Konferenca
Jezikovne tehnologije in digitalna humanistika
Ljubljana, 2024

Conference on
Language Technologies & Digital Humanities
Ljubljana, 2024

GENDER IDEOLOGY: A CORPUS-LINGUISTIC LOOK AT EMERGENT 'ANTI-GENDER' VOCABULARY IN SLOVENIA, CROATIA AND SERBIA

Ksenija BOGETIĆ,[1,2] Vojko GORJANC,[1,3] Jure SKUBIC,[1,3] Alenka KAVČIČ[4]

[1]Institute of Contemporary History
[2]Research Centre of the Slovenian Academy of Sciences and Arts
[3]Faculty of Arts, University of Ljubljana
[4]Faculty of Computer and Information Science, University of Ljubljana

1 INTRODUCTION

The concept of *gender ideology* has been observed as a rising and semiotically shifting linguistic innovation in public discourses across the globe (Borba, 2022; Kuhar & Zobec, 2018; Baran, 2022). Essentially a co-optation of feminist and scholarly terminology by the populist right (Graff, 2020) used to denote a threat to traditional gender/social relations, the phrase has come across issues in theoretical interpretation, leaving linguists to often resort to describing it as an 'empty signifier' with fluid meaning (Gal, forthcoming). Empirical analyses of the phrase's usage in different language contexts are yet to provide a deeper picture of the term's social meanings.

The phrase itself offers an important point of departure for understanding the discourse underpinning the recent rise of the so-called 'anti-gender movements' (Kuhar & Patternotte, 2017), which define themselves as transnational, anti-feminist and anti-LGBTQ social mobilisations against *gender ideology* (Corredor, 2019). Beyond a merely populist term of the far-right, the idea of *gender ideology* as a threat or conspiracy has permeated important political discussions on matters such as reproductive rights and violence over women, providing legitimization for anti-feminist positionings and trivialization of gender-related inequalities. Understanding its use in more 'official' contexts of political debate – such as those of national parliaments – especially as the concept is in many parts of the world yet emerging, is thus an important prerequisite to engaging with anti-gender motions in meaningful

Konferenca
Jezikovne tehnologije in digitalna humanistika
Ljubljana, 2024

Conference on
Language Technologies & Digital Humanities
Ljubljana, 2024

ways.

Corpus data can in this regard provide important insights into processes of semiotic resignification by pointing to specific patterns of usage and association. Multilingual corpora as in the case of the ParlaMint project (Erjavec et al., 2023) are particularly useful for cross-linguistic and cross-national analyses. In particular, such resources may thus help us further understand the circulation of global (anti-gender or other) discourses across geopolitical space (Gal, 2021), and across social settings, among others by allowing comparisons of more general usage and adoption in parliamentary or political discourse in particular.

As a step in this direction, in this talk we will present the first preliminary results of a wider study on the concept of *gender ideology*, looking at three former-Yugoslav states: Serbia, Croatia and Slovenia.

2 BACKGROUND AND SOCIAL CONTEXT

The rise of the so-called anti-gender mobilizations (Kuhar & Paternotte, 2017), has brought forth a range of disturbing events in Europe and beyond, ranging from new motions to limit abortion rights, through bannings of gender studies from curricula, to new waves of violence against LGBTQ+ persons and organizations. Most simply, the notion of anti-genderism is used to refer to transnational social mobilisations against feminism and gender equality, built around the notion of *gender ideology* as a new threat to tradition, nation and the family (Corredor, 2019; Bogetić 2022). The targets of anti-genderism can differ in scope, and are molded to local national concerns, but overall span concepts of LGBTQ+ rights, reproductive rights, sex and gender education, gender studies, and democracy in a broadest sense (Kuhar & Paternotte, 2017).

Gender ideology is a comparatively recent coinage rising to prominence in the 21st century (Kuhar & Paternotte, 2017). Spreading across languages of central and eastern Europe like in many others, the phrase itself, based in a linguistic, semiotic transformation of both gender and ideology is now used to refer to an imported (Western) threat to social relations (Graff, 2020). It has been noted in similar form in different countries and languages across this

Konferenca
Jezikovne tehnologije in digitalna humanistika
Ljubljana, 2024

Conference on
Language Technologies & Digital Humanities
Ljubljana, 2024

space, as e.g. *gendernaia ideologiia* in Russia, *ideologija džendar* in Bulgaria – but also sometimes with differing realizations, as with *teorija spola* and *ideologija spola* as both existing in use in Slovenia (Popič & Gorjanc, 2023) – and as such has become a central point in local populist rhetorics, symbolizing perceptions of social decline stemming from 'liberal' or 'leftist' ideas. Its basis lies in the conviction that society in Western Europe faces imminent downfall because fundamental values like the family, homeland, and nation are no longer being upheld (Verseck, 2018). This in turn implies a need for attention and protection by mobilizing national and/or religious resources.

However, as Paternotte and Kuhar (2017) warned, the automatic entanglement of anti-gender campaigns to the far-right populism might miss on many important regional and national specificities, and the same goes for understanding the language of anti-genderism. Though the term '*gender ideology*' seems to feature in many Slavic-language states in the region, alternative phrases appear in parallel use, and also carry somewhat different connotations, and may also occur in different forms in different public contexts (such as media texts and parliamentary debates). Generally, the area of eastern and central Europe has often been described as a hub of anti-genderism, seen especially with recent developments in Poland and Hungary (Graff, 2020; Peto 2020); other post-socialist areas such as the former Yugoslav space, where gender is long argued to hold a key symbolic role in political life (Bonfiglioli et al. 2015), are less explored.

Also, as some scholars in the post-Yugoslav area have noted the emergence of anti-gender discourses and vocabulary in debates in national parliaments (cf. Kuhar, 2020), parliamentary discourse becomes an important and under-represented discourse setting in which to observe the adoption of the anti-genderist shibboleth of *gender ideology*. This is where recent corpus developments in multilingual, cross-country parliament datasets provide an invaluable point of analysis of language of the parliaments, and discourses of politics and gender. Tracing emergent concepts also allows us to observe some potentials, limitations and challenges in bringing corpus-based analyses together with critical analyses of discourse.

Konferenca
Jezikovne tehnologije in digitalna humanistika
Ljubljana, 2024

Conference on
Language Technologies & Digital Humanities
Ljubljana, 2024

3 AIMS

In this study we aim to investigate the use of the concept of *gender ideology* in Serbian, Croatian and Slovene. Indirectly, we also aim to use the analysis to obtain insights into aspects of anti-genderist rhetoric in the three states of interest. In our preliminary analysis, we find it productive to compare the uses of the term in a web corpus of more general scope and in a corpus of parliamentary discourse, as well as to look across three national contexts once belonging to the same state, now taking some differing political and EU-related positionings. Specifically, our research questions include:

RQ1: How may the frequencies of occurrence compare across the countries and corpus/discourse types?

RQ2: What sets of social/political concepts (sexuality, religion, political affiliation, etc.) may be observed among the top collocates of *'gender ideolog*y*'* in general and parliamentary discourse?

The multiple levels of comparison would of course warrant further and carefully controlled corpus analyses. At this point, we focus on the broadest picture of frequencies and collocation, while finding that a combination of corpus-based analysis and analysis of concordances from a discursive angle is telling of social meanings (see Methodology).

Zooming in on parliamentary discourse as our particular point of interest, we explore two further questions:

RQ 3: As an essentially anti-feminist term, can differences be observed in the distribution of the use of *gender ideology* between male and female MPs?

RQ 4: What may sentiment analysis of the statements on *gender ideology* tell us about the phrase's use, in terms of negative/positive/neutral attitudes?

4 METHODOLOGY AND CORPORA

In our analysis, we draw on two main corpus resources. The first involves CLASSLA-web, as massive web corpora for Slovenian (CLASSLA-web.sl), Croatian (CLASSLA-web.hr) and Serbian (CLASSLA-web.sr), each comprising around 2 billion words, and including texts published up to 2021 (CLASSLA-

Konferenca
Jezikovne tehnologije in digitalna humanistika
Ljubljana, 2024

Conference on
Language Technologies & Digital Humanities
Ljubljana, 2024

web.sl) and 2022 (CLASSLA-web.hr and CLASSLA-web.sr). Collected by crawling primarily the national top-level internet domains (see https://www.clarin.si/info/k-centre/classla-web-bigger-and-better-web-corpora-for-croatian-serbian-and-slovenian-on-clarin-si-concordancers/), CLASSLA-web is aimed to encompass all the texts written on the web, and thus allows large-scale insights into contemporary language use. Second, we use ParlaMint 4.0 (Erjavec et al., 2023), a multilingual set of comparable corpora containing parliamentary debates from 29 parliaments and covering at least the period from 2015 to 2022. For this analysis, we use ParlaMint's latest versions of parliamentary corpora of Serbian, Croatian and Slovene (ParlaMint-SR 4.0, ParlaMint-HR 4.0, ParlaMint-SI 4.0); given the somewhat differing corpus timespans (e.g. going back to 1997 in SR, and late 2003 in HR), we select the 2004-2022 timespan for our analysis.

Using the noSketchEngine concordancer, we perform a search of the node 'language ideology' in the three language corpora (i.e. rodna ideologija Sr / rodna ideologija Hr / ideologija spola / teorija spola Slo), namely CLASSLA-web-sr, CLASSLA-web-hr, CLASSLA-web-slo, and ParlaMint-SR 4.0, ParlaMint-HR 4.0, and ParlaMint-SI 4.0. Collocation analysis (McEnery and Hardie 2011), is performed with the 5L-5R node span (cf. Baker 2010 on discourse-oriented analyses), and in SketchEngine calculated with the logDice measure. This allows comparisons across the corpus/discourse type (general/Web and Parliaments), and across the three states. Qualitative analysis of the concordance liness for the top-ranking collocates provides more insights into these uses, in line with our perspective on combining the corpus-based approach with contextually situated (critical) discourse analysis, while the perspectives can be additionally enriched with sentiment analysis.

5 FINDINGS OVERVIEW

Looking at the CLASSLA-web corpora, the phrase *gender ideology* is found to have different levels of prominence in the three national/language contexts examined. In Croatia, *rodna ideologija* can be seen to represent an established term, with several thousand occurrences across different genres (2.2 hits per mill. tokens). By contrast, in the Serbian corpus, *rodna ideologija* is far less

Konferenca
Jezikovne tehnologije in digitalna humanistika
Ljubljana, 2024

Conference on
Language Technologies & Digital Humanities
Ljubljana, 2024

infrequent, mostly belonging to scholarly discourse; it appears to be a concept only emerging in the past several years, in relation to educational and gender-sensitive language debates in particular (0.1 hits per mill. tokens). In Slovenia, our analysis confirms that two variants exist in use – '*teorija spola*' and '*ideologija spola*' (cf. Popič & Gorjanc, 2020) – with the first strongly preferred (0.6 hits per mill. tokens, over a thousand occurrences), and the latter more sporadic (less than 0.1 hits per mill. tokens). The social meanings and associations of the phrase become clearer in collocation analysis.

While notable similarities can be observed across the collocate lists for the three sub-corpora, in line with the more global anti-genderist discourses, we can also observe interesting differences when looking at the top collocates. For Croatia, for example, these include a more procedural focus on the Instanbul Convention and its ratification, and varied sexuality related terms (*LGBT*, *LGTB*, *homoseksualnost*); in Slovenia, the top collocates refer to activism, with a more specific political focus on 'leftism' and 'cultural Marxism', in line with some recent observations on the role of memory politics in anti-gender discourses in the country (Kuhar & Shevtsova, forthcoming). For Serbia, we note a mix of academic terms, and those suggesting colonization and outside threat to nation and tradition, whose meanings become clearer in discourse-based concordance analysis.

Turning to the ParlaMint corpus, we find that the phrase has entered parliamentary usage, however at very different levels in the differing countries (again by far most prominent in ParlaMint-HR, and almost entirely absent in ParlaMint-SR). Collocation analysis points to some differing foci, which also show differences from the CLASSLA-web general discourse; still, we note some limitations in tracing and comparing the meanings of emergent terms such as the one in question. For Slovenian, thus, the majority of the collocates are grammatical/function words, of little use for interpreting social meanings. Further, using the affordances of SketchEngine and ParlaMint metadata, our investigation into the gendered aspects of the use of our phrase of interest shows a general preference for the term to be used by male MPs over the female MPs. Sentiment analysis, finally, gives information on the attitudes surrounding the usage of the term, which are unsurprisingly negative; still, evaluating the results we also note some issues in automatic sentiment

Konferenca
Jezikovne tehnologije in digitalna humanistika
Ljubljana, 2024

Conference on
Language Technologies & Digital Humanities
Ljubljana, 2024

assignment which may be especially problematic in corpus analyses of emergent terms, coloured by significant levels of irony and sarcasm.

REFERENCES

Baran, D. (2022). 'Rainbow plague' or 'rainbow allies'? TĘCZA 'rainbow' as a floating signifier in the contestation of Poland's national identity. Gender & Language, 16(3).

Bogetić, K. (2022). Politics of resignification: central and eastern European perspectives on language and gender. *Gender & Language*, 16(3).

Bonfiglioli, C., Kahlina, K., & Zaharijević, A. (2015). Transformations of gender, sexuality and citizenship in South East Europe. In Women's Studies International Forum, 49, 43-47. Pergamon.

Borba, R. (2022). Enregistering 'gender ideology': the emergence and circulation of a transnational anti-gender language. Journal of Language and Sexuality 11(1), 57–79.

Corredor, E. (2019). Unpacking 'gender ideology' and the global right's antigender countermovement. Signs: Journal of Women in Culture and Society 44(3), 613–638.

Erjavec, T.; et al., (2023), Multilingual comparable corpora of parliamentary debates ParlaMint 4.0, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, http://hdl.handle.net/11356/1859.

Gal, S. (2021). Gender and the discursive authority of far right politics. Gender and Language 15(1), 96–103.

Ljubešić, N.; Rupnik, P. and Kuzman, T. (2024). Slovenian web corpus CLASSLA-web.sl 1.0, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, http://hdl.handle.net/11356/1882.

Ljubešić, N.; Rupnik, P. and Kuzman, T. (2024). Croatian web corpus CLASSLA-web.hr 1.0, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042,

Konferenca
Jezikovne tehnologije in digitalna humanistika
Ljubljana, 2024

Conference on
Language Technologies & Digital Humanities
Ljubljana, 2024

http://hdl.handle.net/11356/1929.

Ljubešić, N.; Rupnik, P. and Kuzman, T. (2024). Serbian web corpus CLASSLA-web.sr 1.0, Slovenian language resource repository CLARIN.SI, ISSN 2820-4042, http://hdl.handle.net/11356/1931.

Kuhar, R., & Paternotte, D. (eds.) (2017). Anti-gender Campaigns in Europe: Mobilizing against Equality. Lanham: Rowman & Littlefield.

Kuhar, R., & Zobec, A. (2017). The anti-gender movement in Europe and the educational process in public schools. CEPS journal, 7(2), 29-46.

Popič, D., & Gorjanc, V. (2023). Corpus-Linguistic Analysis of Speech Communities on Anti-Gender Discourse in Slovene. Gender a výzkum/Gender and Research, 23(2), 140-166.