

A large, stylized illustration of three muscular, bronze-colored figures in a heroic pose. The figure on the left holds a torch aloft, the central figure holds a large red flag with a yellow hammer and sickle, and the figure on the right holds a large, dark, wing-like object. The background is a bright, hazy sky.

The automatic analysis of linguistic features in Communist propaganda texts

Veronika Vincze¹ - Martina Katalin Szabó^{1,2} - Orsolya Ring²

1. MTA-SZTE Research Group on Artificial Intelligence (RGAI)

2. TK Computational Social Science - Research Center for Educational and
Network Studies (CSS-RECENS)





About the recent project

- Hungarian Communist propaganda texts
- Corpus of Pártélet journal
- The first and the last two years
- Statistical, morphological, syntactic and semantic features

- **Purpose:**

Pártélet: totalitarian language usage

How and to what extent political forces can affect linguistic choices (biases and deception)

To develop an automatic tool

„Barátságunk
erősítésével a dolgozó
milliók álmát valósítjuk
meg”
(Pártélet)

(‘By strengthening our
friendship, we make the
working millions' dream
come true’)



The structure of the presentation

- Background to the recent project
- About the recent study; corpus
- Some of our recent results

About the research project I.



1956-1989: “Kádár era”; an active and decisive period of Hungarian history.

A widely examined topic in sociology, social history, political history, and history (e.g. Pap 2015, 2016; Kovai 2016; Gyáni 2016; Medgyesi 2017; Herczog 2017)

But: traditional toolboxes; a qualitative and manual analysis; time-consuming and costly; large amounts of textual data pose a challenge

Unusual approach: large datasets, quantitative perspective (McClurken 2012; Miller 2013; Sebők 2016)

About the research project II.

There was no digitized textual corpus available

Pártélet: official journal of the governing party, the Central
Leadership of the Hungarian Socialist Workers' Party;
1956-1989

For party members; 51 150 copies

A politikai szervező és nevelőmunka egységéért

A most folyó beszámoló és vezetőségválasztó taggyűléseken a kommunisták felméri azt az utat, amelyet pártszervezetük az utolsó két évben, és különösen a Központi Vezetőség tavalyi márciusi határozata óta, e határozat segítségével megtett. Az országgyűlés legutóbbi ülésén elhangzott beszédekből már ismerjük ennek a fejlődésnek országos eredményeit; a taggyűlések most kimutatják, hogy egy-egy üzem, termelőszövetkezet, gépállomás, állami gazdaság, község, hivatal, vagy intézmény mennyivel járult hozzá az ország általános fejlődéséhez, vagy mennyire maradt el mögötte.

A taggyűlések azt is megmutatják, hogy ezt a fejlődést mindegyik dolgozó népnünk szorgalma, munkakészsége, alkotó energiája vitte előre — az elmúlt két esztendőben ugyanúgy, mint a felszabadulásunk óta eltelt egész időszakban. A taggyűlések egyúttal — mind eredményekkel, mind hiányosságaikkal — bizonyítanak egy másik sokszor hangzatosított igazságot is: azt, hogy a népnünk rejlő hatalmas erő nem önmagától, ösztönszerűen tör fel belőle, hanem annyira bontakozik ki és fejlődik, amennyire a párt feltárja és szervezi. A pártnak ez a tevékenysége — minden sikerünk egyik legfontosabb tényezője.

A pártmunka legfontosabb feladata a dolgozó tömegek, a kommunisták és pártönkivüliek szervezése és nevelése. A párt szervező és nevelőmunkáját elméletileg el lehet határolni egymástól, de ne feledjük el, hogy a gyakorlatban a kettő legtöbbször összefonódik. Egyes pártfeladatok megoldásáért elsősorban politikai szervezőmunkát kell végeznünk, más feladatok végrehajtásáért túlnyomórészt politikai nevelőmunkát — de nincs olyan pártfeladat, amelyet kizárólag szervezéssel, vagy kizárólag neveléssel is teljesíthetnénk.

Aki politikai szervező munkát végez, az a pártszervezetben és a pártönkivüli dolgozók között küzd népnünk életének állandó javításáért, s azért, hogy ennek akadályait legyőzze. Tanácskozik a dolgozókkal, tanácsokat ad nekik, hogyan éljenek és dolgozzanak. Segít nekik helyes kezdeményezéseik megvalósításában, mozgósítja és lelkesíti őket arra a munkára, amelyet ki-kí maga területén mindennap végez, s amelyet összefoglalóan úgy nevezünk: a szocializmus építése. Azon dolgozik, hogy a párt politikája megvalósuljon és a nép ellenségeinek mesterkedései meghiúsuljanak; ennek megfelelően osztja el az erőt a munka



Corpus compilation procedure

1. PDF pages: download and further process (cf. Szabó et al. 2019)
2. Resulting raw data was fed into magyarlanc (<http://www.inf.u-szeged.hu/rgai/magyarlanc>) (Zsibrita 2013): split into sentences; tokenization; remove punctuations and stopwords; lemmatization



Basic statistics of the corpus

33 volumes; 12 issues each year

Full corpus: more than 13 million tokens

Distribution of tokens: quite balanced over the years

Corpus of the recent study

- Single-year periods:
 - January 1957 - December 1958 (17 issues)
 - January 1988 - April 1989 (16 issues)
- Selection: specific historical and sociological criteria:
 1. Directly following the Hungarian revolution (1956); the active starting period
 2. Just before the regime change in Hungary

Rise and fall of the regime





Hypotheses

- **Deception:** conceal facts that were unpleasant
- **Persuasion:** to manipulate beliefs and intentions
- Assumed to be decisive in both periods
- But the linguistic characteristics of the two subcorpora were also assumed to be different



Processing steps of the two sub-corpora

- The magyarlanc toolkit
- Extracted basic statistical, morphological and syntactic features
- bias and deception → uncertain words, sentiment and emotion expressions
- We compared the features
 - ...of the two sub-corpora
 - ...with the research findings of the Institute for Propaganda Analysis (1938)

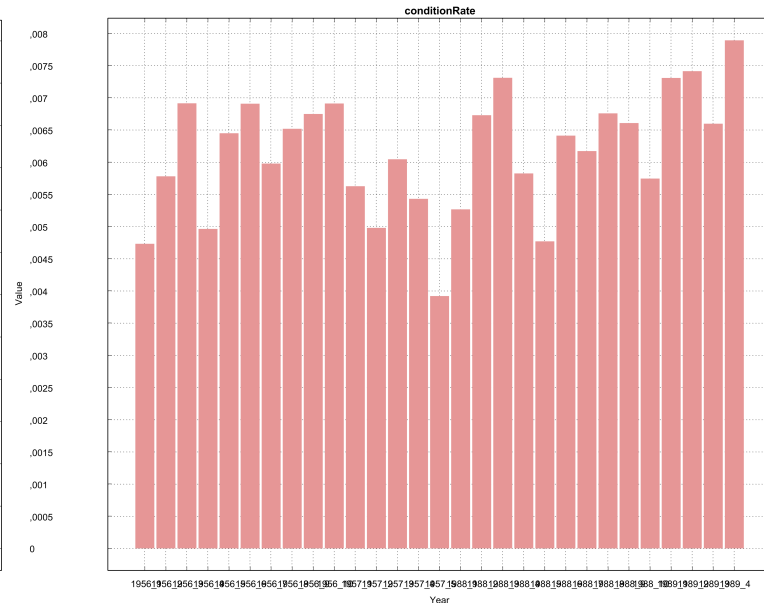
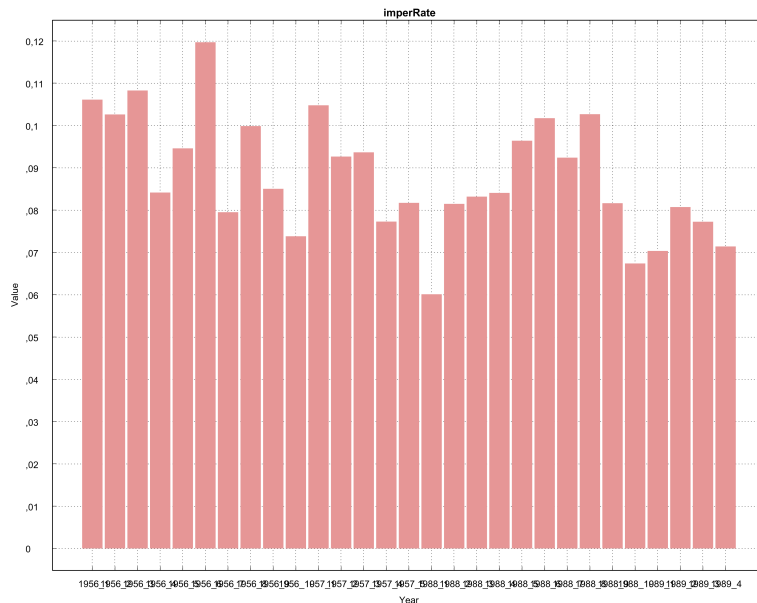


Results

- Significant differences between the two periods
- Some of the most interesting findings

Verb moods

- Implicit orders from the party
- uncertainty and political decline



Verb forms, tenses

- Frequency of Pl1 is notably higher in the 2nd corpus
- Manual analysis:
 - 1956-57: texts address other actors
 - 1988-89: authors discuss the leadership itself
- Cf. past and present tenses



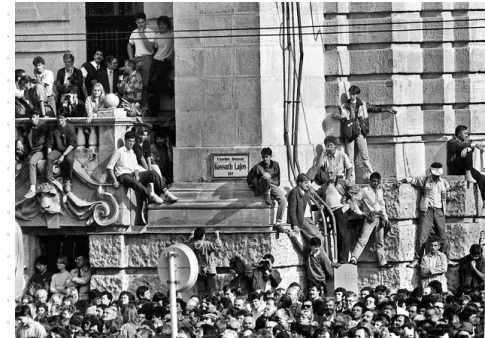


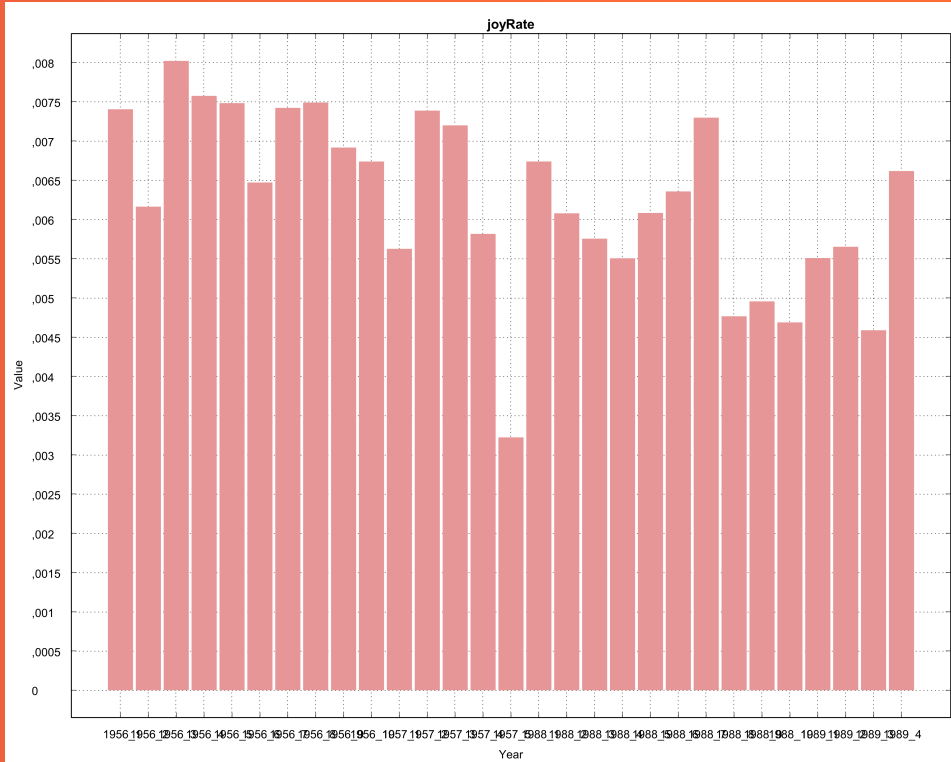
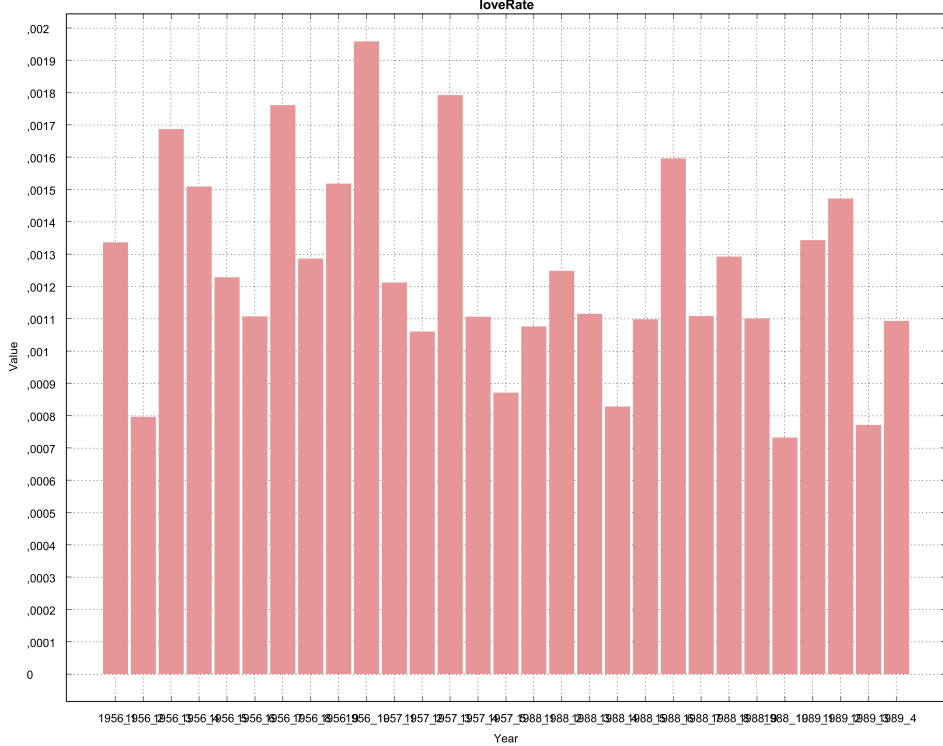
Superlatives

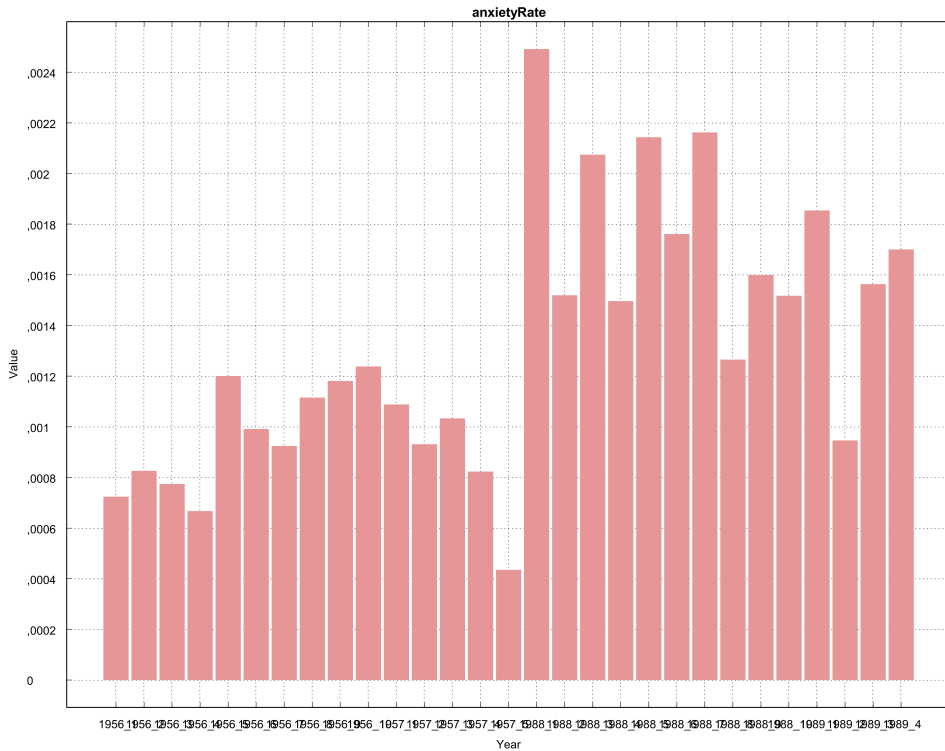
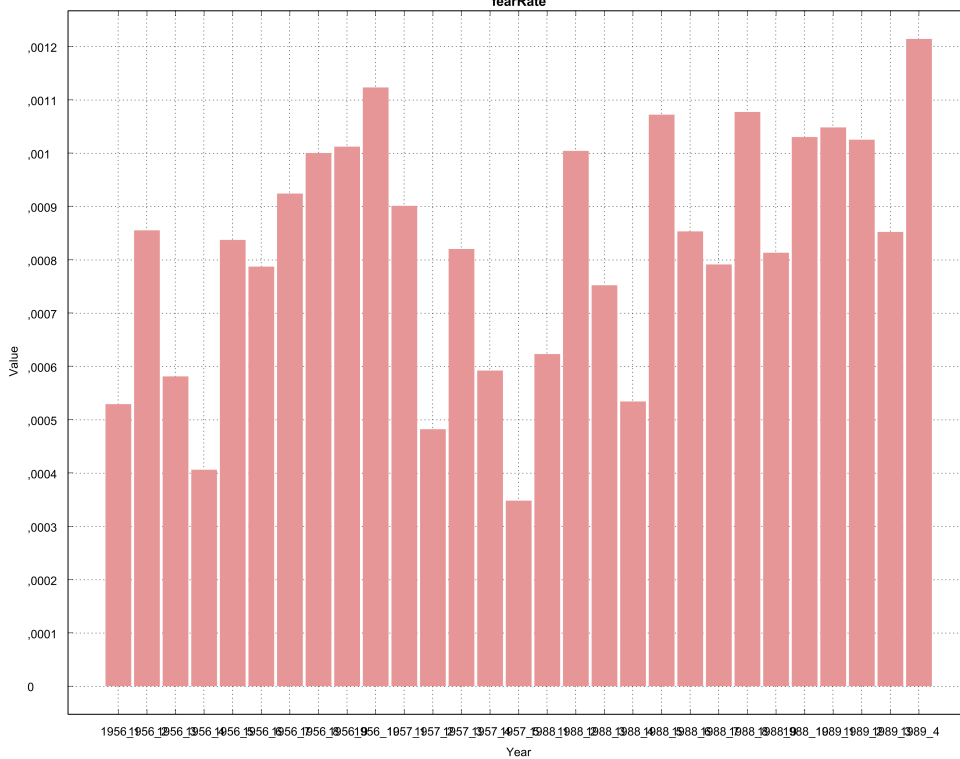
- 1st: high frequency of superlative of adjectives and adverbs (e.g. *legjobb* 'best', *leghatékonyabban* 'most efficiently': power, confidence and unquestionability)
- Cf. findings of the Institute for Propaganda Analysis (1938)

Emotions and sentiments

- **Positive emotions (joy and love):** 1st corpus
- A typical device for propaganda: emotionally appealing phrases (e.g. love of country, desire for peace, freedom etc.) (cf. "glittering generality"; Institute for Propaganda Analysis).
- **Negative emotions (anger, fear and anxiety):** 2nd corpus; anticipate a coming regime change





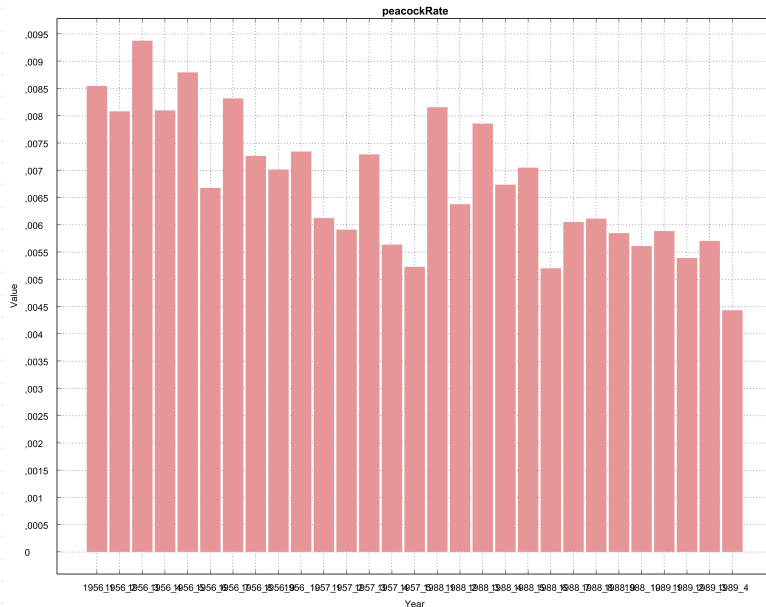




Initial results

Powerful and confident communication → more uncertain and weak discourse

Uncertainty



- Uncertainty of the speaker in the truth content of the information
- **Discourse level uncertainty: 1st:**
 - The lacking information is not related to the propositional content of the utterance but to other factors (e.g. *much, more, often* etc.) ↔ **semantic uncertainty** (e.g. *may, possible, believe* etc.): much less in the 1st
- Bias evoked by discourse level uncertainty: typical feature of propaganda discourse; implicitly more deceptive

Peacock: words that express unprovable qualifications or exaggerations (e.g. brilliant, excellent and best-known) (Vincze 2017)



Most frequent words

- Belonging to certain parts-of-speech
- **Nouns**
 - 1st: total authority and power of the leadership (e.g. *project, plan, fulfillment* etc.)
 - 2nd: possibility of differing opinions and the opportunity to choose between different options (e.g. *possibility, decision, opinion* etc.)



Conclusions and future work

- linguistic features of Hungarian propaganda texts
 - from the beginning and the end of the Kádár era
 - morphological, syntactic and semantic features
 - differences over time
-
- non-propagandistic texts
 - implement an automatic detector of propaganda

Thank you for your attention!

Veronika Vincze - Martina Katalin Szabó - Orsolya Ring



MTA-SZTE Research Group on Artificial Intelligence (RGAI)

MTA TK Computational Social Science - Research Center for Educational and
Network Studies (CSS-RECENS)