Mela, 'Digging Deeper into the Finnish Parliamentary Protocols: Using a Lexical Semantic Tagger for Studying Meaning Change of Everyman's Rights (*allmansrätten*)', *Proceedings of the Digital Humanities in the Nordic Countries Fifth Conference*, CEUR Workshop Proceedings 2612, (2020), pp. 63–80.

⁶S. Simola, 'A Century of Partisanship in Finnish Political Speech', (2020), <https://sites.google.com/site/sallasimolaecon/home/research>.

⁷C. Poyet & T. Raunio, 'Reconsidering the Electoral Connection of Speeches: The Impact of Electoral Vulnerability on Legislative Speechmaking in a Preferential Voting System', *Legislative Studies Quarterly* 46, (2021), pp. 1087–112. doi:10.1111/lsq.12314.

on topical political matters.⁸ Since then debates on topical matters have been able to be held at short notice, and reports and information the government is constitutionally obliged to provide to the *Eduskunta* have been debated in the plenary.⁹ Since our data covers a period consisting of both pre- and post-reform plenary debates it is interesting to see whether the reform has somehow affected the topic structure.

Although we do not deny the impact and influence of the parliamentary culture and rules on plenary debates, our primary interest is in the ‘politics vocabulary’ used in parliamentary debates to discuss topical matters. Here we follow Palonen’s ideas about ‘politicization’ as a speech act marking a change or even a break in relation to politics and opening up a spectrum of opportunities. The core point here is that no phenomenon can be labelled as political by nature, they become political once someone politicizes, i.e. interprets them as political.¹⁰ Departing from this point we understand plenary debates as speech acts of politicization: taking up a question or matter in a plenary debate politicizes it and opens it for opportunities.

In this article, we seek to answer the following research question: *Is LDA (latent Dirichlet allocation) based topic modelling a reliable research tool to explore themes and dynamics of plenary debates?* From this perspective this article paper has two main objectives. First, it seeks to provide evidence that LDA is capable of producing meaningful and reliable results also when applied to non-English textual corpora. With meaningful we mean a topic structure, in which a majority of the topic labels can be easily determined by the most influential content words. With reliable we mean that most topics can be explained by the main political, societal, and historical developments. Here we lean heavily on a recent study of Ahonen and Koljonen¹¹ exploring transitory and resilient topics in Finnish party manifestos by applying LDA topic modelling.

The article proceeds in three steps. In the first section, we discuss our data and the workflow of data pre-processing, as well as describe general questions related to LDA topic modelling and the model validation. The second section is dedicated to the presentation of our empirical results. This section starts with the presentation of our LDA model, followed by a critical assessment and validation of the model. The latter part of this section presents empirical evidence for the reliability and validity of the model by critically reflecting selected topics against the context of Finnish political history. The article is summarized with concluding remarks focusing on the main findings in a theoretical, methodological, and empirical sense.

Data and methods

Dataset and pre-processing

The Finnish parliamentary speech dataset used in this study has been digitized and made computer readable in the Semantic Parliament project (SEMPARL).¹² The full dataset

⁸See <https://www.eduskunta.fi/EN/naineduskuntatoimii/historia/Pages/default.aspx> (Link visited on 21 March 2023).

⁹Poyet & Raunio, ‘Reconsidering the Electoral Connection’.

¹⁰K. Palonen, ‘Politicisation as a Speech Act: A Repertoire for Analysing Politicisation in Parliamentary Plenary Debates’, in T. Haapala and Á. Oleart (eds), *Tracing the Politicisation of the EU: The Future of Europe Debates Before and After the 2019 Elections* (Cham, 2022), pp. 67–90.

¹¹P. Ahonen & J. Koljonen, ‘Transitory and Resilient Salient Issues in Party Manifestos, Finland, 1880s to 2010s’, *Proceedings of the 10th International Conference on Social Informatics Part I*, (2018), pp. 23–37.

¹²<https://seco.cs.aalto.fi/projects/semparl/en/> (last visited: 16 March 2023).

covers all transcribed speeches given in plenary sessions in the Finnish parliament (years 1907–2021 in full at the time of writing) containing almost 1.2 million speeches and over 200 million words. The 1980–2010 subset of speeches modelled in this paper consisted of around 476,000 speeches and 90 million words but were reduced to around 454,000 speeches and 25 million words during pre-processing. The pre-processing consisted of seven main steps: lemmatization and part-of-speech tagging of all words; removing Swedish language sentences; removing parts of speech other than nouns, verbs and adjectives; removing speeches by the Speaker of Parliament; removing speeches shorter than 5 words; removing words with no letters; removing words with corpus frequency less than 5 or over 57,000. Three additional steps were used to create four slightly differing word-sets or dictionaries, referred to later as d1–d4. These additional steps were: removing words with document frequency of less than 5 or over 57,000; allowing some exception words to remain; removing some additional stop-words common to parliamentary speech.

The main pre-processing steps were based on known best practices. Though the stemming of words has been noted not to improve the results of topic models,¹³ lemmatizing gives a more accurate image of speech¹⁴ especially in highly inflectional languages such as Finnish (for example, a single Finnish noun can theoretically have around 2,200 different inflectional forms). Lemmatizing also decreases the list of words used by the model (model dictionary), the model size and the runtime. In this study the speeches were lemmatized as well as annotated with part-of-speech and dependency relation tags using Turku Neural Parser which works extremely well with the Finnish language.¹⁵ Including only nouns, verbs and adjectives in the model was for the sake of clarity and time constraints. Finland has two official languages, Finnish and Swedish, but since multilingual topic models are still under development,¹⁶ only Finnish sentences were added to the final corpus. Languages were automatically determined with Python's *langdetect* library, which is around 93% accurate.¹⁷ Speeches from the Speaker of Parliament were excluded, since most of them can be considered uninformative. Words that did not contain any letters were also left out. Topic models are generally not good with very short texts, so all speeches shorter than five words were removed. Words that occur rarely or extremely often do not have much impact in the model and can be removed to compactify the model.¹⁸ The recommended way to conduct vocabulary pruning is to remove words

¹³A. Schofield & D. Mimno, 'Comparing Apples to Apples: The Effects of Stemmers on Topic Models', *Transactions of the Association for Computational Linguistics* 4, (2016), pp. 287–300.

¹⁴M. Nelimarkka, 'Aihemallinnus sekä muut ohjaamattomat koneoppimismenetelmät yhteiskuntatieteellisessä tutkimuksessa: Kriittisiä havaintoja', *Politiikka* 61, (2019), pp. 6–33; J. Fiva, O. Nedregård, & H. Øien, 'Polarization in Parliamentary Speech', *CESifo Working Paper* 8818, (2021).

¹⁵J. Kanerva, F. Ginter, & T. Salakoski, 'Universal Lemmatizer: A Sequence to Sequence Model for Lemmatizing Universal Dependencies Treebanks', *Natural Language Engineering* 27, (2020), pp. 545–74; J. Kanerva, F. Ginter, N. Miekka, A. Leino, & T. Salakoski, 'Turku Neural Parser Pipeline: An End-to-End System for the CoNLL 2018 Shared Task', *Proceedings of the CoNLL 2018 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies*, (2018).

¹⁶For example: C.-H. Chang, & S.-Y. Hwang, 'A Word Embedding-based Approach to Cross-lingual Topic Modeling', *Knowledge and Information Systems* 63, (2021), pp. 1529–55; F. Lind, J.-M. Eberl, O. Eisele, T. Heidenreich, S. Galyga, & H.G. Boomgaarden, 'Building the Bridge: Topic Modeling for Comparative Research', *Communication Methods and Measures* 16, (2022), pp. 96–114.

¹⁷M. Kostelac, 'Comparison of Language Identification Models', *Model.Predict*, (2021). <https://modelpredict.com/language-identification-survey#benchmark-results>.

¹⁸M.J. Denny, & A. Spirling, 'Text Preprocessing for Unsupervised Learning: Why it Matters, When it Misleads, and What to do about it', *Political Analysis* 26, (2018), pp. 168–89.

with a document frequency under 0.5% or over 99%,¹⁹ but our method was a little less strict: words with a corpus frequency (how many times a word occurs in the entire collection of documents) of less than five or more than 57,000 were removed.

Topic model: latent Dirichlet allocation (LDA)

Several different kinds of topic models exist and more are continuously being developed. The most famous and widely used model is latent Dirichlet allocation (LDA)²⁰ which was also used in this study. The version used here was from a/the Python library *Gensim*²¹ which uses an online variational Bayes algorithm.²² In practice, a corpus of documents is input into the LDA programme and the result is a model where every document is represented by a distribution of topics, and every topic is represented by a distribution of words.

LDA learns the document-topic and topic-word distributions in an unsupervised manner, but these distributions are guided by two hyper-parameters that can be optimized; in Gensim these are called alpha (for document-topic distribution) and eta (for topic-word distribution). The third and perhaps the most important parameter that needs to be optimized is the number of topics that LDA will create, usually represented by K. Gensim can learn good alpha and eta values automatically from the data, but K always needs to be given. Parameter optimization requires creating several different LDA models and comparing them. Altogether around 200 models were created for this purpose and 12 were chosen for closer inspection.

Validation: coherence scores, close reading and classification tasks

The 12 candidate models were chosen based on coherence scores and the best model was chosen based on topic interpretability. Finally, to make certain that the chosen LDA model was adequate, it was validated with two classification tasks.

The most popular method for computational evaluation of topic models used to be perplexity but since it was found not to correlate with topic interpretability²³ it has been replaced by coherence. Newman et al.²⁴ define topic coherence simply as semantic relatedness, but it can be understood as words appearing in similar contexts. Coherence scores are calculated for individual topics and their average can represent the coherence of an entire topic model. Coherence intrinsically favours a smaller number of large topics

¹⁹D. Maier, A. Niekler, G. Wiedemann, & D. Stoltenberg, 'How Document Sampling and Vocabulary Pruning Affect the Results of Topic Models', *Computational Communication Research* 2, (2020), pp. 139–52.

²⁰D.M. Blei, A.Y. Ng, & M.I. Jordan, 'Latent Dirichlet Allocation', *Journal of Machine Learning Research* 3, (2003), pp. 993–1022.

²¹R. Řehůřek, & P. Sojka, 'Software Framework for Topic Modelling with Large Corpora', *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks* (2010), pp. 45–50.

²²M.D. Hoffman, D.M. Blei, & F. Bach, 'Online Learning for Latent Dirichlet Allocation', *Proceedings of the 23rd International Conference on Neural Information Processing Systems* 1, (2010), pp. 856–64.

²³J. Chang, S. Gerrish, C. Wang, J. Boyd-graber, & D. Blei, 'Reading Tea Leaves: How Humans Interpret Topic Models', in Y. Bengio, D. Schuurmans, J. Lafferty, C. Williams, and A. Culotta (eds), *Advances in Neural Information Processing Systems* 22 (Curran Associates, 2009), <https://proceedings.neurips.cc/paper/2009/file/f92586a25bb3145facd64ab20fd554ff-Paper.pdf>.

²⁴D. Newman, J.H. Lau, K. Grieser, & T. Baldwin, 'Automatic Evaluation of Topic Coherence', *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (2010), pp. 100–8.

over a large number of small topics,²⁵ which suits the aim of acquiring relatively broad topics and a compact model.

According to Gensim's documentation, the formulas and pipeline used in Gensim's coherence model are from Röder et al.²⁶ Gensim can calculate four different coherence scores: C_v , C_{uci} , C_{npmi} and C_{Umass} . As explained by Röder, each coherence score is calculated in four steps: segmentation, probability estimation, confirmation measure, and aggregation.²⁷ Segmentation defines which parts of the texts are compared to each other: C_{uci} , C_{npmi} and C_{Umass} segment and compare single word pairs, C_v compares single words to the entire word set using word context vectors. Probability estimation estimates how probable it is for the words or word-sets to co-occur: C_{Umass} simply calculates the number of documents where the word occurs and divides it by the total number of documents (*Boolean document* method); C_{uci} , C_{npmi} and C_v use the *Boolean sliding window* method where a window of n words moves across the entire text one word at a time and creates n -word pseudo documents, for which the probability is calculated a similar way to full documents. The confirmation measure uses probabilities to compute semantic support for the word pairs and sets. There are two kinds of confirmation measures: direct and indirect. The benefit of the indirect measure is that it detects semantic support between words that do not often occur together (for example, 'skiing' and 'cycling'), but do co-occur with similar word-sets (for example, {'sports', 'equipment', 'speed'}). Direct measures are used in C_{uci} (log-ratio), C_{npmi} (normalized log-ratio) and C_{Umass} (log-conditional) while C_v uses an indirect measure (cosine similarity + normalized pointwise mutual information). Aggregation simply combines the confirmation measures into a single coherence score; all scores use arithmetic mean to do this.

C_v coherence has been found to best correlate with human ratings of topics.²⁸ Because C_v was also used by Loukasmäki and Makkonen²⁹ as well as many others³⁰ it was chosen as the main scoring formula to be used in this paper (Gensim's default window size of 110 was used). However, the original implementation of C_v has been contested³¹ so the other three coherence scores were used to complement the results of C_v .

The highest C_v coherence scores usually represent the best models but, as was mentioned previously, coherence tends to favour small numbers of topics. The dataset

²⁵D. Mimno, H.M. Wallach, E. Talley, M. Leenders, & A. McCallum, 'Optimizing Semantic Coherence in Topic Models', *Proceedings of the Conference on Empirical Methods in Natural Language Processing* (2011), pp. 262–72.

²⁶M. Röder, A. Both, & A. Hinneburg, 'Exploring the Space of Topic Coherence Measures', *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining*, (2015), pp. 399–408.

²⁷Röder et al., 'Topic Coherence Measures'.

²⁸Röder et al., 'Topic Coherence Measures'.

²⁹Loukasmäki & Makkonen, 'Eduskunnan puheenaiheet'.

³⁰Newman et al., 'Automatic Evaluation'; G. Uddin, F. Sabir, Y.-G. Guéhéneuc, O. Alam, & F. Khomh, 'An Empirical Study of IoT Topics in IoT Developer Discussions on Stack Overflow', *Empirical Software Engineering* 26, (2021), p. 121; J. Pozsgai-Alvarez & I. Pastor Sanz, 'Mapping the (Anti-)corruption Field: Key Topics and Changing Trends, 1968–2020', *Journal of Computational Social Science* 4, (2021), pp. 851–81; M. Molavi, M. Tavakoli, & G. Kismihók, 'Extracting Topics from Open Educational Resources', in C. Alario-Hoyos, M.J. Rodríguez-Triana, M. Scheffel, I. Arnedillo-Sánchez, & S.M. Dennerlein (eds), *Addressing Global Challenges and Quality Education* (Springer International Publishing, 2020), pp. 455–60; J. Bennett, B. Rachunok, R. Flage, & R. Nateghi, 'Mapping Climate Discourse to Climate Opinion: An Approach for Augmenting Surveys with Social Media to Enhance Understandings of Climate Opinion in the United States', *Plos One* 16, (2021), pp. 1–16.

³¹M. Röder, *Comment: GitHub/Dice-Group/Palmetto: Not Being Able to Replicate Coherence Scores from Paper – #13*, (2018), <https://github.com/dice-group/Palmetto/issues/13#issuecomment-371553052>.

used in this study is quite large and covers a long period of time, so statistically one would expect to find many topics. However, very fine-grained topics would be unsuitable for the purposes of obtaining an overview of the data and a complex model would make further use challenging. In addition, models with high numbers of topics usually include many split and low-quality topics which would make diachronic changes very difficult to interpret.

Human evaluation is still considered the most reliable method of whether or not topics make sense³² and it sometimes differs from the results of automated metrics such as coherence.³³ Thus, coherence alone could not be used to choose the best model, only to select the candidate models. Each topic in each candidate model was then ranked in quality by close reading the 25 most probable words for each topic. Three simple ranking categories were used: A topic was labelled ‘mixed’ if it contained words clearly belonging to two or more topics, and if two or more topics could clearly be combined into one the topics were labelled as ‘split’. The label ‘junk’ was used if no unifying theme could be found or if the topic was considered too general (for example, a topic about times and time changes).

The model that proportionally had the most ‘good’ topics was chosen for two classification tests. Loukasmäki & Makkonen³⁴ partly used the same data as this study and they reported minor differences in topic choices between government and opposition speakers when examining simple sums of topic probabilities in all speeches. Following this, the first test consisted of classifying the speeches accordingly to whether they were government or opposition based as regards the topic distributions. The second test concerned letting a classifier predict topics for speeches from their vector representations. This time the best classified topics were compared to the topics with the highest coherence scores. The idea was that if the topics were of good quality and coherent, they should also be easier to classify.

For the first classification test an XGBoost³⁵ model was trained with speech topic vectors. XGBoost is a gradient-boosted decision tree (GBDT) machine learning library and is currently one of the most powerful and most used machine learning algorithms. It makes its predictions by making many small and weak tree models and combining their results. The version used in this study was the XGBoost API with Python’s *scikit-learn* interface.³⁶ XGBoost requires considerable parameter tuning, which is time consuming with large models, therefore, for the second test a simpler classification tool was used, namely Random Forest³⁷ also from Python’s *scikit-learn*.

³²Röder et al., ‘Topic Coherence Measures’.

³³A. Hoyle, P. Goel, D. Peskov, A. Hian-Cheong, J. Boyd-Graber, & P. Resnik, ‘Is Automated Topic Model Evaluation Broken?: The Incoherence of Coherence’, *Pre-Proceedings of the 35th Conference on Neural Information Processing Systems*, (2021).

³⁴Loukasmäki & Makkonen, ‘Eduskunnan puheenaiheet’.

³⁵T. Chen & C. Guestrin, ‘XGBoost: A Scalable Tree Boosting System’, *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (2016), pp. 785–94.

³⁶F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, & E. Duchesnay, ‘Scikit-learn: Machine Learning in Python’, *Journal of Machine Learning Research* 12, (2011), pp. 2825–30.

³⁷L. Breiman, ‘Random Forests’, *Machine Learning* 45, (2001), pp. 5–32.

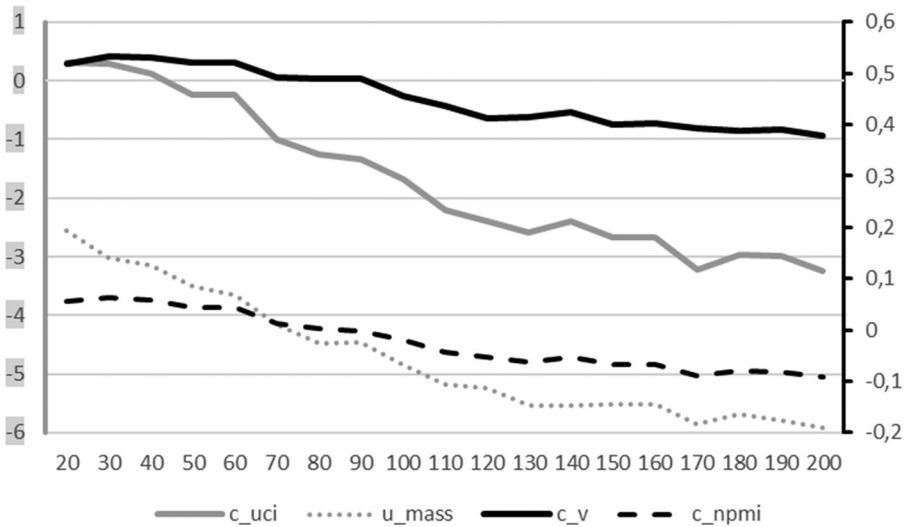


Figure 1. Trends of four different model coherence scores with the number of topics between 20 and 200, using step = 10. Dictionary d5 was used to create the models. The grey lines (C_{uci} and U_{mass}) use the grey Y axis on the left and the black lines (C_v and C_{npmi}) use the black Y axis on the right.

Results

LDA model and validation

The initial tests (shown in Figure 1) showed that models within the range of 20–60 topics had relatively high coherence scores. A point of interest was also found around 140 topics where C_v , C_{npmi} and C_{uci} made a small peak in the graph, but the number of topics was deemed too high for the purposes of this study. We then created 180 models with 15–60 topics using the four dictionaries. Twelve candidate models were selected based on peak coherence scores and their topics were interpreted and labelled as either good, mixed, split or junk. Candidate model details are listed in Table 1. A model with 26 topics made with dictionary d4 (t26d4) was chosen for further analysis as it proportionally contained the highest number of good quality topics (81%): only topics ‘social benefits’ and

Table 1. Selected candidate LDA models, their selection criteria and percentage of quality topics. Each selection criterion score is either the highest (I) or the second highest (II) among models created with the same dictionary.

Model (topic number and dictionary)	C_v	Other criteria	Quality topics
t23_d3	0,57 (I)	–	70%
t25_d2	0,57 (I)	–	68%
t25_d3	0,55	C_{uci} (I): 0,66	76%
t26_d4	0,55 (I)	–	81%
t28_d4	0,55 (II)	–	75%
t30_d3	0,56	C_{npmi} (I): 0,10	53%
t31_d1	0,57 (I)	–	48%
t33_d3	0,57 (II)	C_{npmi} (II): 0,10	52%
t34_d3	0,57 (II)	–	59%
t35_d3	0,57 (II)	–	49%
t38_d2	0,57 (II)	–	55%
t40_d4	0,54	C_{npmi} (I): 0,07	65%

‘social problems’ were considered as two sides of one comprehensive topic ‘social care’, thus labelled as split, and two other topics were considered as junk but interpreted as ‘parliamentary factions’ and ‘general’. The top 10 topic word lists and interpretations of model t26d4 are presented in Appendix 1.

It should be noted that all candidate models – which had the numbers of topics very close to each other – had quite similar topics, just organized and mixed differently. All candidate models had topics that could be interpreted in the following categorisations: ‘taxation’, ‘legislation’, ‘agriculture’, ‘budget’, ‘education’, ‘employment’, ‘energy’ and ‘social and health care’. Some models included some interesting topics that were not present in the chosen model t26d4. For example, several models had topics that could be interpreted as ‘value judgement’ which would be an interesting point of study on its own.

A classical dividing line in parliamentary debates has often been that between the governmental parties and the parliamentary opposition. Here the underlying assumption is that the government vs. opposition polarization in plenary debates reflects the overall party polarization, so that in countries with a strong party polarization the government vs. opposition will also be stronger.³⁸ As the Finnish parliamentary debating culture is rather cautious and polite,³⁹ we expected the debates to reveal no strong polarization between the government and opposition.

Two classification tasks were then carried out in order to validate the chosen topic model. First, document topic vectors were fed into an XGBoost classifier to see if it could differentiate between speeches by the government or the opposition. Speeches with unknown government/opposition status (3.2%) were removed from the data before classification. The resulting classification accuracy for t26d7 with speech topic distributions as data was 0.62, which is somewhat low but can be explained by the long time span which results in ample variation. Usually a low number of topics results in a lower resolution, subsequently the comparison models t40d7 and t140d5, which also had decent coherence scores, were classified in a similar manner to see if a higher number of topics would increase the accuracy. Initially, t40d7 was classified with XGBoost which produced an 0.62 accuracy, which was the same as the chosen 26-topic model. Then all three models were classified with Random Forest, which produced an accuracy score of 0.61 for each model. The consistency suggests that the constricted number of topics should not be an issue in this case. Moreover, the topic differences between the government and opposition observed in Loukasmäki & Makkonen⁴⁰ were not very remarkable, so highly accurate classification could not be expected. More detailed results of the government/opposition XGBoost classifications can be seen in Table 2.

Next, an attempt was made to predict the topics of speeches from the vector representation of the corpus, in which the vocabulary was already pruned and represented as numbers. A Random Forest classifier was used and resulted in an 0.68 accuracy. Some

³⁸F. Wendler. ‘Contesting Europe, or Germany’s Place in Europe? European Integration and the EU Policies of the Grand Coalition Government in the Mirror of Parliamentary Debates in the Bundestag’, *German Politics* 20, (2011), pp. 486–505. doi:10.1080/09644008.2011.606314; K. Auel & T. Raunio. ‘Debating the State of the Union? Comparing Parliamentary Debates on EU Issues in Finland, France, Germany and the United Kingdom’, *Journal of Legislative Studies* 20, (2014), pp. 13–28.

³⁹Poyet & Raunio, ‘Reconsidering the Electoral Connection’.

⁴⁰Loukasmäki & Makkonen, ‘Eduskunnan puheenaiheet’.

Table 2. XGBoost government/opposition classification results of models t26d7 and t40d7.

	t26d7				t40d7			
	Precision	Recall	f1-score	Support	Precision	Recall	f1-score	Support
Government	0.62	0.84	0.72	63227	0.62	0.84	0.72	63227
Opposition	0.59	0.31	0.41	46651	0.59	0.32	0.42	46651
Accuracy			0.62	109878			0.62	109878
Macro avg	0.61	0.58	0.56	109878	0.61	0.58	0.56	109878
Weighted avg	0.61	0.62	0.59	109878	0.61	0.62	0.59	109878

topics were nearly always predicted (f1 score over 0.95): ‘voting’ ‘question time’, ‘law proposals’ and ‘social and health care’. These topics probably consist of vocabulary that are not present in other topics. Other well predicted topics (f1 score 0.67–0.76) were ‘general’, ‘democracy’, ‘development cooperation’ and ‘traffic and transport’. The classifier had problems with the remaining topics. The full classification report is presented in Table 3.

Changes over time were best observed with yearly topic occurrence graphs, which indicate how present each topic was each year. Two slightly different graphs were created for each topic: one illustrating the proportion of speeches where the topic was the most dominant, i.e. the speech was mostly about the topic in question but many other topics could be almost equally present; the other mapping speeches where, in addition to being the most dominant, the topic was at least 30% present or more, i.e. the topic was clearly discussed in the speech. In LDA the document-topic distribution

Table 3. Random Forest classification results for t26d7 topics (most dominant and $\geq 30\%$ present).

	Precision	Recall	f1-score	support
GENERAL	0.62	0.98	0.76	54693
Administration	0.93	0.03	0.06	475
Agriculture	0.89	0.02	0.05	346
Budget	0.77	0.01	0.02	1266
Commerce	1.00	0.00	0.01	331
Crime	0.93	0.09	0.16	435
Democracy	0.92	0.56	0.70	1384
Development cooperation	0.94	0.53	0.68	276
Education	1.00	0.02	0.03	562
Employment	1.00	0.01	0.03	1215
Energy	0.85	0.34	0.49	1446
Foreign and security policy	0.86	0.05	0.09	737
Housing	0.87	0.13	0.22	731
Law proposals	0.95	0.96	0.96	6798
Legislation	0.92	0.05	0.10	6370
Parliamentary factions	0.80	0.03	0.05	2743
Pensions	0.92	0.02	0.04	539
Public sector	0.76	0.31	0.44	20719
question time	0.98	0.97	0.98	3620
Regionality	0.00	0.00	0.00	63
Social and health care	0.98	0.93	0.96	1867
Social benefits	0.95	0.11	0.20	1486
Social problems	1.00	0.02	0.04	349
Taxation	0.98	0.05	0.09	1392
Traffic and transport	0.94	0.52	0.67	480
Voting	0.99	1.00	0.99	3205
Accuracy			0.68	113528
Macro avg	0.87	0.3	0.34	113528
Weighted avg	0.75	0.68	0.61	113528

is controlled by the α hyperparameter, which was in this case automatically adjusted. The chosen LDA model assigned several topics to each speech, which meant that within one speech no topic could be present with a very high percentage. Consequently, some speeches did not have a single topic that was 30% present or more. Using this topic presence threshold resulted in less speeches registered in the yearly topic graphs compared to using no threshold. All topic occurrence graphs with and without the threshold can be seen in Appendix 2.

The ‘general’ topic was the most dominant topic overall (around 30%–55% each year), which was expected. The second most dominant topic overall was ‘public sector’ (12%–27%), and the third was ‘law proposal’ (2%–12%). All other overall topic occurrences remained under 10% each year. Several topics were present less than 1% each year.

Embedding LDA topics to the Finnish contemporary history

Generally speaking the topics identified by LDA were, as the human-decided labels indicate, quite unambiguous and politically meaningful. Further, for most topics it is rather easy to establish connections between the topic and its supposed context. From this perspective, the results from the topic modelling give support to the assumption that topics identified by distant reading are sufficiently reliable as regards describing the main themes of the plenary debates in the Finnish Parliament. Here we strongly rely on the common understanding of plenary speeches as a communicative act aiming at the strengthening of the connection between MPs and citizens, but also as a struggle between different political opinions.⁴¹

In order to better assess the quality of topic modelling we selected three (3) topics for a closer, qualitative analysis. We used the following three (3) criteria for the topic selection:

1. The occurrence of the topic showed a significant variety with clear peaks, thus capturing the ebb and flow typical for parliamentary debates connected to topical issues. In other words, themes come and go in normal society and this is reflected by plenary debates focusing on topical questions.
2. Words with the highest probabilities can be assigned to a certain theme. This means that words used to discuss the topic in plenary sessions can be contextualized.
3. The themes should be relevant for the contemporary history of Finnish politics. Strictly speaking, this is not a computational, but an interpretational criterion. Here the idea was to evidence the usefulness of computational methods for the study of Finnish contemporary history by selecting specific topics considered central to Finnish politics between 1980 and 2010.

The topics selected based on these criteria are topic #2 labelled ‘energy’, topic #7 labelled ‘employment’, and topic #17 labelled ‘democracy’. There were many other topics fitting the above-defined criteria, but we considered these topics well-suited to exemplifying the power of topic modelling with regard to capturing the dynamics of

⁴¹See, for example, K. Palonen. ‘Eduskunnasta puhekunnaksi? Parlamentarismi retorisenä politiikkana’, *Politiikka* 47, (2005), pp. 142–48; L.W. Martin & G. Vanberg. ‘Coalition Government and Political Communication’, *Political Research Quarterly* 61, (2008), pp. 502–16; S.-O. Proksch & J.B. Slapin, *The Politics of Parliamentary Debate: Parties, Rebels and Representation* (Cambridge, 2014).

intensive political debates. The qualitative analysis of each topic follows the same two steps. First, we describe and assess the core vocabulary of the topic by considering it from the perspective of the wider political context. This step is important for demonstrating how well the model captures the main content of plenary debates on the topic in question. Second, we discuss the variety over time by paying strong attention to the peaks. Compared to the step 1, this step focuses on whether the variety identified by LDA is true in the sense that it corresponds to the real intensity of political discourses on the topic. We focus on the peaks on the right-hand side of the occurrence diagrams (Appendix 2) illustrating the annual share of speeches where each speech has at least 30% of its contents allocated to the topic in question. Since many plenary speeches address several issues, we considered this threshold to be the most appropriate.

Energy

The core vocabulary of ten (10) most used words of topic #2, ‘energy’ (see Appendix 1), tackles a wide range of issues not only related to energy production (wood, electricity, nuclear power), but also to issues associated with, for example, climate change (know-how, renewable, emission). Thus, plenary debates on energy questions seem to reflect a re-contextualization of energy debates from purely technical questions to a context framed by global climate change.⁴² Further, these main issues discussed in plenary sessions correspond well with questions related to energy supply in Finland.

An analysis of the intensity of plenary debates on energy issues (see Appendix 2) show clear peaks in the years 1993, 1997, 2002, and 2010. Actually, all of these peaks show a clear connection to global climate politics. The peak in 1993 is most probably linked to the *United Nations Conference on Environment and Development* in Rio de Janeiro in June 1992. This conference, also named the ‘Earth Summit’, was the first global-scale summit on environmental questions and sustainable development.⁴³ The Rio-connection is well exemplified by the speech of MP in Markku Laukkanen (1) representing the Centre Party on 11 March 1993:

(1) ‘Let’s put clear adverse taxes on coal and heavy fuel oil. [So ...] we will also commit ourselves to the Rio Environment Agreement in a genuine way. [...] It is quite clear that once it is signed, we will have to make political actions that show that Finland is genuinely committed to it. A basic energy solution based on coal will certainly destroy Finland’s reputation as a signatory to the Rio Environment Agreement’.

In a similar manner, the ‘Kyoto Protocol’⁴⁴ signed in December 1997 explains the second peak in 1997. The core of the ‘Kyoto Protocol’ consists of market mechanisms based on the trade of emission permits and of emission cuttings below the level of the base year 1990. Both the base year as well the emission cuts were thematised in many plenary speeches as the speech of MP Suvi Lindén (2) (National Coalition Party) on 13 March, 1997, illustrates well:

⁴²K. Elo & J. Karimäki, ‘Luonnonsuojelusta ilmastopoliittikkaan: Ympäristöpoliittisen käsitteistön muutos parlamenttityöskentelyssä 1960–2020’, *Politiikka* 63, (2021), pp. 373–402; I. Massa & R. Sairinen, ‘Suunnistus ympäristösosiologiaan’, in I. Massa and R. Sairinen (eds), *Ympäristökysymys. Ympäristöuhkien haaste yhteiskunnalle* (Helsinki, 1991), p. 68.

⁴³<https://www.un.org/en/conferences/environment/rio1992> (last visited: 28 October, 2022).

⁴⁴https://unfccc.int/kyoto_protocol (last visited: 28 October, 2022).

(2) ‘Restoring emissions to year [19]90 levels cannot be achieved without nuclear power construction, unless the use of coal and peat is completely replaced by natural gas and wood. This is probably not realistic in the time frame set’.

The 2002 peak seems to be connected both to the 2002 United Nations Climate Change Conference in New Delhi in October–November 2002, and the extreme global weather during 2002. Drought conditions in the first half of 2002 were followed by extensive flooding in many locations including Europe (NOAA National Centers for Environmental Information 2003), resulting in intensive debates on global climate change. The plenary discussion revolved around general questions related to climate change and the role of different energy sources in this context. As MP Reijo Kallio (3) (Social Democrats) pointed out on 13 February 2002:

(3) ‘Finland has increased the use of renewable energy sources. [...] But in spite of this, we need coal, natural gas, nuclear power in electricity production. [...] Here some of us have sworn by the blessings of natural gas. [...] Natural gas is certainly a better fuel than coal, but it is fossil and causing emissions’.

The peak in 2010 falls in a period of intensive political debates on global climate change following the global climate summits in Montreal (2005), Bali (2007), and Copenhagen (2009), all of them resulting in a clear intensification of environment-related plenary debates in the Finnish parliament.⁴⁵ The 2010 peak seems also to be connected to the *Foresight Report on Long-term Climate and Energy Policy* adopted by the Finnish government in 2009 to supplement the longer-term ambitions of the *2008 Climate and Energy Strategy*. The Foresight Report is here of central importance, since it ‘reviewed ambitions for sustainable development from a global perspective and outlined possible paths to a low-carbon Finland by 2050’.⁴⁶ One such perspective was raised by MP Johanna Karimäki (4) (Green League) on 11 February 2020, as she advocated for the development of private wind turbines:

(4) ‘In Denmark, for example, small wind turbines installed by individual citizens are supported [by the state]. Their increase is supposed to have positively contributed to the public opinion on wind power’.

Employment

The second topic, topic #7 with the label ‘employment’, is characterized by words typical for a traditional speech on labour force policy. The 10 most important words consists of both labour policy related actors – ‘työn#tekijä’ (employee), ‘työn#antaja’ (employer), ‘työtön’ (unemployed), ‘nuori’ (young person), or ‘työ#voima’ (workforce) – and structural and/or societal concepts, i.e. ‘työ#paikka’ (workplace), ‘työ#elämä’ (work), ‘työttömyys’ (unemployment), ‘työllisyys’ (employment), and ‘palkka’ (pay/salary). Such vocabulary is rather typical and common in labour policy related discourses. Furthermore, the very traditional setting of the core vocabulary might also indicate that the main content of plenary debates on labour policy has remained rather stable and focused on very traditional questions and challenges on labour force policy. This is

⁴⁵Elo & Karimäki, ‘Luonnonsuojelusta ilmastopolitiikkaan’.

⁴⁶IEA, *Energy Policies of IEA Countries: Finland* (2013), <https://www.iea.org/reports/energy-policies-of-iea-countries-finland-2013-review> [accessed 2 February 2022].

quite an interesting observation, especially considering the profound changes the Finnish economy has undergone from 1980 to 2010, including integration to the global economy, the deep recession in the early 1990s, joining the EU in 1995 and the single currency euro in 2002.⁴⁷

Compared to energy-related plenary debates, the timeline of employment-related debates is rather different with one clear peak in 1995–1996. This peak is a distinct aftermath of the great depression in Finland in the early 1990s. The depression was at least partly caused by the extensive de-regulation of the Finnish financial system in the late 1980s and the collapse of Soviet trade in 1991.⁴⁸ During the period from 1992 to 1997 Finland suffered from extremely high unemployment with an average unemployment rate of approx. 17%. The rapid economic collapse demanded extraordinary decisions by the government,⁴⁹ resulting in heated political clashes between the government and opposition in plenary debates. A good example is the speech of MP Esa Lahtela (5) (Social Democrats, government) on 26 April 1995, in which Lahtela demanded clearer facts from the opposition parties about the savings the opposition were insisting on by framing these cuts as an attack on the welfare state:

I would love to have an answer as to where the cuts will now be made. [...] When you talk about local contracting and labour flexibilities, it is worth remembering that an unemployed person is currently the most flexible unit in society.

Overall, considering the centrality of labour politics in debates between the government and opposition, it is no great surprise that the occurrence correlates with the unemployment rate. These concerns were also not only exacerbated by increasing social segregation, a growing share of underprivileged people or social groups being threatened with poverty, but also by government-led crisis management strengthening the role of the executive.⁵⁰ These concerns actually confirm findings from previous studies stressing the correlation between the personal economic situation of citizens and their overall satisfaction with democracy or the political system in general.⁵¹

Democracy

The last topic, topic #17 on ‘democracy’ revolves heavily around constitutional concepts – ‘perustus#laki’ (constitution), ‘perustus#laki#valio#kunta’ (Constitutional Law Committee, often used together with ‘jäsen’ (member)), ‘tasa#valta’ (republic), ‘demokratia’ (democracy), ‘valta’ (power) – and democratic institutions – ‘presidentti’ (president), ‘valtio#neuvosto’ (the [Finnish] Government), ‘vaalit’ (elections). When considered

⁴⁷See e.g. M. Scheinin, *EMU ja Suomen valtiosääntö: yhteiseen rahaan siirtyminen ja Suomen Pankin asema* (Helsinki, 1997); T. Tiilikainen, *Europe and Finland: defining the political identity of Finland in Western Europe* (London, 1997); J. Pekkarinen & K. Kavonius, *Suomi, EU ja maailma: puheenvuoroja talouspolitiikasta* (Helsinki, 2003); M. Kotilainen, *Finland's Experiences and Challenges in the euro zone* (Helsinki, 2006).

⁴⁸J. Tarkka, ‘Ulkoisten tekijöiden merkitys Suomen talouskriisissä’, *Kansantaloudellinen aikakauskirja* 90, (1994), pp. 5–17.

⁴⁹P. Huovinen & H. Piekkola, *Unemployment and Early Retirements of the Aged Workers in Finland*, ETLA Discussion Papers No. 750, (2001).

⁵⁰N. Ellonen & J. Nätti, ‘Job Insecurity and the Unemployment Rate: Micro- and Macro-level Predictors of Perceived Job Insecurity among Finnish Employees 1984–2008’, *Economic and Industrial Democracy* 36, (2015), pp. 51–71.

⁵¹C. Foster & J. Frieden, ‘Crisis of Trust: Socio-economic Determinants of Europeans’ Confidence in Government’, *European Union Politics* 18, (2017), pp. 511–35; M. Matthijs, ‘Integration at What Price? The Erosion of National Democracy in the Euro Periphery’, *Government and Opposition* 52, (2017), pp. 266–94.

from the perspective of the temporal context each of the peaks of 1987, 1994, and 1999 can be relatively easily explained by contextual factors. The elections in 1987 marked a certain turning point in the Finnish political history; this was the first year after the long reign of President Urho Kekkonen and especially the conservatives (6) advocated for a reform of the presidential election system that was based on a direct popular election:

The National Coalition Party pushes for a direct, if necessary, two-stage popular election of the President of the Republic. Such an electoral system would be the most democratic and at the same time the clearest-cut option for the electorate.

The National Coalition Party represented by MP Zyskowitz returned to governmental power in 1987 by forming a pathbreaking coalition government between the conservatives and social democrats under PM Harri Holkeri. Since this coalition was considered unexpected, as a non-socialist coalition was rated more probable, the intervening role of the social democratic President Mauno Koivisto was intensively debated and he was criticized for overstressing his democratic powers. The presidential election reform was at least partly connected to this particular government formation, but also illustrated a longer political development away from a presidential system to a parliamentary democracy.

Plenary debates on democracy during the second peak in 1994 revolved around the approaching referendum on Finland's EU membership. Since it was only the second referendum in Finland after 1917, this questions quite understandably resulted in politically heated discussions; it was also used to politicize questions of representative and direct democracy in Finland, as the speech of MP Heidi Hautala (7) (Green League) on 17 May 1994 exemplifies well:

(7) 'I propose that Finland should, in addition to consultative referendums, start organizing binding referenda as well. [...] Direct democracy is very undeveloped in Finland. The public, after all, has extremely little political influence. Our constitution is highly representative'.

The last peak in 1999 is clearly connected to constitutional reform in Finland. The constitutional reform process was started in 1995, with the expert committee completing its work in 1997, and the Constitutional Law Committee of the Finnish *eduskunta* considering the report in 1998. The Constitutional Law Committee submitted its unanimous report on the bill to parliament in January 1999, followed by a plenary hearing in February 1999 by the outgoing parliament and the final approval by the newly elected parliament in June 1999. The new constitution entered into force on 1 March 2000. Interestingly, debates during this peak continued – as the speech of MP Johannes Leppänen (8) (Centre Party) on 1 June 1999 exemplifies – the earlier plenary debates on presidential elections in 1987 and on the role of referenda and direct democracy in Finland:

(8) 'The basis of our Constitution is representative democracy. [...] New, timely ways to participate must be developed in order to support the representative democracy. Without a functioning civil society, no democracy will prosper'.

Overall, the empirical analysis of the three selected topics establishes evidence for the reliability of the LDA results. Results from the content analysis of randomly selected original plenary speeches allocated to the topics confirm that the labels used to describe the

topics' focus are well in line with the true content of speeches. Further, the model also quite reliably captures the main debates over time. The latter is an important observation, as the model should work with highly volatile and dynamic political discussions. Interestingly, the LDA results do not indicate any significant change possibly related to the major constitutional reform in Finland in 1990s. We do not question the impact of this reform on the role of plenary debates in national politics, however, our results do not indicate any significant change regarding the intensity of plenary debates. In other words, even prior to the reform MPs seem to have considered plenary debates as one forum in which to politicize topical matters.

Further, the empirical results offer support for the interpretation that the model used in this paper is able to identify core vocabularies that are robust over time and to use these as a basis for topic allocation. At the same time, the model is not 'misguided' by context-related vocabularies as the core vocabulary is deeply embedded and, thus, can help the researcher to capture discursive dynamics as well. As we have shown above, the variation of topic occurrence reported as a part of the modelling results is a valuable help to discerning connections between a topic and the underlying political, societal, and historical developments.

Conclusions

In this article we examined the applicability of LDA topic modelling on large parliamentary corpora consisting of three decades of plenary debates (1980–2010) in the Finnish *Eduskunta*. The article's main standpoint is rooted in Kari Palonen's concept of politicization as a speech act to label a question or matter as political, so that it becomes possible to break up with present politics and open up new opportunities. Departing from this understanding our article also considers topics discussed in plenary debates as those having been politicized by MPs, thus reflecting the changing landscape of topical matters.

Our results are significant in three respects. First, the LDA model used to identify topics and their dynamics over time appears to produce reliable, coherent, and analytically meaningful results. It is necessary here to note that the time span is rather long and therefore also subject to natural changes in vocabulary as well as the use of language. The best model of 120 models created was chosen by comparing coherence scores and the interpretability of topics. Since the dataset was new, there was no generally accepted baseline with which to compare the model and because coherence scores depend on context⁵² comparing coherence scores from studies using different datasets was not feasible. Thus, the model was validated by checking that the resulting topic distributions could function as a basis for classification. The results of the topic classification task verified that the topics were reliable and discrete. Although this might be considered a somewhat technical justification, finding a basis for appropriate and meaningful labels fulfils the important task of establishing linkages between the vocabulary of politics used in plenary debates and topic-driver discourses. The fact that such connections could be established

⁵²S. Kapadia, 'Evaluate Topic Models: Latent Dirichlet Allocation (LDA) – A Step-by-step Guide to Building Interpretable Topic Models', *Towards Data Science*, (2019), <https://towardsdatascience.com/evaluate-topic-model-in-python-latent-dirichlet-allocation-lda-7d57484bb5d0>.

based on the results of our LDA topic modelling provides at least one important piece of evidence for the applicability of computational methods on parliamentary speeches.

Second, the relatively low accuracy of the government/opposition classification task is a good reflection of the nature of political discourse in Finland, since compared to many other countries, Finland has relatively low levels of political polarization. This can be attributed to a number of factors, including a strong tradition of consensus-based decision making, a relatively homogeneous population, and a multi-party system that encourages cooperation and compromise. In other words, our findings are not only in line with those from previous non-computational studies focusing on Finnish parliamentary debates, but can also be satisfactorily explained by theories on the Finnish parliamentary system.

Third, the empirical, qualitative analysis of the three selected topics – energy, employment, and democracy – give support to the reliability of the results from LDA modelling. The study shows that topic modelling is a valid tool for finding themes in political discussions and also examining the timelines of the appearances of these themes over a long period of time. Even though the peaks of the topics discussed might already be familiar and well-known to the general populace does not nullify our results. On the contrary, the results presented in this paper are encouraging as they indicate that topic modelling is a reliable method to explore large parliamentary document corpora in order to gain understanding of the ebb and flow of political themes in parliamentary debates. Moreover, LDA seem to be a powerful tool for identifying which topics have been topical in a certain period of time, and also for gaining an understanding of the dynamics of politics vocabulary over time. In other words, LDA seems to be a reliable tool to analyse what topics have been politicized and at what period of time.

Although the results from our unsupervised, computational analysis of plenary debates are encouraging in many respects, we consider it methodologically problematic to completely rely on results produced by computational methods. As an alternative, we encourage and stress the indispensable importance of subject-related knowledge for a proper empirical assessment of the results of a computational, unsupervised analysis. A fruitful dialogue between computational methods, empirical analysis, and political theories is in our opinion the best combination when utilizing the analytical power of computational social sciences in the field of parliamentary studies.

Funding

This article presents research results obtained in the ‘Semantic Parliament: Linked Open Data Service for Studying Political Culture’ research consortium financed by a research grant from the Academy of Finland [grant number: 329969, funding period: 2020–22] and in ‘Research infrastructure for the intersection of law and politics’ infrastructure project funded by the Academy of Finland [grant number: 353569, funding period: 2023–25].

Notes on contributors

Anna Ristilä is a PhD student in Digital Linguistics at the School of Languages and Translation Studies at the University of Turku (UTU). Her dissertation focuses on Finnish parliamentary speeches and explores the underlying topic landscape and its changes over time.

Kimmo Elo, Adjunct Professor, is a senior researcher at the Centre for Parliamentary Studies at the University of Turku (UTU). His research interests include German politics and history since 1945,

theories and politics of European integration, Cold War and post Cold War intelligence, digital parliament studies, theories and methods of network analysis and computational social sciences, as well as knowledge visualization techniques.

ORCID

Anna Ristilä  <http://orcid.org/0000-0003-0236-5401>

Kimmo Elo  <http://orcid.org/0000-0002-3223-5221>

Appendices

Appendix 1

Table A1. Top 10 most relevant words and their translations for all topics in model t26d4.

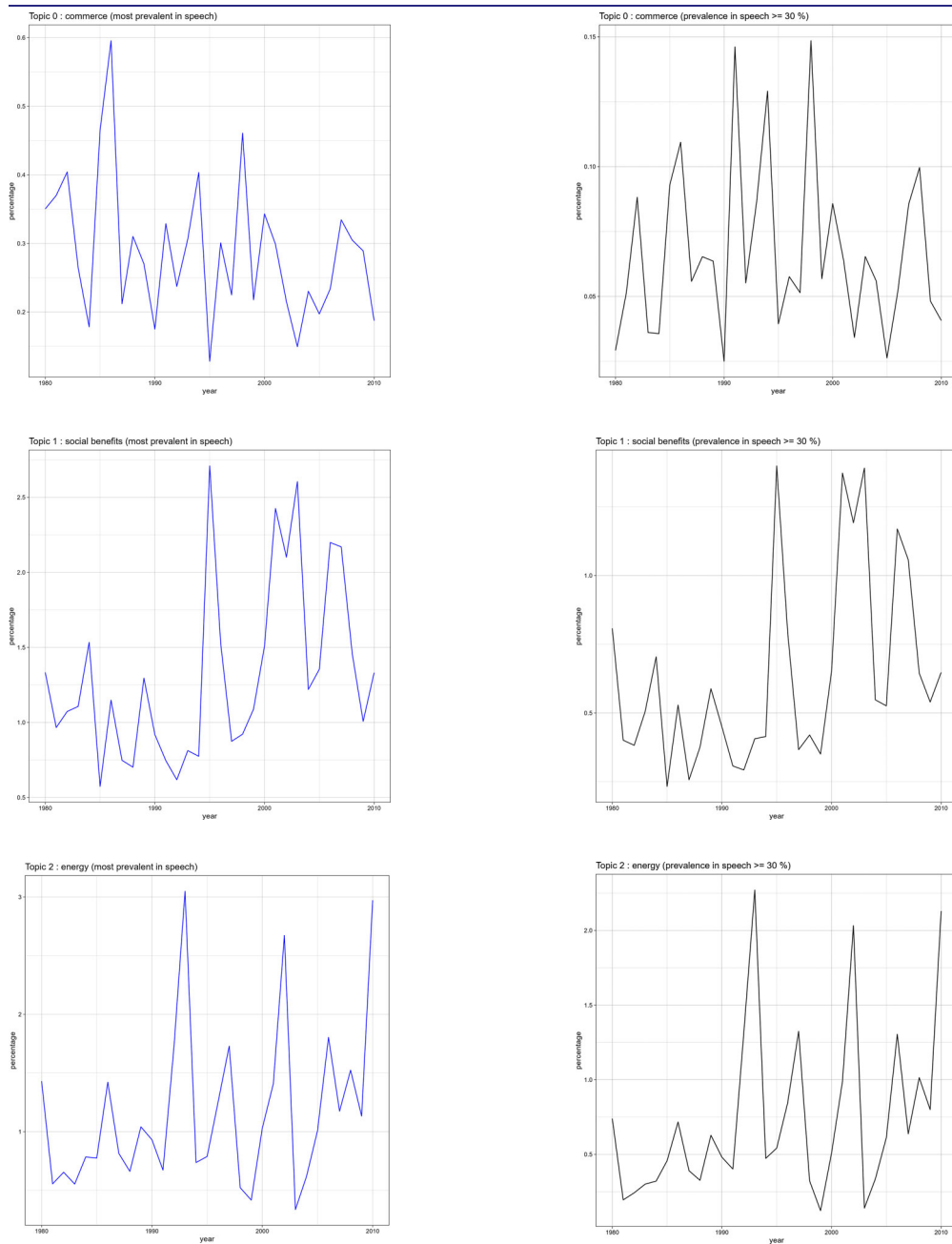
Interpretation	Topic #	Top 10 words
Commerce	0	kauppa, yhtiö, yrittäjä, kilpailu, myydä, markkinat, tuote, ostaa, kuluttaja, ulko#maa 'commerce', 'company', 'entrepreneur', 'competition', 'to sell', 'the market', 'product', 'to buy', 'consumer', 'foreign country'
Social benefits	1	nuori, perhe, tuki, vanhempi, opiskelija, vanhus, lapsi#perhe, elämä, koti, kuntoutus 'young', 'family', 'support', 'parent', 'student', 'elderly person', 'family with children', 'life', 'home', 'rehabilitation'
Energy	2	energia, käyttö, sähkö, uusiutuva, auto, puu, osaaminen, ydin#voima, päästö, metsä 'energy', 'use', 'electricity', 'renewable', 'car', 'wood', 'know-how', 'nuclear energy', 'emission', 'forest'
Regionality	3	maa#kunta, harmaa, kaupunki, asukas, alueellinen, seutu, elin#tarvike, pohjoinen, asua, alue#politiikka 'county', 'grey', 'city', 'inhabitant', 'regional', 'region', 'foodstuff', 'northern', 'to live', 'regional policy'
Education	4	yli#opisto, koulutus, koulu, opetus, kieli, ammatillinen, opiskelija, opettaja, oppilas, nuori 'university', 'education', 'school', 'teaching', 'language', 'vocational', 'student', 'teacher', 'student', 'young person'
Pensions	5	eläke, laki#aloite, aloite, lääke, kansan#eläke, henkilö, kuukausi, korvaus epä#kohta, korjata 'pension', 'legislative proposal', 'initiative', 'medicine', 'social insurance', 'person', 'month', 'reimbursement', 'defect', 'to fix'
Crime	6	poliisi, tieto, rikos, valvonta, viran#omainen, henkilö, laki#valio#kunta, kertomus, oikeus#ministeriö, tuomio#istuun 'police', 'knowledge', 'crime', 'supervision', 'public authority', 'person', 'Standing Committee on Civil-Law Legislation', 'story', 'Ministry of Justice', 'court'
Employment	7	työn#tekijä, työ#paikka, työtön, työn#antaja, työ#elämä, työttömyys, palkka, työ#voima, nuori, työllisyys 'employee', 'workplace', 'unemployed', 'employee', 'working life', 'unemployment', 'pay', 'workforce', 'young person', 'employment'
Legislation	8	lain#säädäntö, käytäntö, peruste, oikeus, laki#esitys, säätää, tarkoitus, käyttö, muuttaa, lausunto 'legislation', 'convention', 'basis', 'justice', 'legislative proposal', 'to degree', 'meaning', 'use', 'to change', 'statement'
Traffic and transport	9	liikenne, määrä#raha, hanke, tie, joukko#liikenne, lisäys, päivä#järjestys, kunto, rata, pää#luokka 'traffic', 'budget', 'project', 'road', 'public transport', 'increase', 'agenda', 'condition', 'track', 'section'
Question time	10	pohja, talous#valio#kunta, yksityis#kohtainen, valtio#varain#valio#kunta, n:o, yleis#keskustelu, laki#ehdotus, sallia, ilmoittaa, debatti 'base', 'Commerce Committee', 'detailed', 'Finance Committee', 'number', 'general discussion', 'legislative proposal', 'to allow', 'to report', 'debate'

(Continued)

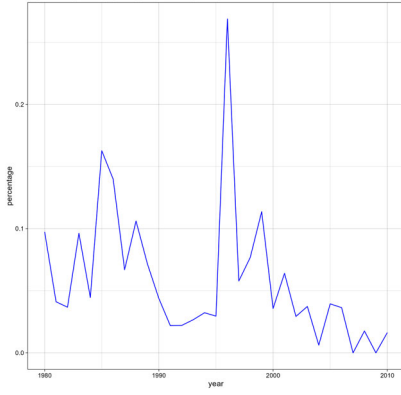
Table A1. Continued.

Interpretation	Topic #	Top 10 words
Administration	11	tehtävä, ministeriö, henkilö, siirtää, hallinto, henkilöstö, viesti, hoitaa, perustaa, järjestää 'task', 'ministry', 'person', 'to move', 'administration', 'personnel', 'message', 'to take care of', 'to establish', 'to arrange'
Public sector	12	palvelu, tulevaisuus, merkittävä, taso, toimen#pide, tarve, jatko, toimi, julkinen, toteuttaa 'service', 'future', 'significant', 'level', 'operation', 'need', 'continuation', 'office', 'public', 'execute'
Foreign and security policy	13	sopimus, kansain#välinen, kansallinen, yhteinen, turvallisuus, linjaus, toimija, selonteo*, selonte*, maailma 'agreement', 'international', 'national', 'mutual', 'security', 'definition of policy', 'operator', 'report*', 'report*', 'the world'
Parliamentary factions	14	kokoomus, linja, pää#ministeri, keskusta, vastuu sosiali#demokraatti, politiikka, edus#kunta#ryhmä, poliittinen, ajaa 'National Coalition Party', 'policy', 'Prime Minister', 'Center Party', 'responsibility', 'Social Democratic Party', 'politics', 'parliamentary group', 'political', 'to drive'
Voting	15	lausuma#ehdotus, äänestys, tyhjä, kone, nuoriso#työ, ehdokas#lista, urheilu, vaali, veikkaus#voitto#vara, maksimi 'draft declaration', 'voting', 'empty', 'machine', 'youth work', 'candidate list', 'sports', 'election', 'profits from Veikkaus', 'maximum'
Law proposals	16	ehdottaa, vasta#lause, päättää, laki#ehdotus, hylätä, ensimmäinen, äänestää, mukainen, sisältö, pykälä 'to suggest', 'objection', 'to decide', 'legislative proposal', 'to discard', 'the first', 'to vote', 'consistent with', 'content', 'section'
Democracy	17	perustus#laki, presidentti, perustus#laki#valio#kunta, jäsen, valtio#neuvosto, valita, vaalit, tasa#valta, demokratia, valta 'constitution', 'president', 'Constitutional Law Committee', 'member', 'Government', 'to choose', 'election', 'republic', 'democracy', 'power'
Development cooperation	18	neuvosto, kulttuuri, jäsen#maa, yhteis#työ, järjestö, ase, kansain#välinen, kokous, pohjois#mainen, kehitys#yhteis#työ 'council', 'culture', 'member state', 'cooperation', 'organization', 'weapon', 'international', 'meeting', 'Nordic', 'development cooperation'
Agriculture	19	maa#seutu, maa#talous, ruoka, tuki, tila, viljelijä, tuotanto, maa#talous#politiikka, metsä#talous, pelto 'rural area', 'agriculture', 'food', 'support', 'estate', 'farmer', 'production', 'agricultural policy', 'forestry', 'field'
Social and health care	20	sosiaali, terveyden#huolto, palvelu, hoito, lähettää, lääkäri, ehdottaa, potilas, puhe#mies#neuvosto, terveys 'social', 'health care', 'service', 'care', 'to send', 'doctor', 'to suggest', 'patient', 'Speaker of Parliament Bureau', 'health'
Taxation	21	verotus, vero, korotus, pieni#tuloinen, eläkeläinen, maksu, tulo, korottaa, nostaa, indeksi 'taxation', 'tax', 'increase', 'person with low income', 'pensioner', 'payment', 'income', 'to increase', 'to raise', 'index'
JUNK	22	kysyä, puoli, ymmärtää, päästä, kertoa, paikka, uskoa, iso, ajatella, viedä 'to ask', 'half', 'to understand', 'to get to', 'to tell', 'place', 'believe', 'large', 'to think', 'to bring'
Housing	23	pankki, asunto, laina, kiinteistö, innovaatio, asuminen, hinta, korko, rakentaminen, asua 'bank', 'apartment', 'loan', 'estate', 'innovation', 'living', 'price', 'interest', 'construction', 'live'
Social problems	24	nainen, mies, alkoholi, nuori, tasa#arvo, väki#valta, isä, suku#puoli, liitto, elämä 'woman', 'man', 'alcohol', 'young person', 'equality', 'violence', 'father', 'gender', 'union', 'life'
Budget	25	talous, budjetti, ensi, kasvu, meno, kasvaa, velka, työllisyys, lama, valtion#talous 'economy', 'budget', 'next', 'growth', 'expense', 'to grow', 'debt', 'employment', 'economic depression'

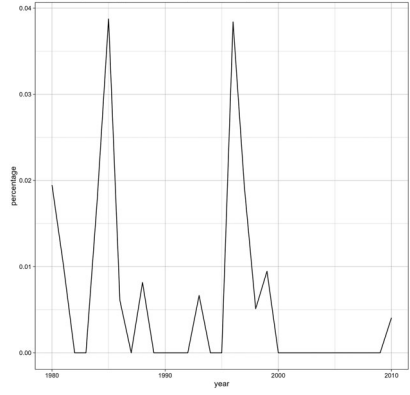
Appendix 2. Topic occurrences over time



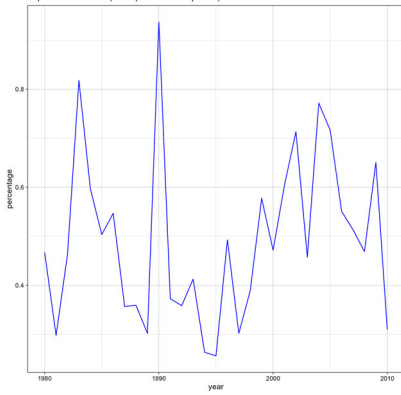
Topic 3 : regionality (most prevalent in speech)



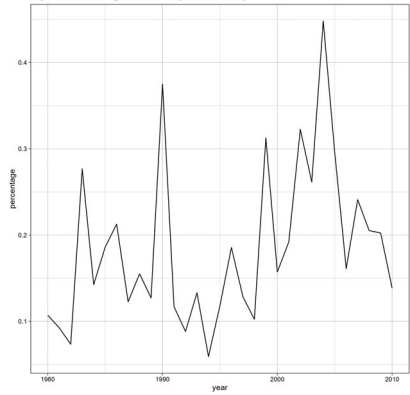
Topic 3 : regionality (prevalence in speech >= 30 %)



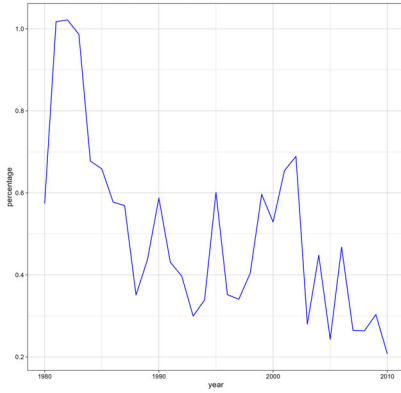
Topic 4 : education (most prevalent in speech)



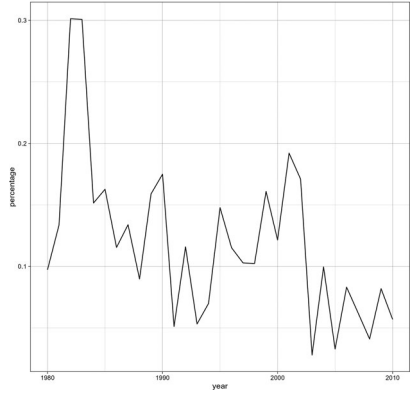
Topic 4 : education (prevalence in speech >= 30 %)



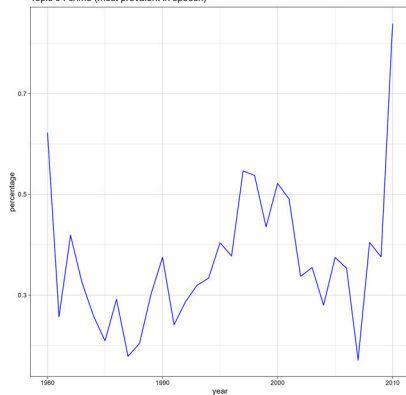
Topic 5 : pensions (most prevalent in speech)



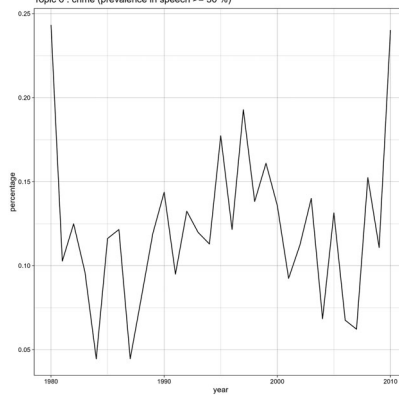
Topic 5 : pensions (prevalence in speech >= 30 %)



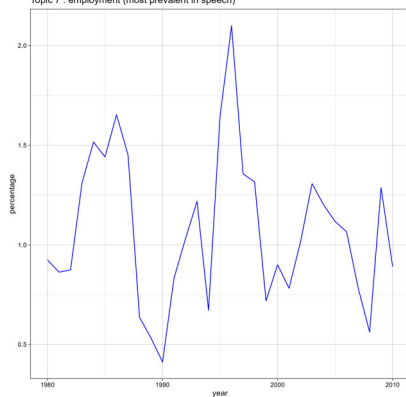
Topic 6 : crime (most prevalent in speech)



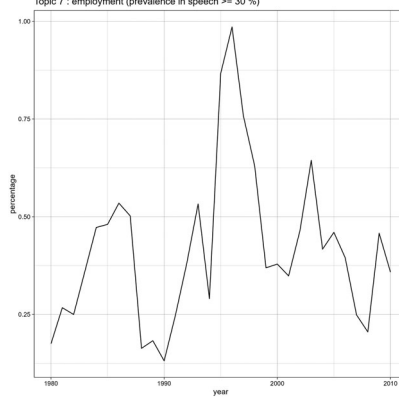
Topic 6 : crime (prevalence in speech >= 30 %)



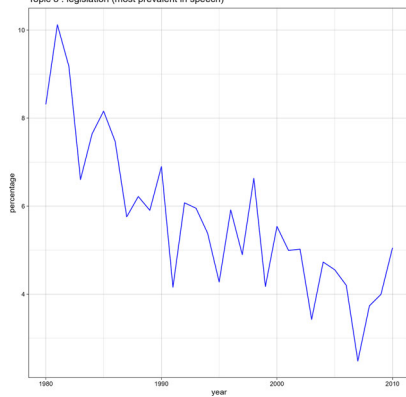
Topic 7 : employment (most prevalent in speech)



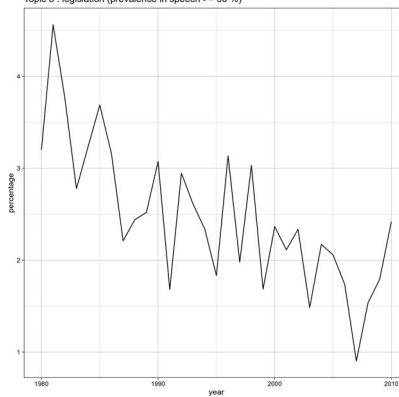
Topic 7 : employment (prevalence in speech >= 30 %)



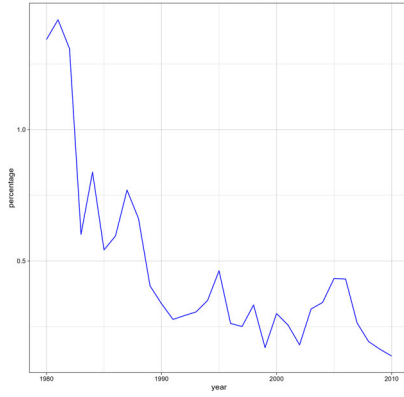
Topic 8 : legislation (most prevalent in speech)



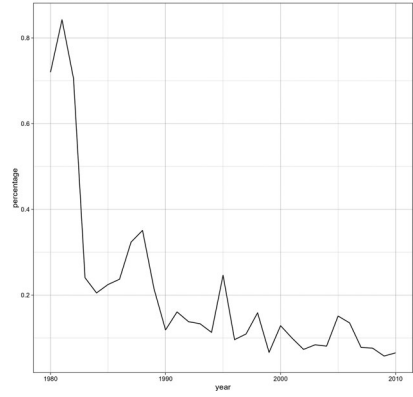
Topic 8 : legislation (prevalence in speech >= 30 %)



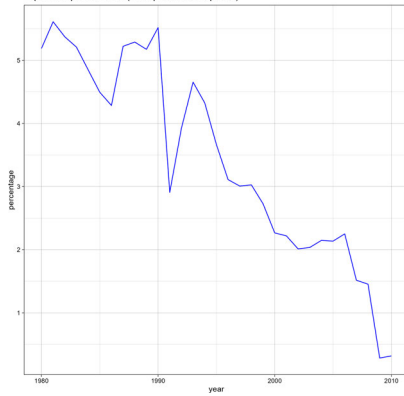
Topic 9 : traffic and transport (most prevalent in speech)



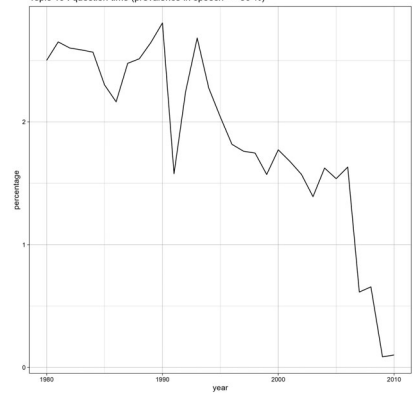
Topic 9 : traffic and transport (prevalence in speech >= 30 %)



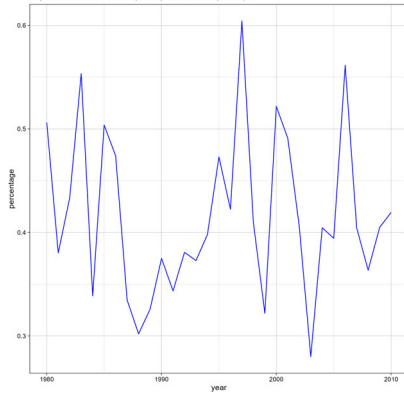
Topic 10 : question time (most prevalent in speech)



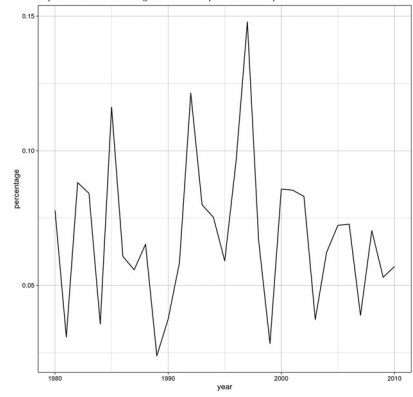
Topic 10 : question time (prevalence in speech >= 30 %)



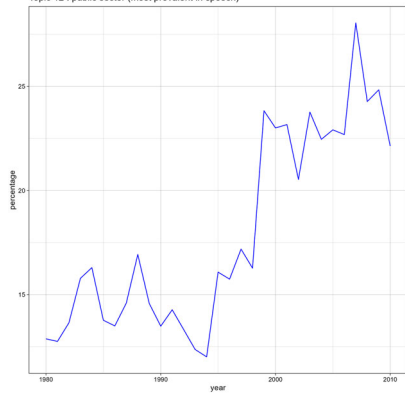
Topic 11 : administration (most prevalent in speech)



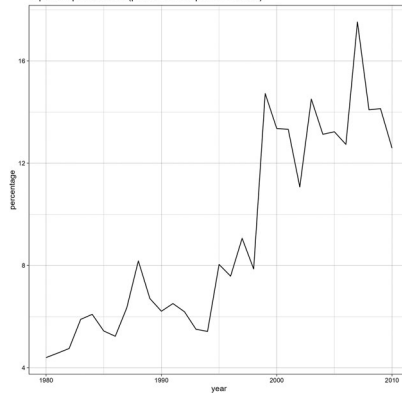
Topic 11 : administration (prevalence in speech >= 30 %)



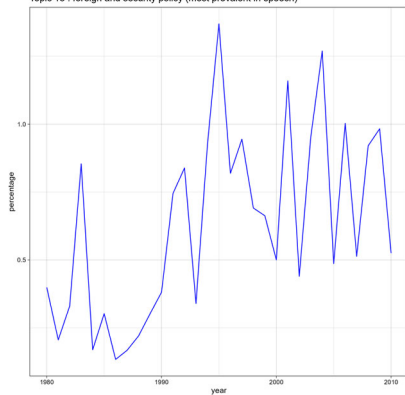
Topic 12 : public sector (most prevalent in speech)



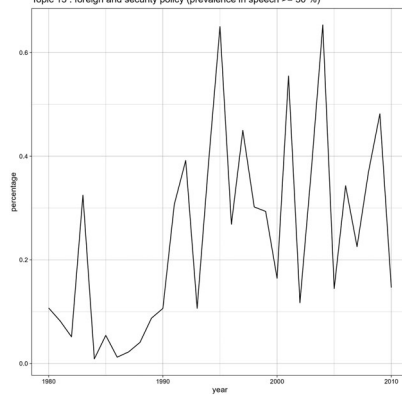
Topic 12 : public sector (prevalence in speech >= 30 %)



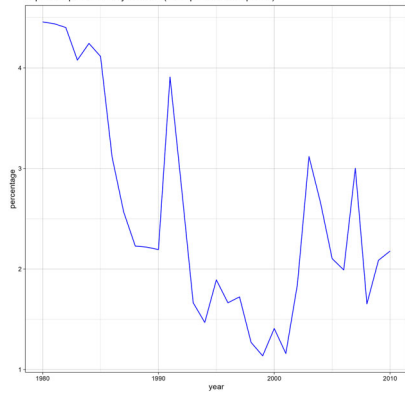
Topic 13 : foreign and security policy (most prevalent in speech)



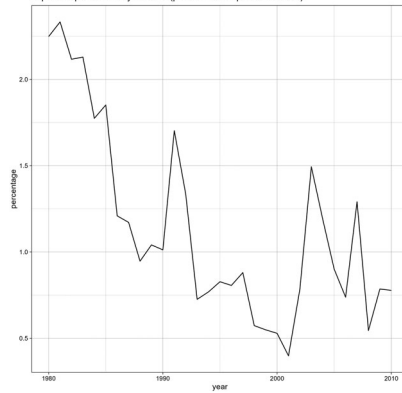
Topic 13 : foreign and security policy (prevalence in speech >= 30 %)

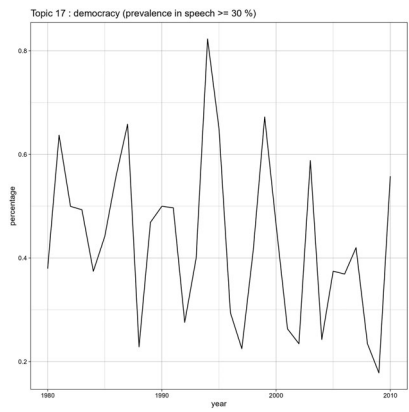
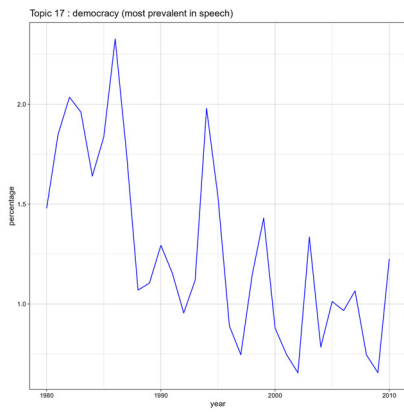
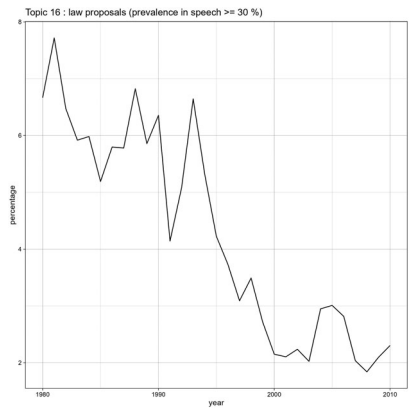
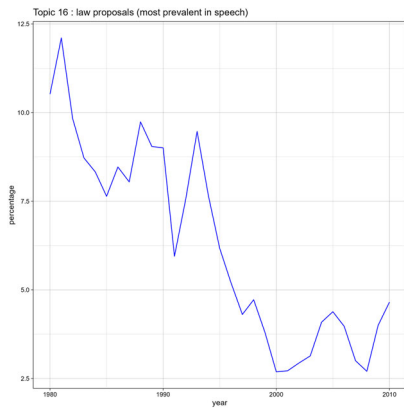
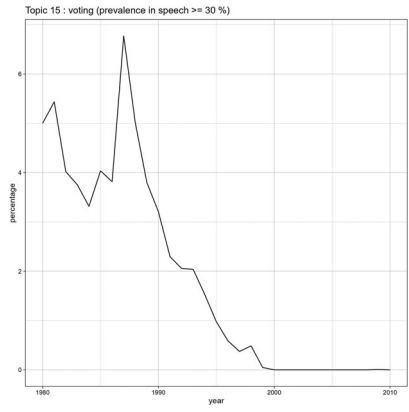
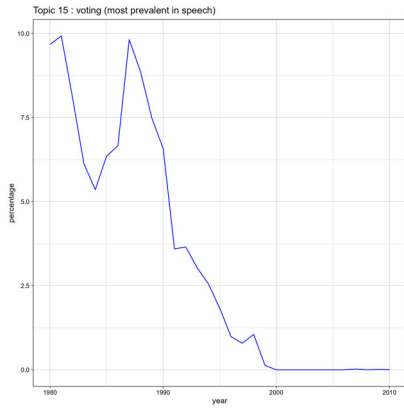


Topic 14 : parliamentary factions (most prevalent in speech)

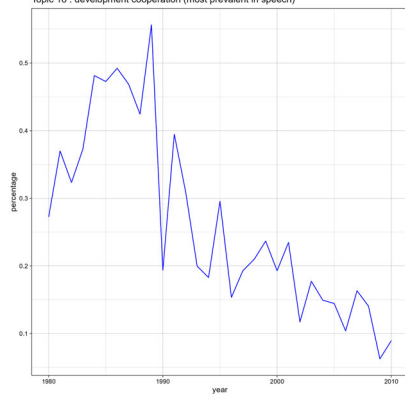


Topic 14 : parliamentary factions (prevalence in speech >= 30 %)

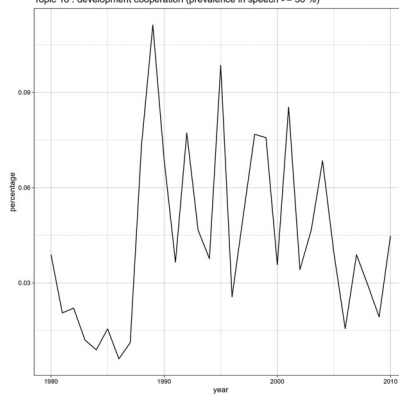




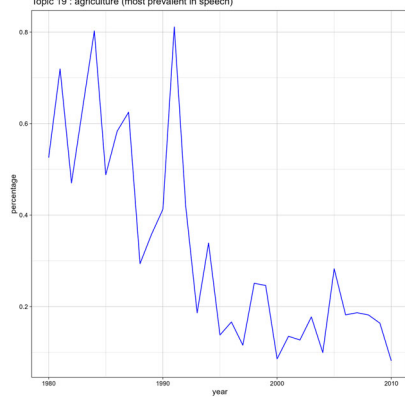
Topic 18 : development cooperation (most prevalent in speech)



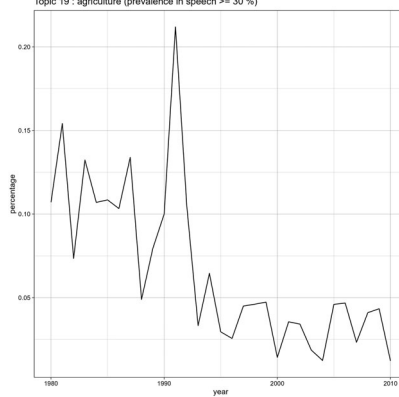
Topic 18 : development cooperation (prevalence in speech >= 30 %)



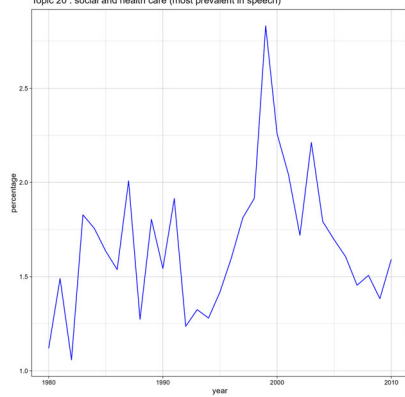
Topic 19 : agriculture (most prevalent in speech)



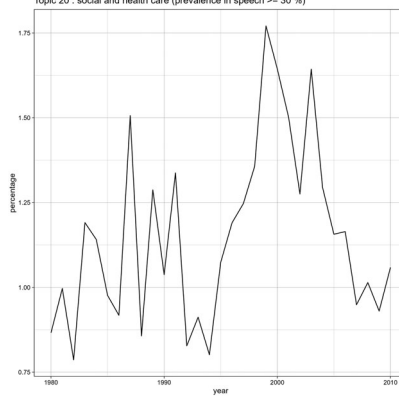
Topic 19 : agriculture (prevalence in speech >= 30 %)



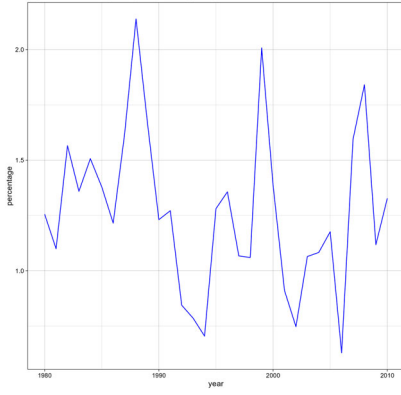
Topic 20 : social and health care (most prevalent in speech)



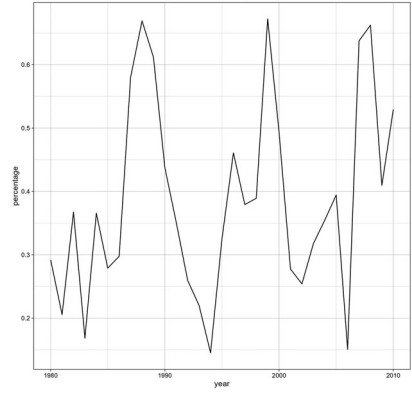
Topic 20 : social and health care (prevalence in speech >= 30 %)



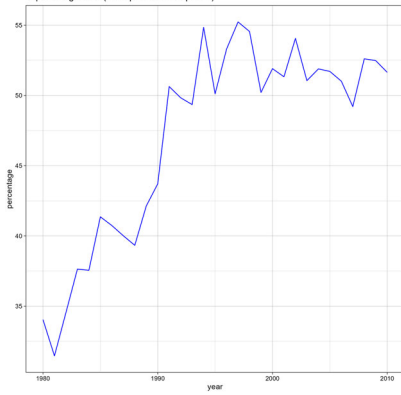
Topic 21 : taxation (most prevalent in speech)



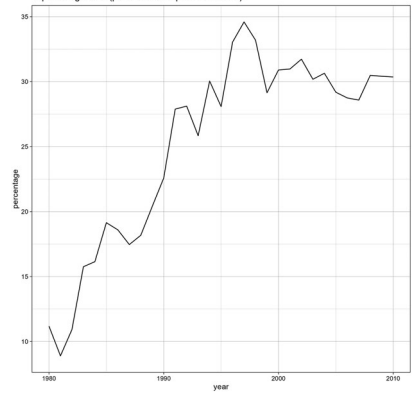
Topic 21 : taxation (prevalence in speech >= 30 %)



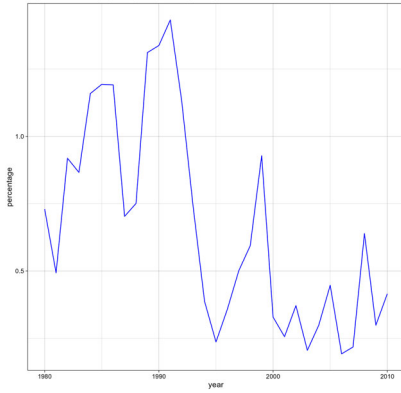
Topic 22 : general (most prevalent in speech)



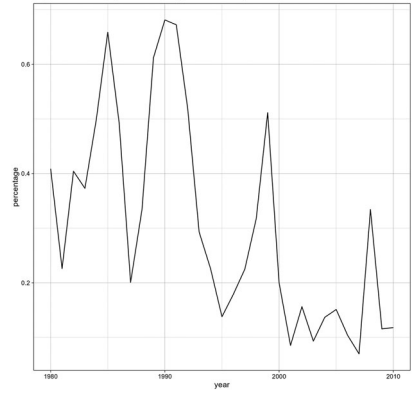
Topic 22 : general (prevalence in speech >= 30 %)



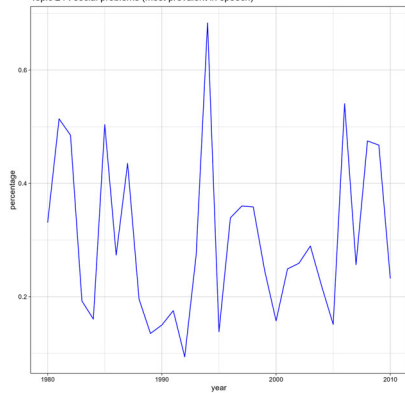
Topic 23 : housing (most prevalent in speech)



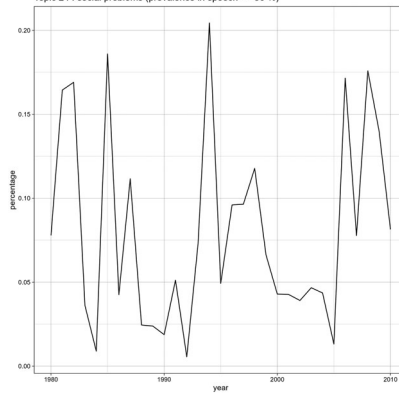
Topic 23 : housing (prevalence in speech >= 30 %)



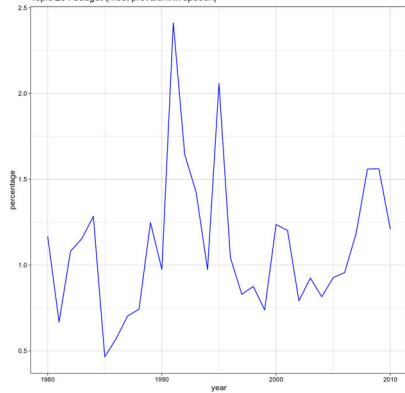
Topic 24 : social problems (most prevalent in speech)



Topic 24 : social problems (prevalence in speech >= 30 %)



Topic 25 : budget (most prevalent in speech)



Topic 25 : budget (prevalence in speech >= 30 %)

