

Analyzing Sentiments towards the European Union in Slovak Parliamentary Speeches (1994–2023)

Jakub Szabó

Comenius University in Bratislava, Slovakia

2024-11-21

Introduction

Ever-increasing availability of online political texts keeps boosting quantitative text analysis in political science and legislative studies (Sebők et al. 2023)

Despite rampant literature on quantitative analysis of political speeches, however, text data analysis of political speeches in Slovakia remains dormant (for a notable exception see Ovádek 2020), often due to lacking data

Consequently, we only have ad hoc or relatively incoherent and scattered understanding of the political élite positions towards the European Union (EU) in Slovakia

Depoliticizing the EU in the Slovak political discourse

Ousting the autocrat (V. Mečiar) in order to be able to proceed with accession negotiations (Schimmelfennig and Sedelmeier 2004)

Omnipresent threat of the repeated exclusion from the accession process effectively depoliticizing the accession negotiations making any criticism of the EU membership illegitimate (Rybář 2011)

Pro-European attitudes and norms adopted by all the main political actors, including Mečiar's HZDS (Malová and Dolný 2016)

All parliamentary parties jointly campaigning in favor of Slovakia joining the EU before the accession referendum in 2003 -> unequivocal victory of the "yes" camp with over 93% (Henderson 2004)

Slovakia as a pro-EU paradise?

For the long time, the EU thus remained a non-polarizing issue

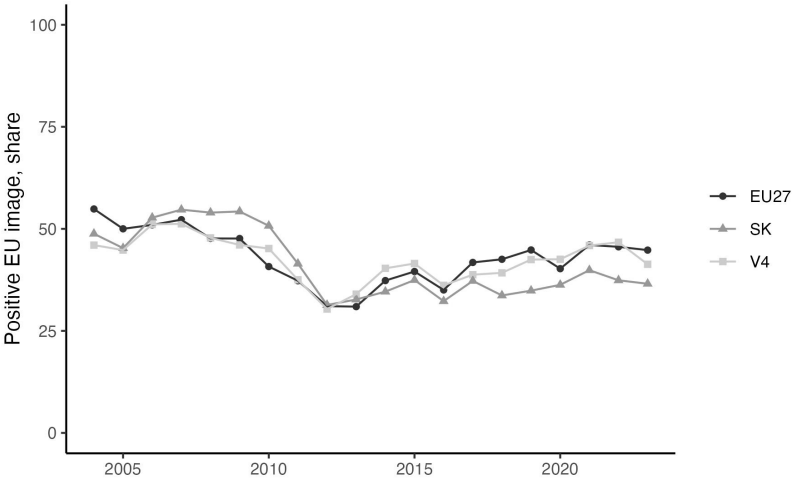
Even the attempts to politicize the “Greek profligacy” and the European Stability Mechanism failed (SaS and the fall of Iveta Radičová government in 2011)

“European integration was equated with national interest” (Henderson 2004, 668) and the only policy area that keeps together heterogeneous domestic interests (Malová and Dolný 2016)

Slovakia had shown, for several years, one of the highest support for the EU (Vilagi and Babos 2024; Gyárfášová and Henderson 2018)

This trend, however, seems to be weakening

Share of Respondents Having a Positive EU Image, 2004-2023



Source: Eurobarometer data, weighted averages

Research Question

Amidst the illiberal Orbán-Kaczynski-Babiš wave in the Visegrad region in the mid-2010s, Slovakia was praised as a “pro-European island in Central Europe” (Fico 2017)

We have, however, little empirical evidence of the evolution of political élite positions towards the EU in Slovakia

Did the political élite positions towards the EU experience similar trend reversal as the public opinion?

To be able to answer observe these trends, a longitudinal approach is needed -> parliamentary speeches from the National Council of the Slovak Republic (hereafter “NRSR”)

NRSR Parliamentary Speeches, 1994-2023

Relying on *rvest* (Wickham 2024) and *RSelenium* (Harrison, Kim, and Völkle 2024) packages, NRSR parliamentary speeches were webscrapped from two online sources:

- ▶ The webpage of the NRSR (NRSR 2024) that provides text transcripts of all parliamentary debates since 2002
- ▶ First two parliamentary terms (1998-2002) extracted from the joint Czech-Slovak digital parliamentary library, which contains all parliamentary documents of legislative bodies in Czech and Slovak lands going back to the late 19th century (Digitálna knižnica 2024)

The transcripts were scrapped with the following information: date, name of the speaker and session number

NRSR Parliamentary Speeches dataset I.

- ▶ **Term:** number indicating parliamentary term (1 - 1994/1998, 2 - 1998/2002, 3 - 2002/2006, 4 - 2006/2010, 5 - 2010/2012, 6 - 2012/2016, 7 - 2016/2020 and 8 - 2020/2023)
- ▶ **Session:** number indicating session number within the parliamentary term
- ▶ **Date:** date when the speech was delivered
- ▶ **Year:** year when the speech was delivered
- ▶ **Name:** name of the speaker who delivered the speech
- ▶ **Age:** achieved age of the speaker in the year the speech was delivered (year of the birth - year)
- ▶ **Gender:** biological sex of the speaker (female/male)

NRSR Parliamentary Speeches dataset II.

- ▶ **Party:** party affiliation of the speaker assigned based on the party ballot on which the speaker was elected into the NRSR or, if applicable, based on the party that nominated the speaker to their position (ministerial nominees)
- ▶ **Party Shift:** flag indicating whether the speaker changed party affiliation during the parliamentary term (change in party affiliation - 1, no change in party affiliation - 0)
- ▶ **Opposition:** flag indicating whether the party is part of the coalition or opposition during the parliamentary term (opposition - 1, coalition - 0)
- ▶ **Minister:** flag indicating whether the speaker is in ministerial position, i.e. part of the government (minister - 1, non-minister/MP - 0)

NRSR Parliamentary Speeches dataset III.

- ▶ **Leader:** flag indicating whether the speaker is a party leader, i.e. chairman of the party (party leader - 1, ordinary member - 0)
- ▶ **Chairman:** flag indicating whether the speaker is either president or vice-president of the NRSR
- ▶ **Left-right:** party left-right scale based on the Manifesto Project data
- ▶ **Conservative-liberal:** party conservative-liberal scale based on the Manifesto Project data
- ▶ **Transcript:** parliamentary speech delivered by the speaker

Pre-processing

The crude dataset contains over 420,000 individual speeches in the NRSR between 1994 and 2023

All speeches that were not delivered by the members of the government or MPs were removed (e.g., MEPs, presidents, general prosecutor, head of the supreme audit office, president of the constitutional court etc.)

All transcripts having less than 30 words were removed (in general, this ensures exclusion of all the cases when the chairman just passes the floor to another speaker)

EU-related Corpus

To only analyse the EU-related speeches, an “EU-filter” was employed (cf. Castanho Silva and Proksch 2021, 780; Rauh, Bes, and Schoonvelde 2020, 407), i. e. only speeches containing some EU-related keywords or its stem in Slovak (“Brusel”, “Európ” and “EÚ”) kept

The EU-related corpus was then automatically translated into English using Google Translate API (Vries, Schoonvelde, and Schumacher 2018) for two reasons:

- ▶ Automatically translated sentiment directory into Slovak by Proksch et al. (2018) proved unreliable. Manual validation revealed that it contains numerous untranslated (English) and unintelligible Slovak words
- ▶ The absence of reliable and validated lemmas and stemming dictionaries for a “peripheral” Slovak language

EU-related Corpus Descriptives

In total, slightly over **12,100** speeches contain some references to Europe (around 3% of all speeches)

Parliamentary Term	No. of Speeches
1st term (1994-1998)	781
2nd term (1998-2002)	1586
3rd term (2002-2006)	1313
4th term (2006-2010)	1672
5th term (2010-2012)	688
6th term (2012-2016)	2554
7th term (2016-2020)	1803
8th term (2020-2023)	1629

Pre-processing EU-related Corpus

As part of the pre-processing, I apply the usual cleaning exercise for the text analysis on the tokenized transcripts (Benoit et al. 2018):

- ▶ Removing numbers, punctuations, symbols, separators and stopwords
- ▶ Lowering the case for all terms
- ▶ Compounding frequently co-occurring words (E.g., European union, European integration or Prime Minister)
- ▶ Adjusting selected tokens to ensure identification of negations and prevention of false sentiment identification

Sentiment Analysis

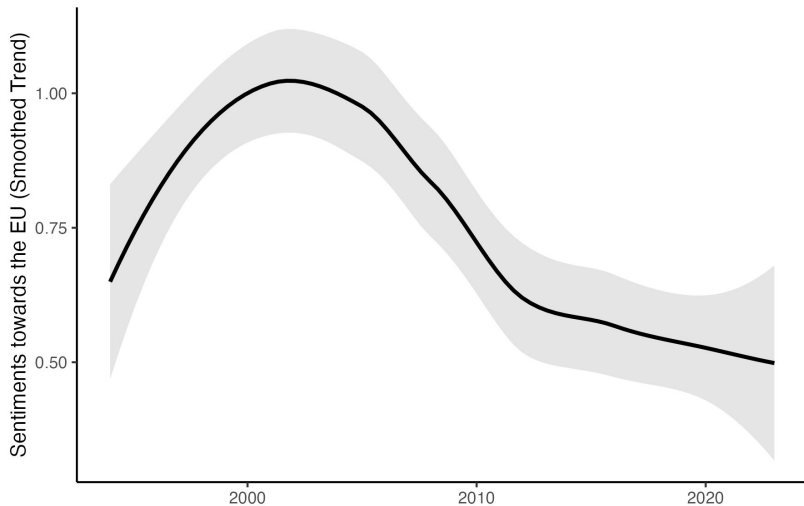
Term-based sentiment analysis using the *Lexicoder Sentiment Dictionary* (LSD)(Young and Soroka 2012) -> one of the most utilized sentiment dictionary in social sciences (Haselmayer and Jenny 2016; Cochrane et al. 2021)

Extracting a window of 10 tokens from each side of the EU-related keyword occurrence and counting the instances positive or negative sentiments from the LSD dictionary are detected (cf. Castanho Silva and Proksch 2021; Rauh, Bes, and Schoonvelde 2020)

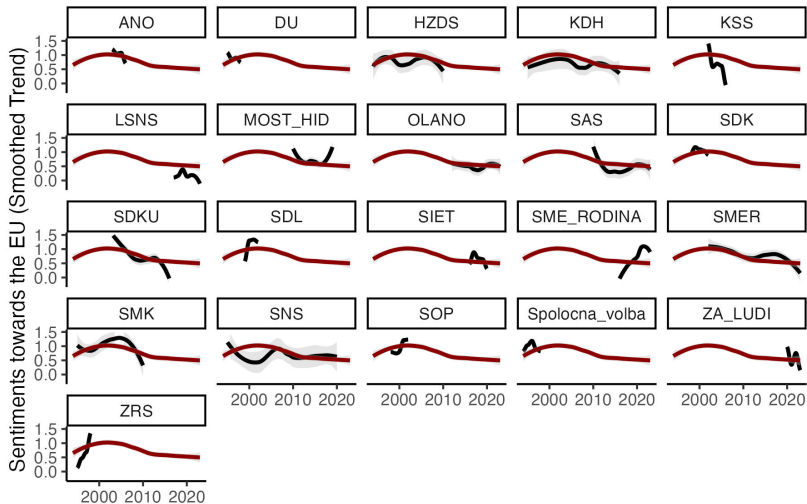
*Validation with 5-token window yields almost identical results

The sentiment per each speech is measured using so-called empirical logit (Proksch et al. 2018, 102)

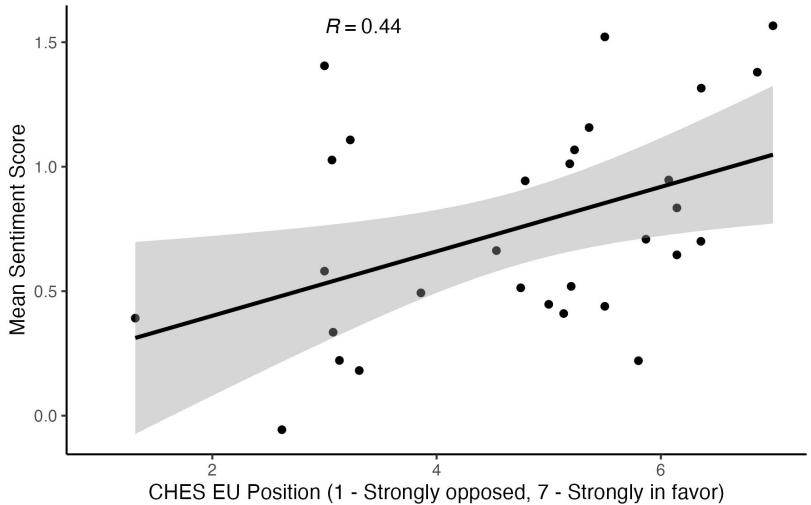
EU-related Sentiments in NRSR, 1994-2023



EU-related Sentiments per Party Affiliation



Validating the Sentiment Measures



Source: Chapel Hill Expert Survey

Discussion

Pro-EU sentiments seem to have peaked around the time of the accession to the EU

Adoption of the EU-friendly discourse as a strategic calculation facing the strong influence of the EU conditionality (Schimmelfennig 2005)

In the aftermath of the accession, we observe a gradual decline in pro-EU sentiments

With the “stick” missing, greater room for politicization of the EU

The pro-EU image of Slovakia likely to be a facade built on fragile foundations

Planned Directions for this Research

Estimating the EU-related (aspect-based) sentiment in NRSR speeches using BERT (e.g., Miok et al. 2024)

Explaining the sentiments with speaker-level and party-level covariates (e.g., age, sex, party ideology, opposition vs. coalition)

Estimating *structural topic model* (Roberts, Stewart, and Tingley 2019) with speech-level and meta-level covariates -> what topics are related with positive and negative speeches?

References

- Benoit, Kenneth, Kohei Watanabe, Haiyan Wang, Paul Nulty, Adam Obeng, Stefan Müller, and Akitaka Matsuo. 2018. "Quanteda: An R Package for the Quantitative Analysis of Textual Data" 3: 774. <https://doi.org/10.21105/joss.00774>.
- Castanho Silva, Bruno, and Sven-Oliver Proksch. 2021. "Politicians Unleashed? Political Communication on Twitter and in Parliament in Western Europe." *Political Science Research and Methods* 10 (4): 776–92. <https://doi.org/10.1017/psrm.2021.36>.
- Cochrane, Christopher, Ludovic Rheault, Jean-François Godbout, Tanya Whyte, Michael W.-C. Wong, and Sophie Borwein. 2021. "The Automatic Analysis of Emotion in Political Speech Based on Transcripts." *Political Communication* 39 (1): 98–121. <https://doi.org/10.1080/10584609.2021.1952497>.
- Digitálna knižnica. 2024. "Spoločná Česko-Slovenská Digitálna Parlamentná Knižnica." <https://www.nrsr.sk/dl/>.