Hidden power structures in Czech printed media

Irene Elmerot

Department of Slavic and Baltic Studies, Finnish, Dutch and German

1. Project overview
These are two examples from a project examining linguistic othering in printed media discourse in the Czech Republic, using the Czech National Corpus (CNC). Journalistic texts after 1989 are used, sometimes with the balanced, representative subcorpus SYN2015 [11]. The project uses a partly corpus-based, partly corpus-driven [2], critical discourse analysis, with some statistical input (e.g. prominence and keyword analyses) [12], DIN ranking [8].

2. ROMA VS. GYPSY – A DISCOURSE & CORPUS ANALYSIS
Article 2016.
Main theory: a parallel to Fidler’s [7] “more automatic mental representation” through the language used about these “others”.

Node words: the Czech lemmata for Roma and Gypsy (Rom and Cikán).
Method: to analyse these lemmata with their adjacent L-1 adjectives.
Question: To what extent do negative and positive adjectives follow the two node words in the discourse of the most popular Czech printed media from 1989 to 2009?
Corpus analysis: a statistical analysis of the frequencies of adjectives adjacent to the two denominations.

Most surprising result: For both denominations, approx. 33% of the adjectives in position L-1 were geographically related (as in “Romanian”, “Czech”, or “local”). See Table 2.1.

More expected results: Negative adjectives twice as frequent with Gypsy compared to the lemma Roma. Neutral adjectives almost twice as frequent with Roma compared to the lemma Gypsy. See Table 2.1.

Useful for: When people discuss whether it is better to use one or the other denomination [cf. 15], to explain that one is closer connected to negativity, and the other more to neutrality.

All numbers are % of absolute frequency

<table>
<thead>
<tr>
<th></th>
<th>Negative adjectives</th>
<th>Neutral adjectives</th>
<th>Positive adjectives</th>
<th>Stereotypical neutral adjective</th>
<th>Geographical adjectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before Cikán</td>
<td>15.9</td>
<td>25.0</td>
<td>6.6</td>
<td>22.5</td>
<td>29.8</td>
</tr>
<tr>
<td>Before Rom</td>
<td>8.25</td>
<td>47.7</td>
<td>0.9</td>
<td>6.0</td>
<td>36.7</td>
</tr>
</tbody>
</table>

Table 2.1: Results from the corpus searches of adjectives associated with the lemmata “Cikán” and “Rom”.

3. AN ANALYSIS OF POLARISING NEWS STRUCTURES
Article 2019 forthcoming.
Main theories: a) In critical reading of news text, several analytical methods should be combined (9); b) Discourses are interpreted differently in a time of crisis [10, 6]; and c) When analysing one text at a time (cf. Tables 3.1, 3.2 and 3.3), patterns emerge more clearly [5].

Method: intertextual analysis of the news structure and wording, keyword analysis and then dispersion analysis of keywords. See Tables 3.1, 3.2 and 3.3.

Most surprising results: Male commentators negatively polarized versus female in the news items, despite this not occurring in original FB posts, whereas women also posted negatively. Word “opinion” [mažor] not widespread despite news items about people’s opinions.

More expected results: Keywords dispersion corroborated the initial result of polarising in groups. “Black man” mentioned only by negative (Table 3.1), “explain” only by experts. The word “racism” not mentioned by negative commentators (see Table 3.1 and 3.2).

Useful for: When people discuss pros and costs of qualitative versus quantitative research in linguistics. When teaching critical thinking in media and communication.

References

Abstract

**What hidden power structures are there in Czech printed media?**

This is an overview of a Language and power-framed project examining linguistic othering in printed media discourse in the Czech Republic, using the Czech National Corpus (CNC). Journalistic texts from 1989 onwards are used, sometimes with the balanced and representative subcorpus SYN2015 (Hnátková et al. 2014). The project uses a partly corpus-based, partly corpus-driven (Baker and McEnery 2015: 47), critical discourse analysis, with the help of statistical input (e.g. prominence and keyword analysis, DIN ranking etc., cf. Fidler and Cvrček 2018: 198–200).

**Example 1:** Subcorpus SYN of the CNC was used to analyse the Czech lemmata for Roma and Gypsy (*Rom* and *Cikán*) with their adjacent (position L−1) adjectives, to see what differences there were in the discourse of the most popular Czech printed media from 1989 to 2009, depending on which denomination had been used.

**Result:** The negative adjectives were about the double when adjacent to the lemma Gypsy compared to the lemma Roma. The neutral words were, on the other hand, almost the double for the lemma Roma compared to the lemma Gypsy. This is relevant to real-life language usage when people ask about the reason for using one or the other denomination (e.g. Zlinská 2019).

**Example 2:** The case of a Lidl advert turned into three news articles. This is an example where the structure of news articles amplify a polarization, which can be viewed both qualitatively and quantitatively. Through corpus software, a keyword analysis using a representative and balanced corpus allows some words or phrases to stand out – both for being there and for being absent. These patterns reflect several contemporary media phenomena, such as ‘click-hunting’, perceived threats to the nation and its traditional citizens, and a lack of indication regarding what in a news item is opinion and what is fact.

**Result:** Keyword and dispersion analyses show present and absent discourse, and clarifies a polarization that might not be obvious from a strictly qualitative research, nor from using a single quantitative method, where groups like men and women or laymen and experts are distinctly separated.

**Keywords:** linguistic othering, discourse analysis, corpus linguistics, Czech

**References**


Hnátková, Milena, Michal Křen, Pavel Procházka, and Hana Skoumalová. 2014. ‘The SYN-Series Corpora of Written Czech’. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC’14)*. Reykjavik: ELRA.