RĪGAS STRADIŅA
UNIVERSITĀTE

Instrument of parliamentary discourse analysis – Saeima debate corpus

Authors: **Aija Kažoka**, Rīga Stradiņš University, programme «Communication culture and multimedia»,
Ilva Skulte, Rīga Stradiņš University asoc.prof.
Latvia

BACKGROUND

The Latvian Saeima debate corpus (http://saeima.korpuss.lv/) [1] represents a special field in the use of language – a political discourse, which includes all the words that members of parliament (MP), ministers and other members of the Saeima session have said on the podium since 1993. The instrument for the analysis of corpus of parliamentary debates in Latvian Saeima was created on the bases of transcripts collected from parliamentary sessions during the period of 25 years (respective, 5th-12th Saeimas after regained independence). Totaling over 13 million words, it offers a convenient set of tools that can be used for parliamentary discourse research (addressing the research of individual sessions, parliamentary term, parties, personalities, age, or gender speeches).

The annotations contain lemma, part of speech, morphological features and syntactic dependencies according to the Universal Dependencies standard format. Texts are automatically tokenized, lemmatized and morphologically analyzed and tagged using CMM based tagger. Syntactic dependencies are inferred by neural transition-based dependency parser trained on Latvian Universal Dependencies corpus version 2.1 [2].

The research work in the cabinet is aimed at obtaining concordances – key words or combinations of words that quote whole sentences or parts where they are used. Additionally, statistics on commonly used word combinations are acquired. The program also allows you to see the name of the neighborhood, in the context of the use of a particular word, which is essential for defining word connections and typical word combinations.

METHODS

The possibilities of the Saeima debate corpus are examined in four research works (with application of critical discourse analysis)

AIM

Show how to critically analyze parliamentary discourse using Saeima debate corpus.

RESULTS

Using the data of four research projects, showed many different uses of language of power in the parliamentary discourse.

CONCSLUSIONS

Saeima debate corpus is a vast and useful resource not only for political discourse research – it can also be applied to analytical journalism to prepare for an interview with a specific MP, check a theory about the process of implementing a policy or some other political process, or to accurately create an assertion, and avoid demagogic generalizations.

The parliamentary debate corpus applies not only to scientific research – it is also a useful tool in analytical journalism.

---

1. Doctor-MPs power of language study finds that physicians of Latvian parliament in their political debates use the same discursive structure of language (the language of power), which is used for communication with a patient; have taken over medical and healthcare topics and dominate them; self-identify with their professional group and use this identification as a tool during arguments. (3, 4)

Saeima corpus analysis – data collection
Table 1.
A summary of data used in this study – in the parliamentary discourse on medical sectors

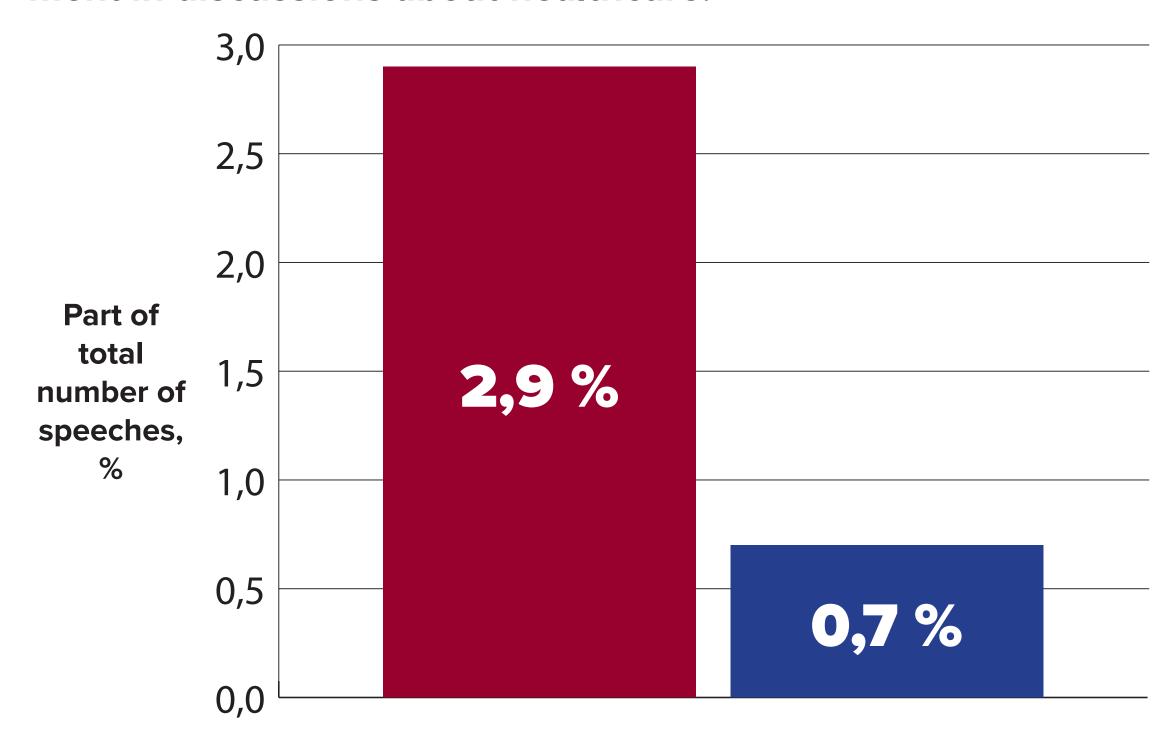| Sector	Number of physician's speeches (total)	The number of speeches between deputies, MPs, government officials (total)	Physicians, %
Diabetes	8	40	20
Cardiology	10	49	20
Onkology	13	84	15
Infectology	16	85	19
Sexually Transmitted Diseases	17	91	19
Gynecology	51	203	25
Family medicine	63	333	19
Hospitals	141	975	14
Total	**319**	**1860**	**17**
All entries in the database	**10 996**	**227 954**	**6**
Part of the total, per cent (%)	2,9	0,7	

Physician-patient language of power use in parliamentary physician's speeches
Table 2.
A summary of data used in this study - in the parliamentary discourse on medical sectors

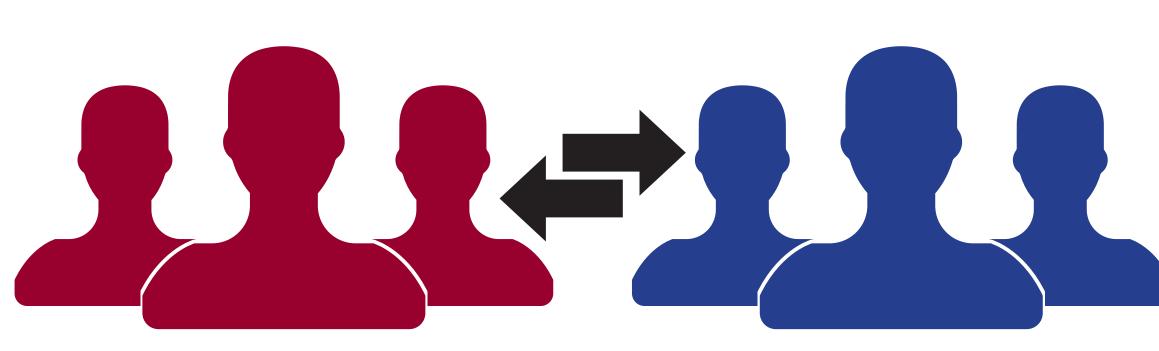|	
„Helping" form	How much funds we will provide, that much we will be able to help the ill. This hall also has a member of the parliament who has such a corona angioplasty.
Technical language	Even now, already patients are being helped by treating cancer, diabetes, hemophilia, phenylketonuria and even myocardial infarction.
„Biomechanical language"	Allocations for free medicines, which must be in accordance with the Cabinet of Ministers regulations, are not foreseen in the budget, and this has caused considerable tension between citizens and doctors, as the possibilities to receive free medicines in accordance with the Cabinet regulations are limited ...
Irony	What has caused bloating for the number of jobs? Maybe we need a cool breeze of wind but, as a doctor, I prescribe an enema.
„Friendly convincing"	For all those who want to see this problem first-hand, I strongly urge you to visit the lung surgery department of Pauls Stradins Hospital.
Guards of the institution	I think it would be better to say that this increase in wages is necessary not for doctors but for our people.
Pointing out inaccuracies	A person who uses phrases like "syphilist", "syphilist injection for 100 lats per piece", of course, is not endowed with meaningful mental capabilities.
Role of God	I address you directly. I ask you to listen, Madam, because I'm telling you!
Issue initiation	How much did one patient pay for the treatment, and how much is it now? How much did they pay for ineffective medicines, Russian antibiotics and so on, and how much now? And how long is it nowadays with a stomach ulcer in a hospital?
Addressing	Sasha! Sorry, Mr Golubov!

Saeima corpus content analysis – data summary

Comparison between physician and other politician involvement in discussions about healthcare.

2,9 % Physician-parliamentarians
0,7 % Other members of the government

■ Physician-parliamentarians ■ Other members of the government

Identification with a professional group

Physicians-parliamentarians → Parliamentarians-non-physicians

"I am a medical professor, I lead a large clinic and I have repeatedly participated in hospital commission work, and I want to say the following..." .

"Honorable medical coryphaeuses, I ask that you do not teach us like students. Do not start lecturing here or spread your propaganda

Usage of first person speech in physician-MP speeches, and the response reaction

---

3. Analysis of the speeches of politicians about resisting the initiative of environmental activists to abandon the use of wildlife in a circus. Politicians argue their view in media, referring to the national tradition, but in the analysis parliamentary speeches, the word "circus" is used with contempt. (4)

How do MPs use (and think!) the word "circus"?
This word appears in Saeima corpus 141 times, out of which only 24 times it's used to discuss legislation.
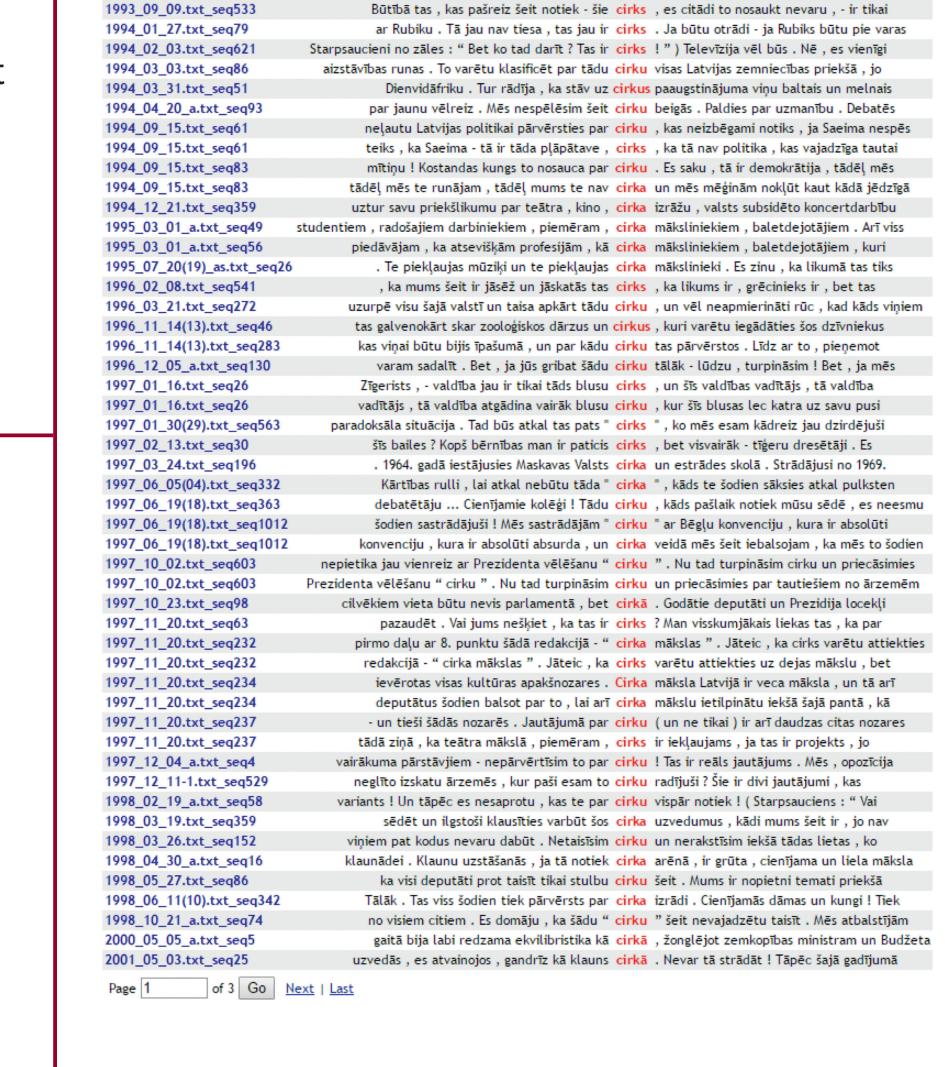
Examples:
government is like a flea CIRCUS and the current government leader – his government is like a flea circus where all the fleas jump in different directions. (61.97)

trying to turn parliamentary sessions into a place dealing with relations, and that is looking like a CIRCUS. (30.4.02)

And now, when you lied to your voters, you want to stop the CIRCUS that was truthfully cultivated within your coalition. (24.11.13)

Observation: the word "circus" is used to describe something as uncouth – an ugly environment filled with conflict; circus is made to be related to absurdity or caricature, never referring to it as a valuable tradition (which is different than the representation of circus in the media).

---

2. While carrying out a comparative analysis of the discursive claims against the state, it was found that in the parliamentary discourse the concept of state is vague. However, it is often compared to the parents, mother, or owner, attributing paternal function. Although in the topical use of the word «state» it does not frequently form a phrase together with the word «welfare», the expectations of the state as the «caregivers» (parents, host, fund manager, or fair distributor of shares) especially in the speeches of politicians, can be communicative about the aforementioned signs of a welfare state crisis. (5)

Topical usage and connotations of the word «state» – definition in dictionaries and topical usage in language corpora

Research employed the method of discourse analysis. Analyzed language corpora included: RSU Communications faculty and UL Artificial Intelligence technological tools for Saeima session transcript corpus analysis: **http://saeima.korpuss.lv/** un **bonito.korpuss.lv/ saeima**, as well as Latvian language corpus located at **nosketch. korpuss.lv**

The topical use of the word «state» in language corpora

Saeima debate corpus (includes all speeches from Saeima sessions in years 1993–2016):
Found **34,652** from 405,074 (8.55%) fragments

Latvian language corpus (without „Saeima" and „Normative acts"):
Found **10,580** (1,888 per million) fragments

Selection principle identifying most common word combinations with the word «state» in language corpora

Frequency of word combinations viewed:
■ by the frequency of numerical use (colocations) excluding punctuation, predilection and conjunctions;
■ one word to the right, one to the left, in all cases and conjugations.

■ **Is the state a big mother or a policeman** who interferes with the self-declared feelings of a person?
   -Ilze Viņķele Vienotība 09/15/2016 regular session
■ Your comparison with a family where the **state is a parent** and merchants are the children who are quite unwise and who grow up and thrive when the state - the parents - cares for them and protects them from acting unreasonably, of course, is beautiful (..)
   -Lolita Čigāne, Vienotība, 10-15-2015 regular session
■ The Latvian **state has been** careful enough like a caring **owner** ...
   -Andrejs Elksniņš, Saskaņa, 13/11/2014 regular session

Words comparatively closest related to the word «state» from 5th to 12th Saeima term

To research collocation, the results are sorted by the logDice function. It reaches peak value when all occurrences of the searched word (i.e. «state») in the text are next to the colocation word , AND all instances of the colocation word are always next to the searched word.
In other words, a high logDice value means that the searched word and colocation word form phrases that repeat themselves regularly.

Nr.	5th Saeima	6th Saeima	7th Saeima	8th Saeima	9th Saeima	10th Saeima	11th Saeima	12th Saeima
1	government	government	government	government	government	democratic	government	revenue
2	budget	revenue	revenue	revenue	president	only	revenue	president
3	revenue	president	language	our	our	language	president	government
4	property	our	our	language	president	president	budžeta	democratic
5	president	president	budget	prezidente	budget	government	our	only
6	our	budget	budget	language	language	budget	budget	language
7	security	budget	Latvia	budget	Amendments	president	control	budget
8	ministrs	Baltic	Baltic	language	control	forest	pension	control
9	reformu	minister	control	democratic	president	pension	Union	president
10	budget	control	Union	budget	budget	governmental	language	government
11	president	minister	preziden	only	revenue	revenue	Baltic	president
12	forest	About	stock	Union	About	secured	Latvia	secret
13	budget	secret	president	security	only	budget	official	pension
14	Baltic	minister	security	Baltic	pension	budget	About	security
15	About	security	budget	social	Amendment	revenue	social	budget
16	radio	language	language	Latvia	democratic	Baltic	others	budget
17	Latvia	Latvia	forest	control	social	budget	social	official
18	company	budget	president	pension	budget	border	control	control
19	minister	minister	About	budget	Latvia	control	probation	About
20	other	from	pension	revenue	civilservice	our	Amendments	Baltic

Conclusions

■ Within the framework of this pilot study it was found that politicians in their speeches deviate from the definition of state, verbalizing it as an alienated, autonomous entity.
■ The most commonly used vocabulary in relation to the word «state» leads to believe that both citizens and politicians expect strict management (also financial) from the «state». It is implemented with the help of an acceptable symbol ("president"), and in it the communication takes place in a defined language, and lastly – which is particularly relevant – being tied to while not being responsible ("our").
■ Comparisons in 12th Saeima transcripts show a desire to see the country as «parent» «mother» or «owner».
■ Although in the topical use of the word «state» it does not frequently form a phrase together with the word «welfare», the expectations of the state as the «caregivers» (parents, host, fund manager, or fair distributor of shares) especially in the speeches of politicians, can be communicative about the aforementioned signs of a welfare state crisis.

---

4. Thinking about health care and examining parliamentary speeches during a single term, it was established that the word "ill" is rarely used for its common meaning, instead – for offensive intonation. (6)

Observation: around 10% of the time, the descriptors "ill" and "illness" are used figuratively, in particular – meaning bad, stupid (data used from 8th Saeima during the process of implementing healthcare reform)

Typical phrases:
Our society is **ill**
Mr.Kalvītis thinks that it's an **illness** that needs to be endured
Ill political system
Ill reaction
I don't know what **illness** is spreading here that you can't even write a document right
That is an **ill** idea.
Don't allow the same **illness** from which the whole government is suffering.
That is an **illness** that is not just an individual case but is common to the whole Saeima.

REFERENCES:
1.The Latvian Saeima debate corpus was created by researchers of the Department of Communication Studies of the Rīga Stradiņš University and the Artificial Intelligence Laboratory of the Institute of Mathematics and Computer Science of the University of Latvia, while implementing the National Research Program Project EKOSOC-LV Nr. 5.2.5.
2. Dargis R., Auzina I., Bojars U, Paikens P., Znotoņš A. „Annotation of the Corpus of the Saeima with Multilingual Standards"
https://www.clarin.eu/sites/default/files/ParlaCLARIN_Session2_2.3.Saeima_Roberts-Dargis_LIREC2018.pdf (viewed 07.09.2018)
3. Kažoka A., „Ārstu vara veiskā parlamentārajā diskursā. Ārstu – politiķu loma medicīnas politikas veidošanā", RSU, 2016
4.Kažoka A., „Civil Society in Action: The Professional Organizations of Medical Doctors"/Puralism Anxiety, Acting Society in Latvia. Edited by S.Kruk, Riga, Rīga Stradiņš University, 2018, 205 p. (125-140 p.)
5. Kažoka A., „Priekšstats par valsti: aktuālā lietojuma konotācijas diskursa analīzes skatījumā", RSU Scientific Conference Publication of Theses, 2018, RSU, p 268.-269. https://www.rsu.lv/sites/default/files/imce/Zin%C4%81tnes%20departaments/2018/RSU_zinatniskas_konferences_tezes_2018.pdf
6. From open course "Kultūras žurnālistikas vasaras skole" (The Summer School of Culture Journalism"), Riga, Latvija, 11.08.2017.
6. Presented at the opening ceremony of the Saeima debate corpus, RSU, 2014