Edited by Darja Fišer and Philippa Smith

THE DARK SIDE OF DIGITAL PLATFORMS:

LINGUISTIC INVESTIGATIONS OF SOCIALLY UNACCEPTABLE ONLINE DISCOURSE PRACTICES

Book series Translation Studies and Applied Linguistics

Ljubljana 2020

THE DARK SIDE OF DIGITAL PLATFORMS: LINGUISTIC INVESTIGATIONS OF SOCIALLY UNACCEPTABLE ONLINE DISCOURSE PRACTICES

BOOK SERIES TRANSLATION STUDIES AND APPLIED LINGUISTICS
ISSN 2335-335X, e-ISSN 2712-3855

Edited by: Darja Fišer and Philippa Smith
Reviewers: Tanja Petrović, Stavros Assimakopoulos
Editorial board: Špela Vintar, Vojko Gorjanc and Nike Kocijančič Pokorn
English language proofreading: Paul Steed
Technical editing: Jakob Lenardič
Layout: Jure Preglau

Published by: Ljubljana University Press, Faculty of Arts
Issued by: Department of Translation Studies
For the publisher: Roman Kuhar, Dean of the Faculty of Arts, University of Ljubljana

Ljubljana, 2020
First edition
Printed by: Birografika Bori d. o. o.
Design: Kofein, d. o. o.
Price: 14.90 EUR
Print run: 150 copies

The publication of the book was funded by the Slovenian Research Agency within the national basic research project "Resources, methods, and tools for the understanding, identification, and classification of various forms of socially unacceptable discourse in the information society" (J7-8280, 2017–2020) and the Slovenian-Flemish bilateral basic research project "Linguistic landscape of hate speech on social media" (N06-0099, 2019–2023).

First e-edition. Digital copy of the book is available on: https://e-knjige.ff.uni-lj.si/
DOI: 10.4312/9789610603061

Kataložna zapisa o publikaciji (CIP) pripravili v Narodni in univerzitetni knjižnici v Ljubljani

COBISS.SI-ID=304562944
ISBN 978-961-06-0307-8

E-knjiga
COBISS.SI-ID=304546304
ISBN 978-961-06-0306-1 (pdf)

Table of contents

TABLE OF CONTENTS

Preface

Few of us could be oblivious to the overwhelming negative behaviours that are now appearing with some voracity online, involving intolerant, abusive and hateful speech. These are not new phenomena, although historically they have been limited to specific off-line and local contexts. However, the advent of digital communication technologies – the anonymity and instantaneity of which coupled with their ever-growing importance as a source of information and communication – has given them an unprecedented boost and a global dimension. Because of this, new interdisciplinary theoretical and analytical methods and approaches are called for to improve our understanding of the shifting patterns of such practices in different parts of the world. In particular, a focus is needed on the ways in which to tackle and respond to the proliferation of these behaviours in the new media, and the radicalization of online space in the contemporary, increasingly multicultural information society.

While intolerant, abusive and hateful speech online has received a lot of attention by researchers in social media and communication studies, its linguistic aspects have yet to be thoroughly investigated. This book contributes to filling this gap by showcasing how a linguistic perspective has much to offer in unravelling exactly what is occurring. The idea for this book emerged from the meeting of researchers selected for a panel "The Dark Side of Social Media" at the 22nd Sociolinguistics Symposium in New Zealand in 2018. With a common goal to interrogate the linguistic aspects of negative online behaviours on different social media platforms and against different targets, these authors approached the phenomenon from a wide range of methodological frameworks. While primarily interested in identifying, describing and understanding intolerant, abusive and hateful speech online thoroughly and comprehensively, they also had a common belief that their work could inform efforts to contain or mitigate the impact of negative online behaviours regardless of where they occur.

As the five quite different, but equally interesting, studies offered in this investigation of the negative behaviours that are manifesting themselves in the digital age go beyond social media, we have extended the scope of the title to encompass the broader landscape of digital platforms. Each chapter interrogates a different communicative practice in different modalities for analysis involving a range of online platforms. It is no coincidence that the majority of this research relates to data accessed via online news websites. While news articles can play a role in the reproduction of negative discourses – sometimes in more subtle and unintentional ways when it comes to the representation of others – the comments sections of online new websites have become popular "hang outs" for networked publics where they can gather and discuss topics of the day. At the same time these sites offer users a ready-made online audience to be on the receiving end of an anonymous commenter's negative discourse or ideology that may involve an attack on or abuse of another individual or community.

Equally, social media platforms such as Twitter, Facebook and YouTube also play significant roles in providing a forum for negative behaviours, particularly when comments can be easily shared and "liked" courtesy of the affordances of digital technologies. While tech companies are attempting to address these issues through their content moderation policies, and some governments are implementing legislation to put a stop to online hate and abuse – negotiating and controlling the "wild west" of the internet is far from straightforward, and the wider implications of such interventions are not well understood either. This is why understanding what is occurring online and how users might respond to negative online behaviours is an important focus for academics and is also addressed in this volume.

The studies that are presented in the following pages offer insights into different quantitative and qualitative approaches when it comes to the complexities of the linguistic analysis of negative online behaviours. They range from corpus-based methods, to content analysis and critical discourse analysis, though some studies also apply a mixed-methods approach. Overall the range of methodological frameworks and theoretical approaches utilized in the explorations into online discourses of racism, misogyny, homophobia/LGBQT+ rights and islamophobia/anti-immigrant discourses presented in this volume contribute towards a comprehensive understanding of the linguistic landscape of online hate. Despite the fact that the studies in this volume analyse data from three very different countries, cultures and languages – Cyprus, Japan and Slovenia – this demonstrates that negative online behaviours are a global problem which cannot be ignored. Therefore, the findings of these studies have much wider implications. We have organized the chapters according to the granularity of the linguistic units investigated in each study, starting with an analysis of orthographic practices and going all the way to culturo-historical contextualization of discourse.

In Chapter One authors Kristina Pahor de Maiti, Darja Fišer and Nikola Ljubešić from the University of Ljubljana and the Jožef Stefan Institute in Slovenia present their findings from a detailed quantitative as well as qualitative analysis of orthographic practices in socially unacceptable comments posted by users on the Facebook pages of three Slovenian mainstream news outlets under articles related to LGBTQ+ issues and migrants. Their results show that nonstandard features of socially unacceptable comments surpass the surface spelling deviations that are commonly found in computer-mediated communication. Their work sheds light on the mechanics of the language used in socially unacceptable comments, and shows that differences between the surface linguistic structure of what is socially unacceptable and acceptable do exist, which is a promising finding that could inform the machine learning community to help curb socially unacceptable discourse practices online.

Chapter Two analyses the discourse of two opposing Twitter accounts which were created for the referendum campaign on LGBTQ+ rights in Slovenia 2015. The authors of the chapter, Vojko Gorjanc and Darja Fišer from the University of Ljubljana, perform comparative content and critical discourse analyses of the two campaigns to uncover the prevalent persistency of heteronormativity and the disruptive perception of any counterforce efforts through a systematic, multimodal utilization of the concepts of "normal" and "natural" as key aspects of gender and sexual identities. Put in a broader socio-political context, the discourses of the referendum campaigns contributed to a heightened sensitivity of the general as well as political public.

Chapter Three presents findings from a study by Zoran Fijavž and Darja Fišer from the University of Ljubljana that examined metaphorical expressions of water flow in online news on the Balkan migration crisis in 2015. They show that metaphorical expressions were ubiquitous in the discourse, both in terms of authorship and timeline. While most frequently found in reported statements, the metaphorical expressions were also used by the journalists directly, which suggests that the metaphorical frame was heavily mediatized and uncritically reproduced by the media.

In Chapter Four Fabienne Baider and Anna Bobori from the University of Cyprus also approach anti-migrantion discourse and use the same theoretical framework, taking an in-depth look at the associated metaphor of sexual threat in their study of an online news article reporting an alleged violent rape in Sweden and the online comments in response to it. This study demonstrates how the discourses surrounding a fake news story can have negative effects on how migrants are perceived by journalists and readers alike. Yet at the same time this investigation of how some commenters attempted to counter these perceptions leads the researchers to propose some recommendations for the creation of effective counter narratives.

The study by Goran Vaage, from Kobe College in Japan, which appears in Chapter Five, concludes the volume by providing a comprehensive culturo-historical and sociolinguistic overview of present-day discourses of organized covert racism as well as overt racist hate speech by the Japanese right-wing organization Zaitokukai that features on YouTube. The study points to the role of digital media in the staging and dissemination of hate speech, but also as a facilitator for such speech outside the virtual realm.

It is anticipated that this book – while focused on studies on the linguistic aspects of socially unacceptable discourse practices online – will also be of interest to others who situate themselves outside of this field because of the contribution it brings in the building of knowledge on topics that also apply in sociology,

anthropology, computer science and internet studies. We also believe that readers beyond academia – governments, organizations and tech companies – will appreciate the content of this issue as they too seek to gain a better understanding of, and solutions to, the impact of the dark side of digital platforms and the role they play in stemming the proliferation of these concerning discourses.

Our work would not have been possible without the dedicated efforts of all the authors who submitted their contributions and without the careful and insightful comments of the reviewers. We would also like to thank the language editor Paul Steed and the technical editor Jakob Lenardič for polishing the manuscripts, and for all the support and good spirits provided by Matevž Rudolf and Jure Preglau from the Faculty of Arts Publishing House.

Darja Fišer and Philippa Smith
Ljubljana, Slovenia and Auckland, New Zealand
11 February 2020

Nonstandard linguistic features of Slovene socially unacceptable discourse on Facebook

***Kristina Pahor de Maiti**, Faculty of Arts, University of Ljubljana, Slovenia*

***Darja Fišer**, Faculty of Arts, University of Ljubljana, Slovenia, Jožef Stefan Institute, Ljubljana, Slovenia*

***Nikola Ljubešić**, Jožef Stefan Institute, Ljubljana, Slovenia, Faculty of Computer and Information Science, University of Ljubljana, Slovenia*

Abstract

Socially unacceptable discourse (SUD) is a topical issue in many areas of social science and humanities, but the analysis of its linguistic dimension is often neglected. This paper explores the surface linguistic features of SUD and aims to determine the level of linguistic standardness of SUD Facebook comments in comparison to its non-SUD counterpart. The analysis was conducted on a dataset extracted from the Slovene part of the FRENK corpus which contains SUD and non-SUD comments from Facebook pages of news outlets on the topics of migrants and LGBT issues. The dataset was manually annotated for nonstandard language features using a custom-built annotation schema. The results of the quantitative analysis show that SUD comments are statistically significantly less standard in comparison to non-SUD comments. Despite the high level of nonstandardness of both subsets of comments, the qualitative analysis proved that nonstandard features of SUD comments surpass the surface spelling deviations that are commonly found in computer-mediated communication.

Keywords: hate speech, socially unacceptable discourse, computer-mediated communication, linguistic standardness, Facebook

1 INTRODUCTION

Human communication has always included discriminatory discourse practices, but it rarely reached a broad public audience. This changed with the development of computer-mediated channels of communication, such as social media. A substantial increase in online hate speech against marginalized groups (immigrants and refugees, the LGBT community, Roma, religious groups, etc.) can now be observed especially in relation to certain (geo)political events (e.g., the EU refugee crisis beginning in 2015) (Motl and Bajt 2016).

In this paper, we use the concept of socially unacceptable discourse (SUD) which goes beyond the harshest forms of legally prosecutable hate speech as it addresses a broader range of practices and encompass a full spectrum of hateful, threatening, abusive and discriminatory speech as well as indecencies and insults. Several authors (Bajt 2016, Ivanou 2017, Tsesis 2002) have argued that not only hate speech but all forms of SUD pave the way towards social actions that can have an aggravating effect on individuals and social groups that are the target of SUD, as well as on the society as a whole. In light of this, it is clear that all manifestations of SUD have the potential to be harmful for the society and should also be treated as such.

Although there is a consensus on the negative effects of SUD on society (Bojarska 2019, Nielsen 2002) and some progress with regard to limiting its proliferation has been made (Jourová 2019),[1] its accurate and timely detection (ElSherief et al. 2018, Vidgen and Yasseri 2019, Zhang and Luo 2019) together with effective prevention of its negative influence (Ullmann and Tomalin 2019) remain a challenge. Moreover, despite its uncontested presence in contemporary society, SUD is still not thoroughly researched. As a societal phenomenon, its scholarly treatment extends over many different disciplines, which is why it necessitates an in-depth analysis in all of the relevant fields as well as a cross-disciplinary scope. For some time now, online hate has received a lot of attention from policy-makers and researchers from various scientific fields, but despite the fact that SUD is nearly always manifested through linguistic means, it is predominantly studied as a societal and psychological phenomenon (Waqas et al. 2019), and even though SUD as a research topic is not unknown to sociolinguistics and discourse studies, it has primarily been treated in the domain of critical discourse analysis, often leaving its linguistic characteristics neglected (Klein 2018). This chapter therefore focuses on a set of surface linguistic features with the aim to identify the general linguistic characteristics of hateful comments in comparison to non-hateful ones. Understanding that the language of computer-mediated communication (CMC language) is strongly characterized by a certain level of nonstandard writing (Crystal 2001), the main challenge of this study will be to distinguish the specific features of SUD from the general

1 For a critique see Bukovská (2019).

characteristics of CMC language. The chapter is an extended version of a short conference paper (Pahor de Maiti et al. 2019) where we presented the initial results of the analysis of surface linguistic features of online hate speech. In this chapter, we elaborate on the presentation of related studies and the dataset preparation process as well as provide a more detailed interpretation of the results.

In the next sections, we will first present the results of related studies (section 2) and the study design (section 3). In section 4, we will move to the presentation and discussion of the results. The final section (section 5) includes an overview of the findings and describes possible next steps for future work.

2 RELATED WORK

2.1 Socially unacceptable discourse

By being a global space that enables immediate and participatory expression of ideas under the guise of anonymous or fake user identity, the internet has, among its many positive impacts, also evolved into a "breeding ground for the phenomenon of *cyberhate*" (López and López 2017) which is produced not only by members of extremist groups but also increasingly by regular internet users. Hate speech disseminates hatred or disqualification of an individual or group based on their race, skin colour, ethnicity, sex, disability, religion or sexual orientation (Nockleby 2000). However, Vehovar et al. (2012) point out that the definition of hate speech should be broadened and also include various forms of offensive speech, such as hurtful, derogatory or obscene comments about someone. In light of this, all such nuances of hate speech should be taken into account as research shows that even the most indirect inappropriate messages can incite harmful action (Baider et al. 2017, Muskat and Assimakopoulos 2017). In this chapter we prefer to apply the term socially unacceptable discourse (SUD) so as not to confuse it with the narrowly defined legally prosecutable hate speech (Banks 2011).

In Slovenia, research on SUD started in the 1990s and has been mainly confined to law and social sciences throughout the course of the early 2000s. From 2006 onwards, we have observed an intensification of research that goes beyond these fields and stretches into those of linguistics and computer science. Slovene linguists have researched SUD in traditional media (cf. Červ and Kalin Golob 2012), language reference books (cf. Gorjanc 2005) and translations (cf. Gorjanc 2012). Until now, however, not many analyses have been conducted on online SUD data and those that have (cf. Čeferin and Mežnar 2014, Jakopič 2013, Merljak 2011) mainly centred around the concept of hate speech, the legal implications of its use and editorial challenges for restricting SUD online. This shows

that the linguistic dimension of (online) SUD is still notably under-researched which motivated us to identify the language features of SUD in order to better understand, detect and prevent undesired communication patterns.

2.2 The language of computer-mediated communication

When investigating the language of online SUD, it is important to consider the general characteristics of CMC language in order to avoid misattribution of their general characteristics to SUD. CMC language is well known to include the use of unconventional spelling and often integrates informality and deviations from the norm on the level of grammar and punctuation (Crystal 2001). CMC language is also characterized by a number of features used by writers in order to overcome the technical limitations of both the communication platform and device used, as well as to fulfil the social need for a quick response (e.g., non-canonical text abbreviation strategies (Bieswanger 2013)), and to fill in for the limited affordances for nonverbal communication (e.g., emojis/emoticons and expressive punctuation (cf. Androutsopoulos 2011)).

Length and time restrictions are often tackled with shortening strategies, which may be realized on different levels of linguistic description, and are language-specific. For Slovene, Goli et al. (2016), who investigated Twitter messages, found that shortening is commonly used on the orthographic and lexical levels, as well as on the syntactic level, where it usually includes omissions of the auxiliary verb *to be*. Emoticons and emojis[2] are a popular feature of CMC language, and from a pragmatic standpoint they fulfil multiple communicative functions. Dresner and Herring (2010) found that emoticons and emojis can be used (1) as emotion indicators mimicking non-verbal communication, (2) as elements conveying non-emotional messages (e.g., a joke), and (3) as elements indicating the illocutionary force of the message (e.g., mitigating or reinforcing the content), especially when used in place of the final punctuation (Amaghlobeli 2012).

Due to certain communicative contexts, such as the instantaneous and informal nature of CMC, language of CMC also exhibits features that were traditionally attributed to speech (cf. Androutsopoulos 2011, Baron 2002). Zwitter Vitez and Fišer (2018) found that elements of modality, used to express opinion, judgement or certainty, represent the strongest connection between spoken discourse and language in Slovene online comments. A connection to spoken language has also been found by Goli et al. (2016) who argue that shortening on the orthographical level reflects the tendency to use phoneticized spelling.

2 An emoticon is a representation of a facial expression with a combination of keyboard characters (e.g., :->), whereas the an emoji represents a facial expression/emotion, or another notion/object in the form of a symbol, icon or a picture.

Beyond technical and communicative factors, research has already proved that socio-demographic factors, such as personality, social status, level of education and age, also play an important role in CMC and impact how people express themselves (Gill 2011, Hilte et al. 2018). Given that SUD is a highly subjective type of discourse, we can assume that personality traits significantly influence this type of discourse as well. The role of personality has been shown in relation to the use of shortening strategies. It was found that shortening is caused not only by the length limitations of some online platforms, but also exhibits the societal and emotional aspect of the message, conveying signs of authorship, community identity or illocutionary force (Goli et al. 2016). Similarly, emoticons/emojis reflect the author's profile on many different levels: they are not only a vehicle to express the emotions of the producer, but also fulfil other pragmatic functions, such as author's intentionality (Spina 2017).

2.3 Linguistic standardness

The notion of standardness defines the optimal form of the written language used for (mass) communication in the public sphere (Skubic 2005). However, private and public communication are closely intertwined on the internet, which causes difficulties when defining the necessary level of message formality, as CMC can be used for many purposes on the long axis from highly formal to fully profane contexts (Splichal 2017). The analysis of nonstandard writing is especially interesting for Slovene due to its highly prescriptive linguistic culture. In this chapter, we understand the notion of linguistic standardness in proportion to the level of author's compliance with the linguistic norm that is prescribed by the normative orthographic and grammar guides. The nonstandard writing practices of Slovene CMC language have been extensively studied in the JANES project[3] and the analysis of Slovene tweets has shown that nonstandard writing (especially on the orthographic level) is common in informal communication on social media (Fišer et al. 2018).

2.4 The characteristics of socially unacceptable discourse

It has been shown that SUD is a frequent phenomenon in online comments on news portals and social media (cf. Lewandowska-Tomaszczyk 2017). The intensity of SUD in online comments actually reinforces proportionately to the discourse used in the initial news post (Bajt 2016), showing the great impact of media discourse on the discussions on online forums (cf. Gorjanc 2005)

3 The JANES project involved the building of a large corpus of Internet Slovene. See: http://nl.ijs.si/janes/english/.

and consequently the moral responsibility of the media in the publication and administration of news posts.

Given that SUD is more often expressed covertly than overtly, it is of key importance that analysis of SUD is not limited only to the level of lexis (e.g., presence of swear words). Muskat and Assimakopoulos (2017) showed that negative opinions are frequently expressed indirectly, which could be due to commenter's need for face-saving and the fear of appearing intolerant. In a similar vein, Červ and Kalin Golob (2012) found that authors of SUD, knowing that their message violates communicative norms, intentionally remove explicit negative evaluations and instead express the intended meaning through irony, rhetorical questions or punctuation marks. Furthermore, an interesting observation has been made by Zwitter Vitez and Fišer (2016), who found that negative comments actually have a more standard orthography, are longer and have a more complex syntax structure than positive comments. They argue that this is linked to the author's desire to present their negative opinion in a neutral linguistic form in order to be acceptable to a broader public (i.e., not focused on a specific socio-demographic community). This could also be explained by the author's tendency to appear as rational and competent as possible given that poorly written texts by the general public are often perceived as a sign of intellectual primitivity, lack of education and low social status, as well as an evidence of the lack or inability to think logically (Jessmer and Anderson 2001, Neustupný and Nekvapil 2003).

3 STUDY DESIGN

3.1 Research questions and hypotheses

The aim of our analysis was to investigate the length, lexical diversity and linguistic standardness of Facebook comments in order to establish whether any specific linguistic characteristic can be observed in hateful comments compared to the features that are typical of CMC in general. Therefore, our research questions and hypotheses for investigation are:

1. Research question 1: Comment length
    1.1. Hypothesis 1.1: SUD comments are shorter than non-SUD comments.
2. Research question 2: Lexical diversity
    2.1. Hypothesis 2.1: Vocabulary diversity is larger in non-SUD comments compared to SUD comments.
    2.2. Hypothesis 2.2: Non-SUD comments contain more emoticons and emojis than SUD comments.

3. Research question 3: Linguistic standardness
   3.1. Hypothesis 3.1: Punctuation to non-punctuation ratio is higher in SUD comments.
   3.2. Hypothesis 3.2: SUD comments are linguistically less standard than non-SUD comments.

3.2 Dataset

In this chapter, we used the Slovene part of the FRENK corpus which contains 6,545 and 4,571 comments about migrants and LGBT issues, respectively, that were posted in response to posts on the Facebook pages of the three Slovene mainstream news media with the most visited websites according to the Alexa service:[4] 24ur.com,[5] SiOL.net.Novice,[6] Nova24TV[7] (Ljubešić et al. 2019).

The FRENK corpus is annotated according to a project-specific two-level annotation schema that identifies the type (see Table 1) and target (see Table 2) of SUD. Comments that do not include socially unacceptable discourse are marked with the "Acceptable speech" label. The rest are assigned two-dimensional labels indicating the type of socially unacceptable discourse and its target. Comments targeting groups/individuals on the basis of their religion, gender, sexual orientation, ethnicity, race, etc. are annotated as "Background". If SUD is aimed at individuals due to their particular group affiliation (professional or political affiliation, etc.), the "Other" category is selected. Comments without any specific target, but that nonetheless contain uncivil language, are marked as "Inappropriate speech".

Table 1: Type of SUD in the FRENK corpus.

Type of SUD	Background – violence
	Background – offensive speech
	Other – threat
	Other – offensive speech
	Inappropriate speech

After the type of SUD is determined, the annotator identifies against whom the comment is directed.

4 https://www.alexa.com/topsites/countries

5 https://www.facebook.com/24urcom

6 https://www.facebook.com/SiOL.net.Novice

7 https://www.facebook.com/Nova24TV

Table 2: Targets of SUD in the FRENK corpus.

Target of SUD	Migrants/LGBT (comments targeting the specific social group)
	Related to migrants/LGBT (comments targeting individuals/groups that express a positive attitude towards the specific social group)
	Journalist/media (comments against the author of the news post or against the news media hosting the news post)
	Commenter (comments against authors of preceding comments to the same news post)
	Other (comments targeting individuals/groups that do not belong to previous categories or express negative attitude towards migrants/LGBT)

For our analysis, we extracted all 520 comments in the FRENK corpus which contained elements of violence and threat regardless of their target for both topics (migrants: 417 comments; LGBT: 103 comments), and a randomized sample of 520 comments (respecting the same share per topic) labelled as "Acceptable speech". The dataset was verticalized, morphosyntactically tagged and lemmatized with the ReLDI tagger (Ljubešić and Erjavec 2016).

3.3 Typology of nonstandard linguistic features

In order to observe the nonstandard linguistic features in SUD systematically, we developed an annotation schema for the tagging of nonstandard linguistic features in Facebook comments, based on the guidelines for normalizing CMC language (Čibej et al. 2016) and the Slovene Normative Orthography Guide (Toporišič 2007). Slovene has a strong prescriptive tradition that stretches far beyond orthography and also covers grammar and lexis. Our definition of standardness is based on the codified standard regarding spelling, lexis and grammar as set forth in the Normative Orthography Guide, but takes into account the specific communicative context of comments under investigation, especially when considering lexis (e.g., offensive words, typical for SUD comments on social media, were not annotated as nonstandard lexis). The annotation schema thus consists of five categories: orthography, lexis, morphology, syntax and word order (see Table 3). The subcategories describe the range of a certain category. They are only meant as support for a precise and reproducible annotation process and do not represent an additional annotation layer. In the initial phase, we tested our annotation schema on a smaller sample in order to consolidate the range of categories and set the categorization process (described in detail in section 3.4). The final annotation schema differed from its first version mainly in a higher number of subcategories (that were not anticipated from the related work on CMC writing practices).

Table 3: Annotation schema of nonstandard linguistic features in CMC language.

Category	Description of a subcategory	Example *ex.* → *standard* /English/
Orthography (O)	Incorrect use of lower/upper case	/…/ *atlantika* → ***A**tlantika* /Atlantic/
	Incorrect punctuation and spacing	*Za vse tiste ki pravijo* → *Za vse tiste**,** ki pravijo* /For all those who say; missing comma/ *nebi* → *ne bi* /should not/
	Typographical errors	*stslišče* → *sta**l**išče* /opinion/
	Regional transformations of standard lexis	*kuj* → *takoj* (immediately)
	Character flooding	*BRAVOOOOOOOOOOO*
	Omission of diacritics	*ce* → ***č**e* (if)
	Alphanumeric words and creative spelling	*Kolikox* → *Koliko**krat*** (How many times)
	Phoneticized spelling	*nč* → *nič* (nothing)
Lexis (L)	Content words from dialects & slang	*lih* → *ravno* (temporal just)
	Nonstandard abbreviations & acronyms	*not found*
	Words in foreign language	*gamad* → *golazen, mrčes* (vermin in Croatian)
	Semantically inappropriate words	*mogli* → *morali* (could instead of should)
Morphology (M)	Erroneous verb/noun affixes	*sprejom* → *sprej**em*** (with spray)
	Incorrect grammatical gender/ number/aspect	*elektrosokerje* /…/ *ali so namenjena samo za nas* → *elektrosokerje* /…/ *ali so namenjen**i** samo za nas* (electroshocker … or are they meant only for us)
Syntax (S)	Incorrect use of grammatical cases	*pod vzglavniku* → *pod vzglavnik**om*** (under the pillow)
	Incorrect use of definiteness	*taglavne* → *glavne* (the main; vernacular particle “ta” added)
	Syntactic ellipsis not justifiable by the context or clearly non-neutral	*pravijo da jih treba ubit* → *pravijo da jih **je** treba ubit* (they say that they need to be killed)
	Inappropriate parts of speech	*noben* (pronoun) → *nobeden* (noun) (no one)
	Nonstandard structures	*na vsake toliko kvatre* → *na vsake toliko* OR *na vsake kvatre* (from time to time; tautology)
Word order (W)	Nonstandard/non-neutral	*gamad voditeljska* → *voditeljska gamad* (vermin leaders)

Our annotation schema does not strictly follow all the codification rules (e.g., we eliminated the subcategory for incorrect use of the supine),[8] but has been designed to serve a dual purpose: (1) to establish whether there are any differences in the level of text standardness between SUD and non-SUD comments, and (2) to identify possible idiosyncrasies of SUD comments. In addition, although some features (e.g., marked word order) might not be traditionally classified as strictly nonstandard (e.g., because they simply fulfil a specific rhetorical function), they were still annotated as nonstandard due to their frequency observed during the annotation test phase.

3.4 Annotation

Manual annotation of the dataset was performed by one annotator who used the following reference guides: SSKJ – Dictionary of the Slovenian Standard Language;[9] Pravopis – the normative orthography guide for Slovene;[10] Slovene grammar (Toporišič 2004); Slovene lexicon SloLeks (Dobrovoljc et al. 2019);[11] Janes-Norm corpus (Erjavec et al. 2016). During the test annotation, we created internal guidelines in order to ensure consistency, which proved especially useful for disambiguation of borderline cases:

1. A token is defined as nonstandard if it can be assigned to at least one category in the typology.
2. The study uses a one-dimensional annotation schema, meaning that even if a token could be attributed to several subcategories of a single category, that category is annotated only once (e.g., *mas* → *imaš* (have) was labelled with the O-category label only once, despite the fact that the token is missing a diacritic (i.e., the first signal for the O-category label) and exhibits a phoneticized spelling (e.g., the second signal for the O-category label).
3. If the form of a token corresponds to multiple categories on the basis of multiple elements, all of the appropriate categories are indicated (e.g., *slovencov* → *Slovencev* (Slovenes) was labelled with the O- and M-category: O-category label for lower case used instead of the upper case, and M-category label for the incorrect suffix).

8 An impersonal verb form used predominantly with verbs of motion instead of which the infinitive is frequently misused, e.g., *go see* → *pojdi pogledat* (correct) vs. *pojdi pogledati* (incorrect). An explanation for this decision is given in section 3.4. Annotation.

9 Dictionary of the Slovenian Standard Language: https://fran.si/130/sskj-slovar-slovenskega-knjiznega-jezika

10 Normative orthography guide: https://fran.si/

11 Slovene morphological lexicon: http://eng.slovenscina.eu/sloleks

4. If the form of a token corresponds to multiple categories on the basis of the same element (e.g., missing the final letter), the O-category label is prioritized. Despite slightly skewing the final results (by boosting the O-category), this approach ensures consistency in the annotation process and still returns realistic results with regard to the level of overall non-standardness of SUD. This approach pertains especially to the following two types of examples:
    4.1. a token representing an infinitive verb which is missing its final letter (e.g., *treba /…/ kastrirat → treba/…/ kastrirati* (needs castrating)) which can be indicative of the incorrect use of supine (therefore the S-category) or of a typographical mistake (therefore the O-category). Tokens of this type were annotated with the O-category label which enabled us to refrain from automatically categorizing this feature as a grammatical mistake and rather assume it was an unintentional mistake.
    4.2. Similarly, the spelling variants of pronouns (e.g., *sma → sva* (we are)) possibly taking the O-category label (due to typographical mistake in an otherwise standard word) or the L-category label (due to it being a regional variant), were again placed under the O-category, so there was no need to differentiate between a typographical mistake and dialectal transformation.
5. We annotated any missing, redundant, excessively repeated or incorrect (combination of) punctuation marks. Spacing, however, was annotated only on the level of words (erroneously written together or apart, e.g., *nebo → ne bo* (will not)) and not on the level of punctuation marks due to the technical limitations that originate from the text verticalization process.
6. In case an emoticon/emoji appeared at the end of the sentence, it was treated as a final punctuation mark (cf. Pertot et al. 2016) mentions and emotional expressions – emojis and emoticons.
7. Word order and syntactic constructions deviating from the norm were indicated with only one W/S-category label at the position of the first token in the nonstandard phrase/sentence.
8. Since Slovene has a free word order and syntactic ellipsis can be interpreted in many ways, we annotated only clearly incorrect and non-neutral word order and omissions (e.g., missing auxiliary verb).
9. Comments written entirely in a foreign language were out of scope of this study.

4 RESULTS AND DISCUSSION

The following sections present the findings of our analysis based on the quantitative results that were gathered relating to comment length, lexical diversity and linguistic nonstandardness of SUD comments. Next, we present insights gathered from qualitative analysis regarding nonstandard orthographic and syntactic features, nonstandard word order, and the use of emoticons and emojis in SUD comments. Our discussion relates these findings to the research questions and hypotheses.

4.1 Basic statistics

Our dataset comprises a total of 19,091 tokens which are divided equally between SUD and non-SUD comments (see Table 4). For our analysis we extracted only the relevant tokens (18,103) and removed all irrelevant ones (988; comments written entirely in a foreign language).

Table 4: Structure of the dataset (number of tokens).

	SUD	Non-SUD	Total
Nonstandard	2,925 (30%)	1,842 (22%)	4,767 (26%)
Standard	6,683 (70%)	6,653 (78%)	13,336 (74%)
Total	9,608	8,495	18,103

4.2 Quantitative analysis

4.2.1 Comment length

The median for comment length, which was calculated by taking into account all tokens in the comments, was 12 tokens per comment for the SUD subset and 11 for the non-SUD subset. This result shows only a minor difference between the subsets, which is not statistically significant. Therefore, due to the similarity between both subsets, we can reject Hypothesis 1.1, which stated that SUD comments are shorter than non-SUD comments as a result of an immediate, emotional response to a newspaper article. Our findings also do not support the finding of Zwitter Vitez and Fišer (2016) that negative comments are longer in comparison to positive ones.

4.2.2 Lexical diversity

First, we calculated the type to token ratio (TTR) for each type of discourse over 100 random draws of 1,000 tokens. TTR is slightly higher for SUD comments (0.61) in comparison to non-SUD comments (0.58). Second, we calculated the content-to-function-word ratio for SUD comments, which was 1.32, and thus again slightly higher compared to that for non-SUD comments, which was 1.25.

These results show that vocabulary diversity is larger in SUD comments and therefore rejects Hypothesis 2.1, which stated the opposite. Assuming that SUD comments are more expressive and emotional, this outcome may not be surprising, given that people tend to use more colourful and creative language for emotionally-charged content. The lexical characteristics of hateful speech are addressed in detail by Franza et al. (2019), but a quick comparison of SUD and non-SUD nouns referring to a person in our dataset shows nine offensive nouns (*idiot, vermin,* …) vs. one general noun (*human*). This limited but meaningful insight directly supports our finding on lexical diversity and the hypothesis about higher language expressivity.

Next, we calculated the relative frequency of emoticons and emojis, which was 0.005 for SUD comments and 0.009 for non-SUD comments. We also counted the number of different emoticons and emojis. SUD comments contained 24 different emoticons and emojis while non-SUD contained 34, 35% of which overlap with those found in SUD comments. To test whether the occurrence of emoticons and emojis in non-SUD comments in comparison to SUD comments is significantly higher, we ran an approximate randomization test with 1,000 iterations, obtaining a p-value of 0.0008. This means that the probability of obtaining the same or greater difference between the two types of comments randomly is below 0.001. Therefore, we can safely discard the null hypothesis that there is no difference between the usage of emoticons and emojis in SUD and non-SUD comments and confirm Hypothesis 2.2, which stated that there are more emoticons and emojis in non-SUD comments.

Less frequent use of emojis in SUD could be explained by the lack of available emoticons and emojis or the relative difficultly of accessing more specific emojis through the emoji keyboard (Bočková 2019) which can be perceived as too time-consuming during the creation of an emotionally-charged comment. In addition, the use of emojis could also be influenced by a communication strategy of the author, with which they try to achieve emotional detachment from the content of the comment. With the absence of emoticons/emojis, the emotional expressivity of the comment is lowered and the comment could be perceived as less emotional/more reasonable and thus more cogent, especially because – as argued

by Micciche (2007, as cited in Laflen and Fiorenza 2012) – emotions can be naïvely perceived as the opposite of reason. Both subsets contain more emojis than the more traditional emoticons but the overlap of the latter is bigger between the datasets. This is not unexpected, as different OS/application providers offer different sets of emojis (with greater expressiveness) in contrast to emoticons, which are limited to the keyboard characters. In addition, the user needs to possess some knowledge to be able to create emoticons, whereas emojis only need to be picked out of the proposed set. Moreover, it has been found that emojis are gradually substituting emoticons while fulfilling the same functions, which is generally message strengthening (Pavalanathan and Eisenstein 2015, Pertot et al. 2016).

4.2.3 Linguistic nonstandardness

By counting all punctuation marks (on the token level) versus all other tokens, we obtained a punctuation-to-non-punctuation ratio of 0.09 for SUD and 0.12 for non-SUD comments. This rejects Hypothesis 3.1, which assumed more punctuation marks in SUD comments due to a possibly higher expressiveness of such comments. The result is not surprising, as non-hateful comments are not necessarily all neutral. We suggest, therefore, that it would be useful to also take into account the sentiment of the comments in future research.

As Table 4 above showed, a total of 4,767 nonstandard elements were identified in the dataset. The share of nonstandard tokens in the non-SUD subset is 22%, whereas the SUD subset contains 30% of nonstandard tokens. In Table 5, the percentage of nonstandard features in SUD and non-SUD comments can be observed with regard to their type (e.g., spelling mistakes were put into the O-category (marked with O), i.e. nonstandard orthography). It should be noted that some tokens have been classified into more than one category (e.g., a token with incorrect spelling (O – Orthography) and grammatical case (S – Syntax)) which is why the final number of nonstandard tokens is higher than indicated in Table 4.

Table 5: Amount of nonstandard tokens per category in the dataset.

	O	L	M	S	W
SUD	2,414 (82%)	298 (10%)	44 (2%)	384 (13%)	156 (5%)
Non-SUD	1,659 (90%)	104 (6%)	17 (1%)	87 (5%)	44 (2%)

Table 5 shows that by far the most prominent category in both types of discourse is Orthography, with slightly over 80% of the annotations in SUD and 90% in

non-SUD comments. The rest of the categories are much less frequent in both types of comments: Syntax represents 13% of the annotations in SUD comments and 5% in non-SUD, the Lexis category was assigned to 10% of SUD comments and 6% of non-SUD, Word order had a 5% share in SUD and 2% in non-SUD, and Morphology 2% in SUD and 1% in non-SUD.

The result of the chi-square test (X^2 (1, N = 18,103) = 178.4, p = 0.0001) on the independence of the variables of linguistic standardness and the social acceptability of the comment showed that we can reject the null hypothesis on the independence of the variables and accept the alternative hypothesis that these two variables are actually dependent. Based on these results we can confirm Hypothesis 3.2, that SUD comments are more nonstandard than non-SUD comments. We should point out however that the study by Zwitter Vitez and Fišer (2016) to a certain extent contradicts our findings on SUD nonstandardness. Assuming that SUD comments are predominantly negative, then according to this earlier study (ibid.) the SUD comments should have been more standard than the non-SUD comments, for which we assume positive or neutral charge. Not knowing the actual sentiment distribution of our comments makes it difficult to reach a conclusion on this matter, but this certainly opens up an interesting next step in our future work.

While the prevalence of the O-category in both subsets was not unexpected, as "CMC language is prototypically known for the use of unconventional, nonstandard spelling" (Verheijen et al. 2017), it is interesting that all the other categories were twice as frequent in the SUD-subset compared to non-SUD comments. This could indicate that nonstandard features in SUD comments are more profound and go deeper than the surface spelling deviations, which are typical of CMC language in general. Another interesting observation is the very low number of irrelevant comments (i.e., those written in languages other than Slovene) throughout the dataset (≤ 1%), except in non-SUD migrants-related comments where the share of irrelevant tokens was 14%. While this is not the focus of our analysis, the fact that authors of non-hateful speech convey their message in different languages could indicate their closer connection with other cultures, which could also be a reason for their more inclusive stance.

4.3 Qualitative analysis

The quantitative analysis showed that differences between SUD and non-SUD comments exist on the level of lexical diversity and standardness, but not on the level of comment length. Following our aim to identify linguistic characteristics of SUD, we proceed with a qualitative analysis of certain aspects that stood out during the annotation process and analysis. In this section, we present the results

of the qualitative analysis that offer solid starting points for future work leading to identification of SUD-specific features.

4.3.1 Nonstandard orthographic features

Our one-dimensional annotation schema does not allow us to further break down the main categories, but looking into the O-category of the SUD subset, we noticed that nonstandard punctuation stands out as an important nonstandard feature, with nonstandard punctuation marks representing 5% of all O-category labels. This mainly includes multiplication of punctuation, such as commas (e.g., ,,,,,,,,,,,,,,,,,,), nonstandard combinations, such as full stops and exclamation marks, (e.g., ..!!!) and frequent use of duplicated punctuation, such as full stops (e.g., ..) which is nonstandard despite its popular use on social media. These elements can also be found in non-SUD subset but are less frequent there. The difference might be due to a higher expressivity of SUD comments, as it has been proved that emotions can also be communicated through expressive punctuation marks (cf. Osrajnik et al. 2015).

4.3.2 Nonstandard syntactic features

In the S-category, we observed a frequent use of informal syntactic structures in SUD comments (in bold):

1. colloquial use of the particle *za* and preposition *od*: *za pred zid ste* ***za****postavit* (they should be lined up against the wall*); ni več kaj* ***za*** *razmišljat (*no need to think any further*); sranje* ***od*** *LGBT (*LGBT are a piece of shit*);*
2. colloquial structure with verb *to give*: ***dajte*** *mi nekaj povedat (*do tell me something);
3. syntactic ellipsis which can be a grammatical error (a missing auxiliary verb *to be*; now added between the slashes): *da jih* ***/je/*** ***treba*** *ubit (*they should /be/ killed*); **treba /je/** spustiti elektriko* (we should let the electricity flow);
4. stylistically non-neutral, nonstandard structure (infinitive structure in the place of auxiliary and predicative verbs; now added between the slashes): ***/ treba/*** *jih* ***/je/*** *kaznovati ali v zaport strpati (*they /should be/ punished or put in jail*); mine* ***/je treba/*** *postavt (*land mines /should be/ planted*).*

According to similar studies on spoken language and shortening phenomenon (cf. Goli et al. 2016, Zwitter Vitez and Fišer 2018), we can argue that the first

three items above (i.e., item 1: colloquial phrase formation,[12] item 2: interaction words, item 3: missing auxiliary verb) represent general CMC language characteristics rather than SUD-specific features, and suggest similarities between CMC in general and spoken language. In contrast, item 4 seems to highlight SUD-specific structures. The examples are especially interesting because they show how SUD comments are often formed as short, powerful calls to action or/and impersonal, infinitive structures which could indicate the author's desire to omit the performer of the action, be it themselves or their government. A similar observation was also reached in other studies (cf. Goli et al. 2016) which showed that shortening can be used to underline the specific illocutionary force of the message.

4.3.3 Nonstandard word order

Despite the fact that the syntax of Slovene language is characterized by free word order, certain placements of words are perceived by Slovene speakers as non-neutral. The analysis showed that such word order was used in more than 70% of the W-category tokens in both subsets. A closer look reveals that this is mainly due to the verbs that are placed at the end of the sentence. There were 45% of such cases in the SUD subset and 61% of such cases in the non-SUD subset (non-neutral position, i. e. final position of the verb in bold):

1. *take kot si ti bi jest na garmadi* ***zazgal*** (I would burn people like you at the stakes); *ste* ***se*** *drugače* ***obnašal*** (you behaved differently); *dam si jih* ***peljite*** (take them home); *v zemljo* ***se zabij*** (bugger off).

In the SUD comments, frequent non-neutral constructions of swear phrases with the noun preceding the adjective were also observed (non-neutral position of the adjective following the noun in bold):

2. *golazen* ***pedrska*** (fag vermin); *vlada* ***naša zablojena*** (our stupid government); *golazen* ***necloveska zblojena*** (stupid inhuman vermin); *vsiljivce* ***teroristične*** (terrorist intruders).

On the one hand, these examples once again point to the similarities of the studied comments with spoken discourse, while on the other they are indicative of the emphasizing role of non-neutral word order. Taking into account that in Slovene the most important information comes last, we could deduce that by placing a verb at the end the author puts the emphasis on the action, whereas in the case

12 With the exception of the colloquial swear phrase formation which is by its nature typical of SUD.

of the adjective at the end of a sentence, the author probably wants to draw the reader's attention to the quality of the headword they modified and use it as a justification for the overall meaning of the comment.

4.3.4 Emoticons and emojis

Emoticons and emojis are not only emotion carriers, but are also used to convey the illocutionary force of the message (Dresner and Herring 2010). We first observed a well-known phenomenon of positive emoticons/emojis being used in SUD comments. The pragmatic function of this strategy is to weaken the illocutionary force of the message (Li and Yang 2018). For example:

3. *metek v glavo:)* (bullet to the head:))*; noter bi jih zaprli pa se naj kurijo ;)* (let them be shut inside and burn ;))*; lahko jim vržem samo ročno granato :P* (all I can do is to throw them a hand grenade :P).

Second, we analysed the overlapping emoticons/emojis, which were found to be of two kinds: thematic symbols (🐶) and facial expression symbols (:D). Negative symbols prevail in SUD comments (:-(, 🔫, 😈, ☹, etc.) while there are only three such symbols in non-SUD comments (:(, 😓, :/). Although positive symbols are used in both subsets, non-SUD comments contain many more different symbols for happiness. We also noted the following subset-specific symbols: in SUD comments we found two symbols of physical and social power: 💪 and 😎, whereas in non-SUD comments we could observe symbols of love and peace: <3, :* and ✌. These findings are not surprising per se as, in general, they underlie the content of the comments, as expected. However, the frequency, diversity, distribution and uniqueness of emoticons/emojis in a certain subset can be used as a good orientation point for further analysis of the underlying beliefs that are used as a basis for arguments in the comments.

5 CONCLUSION

This chapter looked into the basic surface linguistic features and the level of standardness of SUD comments in relation to non-SUD comments from the Slovene part of the FRENK corpus (Ljubešić et al. 2019). The analysis was based on a custom-built annotation schema which was used to manually annotate the dataset. The schema with its well-defined categories and dedicated guidelines is compact enough to ensure accurate annotations, and as such proved suitable for the scope of our analysis. In the future we suggest that the dataset should be

extended with multiple annotators to establish inter-annotator agreement. In addition, a multi-dimensional schema would be useful in order to obtain more detailed insights into the categories.

We conducted a quantitative analysis which revealed statistically significant differences between SUD and non-SUD comments on the level of lexical diversity and nonstandardness, but not on the level of comment length. The analysis showed that SUD comments are of similar length to non-SUD comments, but that SUD comments exhibit greater vocabulary diversity. Furthermore, we observed a lower frequency of emoticons/emojis and punctuation marks in SUD comments. The results also showed that SUD comments are indeed linguistically less standard than non-SUD comments, even though this feature did not prove as characteristic as initially expected, since the nonstandardness of non-SUD comments was also relatively high (30% vs. 22% respectively). SUD comments however exhibit a peculiar tendency towards nonstandard features (namely deviations in syntax and word order conventions) that surpass simple spelling errors (which are typical of CMC in general) and go deeper into the language structure.

In the qualitative analysis we further investigated the most striking features observed during the annotation process and quantitative analysis. The qualitative analysis showed that nonstandard punctuation marks are a characteristic feature of our entire dataset, but that expressive punctuation is more common in SUD comments. We also observed frequent non-neutral and informal syntactic structures and word order. On the one hand, this points to previously established similarities with spoken language, but on the other hand, close reading of those examples also showed that these elements have an emphasizing role in the comments and point to the tendency of the author to use them in order to justify or detach themselves from the content. Lastly, our analysis regarding the use of emoticons and emojis returned results that support findings from related studies. We found that emoticons and emojis are used to support the content and commonly also to weaken the illocutionary force of the message. The use of unique thematic emojis are indicative of the underlying beliefs held by the authors of the comments, and can be a good starting point for further sociolinguistic analysis of SUD.

With our analysis, we were able to address the surface linguistic features, which are an under-researched aspect of SUD, and thereby gain insights into the mechanics of the language used in SUD comments. We showed that differences between SUD and non-SUD exist in the surface linguistic structure, which is a promising finding that could assist machine learning and help better curb SUD online. However, our results can primarily serve as an orientation point due to the scope of the analysis, as indicated at the beginning of this section.

There are a number of areas following this study where we have been able to identify opportunities for future work. We envisage the examination of the variety of punctuation marks used and the characteristics of final punctuation to be one area. We also wish to further investigate the semantics of positive and unique thematic emoticons/emojis in SUD comments, and the semantic role of syntactic ellipsis and impersonal structures. Valuable insights could also be gained by investigating the syntactic complexity, vocabulary and argumentation strategies in SUD comments. Furthermore, it would be interesting to test if the identified SUD features help improve the automatic classification of SUD, as this would contribute to helping flag potential inflammatory discourse that may lead to aggravating effects. For comparative and validation purposes, we also intend to extend the analysis to the English subset of the FRENK corpus.

Acknowledgements

The work described in this chapter was funded by the Slovenian Research Agency within the national basic research project "Resources, methods, and tools for the understanding, identification, and classification of various forms of socially unacceptable discourse in the information society" (J7-8280, 2017–2020) and the Slovenian-Flemish bilateral basic research project "Linguistic landscape of hate speech on social media" (N06-0099, 2019–2023).

References

Amaghlobeli, N. (2012): Linguistic features of typographic emoticons in SMS discourse. *Theory and Practice in Language Studies*, 2 (2): 348–354.

Androutsopoulos, J. (2011): Language change and digital media: A review of conceptions and evidence. In N. Coupland and T. Kristiansen (eds.): *Standard Languages and Language Standards in a Changing Europe*: 145–161.

Baider, F. H., Assimakopoulos, S., and Millar, S. L. (2017): Hate Speech in the EU and the C.O.N.T.AC.T. Project. In F. H. Baider, S. Assimakopoulos, and S. L. Millar (eds.): *Online Hate Speech in the European Union – A Discourse-Analytic Perspective*: 1–6. Cham: Springer.

Bajt, V. (2016): Anti-immigration hate speech in Slovenia. In N. Kogovšek Šalamon and V. Bajt (eds.): *RAZOR-WIRED. Reflections on Migration Movements through Slovenia in 2015*: 51–60. Available at: http://www.mirovni-institut.si/wp-content/uploads/2016/03/Razor_wired_publikacija_web.pdf (28 January 2020).

Banks, J. (2011): European regulation of cross-border hate speech in cyberspace: The limits of legislation. *European Journal of Crime, Criminal Law and Criminal Justice*, 19 (1): 1–13. Available at: https://doi.org/10.1163/157181711X553933.

Baron, N. S. (2002): *Alphabet to email: How written English evolved and where it's heading*. London: Routledge.

Bieswanger, M. (2013): Micro-linguistic structural features of computer-mediated communication. In S. Herring, D. Stein, and T. Virtanen (eds.): *Pragmatics of computer-mediated communication*, vol 9: 463–485. Berlin: De Gruyter.

Bočková, R. (2019): *The Use of Punctuation, Emoji and Emoticons in YouTube Abusive Comments*. Bachelor's thesis. Prague: Charles University, Faculty of Arts.

Bojarska, K. (2019): *The dynamics of hate speech and counter speech in the social media – summary of scientific research*. Centre for the Internet and Human Rights (CIHR). Available at: https://cihr.eu/wp-content/uploads/2018/10/The-dynamics-of-hate-speech-and-counter-speech-in-the-social-media_English-1.pdf (28 January 2020).

Bukovská, B. (2019): *The European Commission's Code of Conduct for Countering Illegal Hate Speech Online – An analysis of freedom of expression implications*. Transatlantic High Level Working Group on Content Moderation Online and Freedom of Expression. Available at: https://www.ivir.nl/publicaties/download/Bukovska.pdf (28 January 2020).

Čeferin, R., and Mežnar, Š. (2014): Online hate-speech and anonymous internet comments: How to fight the legal battle in Slovenia? *Annales. Series Historia et Sociologia*, 24 (3): 477–488.

Červ, G., and Kalin Golob, M. (2012): Sovražni govor kot tip vrednotenja: Jezikoslovni vidiki. *Teorija in Praksa*, 49 (1): 149–170.

Čibej, J., Fišer, D., and Erjavec, T. (2016): Normalisation, tokenisation and sentence segmentation of Slovene tweets. In A. Utka, J. Vaičenonienė, and R. Butkienė (eds.): *Proceedings of Normalisation and Analysis of Social Media Texts (NormSoMe)*: 5–10.

Crystal, D. (2001): *Language and the Internet*. Cambridge: Cambridge University Press.

Dobrovoljc, K., et al. (2019): *Morphological lexicon Sloleks 2.0*. Slovenian language resource repository CLARIN.SI. Available at: http://hdl.handle.net/11356/1230.

Dresner, E., and Herring, S. C. (2010): Functions of the nonverbal in CMC: Emoticons and illocutionary force. *Communication Theory*, 20 (3): 249–268.

ElSherief, M., Kulkarni, V., Nguyen, D., Wang, W. Y., and Belding, E. (2018): Hate lingo: A target-based linguistic analysis of hate speech in social media. In *Proceedings of the Twelfth International AAAI Conference on Web and Social Media*.

Erjavec, T., Fišer, D., Čibej, J., and Arhar Holdt, Š. *CMC training corpus Janes-Norm 1.2*. Slovenian language resource repository CLARIN.SI. Available at: http://hdl.handle.net/11356/1084.

Fišer, D., Miličević, M., and Ljubešić, N. (2018): Zapisovalne prakse v spletni slovenščini. In D. Fišer (ed.): *Viri, orodja in metode za analizo spletne slovenščine*: 124–139. Ljubljana: Znanstvena založba Filozofske fakultete.

Franza, J., and Fišer, D. (2019): The lexical inventory of Slovene socially unacceptable discourse on Facebook. In J. Longhi and C. Marinica (eds.): *Proceedings of the 7th Conference on CMC and Social Media Corpora for the Humanities*: 43–47. Cergy-Pontoise University.

Gill, A. (2011): Language and Personality in Computer-Mediated Communication. *Information Design Journal*, 19: 250–257.

Goli, T., Osrajnik, E., and Fišer, D. (2016): Analiza krajšanja slovenskih sporočil na družbenem omrežju Twitter. In T. Erjavec and D. Fišer (eds.): *Zbornik konference Jezikovne tehnologije in digitalna humanistika*: 77–82. Ljubljana: Znanstvena založba Filozofske fakultete.

Gorjanc, V. (2005): Diskurz slovenskih spletnih forumov–dokončen pokop strpnosti. *Večkulturnost v Slovenskem Jeziku, Literaturi in Kulturi*, 41: 20–29.

Gorjanc, V. (2012): Ideologija heteronormativnosti, prevodna in jezikovna norma. In A. Bjelčević (ed.): *Ideologije v slovenskem jeziku, literaturi in kulturi*, vol. 48: 38–44. Ljubljana: Znanstvena založba Filozofske fakultete.

Hilte, L., Vandekerckhove, R., and Daelemans, W. (2018): Adolescents' social background and non-standard writing in online communication. *Dutch Journal of Applied Linguistics*, 7 (1): 2–25.

Ivanou, U. (2017): Stereotyping Vulnerable Groups. In S. Assimakopoulos, F. H. Baider, and S. Millar (eds.): *Online hate speech in the European Union: A discourse-analytic perspective*: 29–31. Dordrecht: Springer.

Jakopič, K. (2013): Profesionalni razpihovalci sovražnega govora na spletu. *Medijska preža*, 44/45: 25.

Jessmer, S. L., and Anderson, D. (2001): The effect of politeness and grammar on user perceptions of electronic mail. *North American Journal of Psychology*, 3 (2): 331–346.

Jourová, V. (2019): *How the Code of Conduct helped countering illegal hate speech online*. EC Directorate-General for Justice and Consumers. Available at: https://ec.europa.eu/info/sites/info/files/hatespeech_infographic3_web.pdf (28 January 2020).

Klein, G. B. (2018): Applied Linguistics to Identify and Contrast Racist 'Hate Speech': Cases from the English and Italian Language. *Applied Linguisics Research Journal*, 2 (3): 1–16.

Laflen, A., and Fiorenza, B. (2012): Okay, My Rant is Over: The Language of Emotion in Computer-Mediated Communication. *COCOMP Computers and Composition*, 29 (4): 296–308.

Lewandowska-Tomaszczyk, B. (2017): Incivility and confrontation in online conflict discourses. *Lodz Papers in Pragmatics*, 13 (2): 347–367.

Li, L., and Yang, Y. (2018): Pragmatic functions of emoji in internet-based communication – A corpus-based study. *Asian-Pacific Journal of Second and Foreign Language Education*, 3 (1): 1–16.

Ljubešić, N., and Erjavec, T. (2016): Corpus vs. Lexicon supervision in morphosyntactic tagging: The case of Slovene. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, and S. Piperidis (eds.): *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*: 1527–1531. European Language Resources Association (ELRA).

Ljubešić, N., Fišer, D., and Erjavec, T. (2019): The FRENK Datasets of Socially Unacceptable Discourse in Slovene and English. In K. Ekštein (ed.): *Text, Speech, and Dialogue*: 103–114. Dordrecht: Springer.

López, C. A., and López, R. M. (2017): Hate Speech in the Online Setting. In S. Assimakopoulos, F. H. Baider, and S. Millar (eds.): *Online Hate Speech in the European Union – A Discourse-Analytic Perspective*: 10–12. Cham: Springer.

Merljak, S. (2011): Samoregulacija spletnih medijev: Kodeks, moderiranje in celostna registracija uporabnikov. *Medijska preža*, 39: 25.

Micciche, L. R. (2007): *Doing emotion: Rhetoric, writing, teaching. Language and the Internet*. Portsmouth, NH: Boynton/Cook Press.

Motl, A., and Bajt, V. (2016): *Sovražni govor v Republiki Sloveniji: Pregled stanja*. Ljubljana: Mirovni inštitut.

Muskat, R. V., and Assimakopoulos, S. (2017): The Implicit Dimension of Discriminatory Discourse. In S. Assimakopoulos, F. H. Baider, and S. Millar (eds.): *Online Hate Speech in the European Union – A Discourse-Analytic Perspective*: 42–46. Cham: Springer.

Nielsen, L. B. (2002): Subtle, pervasive, harmful: Racist and sexist remarks in public as hate speech. *Journal of Social Issues*, 58 (2): 265–280.

Neustupný, J. V., and Nekvapil, J. (2003): Language management in the Czech Republic. *Current Issues in Language Planning*, 4 (3–4): 181–366.

Nockleby, J. T. (2000): Hate speech. *Encyclopedia of the American Constitution*, 3: 1277–1279.

Osrajnik, E., Fišer, D., and Popič, D. (2015): Primerjava rabe ekspresivnih locil v tvitih slovenskih uporabnikov in uporabnic. In D. Fišer (ed.): *Zbornik konference Slovenščina na spletu in v novih medijih*: 50–56. Ljubljana: Znanstvena založba Filozofske fakultete.

Pahor de Maiti, K., Fišer, D., and Ljubešić, N. (2019). How haters write: Analysis of nonstandard language in online hate speech. In J. Longhi and C. Marinica (eds.): *Proceedings of the 7th Conference on CMC and Social Media Corpora for the Humanities*: 37–42. Cergy-Pontoise University.

Pavalanathan, U., and Eisenstein, J. (2015): Emoticons vs. Emojis on Twitter: A causal inference approach. *ArXiv Preprint ArXiv:1510.08480*. Available at: https://arxiv.org/abs/1510.08480 (28 January 2020).

Pertot, K., Petrovčič, M., and Strojan, N. (2016): #Analiza novih komunikacijskih elementov na družbenem omrežju @Twitter. In T. Erjavec and D. Fišer (eds.): *Zbornik konference Jezikovne tehnologije in digitalna humanistika*: 239–244. Ljubljana: Znanstvena založba Filozofske fakultete.

Skubic, A. E. (2005): *Obrazi jezika*. Ljubljana: Študentska založba.

Spina, S. (2017): Emoticons as multifunctional and pragmatic Resources: A corpus-based Study on Twitter. In *Media Corpora for the Humanities (Cmc-corpora17)*: 25–29.

Splichal, S. (2017): Medijatizacija zasebnosti in javnosti: zla publiciteta, anonimnost in cenzura na internetu. In S. Splichal (ed.): *Zagovor javnosti: med svobodo izražanja in sovražnim govorom*: 9–29. Ljubljana: Slovenska akademija znanosti in umetnosti.

Toporišič, J. (2004): *Slovenska slovnica*. Maribor: Obzorja.

Toporišič, J. (2007): *Slovenski pravopis*. Ljubljana: Slovenska akademija znanosti in umetnosti.

Tsesis, A. (2002): *Destructive messages: How hate speech paves the way for harmful social movements*. New York: NYU Press.

Ullmann, S., and Tomalin, M. (2019): Quarantining online hate speech: Technical and ethical perspectives. *Ethics and Information Technology*: 1–12. Available at: https://doi.org/10.1007/s10676-019-09516-z.

Vehovar, V., Motl, A., Mihelič, L., Berčič, B., and Petrovčič, A. (2012): Zaznava sovražnega govora na slovenskem spletu. *Teorija in Praksa*, 49 (1): 171–189.

Verheijen, L., Fišer, D., and Beißwenger, M. (2017): WhatsApp with social media slang?: Youth language use in Dutch written computer-mediated communication. In D. Fišer and M. Beißwenger (eds.): *Investigating Computer-Mediated Communication: Corpus-Based Approaches to Language in the Digital World*: 72–101. Ljubljana: Ljubljana University Press.

Vidgen, B., and Yasseri, T. (2019): Detecting weak and strong Islamophobic hate speech on social media. *Journal of Information Technology & Politics*, 17 (1): 66–78.

Waqas, A., Salminen, J., Jung, S., Almerekhi, H., and Jansen, B. J. (2019): Mapping online hate: A scientometric analysis on research trends and hotspots in research on online hate. *PLOS One*, 14 (9).

Zhang, Z., and Luo, L. (2019): Hate Speech Detection: A Solved Problem? The Challenging Case of Long Tail on Twitter. *Semantic Web*, 10 (5): 925–945.

Zwitter Vitez, A., and Fišer, D. (2016): Linguistic analysis of emotions in online news comments an example of the Eurovision song contest. In D. Fišer and M. Beißwenger (eds.): *Proceedings of the 4th Conference on CMC and Social Media Corpora for the Humanities* (pp. 74–76). Znanstvena založba Filozofske fakultete. Available at: http://nl.ijs.si/janes/wp-content/uploads/2016/09/CMC-2016_Zwitter_Fiser_Linguistic-Analysis-of-Emotions.pdf (27 February 2020).

Zwitter Vitez, A., and Fišer, D. (2018): Govorne prvine v nestandardni spletni slovenščini. In D. Fišer (ed.): *Viri, orodja in metode za analizo spletne slovenščine*: 254–272.

Twitter discourse on LGBTQ+ in Slovenia

***Vojko Gorjanc**, Faculty of Arts, University of Ljubljana, Slovenia*

***Darja Fišer**, Faculty of Arts, University of Ljubljana, Slovenia, Jožef Stefan Institute, Ljubljana, Slovenia*

Abstract

Since Twitter, as one of the main social networking platforms, plays an important part in forming gender and sexual identities, the aim of this study was to perform a corpus analysis of Twitter discourse pertaining to the LGBTQ+ community in Slovenia. The dataset for our study was extracted from the Janes corpus of Slovenian user-generated content, which contains almost 215 million tokens from Slovenian blog posts, forum messages, news comments, tweets, etc., and is richly annotated with socio-demographic and linguistic metadata. Our analysis is mainly qualitative and based on content analysis, followed by a critical discourse analysis examining how much control one social group imposes over another and tries to limit the freedom of other people's actions, using the concepts of "normal" and "natural" as key aspects in forming gender and sexual identities. The results of the analysis show the persistence of heteronormativity; so much so that it becomes naturalized, whereas any counterforce is seen as disruptive.

Keywords: corpus analysis, critical discourse analysis, computer-mediated communication, queer linguistics, heteronormativity

1 SOCIO-CULTURAL CONTEXT

After the decriminalisation of homosexuality that took place in Slovenia in 1977, the question of social inequality based on sexual orientation became a relevant political topic in the 1980s, and was part of an extensive discussion in civil society movements (Mencin Čeplak and Kuhar 2010, 284). At that time, most of the political demands for social equality made by the LGBTQ+ community had been addressed (Kuhar and Mencin Čeplak 2016, 147). From the late 1970s to late 1980s, the Slovenian political space was marked by cultural and political civil society movements, which, among other demands, called for the elimination of discrimination based on sexual orientation through a series of citizens' initiatives. These issues were gradually overshadowed by Slovenia's struggle for political independence from Yugoslavia, and thus the initiatives advocating the rights of national, gender-related and sexual minorities slowly faded away from the political discourse (Mencin Čeplak and Kuhar 2010, 284–285). The initiatives for the elimination of gender inequalities were not abandoned in the course of drafting the Slovenian constitution – but after its adoption in 1991 they were considered less important compared to the 1980s, and were no longer at the forefront of political discourse along with concrete policy initiatives for the non-discriminatory treatment of gays and lesbians in legislation (Mencin Čeplak and Kuhar 2016, 153).

Discussions on the regulation of same-sex partnerships and initiatives for the adoption of a new Family Code have not led to any changes in legislation since the mid-1990s. As a matter of fact, the introduction of the Civil Partnership Registration Act in 2005 provided no appropriate solutions to the complaints of the LGBTQ+ community, but actually codified discrimination based on sexual orientation (Mencin Čeplak and Kuhar 2010, 288–289). All further attempts to eliminate discrimination were dominated by heated discussions throughout society in Slovenia. To a large extent, this was caused by a decision that such legislative amendments could be decided by a referendum, which shifted the focus of the discussion to the problems of the LGBTQ+ community but, at the same time, promoted a socially unacceptable public discourse.

The first referendum in 2011 was followed by a second in 2015, but no amendments to the legislation were made based on the results. The breakthrough in the legislation only happened after the Amendment to the Marriage and Family Relations was rejected in the 2015 referendum, which sought to equate the rights of homosexual partnerships with heterosexual ones simply by replacing the words *moški* (Eng. man) and *ženska* (Eng. woman) with the word *oseba* (Eng. person). Because the referendum amendment was rejected, the government worked out a compromise version of the legislation by adopting the Civil Union Act in 2016 (effective in 2017), which equates the rights and duties of heterosexual and

homosexual couples, with the exception of the right to joint adoption of children and the right to assisted reproduction.

2 THE STUDY

The analysis presented in this study will focus on two referendum campaign accounts on the social media platform Twitter in the period before the referendum on the Civil Partnership Registration Act in 2015, i.e. the main opponents of the amendment act with the *Za otroke gre* account (ZOG, Eng. For the children) and the proponents' main account *Čas je ZA* (ČJZ, Eng. It's time FOR).

Twitter and other social networks have become important media in political campaigns (Vonderschmitt 2012, Kumar and Natarajan 2016). Therefore, it should be interesting to see how they function in the Slovenian cultural environment, as the aforementioned campaigns can also be understood as part of a broader Slovenian political discourse. Though formally independent, both referendum campaigns were directly linked to Slovenian party politics: ZOG with the Slovenian right, including extreme right, and ČJZ with the left.

In this study, a corpus analysis will be combined with a critical discourse analysis, which has proven successful as an analytical procedure in related studies (Baker et al. 2008). Discourse as a social interaction will be observed in the specific context of two Twitter campaigns from the viewpoint of establishing social power, its distribution, maintenance and reproduction (van Dijk 2001, 352). Another goal will be to determine the role of discourse in establishing inequality in society and its maintenance in the process of domination (Fairclough 2001).

Discourse analysis therefore addresses some specific societal issues, as discourse is what constitutes culture and society (Fairclough and Wodak 1997, 271–280), while new social practices, which maintain and constantly reproduce the traditional notions and established balances of power, are being constructed in the framework of communication on social networks (Baider and Kopytowska 2018). This is a typical characteristic of Twitter, the social network in which the dominant discourse has the ability of being reproduced by forwarding or sharing posts on other social networks. Lately, tweets have been published by traditional mass media as well, and it has been shown that they often reproduce tweets in a completely non-critical manner. Continuous exposure to a particular normative model thus only reinforces "normal" and "straight" identities, including gender and sexual identities (Motschenbacher 2010, 2011). Groups with less social power are therefore faced with a challenge to formulate strategies so that they can function in the context of the prevailing hegemonic

discourse, perhaps even by trying to change the differences in the social power (Gramsci 1971, Gorjanc 2017).

This research will therefore focus on determining how the question of hegemonic homosexuality was addressed in a highly heteronormative discourse in one of the campaigns (ZOG), and what discourse strategies were used to preserve the "natural and social order", and, consequently, an order that perpetuates inequalities and injustices in society (Fairclough 1985, 1989). In the campaign led by the advocates for change (ČJZ), the focus will be placed on the discourse strategies addressing this normative social order with a desire to change the social balances of power, assuming that any opposite reaction in the hegemonic heteronormative discourse is doomed to be recognized by the majority in society as the one destroying the established social order.

Gender and sexual identities are important parts of a society's make up, its normative models and value systems, and in that respect heterosexuality as a social construct is based on very strict social norms that dominate both the societies and discourses of the West (Coates 2013). As a set of cultural patterns, social and legal norms that establish heterosexual hegemony, heteronormativity conditions discourse practices with an unmarked heterosexual identity, which is always assumed in the discourse until it is actively denied (Yep et al. 2003, Lovaas and Jenkins 2007, Koch 2008). In Western societies, the discourse of heteronormative values, images and behaviours is also linked to the historical discourse of sexuality as a discourse used for shaping national identities (Motschenbacher 2013). For this reason, the aim of this study will be to determine how the discourse of sexuality and gender gets incorporated into national and nationalistic discourses when such models are subject to problematization, as also happens in some other cultural environments (Baider 2018, Baider and Kopytowska 2018).

2.1 Related work

This study establishes itself within research on computer-mediated communication (CMC), i.e. one of many different forms of interaction that people have with each other using computers as a means of communication. This body of research initially focused on building specialized corpora (Beißwenger et al. 2014) and later enabled research into diverse topics based on corpus data (Fišer and Beißwenger 2017, Fišer 2018), including sociolinguistic and discourse studies (Coats 2017, Verheijen 2017, Reher and Fišer 2018). Our study contributes to this body of research, enriching it with several specific aspects of the subject matter and embedding it in the Slovenian linguistic and cultural environment.

The paper is also situated within the framework of queer linguistics, i.e. an interdisciplinary approach within post-structural critical theories based on the analysis and deconstruction of traditional ideas, values and established social relations of power when it comes to issues of identity and the concepts of gender and sexuality (Jagose 1996, Koch 2008, 20). This approach also seeks to help understand the discourse construction of identity categories as ideological social constructs that have been characterized by heteronormativity throughout history (Koch 2008, 38–39). Numerous studies in the field of queer linguistics are based on the corpus-linguistic methodology (Baker 2005). They generally analyse heteronormativity as the dominant feature of discourse (Baker 2005) and related social constructs, such as a corpus study in which discourse shows greater social support for the concept of marriage than, for example, the concept of partnership or a single person (Baker 2008). A corpus-based analysis examining the Cypriot LGBTQ+ community examines the way in which the discourse constructions of sexuality are expressed within the hegemony of heterosexuality and heteronormativity (Baider 2018), whereby the relationship between quantitative and qualitative analysis favours the latter, as it does in our research and in most other discourse studies with a corpus-based approach (Kulick 2005).

Our research contributes to the research on the dominance of heteronormativity with its specific focus on the Slovenian perspective. It integrates the queer approach into the research field of CMC by adding a specific communication platform to other studies observing the balance of power among social groups, such as the extensive research into the discourse used on web forums by white supremacists and their argumentation strategies to establish a communal identity as the norm, which allows platform users to observe deviations from the established normative system and understand everything non-normative as deviant, especially when it comes to national, gender or sexual identity (Brindle 2016).

The methodological integration of a corpus-based approach with critical discourse analysis, which was used in this research, has been found to be effective for the study of discourse characteristics (Baker et al. 2008). The integration has been most useful for the analysis of discourse related to socially marginalized individuals and groups, or vulnerable groups in a society in general, such as refugees and asylum seekers (Baker and McEnery 2005, Gabrielatos and Baker 2008), or when examining the homophobic, xenophobic or hateful narratives aimed at what is other and different in general (Assimakopoulos and Vella Muscat 2017, Baider 2018, Baider and Kopytowska 2018). We have performed similar research (see Gorjanc and Fišer 2018), and with this study we fill the gaps identified in our previous work.

2.2 Research datasets

The research datasets for our analysis are extracted from the Janes[1] corpus of Slovenian user-generated content of almost 215 million tokens, such as tweets, forums, news and corresponding comments, blog posts and corresponding comments, as well as user and discussion pages on Wikipedia, posted in the period 2001–2015. In addition to the origin of the text and time of publication, the Janes corpus contains a number of manual and automatic annotations, e.g. the account type (private, corporate), author's gender (male, female, unknown), and region of publication, the language of the text (Slovenian, English, Bosnian/Croatian/Serbian, other), degree of text standardness (standard, partially non-standard, very non-standard), and sentiment (positive, negative, neutral). The corpus is also richly linguistically annotated, i.e. tokenized, normalized, lemmatized and morphosyntactically tagged (Fišer et al. 2016, 68–69; Fišer et al. 2018).

For the purpose of this research, the following datasets were extracted from Janes:

a) tweets from the official ZOG campaign account, the main opponents of the amending act and the referendum initiators, and

b) tweets from the official ČJZ campaign account, the main supporters of the amendment act.

The ZOG account was set up in February 2015, and the ČJZ account as late as November 2015. ZOG had already called on its followers to help in getting the signatures needed for a referendum, whereas the ČJZ account only served as a part of the campaign, and thus was set up just before the referendum in December 2015. Neither of the accounts have been active since the referendum, but they both remain online with all published material still available.

In total, the datasets comprise a little over 21,500 tokens. The datasets of both official campaigns, ZOG and ČJZ, are comparable in terms of quantity: 150 tweets made by ZOG and 175 tweets made by ČJZ, with 2,308 and 3,275 tokens, respectively. Nonetheless, these two datasets are fundamentally different with regard to the function of the communication channel and their ways of communicating: the ČJZ account is much more embedded in the social network as it has a far greater number of followers (851) and friends (332) compared to ZOG, with 208 followers and only 26 friends. ČJZ also has a higher average number of likes (244) and forwards of tweets (4.19), thus creating a greater Twitter discourse community than ZOG, with an average of 11 likes and 1.81 forwards of tweets. ZOG is therefore more of a one-way communication channel, one of the basic functions being to provide links to media coverage (Gorjanc and Fišer 2018, 279).

1 Available on https://www.clarin.si/noske/run.cgi/corp_info?corpname=janes

However, because one of our previous analyses showed that the official campaign's account (ZOG) was mainly used for directing readers to websites or posts in ideologically related media (Gorjanc and Fišer 2018, 480), another ad hoc corpus of texts, the ZOG MED corpus, was compiled for the purpose of content analysis using the Sketch Engine tool. This corpus, which is based on the texts to which ZOG directed their followers with hyperlinks, comprises just over 52,900 tokens.[2]

The corpus contains 28 texts from 4 sources. Most of the texts are taken from http://24kul.si. This is the webpage of *Zavod za družino in kulturo življenja* ("The Institute for Family and the Culture of Life"), which is closely affiliated with the Catholic Church in Slovenia. The other 3 sources correspond in a roughly equal amount to texts taken from the web portal of the national broadcasting company https://www.rtvslo.si, texts on the webpage https://publishwall.si and texts on the webpage of Ognjišče, https://radio.ognjisce.si, which is a Catholic radio. In terms of token number, the posts taken from Catholic webpages make up 80.5% of the entire corpus. The corpus was tagged for parts of speech and lemmatised with the Sketch Engine tool.

2.3 Methodology

Because the datasets are relatively small in size and because of the limits of the corpus approach when it comes to revealing subtle discourse meanings (Stubbs 1994), we primarily focused on qualitative research. Having said that, some interesting quantitative data were obtained from the datasets as well. Full tweets from both campaigns have already been used in content analysis of the discourse (Gorjanc and Fišer 2018). In order to check the extent to which the official ZOG campaign on Twitter continued in the media coverage to which the campaign directed its followers, the ZOG MED corpus was used for content analysis in this study. The comparison of the ZOG datasets and the ZOG MED corpus was the starting point for determining which types of content influenced the selection of non-verbal communication elements of the opponents of the Act to reinforce their opposition to introducing legislative changes.

3 ANALYSIS

In order to determine which topics were predominant in the discourse of the two campaigns, each tweet was manually annotated for its main topic, with these topics then clustered into broader categories. The results show that tweets from

2 All the texts harvested from the hyperlinks posted in the ZOG tweets, active on 31 January 2019, are included in the corpus.

the two accounts are very different in terms of the topics they address. It was significantly easier to classify the tweets posted by ČJZ than those by ZOG, which is primarily attributed to the clearer focus of ČJZ's entire campaign. This will be elaborated on in more detail in the following two subsections, where the results of the analysis are presented for each campaign.

3.1 The ZOG campaign

The ZOG campaign did not start and end with tweets that would pragmatically address potential followers in any way. Instead, it shared links that directed readers to news articles. Moreover**,** the campaign tweets were extremely heterogeneous in terms of content, and in certain cases even unrelated to the amendment that was the focus of the referendum. This is also indicated by the large share of the posts (15%) that we were unable to classify into any of the topics because the content of such posts was heterogeneous and largely unrelated to the referendum campaign. The analysis also showed that by choosing a diverse range of content this campaign sought to engage with those segments of society that generally would not have dealt with this kind of topic directly, but were more open to topics such as WWII, post-communism, and abortion.

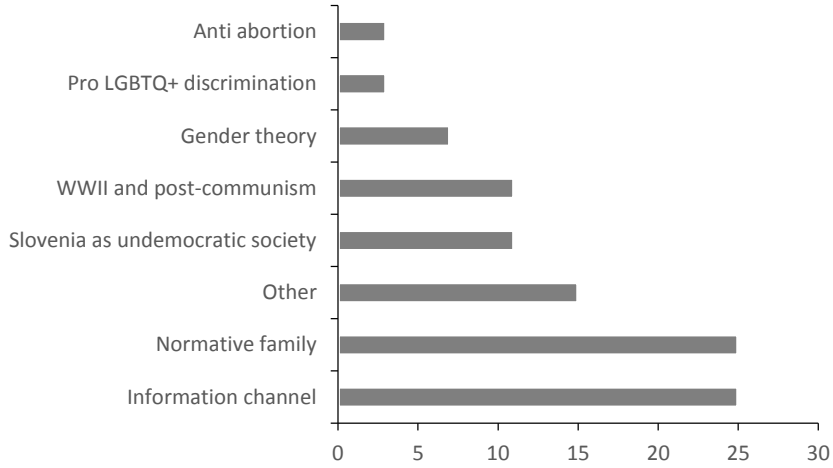

Figure 1: Twitter topics of the ZOG campaign (in %) throughout the period of activity of the ZOG account between 5 March and 3 November 2015 and prior to the referendum.

As can be seen in Figure 1, the most frequent topics we identified were the information channel, which was primarily used for organizing the community in order to gain enough signatures to call a referendum, and the normative family, each

representing a quarter of the posts. The authors of the ZOG campaign used the strategy of reproducing the traditional concept of the family and relied on the fact that this is recognized as the natural order within the hegemonic heterosexual discourse. This hegemonic heterosexual discourse, which follows the formula "family = mother, father and children", appeared to be successful as it dominated the other topics within the ZOG campaign Twittersphere. This norm has historically been accepted as society's preferred way of thinking, particularly from a religious perspective. At the same time, this strategy preserves the natural and social order and is also one of the most successful approaches for maintaining social inequality (van Dijk 1993, Fairclough 1985). However, it should be noted that ZOG limits the social model of the family with the concept of the Catholic normative model. Most of the tweets on this subject are directly related to this understanding of the family, as illustrated in the tweet "*Družina je Božji dar za uresničitev moškega in ženske, ki sta ustvarjena po Njegovi podobi*" (Eng. "Family is a gift from God for the realization of a man and a woman created in His image"), while Pope Francis is referred to as the ultimate authority of this model (Figure 2).

Figure 2: A ZOG tweet referencing the ultimate authority on the normative family and the accompanying text: "'You, the youth, should not be led astray by the harmful transient ideologies. Family is an anthropological fact.' Pope Francis".

In terms of the number of tweets, the Catholic narrative within the normative family topic is immediately followed by the topic in which Slovenia is described as an undemocratic society (11%), with tweets discussing the dysfunction of democracy, the people's "fight" with Parliament, and undemocratic imbalances in media reporting. This seems surprising at first glance. However, in this context, the campaign also engages in activities aimed at the democratization of the judiciary with reference to one of the key players in ZOG (all instances address a court case against the then opposition leader Janez Janša). This topic is followed by tweets about World War II, communism and postcommunist society (11%). These tweets

are mainly linked to the national television broadcasts of the programme *Pričevalci* (Eng. *Witnesses*), thereby positioning ZOG on the "non-communist" ideological side during and immediately after World War II. Because tweets on these two topics are nearly as frequent as those on the family, one can identify a clear strategy for addressing potential followers who are not interested in the content of the amendment itself, but can ideologically align with other topics with the same ideological underpinnings that simultaneously connect with national cohesiveness in a similar way to in some other settings (Baider and Kopytowska 2018). This is particularly noticeable in Slovenia, because of the connection of national cohesiveness with national division, as outlined by the experience of World War II, i.e. the reproduction of a social division that has been present in Slovenian society since the 1950s and resulted from a split between the traditional Catholic society and the liberal movement that culminated during and immediately after World War II.

Posts on the topic of gender theory (7%) cover a significant share of the data, and include tweets that are a direct opposition to progressive gender equality. The topic places ZOG in the framework of religious organizations such as the Roman Catholic Church and other movements on the right-wing of the political spectrum in Europe that use the academic concept of gender as mobilizing tool for social movements. They form an opposition to progressive gender equality and address questions of marriage equality, abortion, reproductive technologies, gender mainstreaming, sex education, sexual liberalism, transgender rights, antidiscrimination policies, etc. (Kuhar and Paternotte 2018). In the ZOG campaign, the topic of gender theory is related to the idea that it allows for a free choice of gender, promotes sexual diversity, and supports different sexual orientations. The last topic within the ZOG campaign covers anti-abortion tweets, which points to a potential new focus of the community and the political parties connected with it, one that diverges from discussions about homosexuality and takes on an issue that they consider a vital part of their politics.

Our analysis of the ZOG corpus indicated the large presence of hyperlinks in ZOG tweets (the average number of hyperlinks per tweet was 0.81) that directed readers to news articles. We followed these hyperlinks to establish their content as part of the overall discourse relating to the ZOG campaign, and compiled a corpus of the hyperlinked news articles. We analysed the ZOG MED corpus by extracting the key multi-word units based on their relative frequency in the corpus compared to the reference Slovenian Web 2015 corpus (Erjavec et al. 2015), since keywords tend to be a good indicator of the overall content of the texts (Scott and Tribble 2006). Our basic keyword extraction procedure, which was performed in the SketchEngine concordancer, lists all the identified n-gram candidates regardless of the scope of a given multi-word unit. This means that the same multi-word unit can be represented in several keywords (e.g. *problematika družine,* Eng. *problems related to the family* and *problematika družine in življenja,* Eng. *problems related to the family and life*), which is important to take into account in the interpretation of the results.

Table 1: 30 top-ranking key multi-word units in the ZOG MED corpus.

Slovene key multi-word expression	English translation
obveščanje o delovanju	communication on the operation
obveščanje o delovanju zavoda	communication on the operation of an institution
problematika družine	problems related to the family
problematika družine in življenja	problems related to the family and life
množično grobišče	mass grave
komunistično obdobje	communist era
vojni grob	war grave
kulturna dediščina	cultural heritage
delovanje zavoda	operation of an institution
človekova pravica	human right
človeško življenje	human life
aktivno državljanstvo	active citizenship
vojno grobišče	wartime mass grave
civilna iniciativa za družino	civic initiative for the family
družina in pravica otrok	family and children's rights
iniciativa za družino	initiative for the family
koalicija za otroke	coalition for children
družina in kultura življenja	family and life culture
državna uprava	public administration
katoličan pred diskriminacijo	a Catholic facing discrimination
kultura življenja	life culture
namen pospeševanja	aim of promotion
pospeševanje temeljnih vrednot	promotion of fundamental values
stigmatiziran v državni upravi	stigmatized in public administration
svoboda in aktivno državljanstvo	freedom and active citizenship
zavod za družino	institute for the family
zaščititi katoličane	to protect Catholics
interna spletna stran zavoda	the institute's internal website
najbolj preganjana verska skupnost	the most persecuted religious community
preganjana verska skupnost	persecuted religious community

The 30 top-ranking key multi-word units which can be seen in Table 1 give a good insight into the content of the articles. If we disregard the first two examples related to communication, the other key words mainly match and reinforce the basic content categories of the Twitter campaign. The topic of *družina* (Eng. family) particularly stands out, and is closely linked to the concept of *kultura življenja* (Eng. life culture) that, in media coverage, closely connects the keywords of *problematika družine* (Eng. problems related to the family), *civilna iniciativa za družino* (Eng. civic initiative for the family), *družina in pravica otrok* (Eng. family and children's rights), and *iniciativa za družino* (Eng. initiative for the family) with a topic

that is only briefly mentioned towards the end of the Twitter campaign, i.e. the right to decide freely about the birth of a child or the issue of abortion: *problematika družine in življenja* (Eng. problems related to the family and life), *človeško življenje* (Eng. human life), *družina in kultura življenja* (Eng. family and life culture), and *kultura življenja* (Eng. life culture).

The media coverage in the links shared by the ZOG campaign places the referendum within the context of Catholicism, with a narrative suggesting that marriage equality is a threat to the religious community. This strategy attempts to engage ZOG supporters in terms of active citizenship to protect Catholicism against this attack: *aktivno državljanstvo* (Eng. active citizenship), *svoboda in aktivno državljanstvo* (Eng. freedom and active citizenship), *zaščititi katoličane* (Eng. to protect Catholics), *najbolj preganjana verska skupnost* (Eng. the most persecuted religious community), *preganjana verska skupnost* (Eng. persecuted religious community), and *katoličan pred diskriminacijo* (Eng. a Catholic facing discrimination).

The topic from Figure 1 which is strongly present in the media coverage that the ZOG campaign linked to in its tweets addresses potential supporters who ideologically associate themselves with the non-communist recent history, that is "World War II and post-communism". In this case, the key narrative specifically targets the emotions by referring to the horrific consequences of World War II: *množično grobišče* (Eng. mass grave), *komunistično obdobje* (Eng. communist era), *vojni grob* (Eng. war grave), and *vojno grobišče* (Eng. wartime mass grave).

Although the connection to the various European movements that utilize the concept of gender theory is an important part of the posts in the Twitter campaign, this topic is not as present in the linked media coverage. It seems that it is easier for the campaign to address and potentially persuade or manipulate supporters with familiar concepts they understand, and particularly with a narrative that preserves the social order and causes no instability, but nevertheless creates enough uncertainty for the community to ensure that its members stay active.

3.2 The ČJZ campaign

The ČJZ campaign started with its first tweet, which addressed potential followers and introduced the campaign's hashtags: "*Pozdravljeni, vsi enakomisleči državljani in državljanke, Slovenija, čas je ZA #casjeZA"* (Eng. Welcome, all like-minded citizens, Slovenia, it's time FOR #timeFOR). The last tweet concludes the campaign by thanking the followers: "***Čas ZA našo zgodbo še prihaja**. Kmalu. Hvala vsem **za** izkazano podporo na voliščih.*" (Eng. **The time FOR our story is yet to come**. Soon. Thank you all **for** the support at the polls. #timeFOR.)

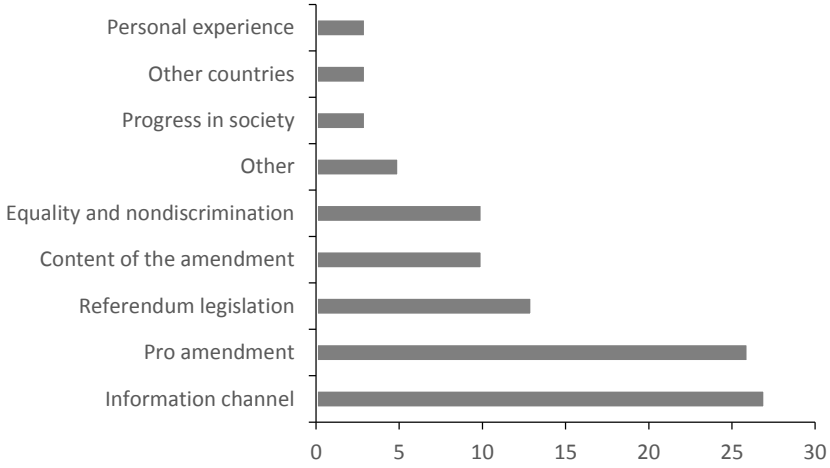

Figure 3: Twitter topics of the ČJZ campaign (in %) throughout the period of activity of the ČJZ account, which was between 22 November and 10 December 2015, during the referendum campaign.

In contrast to ZOG, the ČJZ campaign focused specifically on the referendum itself rather than on wider aspects of society, as can be seen from the list of topics that were identified in the analysis of the corpus of ČJZ in Figure 3. The majority of the tweets (53%) were a combination of pro-amendment comments (26%) and information for supporters about related events (27%). Next are tweets that sought to explain the referendum legislation and its interpretation (13%) and the content of the amending act (10%). The tweets explaining the content of the act, which also responded to disinformation in the mainly right-wing media about it, were an important part of the campaign. It is also clear that one of the campaign's strategies was to get voters to the polls, because a substantial portion of the tweets urged people to participate in the referendum, as we can see in Figure 4 where we have the well-known public figure

Figure 4: A tweet calling for a high voter turnout.

and entertainer Vid Valič and a tweet: "Are you FOR equality? Then it's off to the referendum on Sunday. Yes, your vote counts! @VidValic #casjeZA".

This campaign also served as an information channel (27%), providing important information about and invitations to events. From this we can see that the Twitter campaign played an important role in building a community both online and offline that would support the referendum.

The individual topic of the ČJZ campaign reveals discourse strategies that try to change the established social order as well as the social balance of power. In doing so, the campaign also the used personal experiences of LGBTQ+ community members, although such tweets accounted for a very small percentage of the total (3%). Such tweets aimed to increase understanding of the real life experiences of the LBTQ+ community, and how the referendum would affect them. The same percentage of tweets (3%) linked the campaign with foreign practices and supporters from abroad (3%). By referring to other nations the campaign linked the amendment to the wider discourse on social change and development, which it suggests should now become part of legislation.

Neither campaign refers to the other directly in their tweets. There are, however, subtle references, such as the ČJZ tweet quoting the writer and director Miha Mazzini (Figure 5): "May anyone who so wishes get married. God bless them, says Miha Mazzini." This is one of the rare tweets that subtly address the normative topic of the Catholic concept of marriage.

Figure 5: A tweet referencing the concept of the Catholic norm of the family.

The ČJZ campaign is thus activist, clearly based on the ideological concepts of human rights, human dignity, freedom, justice, equality, and non-discrimination. In other words, everything is understood in the context of social progress as seen in a tweet saying "*Enakopravnost nam je zapisana v koreninah. Čas je, da jo prevedemo še v zakone.*" (Eng. Equality is rooted in all of us. The time has come to incorporate it into legislation). Furthermore, *#casjeZA* (Eng. #timeFOR) carries one of

the strongest non-verbal messages, as the supporters of the amendment were photographed under the monument to the poet France Prešeren in one of the main squares in the capital, Ljubljana (see Figure 6). Prešeren is the author of the lyrics of the Slovenian national anthem, celebrating equality among nations and individual freedom: "*Živé naj vsi naródi, // ki hrepené dočákat› dán, // da kóder sónce hódi, // prepír iz svéta bó pregnán, // da roják, // prost bo vsák, // ne vrág, le sósed bo meják.* " (Eng. "God's blessing on all nations, // Who long and work for that bright day, // When o'er earth's habitations // No war, no strife shall hold its sway; // Who long to see // That all men free // No more shall foes, but neighbours be.").

Figure 6: A tweet showing ČJZ supporters by the Prešeren Monument in Ljubljana (Caption: "Equality is rooted in all of us. The time has come to incorporate it into legislation. #timeFOR").

The tweet symbolically connects the content of the anthem, the place where the photo was taken and the supporters of changing legislation in all their diversity, as we can see activists, celebrities, politicians and members of the general public.

Both campaigns play with the concept of nationality, but use it in two very different ways. The ZOG campaign falls within the populist discourse of the right, which contextualizes nationality as a "natural" part of shaping and preserving a nation. It focuses on who belongs to a nation, what they should be like, and construes the Other as a threat to our nation, on cultural, religious, sexual, ethnic or other grounds (Kuhar 2015, 122). On the other hand, the ČJZ campaign addresses those aspects of national identity that historically emphasized freedom as the key element of cohabitation, and perceived national liberation in the context of the individual's unconditional freedom.

4 CONCLUSION

The aim of this mainly qualitative corpus critical discourse analysis was to observe the discourse of two referendum campaign Twitter accounts on the Amendment to the Marriage and Family Relations Act in 2015, one belonging to ZOG, the main opponents of the amendment and initiators of the referendum, and the central account of the amendment supporters belonging to ČJZ.

All the tweets published by both campaigns were included in the analysis. The findings showed that the ZOG campaign was highly heterogeneous and covered a number of subjects that were not directly related to the amendment, while the ČJZ campaign was more focused in terms of the time and content of the referendum. The ZOG campaign used the strategy of reproducing the traditional concept of the family within the hegemonic heterosexual discourse as the main discourse strategy, but at the same time restricted it with the concept of the Catholic normative model. The use of gender theory as a topic is an important segment of the campaign, which explicitly places itself within the ideological framework of mobilization movements against equality that are seen across the right-wing in Europe. Interestingly, this was not a particularly common topic in the related media articles, which used well-known concepts, mainly with a narrative of the family and family values that preserve the social order and stability, but also worked to trigger feelings of uncertainty and threat with regard to the Catholic community, thus inspiring more action among its members. Tweets are often linked to the issue of abortion, which was only mentioned towards the end of the official ZOG Twitter campaign. This campaign also heavily used the strategy of addressing potential followers who are not interested in the content of the amendment, but can relate to other topics with the same ideological basis: World War II, communism and the post-communist society.

The ČJZ campaign, on the other hand, clearly focused on the referendum legislation and the content of the amendment, but also responded to the disinformation in right-wing media about the referendum. Its campaigners tried to introduce changes to the established social order and social balances of power with the use of celebrities and professionals from Slovenia and abroad, as well as by referencing foreign practices and using the discourse on social change and social development, which in their view also has to be included into legislation. The campaign was also activist, explicitly based on the ideological concepts of human rights, human dignity, freedom, fairness, equality, and non-discrimination.

Although all political campaigns around the world, at least from 2008 onwards, have been aware of the influence of social media and thus exploited such platforms (Vonderschmitt 2012, 3), we find it hard to judge the real impact of both

Twitter accounts with regard to the outcome of the referendum. This is especially so because the Twitter account of the ZOG initiative, which triggered the referendum in the first place, was mainly active in the period until the referendum was called, but not during the official referendum campaign. The referendum was unsuccessful as the proposed legislative measure was not adopted, but the entire discourse on the matter of LGBTQ+ rights in the context of equality of marriage – including the discourses of both campaigns – certainly contributed to a heightened sensitivity of the general public to such issues, and in some respects also that of the political public. The latter, in 2016, a year after the referendum, adopted the Partnership Act, which includes all rights of marriage, except adoption and in-vitro fertilization.

This analysis has given us a good general insight into the discursive strategies of both campaigns; however, for a more comprehensive overview in the future, the qualitative analysis should be extend to take the non-verbal aspects of the posts into account as well, given that non-textual forms of discourse are increasingly becoming one of the most important features of social media. In future work, it would be interesting to extend our analysis to include Tweets tagged with the hashtags of the campaigns, and to include those Tweets that made use of other hashtags linked to topics associated with both campaigns of the referendum. This way, we will be able to arrive at a more comprehensive understanding of Twitter discourse pertaining to the LGBTQ+ community in Slovenia at the time before the referendum.

Acknowledgements

The work described in this paper was funded by the Slovenian Research Agency within the national basic research project "Resources, methods, and tools for the understanding, identification, and classification of various forms of socially unacceptable discourse in the information society" (J7-8280, 2017-2020), and the bilateral basic research project "The linguistic landscape of hate speech on social media" (N6-0099, 2019-2023).

References

Assimakopoulos, S., and Vella Muskat, R. (2017): Exploring xenophobic and homophobic attitudes in Malta: Linking the perception of social practice with textual analysis. *Lodz Papers in Pragmatics*, 13 (2): 179–202.

Baider, F. (2018): "Go to hell fucking faggots, may you die!" Framing the LGBT subject in online comments. *Lodz Papers in Pragmatics*, 14 (1): 69–92.

Baider, F., and Kopytowska, M. (2018): Narrating hostility, challenging hostile narratives. *Lodz Papers in Pragmatics*, 14 (1): 1–24.

Baker, P. (2005): *Public Discourses of Gay Men*. London: Routledge.

Baker, P. (2008): *Sexed Texts. Language, Gender and Sexuality*. London: Equinox.

Baker, P., and McEnery, T. (2005): A Corpus-based Approach to Discourses of Refugees and Asylum Seekers in UN and Newspaper Texts. *Journal of Language and Politics*, 4 (2): 197–226.

Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., and Wodak, R. (2008): A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19 (3): 273–306.

Beißwenger, M., Oostdijk, N., Storrer, A., and van den Heuvel, H., eds. (2014): *Building and Annotating Corpora of Computer-Mediated Communication: Issues and Challenges at the Interface of Corpus and Computational Linguistics*. Special issue of *Journal for Language Technology and Computational Linguistics*, 29 (2).

Brindle, A. (2016): *The Language of Hate. A Corpus Linguistic Analysis of White Supremacist Language*. London and New York: Routledge.

Coates, J. (2013): The discursive production of everyday heterosexualities. *Discourse & Society*, 24 (5): 536–552.

Coats, S. (2017): Gender and grammatical Frequencies in social media English from the Nordic countries. In D. Fišer and M. Beißwenger (eds.): *Investigating Computer-Mediated Communication: Corpus-based Approaches to Language in the Digital World*: 102–121. Ljubljana: Znanstvena založba Filozofske fakultete UL.

Erjavec, T., Ljubešić, N., and Logar, N. (2015): The slWaC corpus of the Slovene Web. *Informatica*, 39 (1): 35–42.

Fairclough, N. (1985): Critical and Descriptive Goals in Discourse Analysis. *Journal of Pragmatics*, 9 (6): 739–763.

Fairclough, N. (1989): *Language and Power*. London: Longman.

Fairclough, N. (2001): *Language and Power.* London and New York: Longman.

Fairclough, N., and Wodak, R. (1997): Critical discourse analysis. In T.A. van Dijk (ed.): *Discourse Studies. A Multidisciplinary Introduction. Vol. 2. Discourse as Social Interaction*: 258–284. London: Sage.

Fišer, D., and Beißwenger, M., eds. (2017): *Investigating Computer-Mediated Communication: Corpus-based Approaches to Language in the Digital World*. Ljubljana: Znanstvena založba Filozofske fakultete UL.

Fišer, D., ed. (2018): *Viri, orodja in metode za analizo spletne slovenščine*. Ljubljana: Znanstvena založba Filozofske fakultete UL.

Fišer, D., Erjavec, T., and Ljubešić, N. (2016): JANES v0.4: korpus slovenskih spletnih uporabniških vsebin. *Slovenščina 2.0*, 4 (2): 67–99.

Fišer, D., Ljubešić, N., and Erjavec, T. (2018): The Janes project: language resources and tools for Slovene user generated content. *Language resources and evaluation*, 52 (3): 1–24.

Gabrielatos, C., and Baker, P. (2008): Fleeing, Sneaking, Flooding: A Corpus Analysis of Discursive Constructions of Refugees and Asylum Seekers in the UK Press 1996–2005. *Journal of English Linguistics*, 36 (1): 5–38.

Gorjanc, V. (2017): *Nije rečnik za seljaka*. Beograd: Biblioteka XX vek.

Gorjanc, V., and Fišer, D. (2018): Twitter in razmerja moči: diskurzna analiza kampanj ob referendum za izenačitev zakonskih zvez v Sloveniji. *Slavistična revija*, 66 (4): 473–495.

Gramsci, A. (1971): *Selections from the Prison Notebooks*. New York: International Publishers.

Jagose, A. (1996): *Queer Theory: An Introduction*. New York: New York University Press.

Koch, M. (2008): *Language and gender research from a queer linguistic perspective: A critical evaluation*. Saarbrücken: VDM.

Kuhar, R. (2015): Konec je sveta, kakršnega poznamo: populistične strategije nasprotnikov Družinskega zakonika. *Časopis za kritiko znanosti*, 43 (249): 118–132.

Kuhar, R., and Mencin Čeplak, M. (2016): Same-sex partnership debate in Slovenian: between declarative support and lack of political will. In K. Slootmaeckers, H. Touquet, and P. Vermeersch (eds.): *The EU enlargement and gay politics: the impact of Eastern enlargement on rights, activism and prejudice*: 147–172. London: Palgrave Macmillan UK.

Kuhar, R., and Paternotte, D. (2018): *Anti Gender Campaigns in Europe: Mobilizing against Equality*. Lanham/New York: Rowman & Littlefield International.

Kulick, D. (2005): The importance of what gets left out. *Discourse Studies*, 7 (4–5): 615–624.

Kumar, A. K., and Natarajan, S. (2016): Role of Social Media in Political Campaigning and its Evaluation Methodology: A Review. *Sona Global Management Review*, 10 (3): 1–12.

Lovaas, K.E., and Jenkins, M.M., eds. (2007): *Sexualities and Communication in Everyday Life: A Reader*. London: Sage Publications.

Mencin Čeplak, M., and Kuhar, R. (2010): Boji za enakost: od diskriminacije homoseksualnosti do redifinicije družine. *Socialno delo*, 49 (5–6): 283–298.

Motschenbacher, H. (2010): *Language, Gender and Sexual Identity: Poststructuralist Perspectives*. Amsterdam, Philadelphia: John Benjamins.

Motschenbacher, H. (2011): Taking Queer Linguistics further: sociolinguistics and critical heteronormativity research. *International Journal of the Sociology of Language,* 2011 (212): 149–179.

Motschenbacher, H. (2013): "Now everybody can wear a skirt": Linguistic constructions of non-heteronormativity at Eurovision Song Contest press conferences. *Discourse & Society*, 24 (5): 590–614.

Reher, Š., and Fišer, D. (2018): Kodno preklapljanje v objavah slovenskih uporabnikov Twitterja. In D. Fišer (ed.): *Viri, orodja in metode za analizo spletne slovenščine*: 294–323. Ljubljana: Znanstvena založba Filozofske fakultete UL.

Scott, M., and Tribble, C. (2006): *Textual Patterns: Key words and corpus analysis in language education*. Philadelphia: John Benjamins.

Stubbs, M. (1994): Grammar, Text, and Ideology: Computer-assisted Methods in the Linguistics of Representation. *Applied Linguistics*, 15 (2): 201–223.

van Dijk, T.A. (1993): Principles of critical discourse analysis. *Discourse & Society*, 4 (2): 249–283.

van Dijk, T.A. (2001): Critical discourse analysis. In D. Schiffrin, D. Tannen, and H.E. Hamilton (eds.): *Handbook of discourse analysis*: 352–371. Oxford: Blackwell.

Verheijen, L. (2017): WhatsApp with social media slang? Youth language use in Dutch written computer-mediated communication. In D. Fišer and M. Beißwenger (eds.): *Investigating Computer-Mediated Communication: Corpus-based Approaches to Language in the Digital World*: 72–101. Ljubljana: Znanstvena založba Filozofske fakultete UL.

Vonderschmitt, K. (2012): The Growing Use of Social Media in Political Campaigns: How to use Facebook, Twitter and YouTube to Create an Effective Social Media Campaign. *Honors College Capstone Experience/Thesis Projects. Paper 360*. Available at: http://digitalcommons.wku.edu/stu_hon_theses/360 (5 February 2020).

Yep, G.A., Lovaas, K.E., and Elia, J.P., eds. (2003): *Queer Theory and Communication: From Disciplining Queers to Queering the Discipline(s)*. New York: Harrington Park Press.

Corpus-assisted analysis of water flow metaphors in Slovene online news migration discourse of 2015

***Zoran Fijavž**, Faculty of Education, University of Ljubljana, Slovenia*

***Darja Fišer**, Faculty of Arts, University of Ljubljana, Slovenia, Jožef Stefan Institute, Ljubljana, Slovenia*

Abstract

The chapter examines metaphorical expressions of water flow in online news from the national public broadcasting service during the period of increased migration through Slovenia in 2015. Existing critical discourse research on migration in Slovenia from that period shows that water flow metaphors were used to highlight the threatened safety of citizens and state, while occluding concerns for the safety of the refugees and serving as a justification for border militarization. With theoretical grounding in conceptual metaphor theory, this chapter uses corpus-based critical discourse analysis to examine a corpus of 215 news articles from a public broadcaster using the Pragglejazz metaphor identification method. The results show that the words in the identified metaphorical expressions were nearly never used in the literal sense and that metaphorical expressions were ubiquitous in the discourse. They were used by most journalists and increased in frequency throughout the examined period. While most frequently found in reported statements, water flow metaphorical expressions were also used in journalists' own voice, suggesting that the metaphorical frame was mediatized and reproduced without significant contestation by the broadcaster. The chapter also compares the use of metaphorical expressions across various terms used for those migrating, finding that while all terms were used within metaphorical expressions, *migrant* was not used in non-metaphorical descriptions of the people arriving.

Keywords: corpus-assisted discourse analysis, migration, metaphor, online news, mediatization

1 INTRODUCTION

Cognitive linguistics research suggests that the preferred solutions to social issues can change based on the metaphors used to describe the issues (Thibodeau and Boroditsky 2011). Discourse analysis studies from various countries and time periods also show that migration-related discourse often involves metaphors comparing those migrating[1] to animals, criminals, goods, natural disasters, or warfare, which frames the perceived cause of migration and responses to the related social issues (O'Brien 2003, van der Valk 2003, Santa Ana 1999). In the context of Slovenia, Pajnik (2016) notes that the media coverage on the increased arrival of refugees in 2015 was profoundly negative, mostly reproducing the language used by politicians, including the use of euphemisms, (the razor wire as a *technical barrier*) and metaphorical expressions (*floods of migrants*).

This chapter builds on those observations and attempts to quantify them by providing an extended analysis of the use of water flow metaphors in the Slovene discourse on migration in 2015, using critical discourse analysis operationalized through insights from cognitive and corpus linguistics. The chapter analyses water flow metaphorical expressions (WFMEs) in online news published between August and December 2015 by MMC RTV Slovenia, the national public broadcaster.

In this chapter, we first provide the political context of the increased movement of refugees and the responses to this in Slovenia in 2015. We then lay out the theoretical and methodological framework. Specifically, we work on the premise that metaphors are used in migration discourse to achieve specific aims by the speakers using them, and that such language use is further reproduced by the media through citation or incorporation. We then provide a methodology for metaphor identification and describe how the migration corpus under investigation in this chapter was created.

In the empirical part of this chapter we analyse water flow metaphors in the compiled migration corpus, describing the frequency of the metaphorical expression across the time period, and their literal or metaphorical usage. Next, we examine if the metaphorical expressions were present in the news mainly through citations or were also used by journalists when not reporting statements. In addition, we examine if any notable difference of the use of water flow metaphors can be observed among individual journalists. We also examine if various namings for those migrating interacted with the use of metaphorical expressions. Lastly, we discuss the conclusions of our findings and provide ideas for future research.

1 In this chapter we use the neutral terms *those arriving*, *those migrating*, *refugees*, and *people* (when the refence is clear) for people who were migrating through the Balkan route in 2015, not implying their legal protection status.

2 THE 2015 MIGRATION IN SLOVENIA

Due to increased socio-political unrest in the Middle East and North Africa, closures of alternative routes through Italy, France and Spain and pressures in Turkey, Europe saw historic migration through practically the only channel enabling the arrival of migrants from Arab and African countries, mostly from Syria, Iraq and Afghanistan, through the so-called Balkan Route in the period from July 2014 to March 2016. The route consisted of the so-called bridge countries (Greece, Macedonia, Serbia, Croatia and Slovenia) that migrants crossed in order to reach their destination countries (Germany, Sweden).

The peak of the Balkan Route was late 2015, when nearly one million refugees reached Germany, over 300,000 of whom crossed through Slovenia. Due to insufficient institutional preparations, measures taken during the period to provide basic services to refugees were inadequate at worst and erratic at best. As the capacities for provision consolidated and improved over the period (and new arrivals also decreased towards December 2015, due to international measures taken towards closing the Balkan route by March 2016), migration became increasingly securitized, which was symbolized by the erection of a razor-wired border fence on the Croatian border and the transfer of police powers to the army.

Ladić and Vučko (2016) provide an extensive timeline of the events from July to November 2015.[2] In July 2015, the Slovene government adopted a contingency plan in case of an increased number of arrivals, yet focusing only on asylum seekers, despite the fact that in neighbouring Hungary and Croatia most refugees did not apply for asylum but only transited through them. Between September and December 2015, out of more than 300,000 people transiting only 100 lodged an application for asylum and half of those left Slovenia in 2016 without waiting for the asylum procedure to end.

In September 2015 Hungary completed fence construction on its Serbian border and closed it. After the first larger group of refugees arrived in Slovenia from Croatia, the Slovenian Police tried to return them in accordance with the contingency plan, but Croatia refused. People were left stranded between the two countries without basic necessities, sparking protest from refugees and activists, during which the police also used tear gas. Afterwards, Slovenia pursued the policy of a humanitarian corridor, opening the border crossings and enabling people to continue en route to Austria. In mid-October 2015, Hungary closed its border with Croatia, rerouting the main West Balkans

2 The specific sources of the statements in this overview of events, such as government body statistics or newspaper articles, can be found in Ladić and Vučko (2016). The Peace Institute, the publisher of this report was one one of the non-governmental ogranizations with the capacity to perform monitoring inside short-term reception and long-term accomodation centers with full access.

migration route through Slovenia. Initially, the conditions for refugees entering Slovenia were poor due to a lack of coordination with Croatian authorities as well as within short-stay reception centres; e.g. at the Brežice reception centre, refugees resorted to burning wood, blankets, and trash for heat, as no heated tents were available.

The provision of basic necessities at the reception centres improved over time, but varied greatly by location and day, depending on the coordination between a multitude of actors. Information on the situation was spread mainly through social media, as the press was barred from the camps and only non-governmental organizations who were part of a monitoring group had access. Throughout this, the movement of refugees was restricted, making them rely on state-provided services by necessity. Accounts on the conduct of the police and army include humiliating remarks, pushing and yelling, which in some critical moments even resulted in the separation of families to different locations. For at least three weeks and without explanation, 79 Moroccans were detained in the Aliens Center, which accommodates foreigners undergoing deportation, while new arrivals from Morocco were not separated from the rest of the refugees.

The period was also marked by increasing securitization of migration, despite police records showing an insignificant number of interventions involving refugees. Newspaper polls reported a shift in public opinion on a border fence, from heavy opposition in September 2015 to wide support in November 2015 when the Slovene government started erecting a razor-wire fence, in spite of a decreasing number of arrivals, no reported border crossings outside of agreed entry points, and greatly improved coordination with Croatian authorities. In October 2015, the Defense Act was amended through a fast-track procedure and granted some police powers to the army, reportedly to buttress insufficient police capacities, despite the fact that the army was already supporting the police on the ground under pre-amendment legislation. Furthermore, the amendment was not limited to 2015, but could be activated at any time by a National Assembly supermajority vote. The power transfer also lacked a complaint mechanism in cases of civilians seeking recourse against the conduct of army members.

3 CONNECTING METAPHORS, CORPORA, AND CRITICAL DISCOURSE

In this section, we lay out the theoretical framework for this chapter from the perspective of conceptual metaphor theory, critical discourse analysis, and corpus linguistics.

3.1 Conceptual metaphor theory

Lakoff (1993) foregrounds that metaphors are a fundamental way of organizing human experiences, a part of daily language use, and a basic feature of conceptual systems. Metaphor in this sense is defined as the conventionalized ontological correspondence between a source and target domain. The term *metaphor* is distinguished from the term *metaphorical expression*; the former refers to a cross-domain correspondence (e. g. TIME IS MONEY, capital letters serve to indicate a comparison of domains, not just the two particular words used to name the domains), the latter refers to the language used to express the former (e.g. *to waste time*). While Lakoff and Johnson (1999) assume that metaphors are grounded in embodied human experiences (through direct sensimotor experience or other embodied concepts), Kövecses (2007) argues for a multi-layered understanding between embodiment and specific conceptual metaphors, which allows for the possibility of the social environment and cognitive processes to alter or diminish a metaphor rooted in the embodied experience.

3.2 Critical discourse analysis

Assuming metaphors are intertwined with social and cultural experience, that relationship can be subjected to critical discourse analysis (CDA), defined by van Dijk (1993) as the discipline that analyses how dominant social groups influence public discourse and the impacts thereof on other social groups. While the field does not have a unitary understating of what *discourse* encompasses, in this chapter we will use the term *discourse* as a collection of meanings, metaphors, representations, pictures, narratives, and statements that form a particular perspective about a person, object or event (Baker 2006). Language use in the context of discourse is understood as social action, which both shapes and is shaped by societal and historical circumstances (Austin 1962, Levinson 1983). Discourses may also achieve a state of *naturalization*, which establishes power relations, ideologies, social roles and knowledge systems as self-evidential (Fairclough 2010).

With regard to applying CDA to media content, Fairclough (1995) proposes three levels of analysis: (1) the analysis of socio-cultural practices of a communicative event, which examines the economic, political and cultural context of the communicative acts with various degrees of abstraction; from the immediate situational context to the broader institutional context or the yet wider frame of society and culture, (2) the analysis of discursive practices which uncovers processes of text creation and text reception, and (3) the analysis of a text that investigates traditional linguistic categories, such as vocabulary, semantics, orthography and

phonetics. CDA approaches can sometimes be criticized for a lack of rigorous methodology and a relatively small number of texts which are included in the analysis (Fairclough 1995). The question of operationalization of discursive theories thus becomes vital, and the next two sections propose that corpora may offer a suitable way forward.

3.3 Corpus linguistics

Corpus linguistics is the study of language through examples of its use (McEnery and Wilson 1996). It is distinct from purely qualitative research on text because it employs corpora, large collections of authentic texts in machine-readable form that enable the use of qualitative as well as quantitative methods in order to describe linguistic phenomena (Baker 2006). In the context of CDA, corpora can expand the number of texts in the analysis, reduce researcher bias, make the research more easily replicable, and enable comparative corpus analysis, for example across time or text sources (Baker 2006).

The basic techniques of corpus analysis are the use of concordances, which show how a queried word is used in context, and frequency lists which give a quantitative summary of the phenomena under investigation. Co-occurrence frequencies are used to determine collocations, words or phrases that are typically used with the queried word (Baker 2006). These techniques are provided by most corpus querying tools called concordancers, such as AntConc (Anthony 2014), which is used in this study. However, we use the Sketch Engine concordance (Kilgarriff et al. 2014) as well, which in addition to the basic corpus analysis techniques also provides Word Sketches, which are summaries of a word's collocations organized in lists according to the syntactic relationship between the key word and its collocates, for example all subjects of the verb *to fly*. Krek and Kilgarriff (2006) adapted the tool for Slovene. In this study, the Word Sketch feature is used for the analysis of semantic preferences of a word.

4 RELATED WORK

This section summarizes research relevant to the design and interpretation of the empirical analysis of this chapter. It examines the impact of metaphorical framing of social issues, the metaphors appearing in the discourses on migration, the specific work done on Slovene migration discourse, and lastly presents the metaphor identification methodology we used in our study.

4.1 Metaphors on migration and their impacts

O'Brien (2003) reports the metaphors used in the discourse on migration in the United States during the early 20th century, a time when the country witnessed increasingly restrictive immigration policies. Drawing on historical sources, the primary metaphor to describe the US was that of a body (a country as a body), while migrants were described in the conceptual frames of unhealthy organisms, resources, natural disasters, war, animals and sub-humans. Van der Valk (2003) analysed the discourse on migration in French parliamentary debates between 1996 and 1997 and found that the French state was referred to in metaphors related to a house or a home, and that migrant-related metaphors were those of war, water flow and traffic. Santa Ana (1999) examined the use of metaphoric language in the media during the Proposition 187 initiative in California, which restricted access to public services for undocumented migrants and introduced stricter citizenship checks. The most commonly used metaphors compared migrants to animals. Other metaphors referred to criminals, plants (weeds) and goods.

Metaphorical expressions in migration discourse, however, do not always entail negative evaluations. Salahshour (2016) identified examples of water flow metaphors used positively to describe the effects of migration on the economy in the analysis of *New Zealand Herald* articles published between 2007 and 2008. Furthermore, the metaphors seem to be naturalized, as suggested by the analysis of Baker and McEnery (2005) that identified water flow, vermin and attacker metaphors not only in the 2003 British newspapers corpus, but also identified water flow metaphors in the corpus of the website of the Office of the United Nations High Commissioner for Refugees.

Most similar to the empirical work of this article is the analysis of metaphors in the Croatian newspaper *Večernji list* by Šaina (2016), that compared metaphor use in October 2015, as an increased number of people migrated through the region, and again in February 2016. The metaphors compared migration to water flows and the country to a home. The discourse was thus subject to securitization along the lines defined by Buzan, Wæver, and Wilde (1998): the process of political actors casting situations primarily as a threat and thereby enabling non-conventional responses.

That metaphorical expressions may frame discourses on social issues is suggested by a series of experiments by Thibodeau and Boroditsky (2011) showing that metaphors may shape opinions on social issues in experimental conditions, such as judgements on which societal response is the most appropriate towards acts of crime. However, Marshall and Shapiro (2018) show that such effects likely depend on the prior beliefs of readers. In their experiments, participants with a high degree of identification with the US were more prone to experience disgust and support restrictive policies when subjected to texts using vermin metaphors to describe migrants.

4.2 Slovene migration discourse

In a comprehensive critical discourse analysis of 965 news articles from the Slovenian Press Agency, Tumas (2017) identifies three predominant discourses regarding those migrating: those of management, crime, and war. Firstly, the arrival of people was framed as an amorphous and difficult logistical burden on Slovenia that needed to be managed efficiently, entailing that maintaining control was the goal. The bearers of the burden were also exclusively Slovene citizens or government agencies: what burdens those arriving might face due to these political responses was not discussed. The managerial discourse was also reflected in bookkeeping-like reports on how many people were moved where each day, and various ways of subdividing them into more and less desirable categories. Another way the media emphasized the burden on the state and need to control refugees was through equating refugees with a natural disaster, such as a flood or wave: Slovenia was at risk of "sinking in this migration wave" and "grappling to manage the surging wave of refugees pouring into the country" (Tumas 2017, 35–36). Secondly, Tumas identified the discourse of criminality. Criminological and detective terminology can be found in the texts (the police *catching* refugees after *detecting* them despite *deterrent* measures). The texts also featured generalizations on the unwillingness of refugees to cooperate with the authorities (e. g. through stories of refugees *breaking through* restrictive fences without any further context to the story, such as the treatment of police, long processing times or other factors). Extraordinary incidents were given explanatory power for the entire group of people, and were juxtaposed with calls from various actors for the expansion of the authority of the military. Condemnations of illegal migration also cast refugees as criminals, as in many cases no strictly legal paths existed for those migrating (Kogovšek Šalamon and Bajt 2016). Thirdly, Tumas (2017) observed a war discourse, reflected in war-related terminology (*rapid interventions, convoys of refugees*, *making peace in refugee camps*), as well as the fact that the detractors of handing police powers to the military were dismissed as *ideological, with partial interests*, and *unpatriotic*.

Pajnik (2016) provides a media analysis for the period between August and December 2015 through the lens of the mediatization of society (Mazzoleni and Schulz 1999). This concept describes the transformation of the media from a place of public deliberation into a space for displays of power by elite publics in front of the people. In doing so, the media redefines the public (as mere recipients of the government's discourse) as well as politics (as party politics exclusively). Pajnik (2016) observes that this dynamic was in place throughout the reporting in the time period. "Strong publics" such as the government, police and army were reflected in the source selection, while "weak publics" such as non-governmental organizations, intellectuals, activists and volunteers on the

ground found no or marginal media space. This selection led to an uncritical reproduction of the language of the former, reflected in parallelisms between the discourse used by the media and by the political elites. One example of this is the use of the term *refugee* by government representatives when talking about supporting those migrating and the terms *economic migrant* and *potential terrorist,* when talking about increased border control. This form of terminological shifting was reproduced and left uncontested in the media reports. The framing of migration as a natural disaster, thereby casting migration as a threat, was also reproduced through the use of water flow metaphorical expressions, such as *refugee flood* and *stopping the refugee wave*. Lastly, the media discourse largely replicated the use of euphemisms for the razor wire placed on the Slovene-Croatian border by the authorities, by referring to it as *a tool to direct migration, a technical obstacle*, *a professional tool*, and *a wire that is actually not a wire, but only appears as one* without contestation.

4.3 Overview of metaphor identification procedures

In related work, many authors have not explicitly mentioned a metaphor identification methodology, which likely means they were determined by the authors' judgement. Baker and McEnery (2005) used concordances of a target word (e. g. *refugee*) which were sorted alphabetically according to the letter of the following word. Salahshour (2016) used collocations from a 3-word window to the left and right of the key word (*migrant*, *immigrant*, *Asian* and the plural forms). The corpora used by Baker and McEnery (2005) and Salahshour (2016), were not lemmatized or tagged for parts of speech. Semino et al. (2017) used manual annotators using the Pragglejazz (Steen et al. 2007) metaphor identification procedure, which we will also apply in this chapter. It includes four steps:

1. Reading and comprehending the texts and discourse.
2. Identification of individual lexical elements.
3. Checking for each lexical unit in text:
    a. its contextual meaning in the text;
    b. its basic meaning in other contexts, where basic meaning is presumed when it is more concrete, related to bodily activity, more specific, or historically older than the contextual meaning;
    c. if the basic and contextual meaning differ and if the contextual meaning can be understood through the basic meaning.
4. If the lexical unit fulfils the criteria in step 3, it is considered metaphorical.

5 CORPUS ANALYSIS OF WATER FLOW METAPHORS

This section describes the process for selecting the material that was included in the migration corpus and the method for determining metaphorical expressions. Different methods from corpus linguistics then provide insights into the use of water flow metaphorical expressions (WFMEs) in the migration corpus. The conceptual theory of metaphor assumes that metaphors are a cornerstone of language use, on the basis of which we hypothesize that WFMEs in the migrations corpus were used mostly in a metaphorical sense. Since the metaphor of migration as a water flow has already emerged in related research (Salahshour 2016, Baker and McEnery 2005, Šaina 2016), we hypothesize that the use of WFMEs in the migration corpus was naturalized and persistent across journalists and the months in the period examined in 2015. Pajnik (2016) highlights the parallelism between the discourse of politicians and the media, so we hypothesized that WFMEs were mostly used in reported statements of other speakers and not by the journalists themselves. Kövecses (2007) notes that metaphors also relate to social experience, so an additional hypothesis was that WFMEs were used differently in relation to the semantic preference of the terms referring to those migrating (e. g. *refugee* vs. *migrant*).

5.1 Research dataset and methodology

5.1.1 Text selection and migration corpus creation

The time frame for the analysis was selected to begin with the increased arrival of refugees (mostly from North Africa and the Middle East) to Slovenia from August 2015 to December 2015. We analysed articles published on the web portal of the national public broadcasting service MMC RTV Slovenia under the news category Slovenia. We chose this news portal because it is a wide-reaching[3] public service broadcaster and considered to be (relatively) ideologically and regionally neutral in comparison to other news sites. All articles from the news site from the aforementioned period were manually reviewed, and only those that directly related to the arrival of refugees were included. We also discarded all interviews, since they tend to be longer texts focusing on a single speaker, which could affect the analysis of frequencies. Article subheadings, images, links to other articles, editorial notices and reader comments were not included in the corpus. Journalists' initials were retrieved as metadata.

3 For December 2019, the website was ranked the 6th most visited website in Slovenia, according to data commissioned by the Slovenian Chamber of Advertising (https://www.moss-soz.si/rezultati/, accessed 13 January 2019).

The selected texts were saved in a .txt file and imported into the open-source concordance software AntConc (Version 3.4.4; Anthony 2014). Metadata was also retrieved from article links using the Python package *beautifulsoup* and included the journalists' initials and publication date.

5.2 Description of the migration corpus

After applying the selection criteria, 215 articles were included in the corpus and analysis. The number of articles and tokens[4] per month are summarized in Table 1.

Table 1: Number of articles and tokens per month.

Month	Num. of articles	% of all articles	Num. of tokens	% of all tokens
August	8	4	4,971	3.99
September	58	26	38,901	31.22
October	63	30	37,836	30.36
November	59	27	31,361	25.17
December	27	13	11,543	9.26
Total	**215**	**100**	**124,612**	**100**

There were 124,612 tokens and 15,597 token types in the entire corpus. The average article length was 360 tokens, spanning from 86 to 1,998 tokens, which suggests the corpus included a wide array of text types. In total, there were 24 different journalists, but nine journalists authored approximately 80% of the articles. In case of multiple authors, each journalist was given a share (e.g. for two authors of an article with 86 tokens, each was considered to have contributed 43 tokens).

We first checked whether the corpus reflected actual events during the selected period. It was expected that the use of words to describe obstacles at the border would increase in November, when Slovenia began to set up a fence at the Croatian border. First we conducted a concordance search of the word *build* (*postaviti*) and sorted the first words right of the keyword to get a list of words referring to border obstacles, which were *obstacle* (*ovira*), *barrier* (*pregrada*), *technical means (tehnična sredstva*), *fence* (*ograja*), and *wire* (*žica*). A second concordance was done with the new key words and non-border-related senses of the key word were removed. As summarized in Table 2, the migration corpus clearly reflected the media discourse relating to border descriptions, since it is much higher in November than October.

4 The smallest unit in a corpus. Typically each word and punctuation mark is a separate token. Spaces between words are not tokens.

Table 2: Term frequency of border obstacle expressions.

Term	August	September	October	November	December
obstacle	0	1	13	36	15
barrier	0	3	1	1	2
technical means	0	0	0	10	0
fence	4	24	58	124	121
wire	0	24	1	21	49
Total	**4**	**52**	**73**	**192**	**187**
Num. of tokens	4,971	38,901	37,836	31,361	11,543
Rel. freq. per 1000 tokens	0.8	1.3	1.9	6.1	16.2

5.3 Metaphor identification process

WFMEs in the corpus were identified through the Pragglejazz metaphor identification process (Steen et al. 2007). The authors of the procedure recommend the use of reference sources in identifying basic and contextual meanings, which was done through the Word Sketch feature in Sketch Engine (Kilgarriff et al. 2014) with the Slovene reference corpus Gigafida (Logar Berginc et al. 2012) as a source.

The metaphor identification process in this chapter was performed in the following steps:

1. All articles were read during the selection process, familiarizing the researcher with the texts and discourse.
2. A frequency list of all types was inspected and candidate words that could be a part of the metaphor *MIGRATION IS A WATER FLOW* were included in the third step. The corpus was not lemmatized, so inflected forms of the same words were skipped.
3. The migration corpus was concordanced for all the words identified in the previous step and the forms in which metaphorical expressions occurred were noted. The Word Sketches in the Sketch Engine tool were then used (using the Gigafida reference corpus), and the collocations that had the same syntactic form as the potential WFMEs were examined.
4. If at least two collocates in the Word Sketch suggesting a water flow metaphor were observed, the word was added to the list of identified WFMEs.

5.3.1 Results of the identification procedure

WFMEs occurred at least once in 45% of the articles in our corpus (99 articles). Table 3 presents the WFMEs identified in the Migration Corps along with examples of use from the migration corpus, all identified forms in the migrant corpus, and the semantic preference of the same term in the Gigafida reference corpus. Semantic preference was determined via Word Sketches in Sketch Engine in the Gigafida corpus, examining only the collocations that had the same syntactic properties as the identified forms in the migrant corpus. We grouped these collocations into semantic fields.

Table 3: Results of the identification procedure.

WFME	Example of use	Identified forms in the migration corpus	Semantic preference from Gigafida corpus
flow (tok)[5]	Slovenia is progressing and has economic growth, so the migrant flow has not caused much damage to the country, said the Prime Minister in one of his answers.	migration flow, refugee flow, migrant flow, flow of migrants, flow of refugees (migracijski tok, begunski tok, prebežniški[6] tok, migrantski tok, tok migrantov, tok beguncev)	physical quantities, bodies of water, goods
flow (pretok)	Interior Minister Vesna Györkös Žnidar noted that the flow of migrants across the external borders is too high /…/	migration flow, refugee flow, flow of people, flow of refugees (pretok migracij, pretok prebežnikov, pretok ljudi, pretok prebežnikov)	people, goods and services, water flow, fluids, thoughts, physical quantities.

5 In this paper, we use flow as a translation equivalent for the Slovene expressions tok *and* pretok, *and* influx *for* dotok *and* pritok. However, the analysis was performed on the individual original Slovene expressions.

6 Both *begunec* and *prebežnik* were translated as refugee in this paper. However, as we further analyse in 4.4.4., they carry different connotations regarding the (il)legal status of people. Begunec readily translates to refugee, but *prebežnik* is more complex. The Slovene reference dictionary (Slovar slovenskega knjižnega jezika) defines the word as "those fleeing over". In the context of migration studies (Pajnik, Lesjak-Tušek, and Gregorčič 2001), prebežnik denotes people who severed economic and social ties with their previous environment due to various circumstances, crossed the border without documents and were detained in Slovenia and it is an informal term rather than a legal category (such as *refugee* or *asylum seeker*). Adding to the complexity, the term *immigrant* was used as the English equivalent in the aforementioned research (Pajnik et al. 2002), yet *immigrant* in Slovene would likely have a distinct profile of collocations from *prebežnik* as well as a different frequency of occurence as a loanword. The analysis, however, was performed on the individual original Slovene expressions.

WFME	Example of use	Identified forms in the migration corpus	Semantic preference from Gigafida corpus
influx (dotok)	And until the factors that push them from their home countries are settled, the influx of refugees will not stop, said [Bojan Dobovsek].	influx of refugees, influx of migrants. (dotok beguncev, dotok migrantov, dotok prebežnikov)	goods, immigrants, water, substances and physical quantities
influx (pritok)	Although the influx of refugees who arrived to Slovenia at the end of last week has decreased /.../.	Influx of people, influx of refugees, influx of migrants (pritok ljudi, pritok beguncev, pritok migrantov, pritok prebežnikov)	river names, goods
wave (val)	Bojana Muršič from SD said that the estimates of the government's unpreparedness to the refugee wave were only approximate.	Immigrant wave, [second] refugee wave, migration wave, waves of migrants and refugees, wave of refugees, wave of migrants, second wave of the refugee crisis, wave of economic migrants (imigrantski val, [drugi] begunski val, prebežniški val, selitveni val, valovi migrantov in beguncev, val prebežnikov, val beguncev, val migrantov, drugi val begunske krize, val ekonomskih migrantov)	radio stations, physical phenomena (e.g. gravity wave), destructive natural phenomena (e.g. flood wave), sea, protest, violence, immigrants and refugees, emotions (positive and negative)
to dam (zajeziti)	Failure to comply with the commitments will prove that European policy is not working and Slovenia will need to step up border measures and measures to dam the flow of refugees, /.../.	to dam migrations, to dam the refugee flow (zajeziti migracije, prebežniški tok)	bodies of water, criminal acts

WFME	Example of use	Identified forms in the migration corpus	Semantic preference from Gigafida corpus
damming (zajezitev)	On the Austrian side of the Špilje border crossing with Slovenia, a wire fence was set up for the damming of the refugee flow.	damming the flow of refugees, damming the refugees (zajezitev toka beguncev, zajezitev toka prebežnikov, zajezitev prebežnikov)	migration, water flow, impurities, crime
river (reka)	The river of refugees on an endless walkway on a cold rainy day is reminiscent of a line of prisoners.	river of refugees (reka beguncev)	river names
to splash (pljuskati)	[Janez Janša] also pointed his finger at Germany /.../, and now that this problem is splashing across Slovenia, Germany should enable Slovenia to catch its breath.	to splash [over] (pljuskati [čez])	bodies of water
to dry up (usahniti)	In the near future, the refugee flow that Europe witnessed this fall will dry up /.../.	[the refugee flow] dries up (begunski tok usahne)	water sources, emotions
throughput [of rivers, borders] (pretočnost)	In the event of a reduction of the throughput at the northern border, our country should also increase the control at the southern Schengen border / ... /.	the throughput of the Slovene-Austrian border, sufficient throughput (pretočnost slovensko-avstrijske meje, dovoljšna pretočnost)	water flow, traffic, blood vessels
to overflow (preplaviti)	/.../ because we cannot allow Slovenia to be overflowed by an uncontrolled number of refugees or migrants /.../.	uncontrolled numbers of refugees or migrants overflow [Slovenia] (nenadzorovano število beguncev ali migrantov preplavi)	emotions, bodies of water, larger groups of people

5.4 Quantitative analysis of metaphorical expressions

The identified WFMEs were further analysed quantitatively. We examined if the WFMEs appeared in the corpus in their basic sense, the naturalization of the WFMEs through examining their occurrence over the months and across different journalists, the use of WFMEs in reported speech as opposed to the journalists' own statements and lastly the interplay between the naming of those arriving and the use of WFMEs.

5.4.1 Proportion of metaphorical uses

In this subsection, we analyse whether the identified WFMEs in the migration corpus appeared solely in the metaphorical sense. A concordance was queried for each of the WFMEs, using word roots and a wildcard to find all inflections of the search term. We then manually removed all irrelevant hits from the number of hits and divided the results into metaphorical uses related to refugee groups, metaphorical phenomena not related to refugee groups, and non-metaphorical occurrences. The results of the analysis are shown in Table 4 in order of frequency.

5.4.2 Naturalization analysis

We tested the hypothesis that the use of WFMEs in migration discourse is stable via frequency analysis by months and authors. The data in Table 5 shows that the relative frequency of WFMEs in the articles increased from August to December 2015.

Table 4: Metaphorical and non-metaphorical uses of WFMEs.

Word	Count (all)	Metaph. use (related to migration)	Metaph. use (other)	Nonmetaph. use	% of metaph. use (related to migration)
tok (flow)	108	105	3	0	97.2
val (wave)	101	98	2	1	97.01
pretok (flow)	17	15	2	0	88.24
dotok (flow)	16	14	2	0	87.5
pritok (flow)	6	6	0	0	100
zajezitev (to dam, nominalized)	5	5	0	0	100
pretočnost (throughput)	4	4	0	0	100
zajeziti (to dam)	2	2	0	0	100
usahniti (to dry up)	2	2	0	0	100
pljuskati (to splash)	1	1	0	0	100
preplaviti (to overflow)	1	1	0	0	100
reka (river)	21	1	0	20	4.76
Total	**284**	**254**	**9**	**21**	**0.89**

Table 5: WFMEs use by month.

Month	Num. of tokens	Absolute frequency of WFMEs	Num. of WFMEs types	Rel. freq. of WFMEs per 10000 tokens
August	4,971	8	4	1.61
September	38,901	67	6	1.72
October	37,836	68	7	1.82
November	31,361	79	7	2.52
December	11,543	32	5	2.77
Total	**124,612**	**254**	**12**	**2.05**

We also analysed the use of WFMEs across different authors. The analysis is based on two pieces of information: the relative proportion of articles written by each journalist and the relative frequency of WFMEs in each journalist's texts. The full results are shown in Table 6.

Table 6: WFMEs use by journalist.

Journalist	WFMEs (English[7])	WFMEs (Slovene)	Num. of WFMEs	Total tokens by author	Rel. freq. of WFMEs by author (per 1000 tokens)	% of written articles
T. H.	to overflow, damming, flow, wave, to dam, influx	tok, preplaviti, zajezitev, pretok, val, zajeziti, pritok, reka	60	21,458	2.80	10.5
A. Č.	flow, damming, to splash, wave, to dam, influx,	tok, zajezitev, pretok, dotok, pljuskati, val, zajeziti, pritok,	71	31,383	2.26	16.39
G. K.	flow, influx, wave, to dam, throughput	tok, pretok, dotok, val, pretočnost, zajeziti	56	25,091	2.23	8.2
B. V.	flow, influx, wave	tok, dotok, val, pritok	27	13,251	2.04	4.28
G. C.	flow, wave, to dry up, influx	tok, pretok, dotok, val, usahniti, pritok	69	34,995	1.97	19.01
Al. Ma.	flow, damming, influx, wave, throughput	tok, zajezitev, dotok, val, pretočnost	49	25,148	1.95	9.53
Sa. J.	flow, influx, wave, throughput, to dam	tok, pretok, dotok, val, pretočnost, zajeziti	52	27,825	1.87	7.73
L. L.	influx, flow, wave	pretok, dotok, tok, val	13	7,191	1.81	3.84
B. T.	wave, influx, flow, throughput	val, dotok, tok, pretočnost	16	9,128	1.75	2.3

7 The number of unique WFMEs differs in the English and Slovene column because both tok *and pre*tok were translated as flow and both dotok *and p*ritok were translated as influx. All the figures in the table, however, correspond to the Slovene original expressions on which the analysis was performed.

Journalist	WFMEs (English[7])	WFMEs (Slovene)	Num. of WFMEs	Total tokens by author	Rel. freq. of WFMEs by author (per 1000 tokens)	% of written articles
G. V.	flow	tok	4	2,597	1.54	1.82
A. K. K.	flow, wave, influx	tok, val, pritok	9	12,120	0.74	4.66
K. T.	wave	val	3	6,453	0.46	2.25
B. R.	none	none	0	2,574	0	2.42
A. P. J.	none	none	0	2,502	0	0.39
J. B.	none	none	0	2,437	0	2.33
To. G.	none	none	0	1,719	0	0.7
M. R.	none	none	0	968	0	0.47
P. P.	none	none	0	689	0	0.15
S. B. L.	none	none	0	584	0	0.47
M. N.	none	none	0	569	0	0.23
Ka. Br.	none	none	0	468	0	0.12
A. M.	none	none	0	386	0	0.47
E. Š.	none	none	0	308	0	0.23
K. Št.	none	none	0	238	0	0.15

Two observations arise from this analysis: the first is that WFMEs were a prevalent feature of the texts, and second, none of the journalists with a high share of the articles had a markedly lower relative frequency of WFMEs. From this it can be concluded that the use of WFMEs in the migration corpus was a discursive norm and not a feature of any individual journalist.

5.4.3 WFMEs in reported statements

In this section, we tested the hypothesis that WFMEs in the articles were mostly produced when journalists reported statements from their sources. WFMEs were classified into three categories according to their source: unknown, journalists, and others (politicians and officials),[8] the latter including WFMEs in direct or indirect reported speech or structures with a similar function (e.g., "According to the mayor, the flow needs to be controlled"). The source of WFMEs was marked unknown when it could not be precisely determined. The results in Table 7 show that the majority of WFMEs in the articles did indeed stem from reported statements.

8 Some of the authors of the statements were MPs, ministers (especially the Interior Minister at the time, Vesna Györkös Žnidar) and the Prime Minister (Miro Cerar at the time).

Table 7: WFMEs by source in order of occurrence.

WFMEs	Num. of occurrences	Journalists	Others	Unknown
flow (tok)	105	22	65	18
flow (pretok)	15	1	12	0
influx (dotok)	14	8	6	0
influx (pritok)	6	4	2	0
wave (val)	98	26	63	11
zajeziti (to dam)	2	0	2	0
damming (zajezitev)	5	4	1	0
Throughput (pretočnost)	4	1	3	0
to dry up (usahniti)	2	1	1	0
river (reka)	1	0	1	0
to splash (pljuskati)	1	0	1	0
to overflow (preplaviti)	1	0	1	0
Total	**254**	**67**	**158**	**29**
%	**100**	**26.4**	**62.2**	**11.4**

5.4.4 WFMEs in relation to naming practices

In this section we examine the occurrence of WFMEs in relation to the semantic preference of the terms used to designate people arriving. Through the familiarity established through reading the text during corpus creation, we decided to further examine three nouns used to refer to those migrating: *begunec* (*refugee*), *prebežnik* (*refugee*), and *migrant* (*migrant*).[9] We examined the semantic preference of these terms through observing adjectival collocations by using Word Sketches in Sketch Engine (Kilgarriff et al. 2014) from the Gigafida reference corpus. From the results in Table 8, two observations were further made. Firstly, the adjective *illegal* (*illegalen*) was most strongly associated with *prebežnik*, followed by *migrant*, and lastly *begunec*. It should be noted that documents containing the phrase *ilegalni prebežnik* (*illegal refugee*)[10] were written from 1997 to 2011, however 459 out of 1,125 occurrences of the phrase date to 2001 (further investigation of which is however beyond the scope of this chapter). Secondly, a difference was observed in the collocation of these nouns with the adjective *ekonomski* (*economic*) which was associated with *migrant*, but not with *prebežnik* and *begunec*.

9 In this section we first provide the Slovene terms followed by a translation to avoid confusion due to the translation of *prebežnik* (see Footnote 5).

10 This oxymoron is due to the lack of a better translation of *prebežnik*.

Table 8: Comparison of adjectival collocates between *begunec*, *prebežnik* and *migrant*.*

	begunec/migrant	prebežnik/migrant	begunec/prebežnik
	Kosovar, Palestinian, Bosnian, Bosniak, Istrian, Albanian, Afghani, Chechen, Hague, accommodated, Kurdish, Dalmatian, temporary, Vič,[11] Tibetan, Serbian, North Korean	detained, North Korean, Cuban, Iranian, Albanian, Chinese, foreign, political	Kosovar, Palestinian, Bosnian, Bosniak, Istrian, Afghani, Chechen, Hague, accommodated, Kurdish, Dalmatian, temporary, Vič, Tibetan, Serbian
	numerous	illegal	Albanian, political
	African	unlawful	Cuban, North Korean
	economic, illegal		Illegal
	working, unlawful, crossborder, daily	numerous, working, African, economic, cross-border, daily	Foreign, Chinese, Iranian, unlawful, detained.

* Adjectives in the dark red row only occurred with the first noun in the column. Adjectives in the dark grey row only occurred with the second noun in the column. Light red and light grey denote a partial preference and white denotes no preference for either noun in the compared pair in each column. The minimum frequency for a collocation to be included was set to 20 occurences in the Gigafida reference corpus.

Next we designated a non-metaphorical expression to serve as a comparison to the commonly observed noun phrases of *refugee* or *migrant waves* and *flows* and decided upon the word *arrival* (*prihod*). We performed the same procedures as for candidates for metaphorical expressions with Sketch Engine's Word Sketches from the Gigafida corpus. Genitival collocations were selected from the reference corpus because those were the forms retrieved in a concordance of *arrival* in the migration corpus. All genitival collocations were related to people, in stark comparison to *wave* and *flow* (shown in Table 9).

We then examined all combinations of the words *arrival* (*prihod*), *wave* (*val*) and *flow* (*tok*) and the words *begunec, prebežnik*, and *migrant*. The combination count included the namings for people in both nominal and adjectival forms.[12] The results are summarized in Table 10. Of all possible combinations, the phrase *prihod migrantov* (*arrival of migrants*) did not appear under the given criteria.

11 Vič is a district in Ljubljana which has an asylum home. In this context, viški begunci refers to refugees staying at the asylum home in Vič.

12 That is both *begunski val* and *val beguncev* were counted, translating into *refugee wave* and *wave of refugees*.

Table 9: Semantic preferences of *arrival*, *flow*, and *wave*.

Word	Semantic preference from Word Sketches
arrival (prihod)	government services, vehicles, seasons, different persons
flow (tok)	physical quantities (e.g. electricity), water bodies, goods
wave (val)	radio stations, physical phenomena, destructive natural phenomena (e.g. flood wave), sea, protests, violence, immigrants and refugees, emotions (positive and negative)

Table 10: Co-occurrence frequencies of refugee or migrant and *wave*, *flow*, and *arrival*.

Target word	Naming used	Num. of occurrences
wave (val)	begunec (refugee)	49
	prebežnik (refugee)	17
	migrant (migrant)	11
flow (tok)	begunec (refugee)	23
	prebežnik (refugee)	23
	migrant (migrant)	13
arrival (prihod)	begunec (refugee)	23
	prebežnik (refugee)	33
	migrant (migrant)	0

Finally, we compared the occurrence of WFMEs in phrases with the three terms, sorting them along the axis of their collocation with *illegal* and separately for their collocation with *economic* (shown in Tables 11 and 12). We could not determine a clear interplay of semantic preference of (il)legality and WFMEs; assuming *prebežnik* is more strongly semantically marked with illegality than *migrant* (as suggested by the reference corpus data), one would expect that the nonmetaphorical *prihod migrantov* (*migrant arrival*) would be at least as common as its metaphorical counterpart *val migrantov* (*wave of migrants*), which was not the case.

Table 11: Occurrence of combinations by semantic preference (*illegal*) and the-presence of WFMEs.

	Semantic preference: illegal - no (*begunec*)	Semantic preference: illegal - partly (*migrant*)	Semantic preference: illegal - yes (*prebežnik*)
Metaphorical noun (*wave, flow*)	72	24	40
Non-metaphorical noun (*arrival*)	23	0	33
Total:	**95**	**24**	**73**

Table 12: Occurrence of combinations by semantic preference (*economic*) and the presence of WFMEs.

	Semantic preference: economic - yes (*migrant*)	Semantic preference: economic - no (begunec & prebežnik)
Metaphorical noun (*wave, flow*)	24	112
Non-metaphorical noun (*arrival*)	0	56
Total:	**24**	**168**

However, the interplay between semantic preference for *ekonomski (economic)* and WFMEs indicates that while *migrant waves* did occur, *migrant arrival* never did. This may be explained by the lower frequency of *migrant* in general, as any possible collocation with a low-frequency word is less likely. On the other hand, the corpus covered all of the migration-related online news of a widely viewed news platform and can thus be considered representative of the language used to report on migration during the surveyed time period.

6 CONCLUSION

In this chapter, we have presented an approach to the analysis of metaphorical expressions of water flow in migration discourse from a theoretical framework of cognitive and corpus linguistics, as well as critical discourse analysis. We analysed a corpus of 215 articles between August and December 2015 from the MMC RTV Slovenia news portal on the topic of refugee arrivals. The research presented in this chapter complements previous studies with the use of corpora and quantitative measures, which can aid qualitative analysis by enabling concordancing through the texts for key words, as well as comparative analysis between sources (e.g. comparison of key words used in different sources relative to a reference).

We found that metaphorical expressions of water flow were used almost exclusively in a metaphorical sense, in line with the conceptual metaphor theory (Lakoff 1993), which puts metaphors from the position of mere linguistic ornaments into the position of a fundamental cognitive mechanism to organize experiences. We observed that the frequency of metaphorical expressions of water flow increased over the examined time period and the metaphorical expressions were identified in 45% of the analysed articles. Previous discourse analyses of the same period in Slovenia pointed to the use of *refugee waves* in the framing of refugees as a threat (Pajnik 2016), as well as a way to define the related recipients of harm

(Tumas 2017). Furthermore, in a response elicitation study, Boroditsky and Thibodeau (2011) report that single metaphorical expressions at the beginning of a text affected the readers' immediate preferences for the responses to criminality. It is also conceivable that the actual policies enacted through the lens of securitization further reinforce the discourse that enabled them in the first place. As Tumas (2017, 37) writes: "The presence of the army and erection of a fence made that need for control given."

The next set of findings from this chapter reveals more on the role MMC RTV Slovenia played in the migration discourse in the period. All the journalists appeared to use the metaphorical expressions consistently. Metaphorical expressions of water flow were used primarily in the reported statements from various sources, such as politicians and officials. Only 26.4% of all metaphorical expressions of water flow could be unambiguously attributed directly to the journalists. This finding ties into the analysis of mediatization of the events during the period described by Pajnik (2016); if the metaphor in this context was indeed used to frame migration as a security crisis, the journalists reproduced it nevertheless using their own voice as well as when reporting statements by sources. It seems appropriate to label this as a naturalization and successful resignification of migrations as a threat over the period (to the extent it was not the case prior) in the discursive arena of power relations, ideologies, social roles and knowledge systems in the sense proposed by Fairclough (2010).

Lastly, while looking at the interaction between the naming of those arriving and water flow metaphorical expressions, we did not find a clear connection between the connotation (*illegal* and *economic*) of various ways of referring to the people arriving to Slovenia and the frequency of their co-occurrence with water flow expressions. However, we did find that the phrase *arrival of migrants* never occurred in the corpus.

While examining the terms used for people migrating, we found that the occurrence of *illegalni prebežniki* (*illegal refugees*) in the reference corpus peaked in 2001, accounting for around 40% of all occurrences of the phrase in the entire reference corpus. This raises interesting methodological questions about the notion of the point of reference, the importance of sampling criteria for reference corpus construction and the implications these two factors have on a common claim that corpora reduce researcher bias (cf. Baker 2006). If a word used for people fleeing was strongly connected to illegality within a single year, then corpora will also reflect that, as many reference corpora strongly draw on newspaper texts. The presence of such biases in language technologies is not a new finding (Caliskan et al. 2017), yet there are two important implications for further research using corpus-assisted critical discourse analysis in particular. Firstly, comparative corpus analysis may only succeed in quantifying the differences between two text collections, but cannot account for pre-existing bias within a chosen

reference corpus. It is, for instance, conceivable that a keyword analysis comparing mainstream sources reporting on migration with a collection of texts from anti-immigration sources would render the inherent biases in mainstream sources invisible. Secondly, it does not seem likely this limitation can be addressed by merely increasing the scale of the text collection, which is one of the most promising features of corpus-based approaches.

This chapter also raises new research questions. One line of further work could assess the impact of metaphorical framing in experimental contexts for water flow metaphors, similar to Thibodeau and Boroditsky (2011) and Marshall and Shapiro (2018). Secondly, the metaphorical expressions we examined in this chapter were used to securitize migration, yet it is conceivable they might also have been subject to subversive resignification as a form of counter-speech against securitization. Was there such use, by whom, and did such use circumvent the destructive and arguably dehumanizing frame that *refugee waves* entails? Thirdly, the research presented in this chapter could be further expanded in scope in various ways, from incorporating other news websites, other genres, and social media texts, to extending the analysis to visual materials and speech. Last but not least, studies similar to our work in this chapter have already been performed across a variety of EU national contexts in response to the 2015 events (for a collection of examples, see Krzyżanowski, Triandafyllidou and Wodak 2018), yet valuable research could be done in synthesis and comparison of research within and across countries. In addition to providing insights on the various methodologies used, cross-lingual studies could compare the discourse on migration in destination countries compared to transit countries (e. g. Slovenia as a transit country never had a discursive struggle over the maximum number of arrivals, a central topic in Austria, a neighbouring destination country). In conclusion, such work could further articulate how and why the events in 2015 shifted speech and action about those migrating into the territory of security, and better understand under which conditions the discourse will contain narratives of people, not natural disasters.

Acknowledgements

The work described in this chapter was funded by the Slovenian Research Agency within the national basic research project "Resources, methods, and tools for the understanding, identification, and classification of various forms of socially unacceptable discourse in the information society" (J7-8280, 2017–2020) and the Slovenian-Flemish bilateral basic research project "Linguistic landscape of hate speech on social media" (N06-0099, 2019–2023).

References

Anthony, L. (2014): AntConc: Version 3.4.4. Tokyo, Japan: Waseda University. Available at: https://www.laurenceanthony.net/software (10 July 2017).

Austin, J. (1962): *How to do things with words*. Oxford: Oxford University Press.

Baker, P. (2006): *Using corpora in discourse analysis*. London: Continuum.

Baker, P., and McEnery, T. (2005): A corpus-based approach to discourses of refugees and asylum seekers in UN and newspaper texts. *Journal of Language and Politics*, 4 (2): 197–226. Available at: http://doi.org/10.1075/jlp.4.2.04bak.

Buzan, B., Wæver, O., and Wilde, J. (1998): *Security: A new framework for analysis*. London: Lynne Rienner Publishers.

Caliskan, A., Bryson, J., and Narayanan, A. (2017): Semantics derived automatically from language corpora contain human-like biases. *Science*, 356 (6334): 183–186. Available at: http://doi.org/10.1126/science.aal4230.

Fairclough, N. (1995): *Media discourse* (1st ed.). London: Bloomsbury Academic.

Fairclough, N. (2010): *Critical Discourse Analysis, The Critical Study of Language* (2nd ed.). Edinbrugh: Pearson.

Gigafida. Available at: http://www.gigafida.net (20 August 2017).

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., and Suchomel, V. (2014): The Sketch Engine: ten years on. *Lexicography*, 1 (1): 7–36. Available at: http://doi.org/10.1007/s40607-014-0009-9.

Kogovšek Šalamon, N., and Bajt, V. (2016): Legal Implications of the "Humanitarian Corridor". In N. Kogovšek Šalamon and V. Bajt (eds.): *RAZOR-WIRED, Reflections on Migration Movements through Slovenia in 2015*: 39–49. Ljubljana: Mirovni inštitut. Available at: http://www.mirovni-institut.si/wp-content/uploads/2016/03/Razor_wired_publikacija_web.pdf (31 January 2020).

Kövecses, Z. (2007): *Metaphor in Culture: Universality and Variation*. Cambridge: Cambridge University Press.

Krek, S., and Kilgarriff, A. (2006): Slovene Word Sketches. In T. Erjavec and J. Žganec Gros (eds.): *Proceedings of the 5th Slovenian/First International Languages Technology 2006 Conference: IS-LTC'06*: 62–67. Ljubljana: Institut Jožef Stefan. Available at: http://nl.ijs.si/is-ltc06/proc/12_Krek.pdf (31 January 2020).

Krzyżanowski, M., Triandafyllidou, A., and Wodak, R. (2018): The Mediatization and the Politicization of the "Refugee Crisis" in Europe. *Journal of Immigrant & Refugee Studies*, 16 (1–2): 1–14. Available at: http://doi.org/10.1080/15562948.2017.1353189.

Ladić, M., and Vučko, K. (2016): Slovenia's Response to Increased Arrivals of Refugees: We Don't Want Them, But We Also Don't Understand Why They Don't Want to Stay. In N. Kogovšek Šalamon and V. Bajt (eds.): *RAZOR-WIRED, Reflections on Migration Movements through Slovenia in 2015*: 15–29. Ljubljana: Mirovni inštitut. Available at: http://www.mirovni-institut.si/wp-content/uploads/2016/03/Razor_wired_publikacija_web.pdf (31 January 2020).

Lakoff, G. (1993): The contemporary theory of metaphor. In A. Ortony (ed.): *Metaphor and Thought*: 202–251. Cambridge: Cambridge University Press. Available at: http://doi.org/10.1017/CBO9781139173865.013.

Lakoff, G., and Johnson, M. (1999): *Philosophy in the flesh: The embodied mind and its challenge to Western thought.* New York: Basic Books.

Levinson, S. (1983): *Pragmatics*. Cambridge: Cambridge University Press.

Logar Berginc, N., Grčar, M., Brakus, M., Erjavec, T., Arhar Holdt, Š., and Krek, S. (2012): *Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in ccKRES: gradnja, vsebina, uporaba*. Ljubljana: Trojina, zavod za uporabno slovenistiko; Fakulteta za družbene vede.

Marshall, S., and Shapiro, J. (2018): When "Scurry" vs. "Hurry" Makes the Difference: Vermin Metaphors, Disgust, and Anti-Immigrant Attitudes. *Journal of Social Issues*, 74 (4): 774–789. Available at: http://doi.org/10.1111/josi.12298.

Mazzoleni, G., and Schulz, W. (1999): "Mediatization" of Politics: A Challenge for Democracy?. *Political Communication*, 16 (3): 247–261. Available at: http://doi.org/10.1080/105846099198613.

McEnery, T., and Wilson, A. (1996): *Corpus Linguistics*. Edinburgh: Edinburgh University Press.

O'Brien, G. (2003): Indigestible Food, Conquering Hordes, and Waste Materials: Metaphors of Immigrants and the Early Immigration Restriction Debate in the United States. *Metaphor and Symbol*, 18 (1): 33–47. Available at: http://doi.org/10.1207/s15327868ms1801_3.

Pajnik, M. (2016): Migration in the Mirror of Mediatized Anti-politics. In N. Kogovšek Šalamon and V. Bajt (eds.): *RAZOR-WIRED, Reflections on Migration Movements through Slovenia in 2015*: 63–70. Ljubljana: Mirovni inštitut. Available at: http://www.mirovni-institut.si/wp-content/uploads/2016/03/Razor_wired_publikacija_web.pdf (31 January 2020).

Pajnik, M., Lesjak-Tušek, P., and Gregorčič, M. (2001): *Prebežniki, kdo ste?*. Ljubljana: Mirovni inštitut.

Pajnik, M., Lesjak-Tušek, P., Gregorčič, M., and Vuković, O. (2002): *Immigrants, who are you?*. Ljubljana: Mirovni inštitut.

Šaina, V.-F. (2016): *Metafore »begunske« krize: kritična analiza diskurza v Večernjem listu*. Bachelor's thesis, University of Ljubljana, Slovenia. Available at: http://dk.fdv.uni-lj.si/diplomska_dela_1/pdfs/mb11_saina-veronika-foska.pdf (31 January 2020).

Salahshour, N. (2016): Liquid metaphors as positive evaluations: A corpus-assisted discourse analysis of the representation of migrants in a daily New Zealand newspaper. *Discourse, Context & Media*, 13: 73–81. Available at: http://doi.org/10.1016/j.dcm.2016.07.002.

Santa Ana, O. (1999): "Like an Animal I was Treated": Anti-Immigrant Metaphor in US Public Discourse. *Discourse & Society*, 10 (2): 191–224. Available at: http://doi.org/10.1177/0957926599010002004.

Semino, E., Demjén, Z., Demmen, J., Koller, V., Payne, S., Hardie, A., and Rayson, P. (2017): The online use of Violence and Journey metaphors by patients with cancer, as compared with health professionals: a mixed methods study. *BMJ Supportive & Palliative Care*, 7 (1): 60–66. Available at: http://doi.org/10.1136/bmjspcare-2014-000785.

Slovar slovenskega knjižnega jezika (1970–1991/web version: 2000). Ljubljana: ZRC SAZU. Available at: http://bos.zrc-sazu.si/sskj.html (20 January 2019).

Steen, G. J., Cameron, L. J., Cienki, A. J., Crisp, P., Deignan, A., Gibbs, R. W., Grady, Z., Kövecses, Z., Low, G. D., and Semino, E. (2007): MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol*, 22 (1): 1–39. Available at: http://doi.org/10.1080/10926480709336752.

Thibodeau, P., and Boroditsky, L. (2011): Metaphors We Think with: The Role of Metaphor in Reasoning. *PLOS One*, 6 (2). Available at: http://doi.org/10.1371/journal.pone.0016782.

Tumas, J. L. (2017): *Prisons and the border: Foucault's heterotopia in the context of two national crises*. Master's thesis, University of Ljubljana, Slovenia. Available at: http://dk.fdv.uni-lj.si/magistrska_dela_2/pdfs/mb22_tumas-jenny.pdf (31 January 2020).

van der Valk, I. (2003): Right-Wing Parliamentary Discourse on Immigration in France. *Discourse & Society*, 14 (3): 309–348. Available at: https://www.jstor.org/stable/42888571.

van Dijk, T. A. (1993): Principles of Critical Discourse Analysis. *Discourse & Society*, 4 (2): 249–283. Available at: https://doi.org/10.1177%2F0957926593004002006.

Vezovnik, A. (2017): Securitizing Migration in Slovenia: A Discourse Analysis of the Slovenian Refugee Situation. *Journal of Immigrant & Refugee Studies*, 16 (1–2): 39–56. Available at: http://doi.org/10.1080/15562948.2017.1282576.

Mitigating the frame SEXUAL THREAT in anti-migration discourse online

Fabienne H. Baider, *University of Cyprus, Cyprus*
Anna Bobori, *University of Cyprus, Cyprus*

Abstract

Studies of online comments focused on migration have mainly revealed the negative stereotyping of migrants – a stereotyping that involves different discourse strategies. One such strategy is the use of cognitive shortcuts called frames; the THREAT metaphor was identified as the most frequent frame for migrants (Stephan et al. 1999, Musolff 2015, Baider and Kopytowska 2017). This article examines the strategies used to construct migrants as a threat, specifically, as perpetrators of sexual violence and abuse, and the strategies used to counter this discursive construction. Our theoretical framework is interdisciplinary, we make use of frame analysis (Fillmore 1982), theories of sexual violence (Yuval-Davis 1997) as well as counter-narrative concepts (Braddock and Horgan 2016). Our data comprises online posts and articles shared *before* the Cologne New Year's Eve attacks in 2015 and therefore could not be attributed to those events. The data was collected within the scope of an EU-cofounded program, C.O.N.T.A.C.T., which focused on hate speech. From our analysis we conclude that to make an impact on social attitudes, interventions relying on counter-narratives must be well organized and well thought through because a spontaneous strategy does not seem to be the answer, but can even be counterproductive.

Keywords: hate speech, migration, sexual assault, frame analysis, counter-narrative

INTRODUCTION

Research focused on migration discourse in the press, by politicians, or in online comments has emphasized the negative stereotyping of migrants as involving different discourse strategies. One strategy uses cognitive shortcuts called frames (Fillmore 1982), and the most common frame related to migration is the THREAT metaphor (Stephan et al. 1999, Musolff 2015, Baider and Kopytowska 2017). This paper examines the argumentation strategies used in anti-migration comments that construct Muslims as sexual predators, and discusses how such a frame can be mitigated with counter-arguments. Migrants as sexual violence offenders is not a topic confined to online comments, since the President of the United States declared in January 2019 that, "thousands of illegal immigrants are in Texas jailed for child sex crime".[1] This statement was made, according to the newspaper reporting it, in order to gather support for the wall the president wanted to be built between Mexico and the US to curb or prevent migration from the Central and South American countries. Even though such arguments in anti-migration discourse are unsupported by the facts, few studies have investigated the argumentation strategies used either to construct or debunk the sexual threat frame, and thus this paper aims at addressing this issue. Our data was retrieved online in December 2015, before the 2015 attacks in Köln and in other German cities; they were collected within an EU-cofounded program, C.O.N.T.A.C.T.,[2] focused on hate speech. They consist of comments posted under a Greek Cypriot article focused on specific incidents of sexual assaults perpetrated by migrants in Sweden, among which we concentrate more specifically on comments from the most extensive thread. We first contextualize the data specifying the macro, meso and micro levels (Fairclough 1995) of the topic of migration perceived as a security and sexual threat, specifically in European countries including Cyprus. Our second section presents and analyses the data in order to assess how successful counter-arguments to deconstruct some of the racist claims are. In our third section we propose /suggest strategies to improve counterattacks. Our theoretical framework focuses on frame analysis (Fillmore 1982) as well as theories of sexual violence (Yuval-Davis 1997).

1 FRAMING MIGRATION AS A SECURITY ISSUE

In this first part of our paper we review the THREAT frame associated with migration; in general, this frame constructs migrants as potential criminals, while more specifically it constructs them as sexual predators.

1 The *Sunday Times*, 13th January 2019 https://www.thetimes.co.uk/article/trump-claims-thousands-of-migrants-are-child-sex-offenders-zx5cdw05l

2 http://reportinghate.eu/en/about-us/

1.1 Framing migrants as a criminal threat

That migration is perceived as a security issue has long been a concern among specialists working in the field of migration studies; moreover, this view has intensified since the end of the 1980s, according to some researchers:

> International migration is probably one of the most cited (….) of the so called new security agenda[s], which emerged at the end of the Cold War, and resulted in a broadening and deepening of our understanding of what constitutes a security threat or challenge. (Wohlfeld 2014, 63)

Numerous polls and surveys have recorded that migration is seen as a potential threat to both "national sovereignty and human security" (Thompson 2013).[3] However, distorted beliefs and exaggerated overestimated numbers have caused an overblown perception of the migrant threat. In fact, according to a EUROSTAT report in 2013, only 4% of the residents of EU countries are non-EU nationals, while an Ipsos Mori opinion poll (Wohlfeld 2014, 63) revealed a significant overestimation of the number of migrants by Italians, who believed they represented 30% of the population while the official figure was only 7%. Such distorted perceptions have been explained by the ways in which the media report on migration, in narratives as well as images and how these reports are discussed online. Indeed, the media role in construing the public imaginary cannot be underestimated: they greatly shape public perceptions, opinions, and the 'collective consciousness' (Lippmann 1922). The media sets an agenda, frames a debate, interprets an issue—in terms of migration usually through negative framings, such as the proximisation of fear[4] (Cap 2017) and biased contextualization (Bauder 2011). For example, because the choice of words will have a psychological impact, references to migrants, refugees or asylum seekers as 'illegal migrants', which is the case in far-right articles, but also in some mainstream press, feed into negative feelings associated with illegal / clandestine migration, which is considered as a high security risk (Wohlfed 2014). Even the use of the EU-preferred term 'irregular migration' in the mainstream news (van Dijk 1995) intensifies anxiety among the public. Moreover, research has found that racial considerations play an important role in attitudes towards migrants, non-white people seen as a greater threat; indeed news stories that provide ethnic details of a foreign offender, while withholding this information

3 This dual perception of the threat is important, as it will become clear when we examine the sexual violence topos.

4 *Proximisation* is defined by Cap as a "systematic rhetorical arrangement" by which the speaker aims to "picture the occurring events and their actors as directly affecting the addressee" (Cap 2008, 29). For example when proximizing spatially, the theatening entities are depicted as encroaching *physically* the other entities (Cap 2017, 2). With temporal proximization the conflict is represented not only as imminent but also "as momentous, historic and thus of central significance to the discourse addressee, as well as to the speaker"(Cap 2010, 70). Both the temporal and spatial proximization are used in antimigration discourse Therefore it requires an immediate preventive or neutralizing action. We used that concept to analyse anti immigration discursive strategies in far right discourse (Baider and Constantinou 2017).

for a local miscreant, legitimate and intensify these biases (cf. section 5.2 and the Cypriot press).

Considering the Greek situation,[5] Karydis (2004) explains that Greece has "a repressive legal framework regarding migrants" (2004, 352) because of migration's association with social and economic problems, where the expression 'social problems' is understood as 'serious criminality'. However, yet again the 'facts and data belie them' (ibid). Earlier surveys in Greece (Voulgaris et al. 1995) indicated that between 1989 and 1992 the perception of migrants changed drastically: these surveys recorded an annual 29% increase in feelings against migration and by 1995, a record 69% thought there were too many foreigners. In fact, the number of migrants had remained more or less stable during that period. The authors argued that this perception was due to the changed imaginary, in which the migrant was seen as a threat to public order, an idea fed primarily by the mass media.

While a number of studies have shown that the legal and social imaginary link migrants with crime in Europe, the sexual violence phenomenon that is also part of this imaginary is an understudied phenomenon.

1.2 Framing migrants as a sexual threat

In her study on public discourse and sexual/gendered violence, Keskinen (2011, 107) concluded that race or ethnicity often frame such violence; once again, the media plays a central role in this perception, especially with its focus on the forced marriage, honour killing and rape that occurs in some communities. Gendered and sexual violence become associated with (if not described as a characteristic of) the newly arrived population or with a specific religious / cultural community, accentuating differences between the host population and newcomers. This focus conforms to what is called the 'ideological square', which is defined by the opposition of *us* (right values, good behaviours and correct beliefs) against *them* (wrong values, bad behaviours and distorted beliefs). Van Dijk (1995) described such a square as being the basis for racism; Langton (2017) suggested it to be the basis of hate speech, and we suggest it is also the basis for what we have called "alienating speech" whereby negative feelings such as contempt gradually develop to the point of becoming hate speech (Baider 2020).

Even though unlawful actions and (sexual) violence are also inflicted on migrants of both sexes and all ages, both during their journey to Europe and upon their

5 Although the present article focuses on Cyprus, the Cypriot press and the Greek press run many of the same articles, and comments are posted by mainland Greek speakers as well (Baider and Constantinou 2017).

arrival in the host countries (Women's Refugee Commission report), this (sexual) violence does not make the news.[6]

In short, the crime of a migrant is more likely to make front-page news than a crime against a migrant.[7] The Cypriot press is no exception in this.

1.3 Framing migration in Cypriot public discourse

Like other Mediterranean countries, such as Greece, Spain and Italy, Cyprus has experienced major changes in patterns of inward migration over the last few decades. Migration flows reversed in the last third of the twentieth century from emigration to immigration, resulting in a major increase in immigration levels, beginning in 2004 when Cyprus joined the EU. In 2012, migrants made up 30% of the total population; the number went down with the 2013 economic crisis and stagnated at 16% in 2018, a percentage still far higher than the global or US rates.[8] Refugees represent only a small part of these migrants, a mere 10% (of 16%); others are contract workers who until recently[9] were required to return home once their permit expired (usually after five years) or was revoked.

Foreigners are rarely part of the news on any media--they are not usually interviewed, for example. However, when they do appear in the news it is in relation to negative events, and how they bring problems and no benefits (Baider and Kariolemou 2014, Demetriou 2014). Examples given by the ECRI report (2016) include the use of the derogatory term "illegal immigrant" for any and all migrants (cf. our previous section) and therefore implying a link between unlawfulness and migration; making reference to the citizenship of only suspected *foreign* perpetrators of crime; disproportionate reporting of unlawful actions committed by foreigners (e.g., abuse of refugee allowances, thus stereotyping them as scroungers). Cypriot politicians also promote and legitimize anti-foreigner sentiment in a number of ways:

> Prior to the presidential elections in early 2013 (…), some politicians and public figures blamed migrants for unemployment and portrayed them as receiving higher state benefits than Cypriots, while Chinese and Vietnamese migrant women were stereotyped as prostitutes. (Demetriou 2014)

6 https://www.infomigrants.net/en/post/15920/study-sexual-violence-against-migrant-men-and-boys-on-the-way-to-italy-widespread

7 https://www.theguardian.com/news/2019/aug/01/media-framed-migrant-crisis-disaster-reporting

8 https://in-cyprus.com/cyprus-has-higher-percentage-of-migrants-in-population-than-the-u-s/

9 The EU put pressure on Cyprus to change the law.

Such practices "racialise crime and convey messages about the threat posed by migrants to the country's security, health, welfare and culture" (ECRI 2016), and are likely to trigger anxiety and anger towards foreigners – which can then result in violence. At the same time, we have noted a positive self-representation, which facilitates exclusionary behaviour towards migrants as well as foreigners. Regarding discourse analysis, there have been several studies focused on Cypriot online comments and migration that identified a similar correlation between words relating to foreigners and words expressing fear (fear about crime, fear about job losses, fear of erosion of cultural and religious identity) (Baider 2013, 2017, Baider and Constantinou 2017). Because of the specific political context and the division of the island that has resulted in a Turkish presence in the northern region, Muslim migrants are especially targeted. The discursive frame of violence includes the words *barbarians* or *jihadists* (Baider 2019). This language choice creates cognitive associations between *any* migrant or *any* Muslim with terrorism and violence, i.e. the threat metaphor can be associated with all migrants. The aim of the present article is to examine the argumentation that is built to create the *sexual* threat as part of the anti-migration rhetoric and what counter-argumentation is used to debunk such a claim.

To do so our present study explores a set of Greek Cypriot online comments that were posted under an article describing two rapes in Sweden committed by immigrants.

2 THEORETICAL FRAMEWORK, METHODOLOGY AND DATA

2.1 Framing hate speech

In our analysis we make use of the concept of frame as defined by Fillmore (1982):

> *any system of concepts* related in such a way that to understand any of them you have to understand the whole structure in which it fits; when one of the things in such a structure is introduced into a text or into a conversation *all of the others are automatically made available.*[10] (1982, 11, our italics)

The notion of "frame" is grounded in our cognitive and epistemological knowledge, which itself has been formed by our previous (linguistic and non-linguistic) experiences (Fillmore 1982, 112, Barsalou 1992). Frames can be used to

10 This concept can be also defined, as mentioned by Fillmore himself (1982, 11), with the words *schema, script, scenario* or *cognitive model*. However we prefer to use the word *frame* itself because of the rigidity and square-ness it evokes, which captures the constraints imposed on the bodies and behaviors.

understand complex issues on the basis of what can be called "common ground" (Kecskes 2013), i.e., well-known cultural references and values specific to the socio-cultural context that tend to elicit strong emotions. For example, migration framed as an invasion could implicitly evoke Cypriot history, i.e., the division of the island after the 1974 Turkish invasion (Baider 2017), would target the way Cypriots have experienced the 1974 events-e.g., losses human and material, painful memories, intense feelings of threat. Thus, if migrants are depicted in the press as invaders, the reaction is likely to be fear and a demand to control their "invasion". Such linguistic choices impact social relations (cf. van Dijk 1998), since feeling afraid and threatened could be an incitement to violence in self-defence (Allport 1954, Stephan et al. 1999).

Understanding the ways such frames are constructed can help determine the most appropriate counter-argumentation, counter-narratives or alternative narratives (Gemmerli 2016, Braddock and Horgan 2016, Speckhard et al. 2018): "By gaining an increased understanding of user behaviour on Facebook and other social media platforms we target online, we are able to more carefully create and target our counter-narrative content" (Speckhard et al. 2018).

2.2 Data and general comments

Our data comprises a set of migration-related comments posted on Facebook collected by a Greek Cypriot team (Millar et al. 2017). The methodology – initially employed for a 2015 EU project – has been described in detail in recent publications (Baider et al. 2017, Assimakopoulos and Baider 2019).[11] We identified keywords related to migration (e.g., *migrant*, *refugee*, *foreigner*, etc.) and collected all local news stories containing these that were published online over two three-month periods (April to June 2015 and December 2015 to February 2016) so as to avoid the 2015 summer period when the focus was on the influx of migrants. Even though Cyprus was not affected by the refugee crisis at that time, the Greek Cypriot press offered broad coverage of this phenomenon because of the situation in Greece (Baider and Constantinou 2018, 201). Our data found very few comments posted online for Cypriot articles compared to other countries in the consortium. Only one article received many comments (15 pages of comments, 150 comments, 8,463 words), and the discussion was focused on sexual violence and murder. Forty different people participated in the discussion. They appear to be mostly young people, with pseudonyms referring to some universities, for example.

11 This dataset is from studies conducted as part of a European project on hate speech (reportinghate.eu). All data are under the EU research program C.O.N.T.A.C.T. copyright. Comments were collected and annotated by the University of Cyprus for the Cypriot data; please refer to the website reportinghate.eu.

The article was posted on numerous news sites, including websites of far-right political parties like the Greek neo-Nazi party Golden Dawn; a Google search produces almost 4,000 hits.[12] It was still circulating as news in 2016,[13] although the story initially ran on 11 December 2015. The title of the article reads: Νεαρή γυναίκα βιάστηκε μέχρι θανάτου από μουσουλμάνο λαθρομετανάστη[14] (Young woman raped to death by Muslim illegal immigrant).

Both the photo, which can still be seen in the link (footnotes 9 and 10), and the article are quite graphic. The photo shows a woman screaming (we see only the head and shoulders), with dark hands circling her throat; the article vividly describes two brutal rapes committed by foreigners. One migrant is said to be from Pakistan and the other to be "Muslim", both are described as being illegal migrants. In fact, the caption for the article uses the word illegal twice, thereby emphasizing a security threat (cf. the high anxiety about illegal immigration described in 1.2):

> SHOCK IN SWEDEN. Young woman raped to death by Muslim illegal immigrant. Shock in Sweden reveals that a 34-year-old Swedish illegal immigrant raped a young girl in a luxury hotel garage when she went to pick up her car. (Our italics)

The article later asserts that a "huge rape crisis" perpetrated by Muslims has hit Sweden and Norway, and that there had been 1,000 rapes in the first seven months of 2015. While "studies" showed that Muslim were the perpetrators, there are no specific references or details provided for these. The conclusion incriminates a "they" who do not take action, live in a state of denial, and implicitly expose the population to danger:

> (1) Those numbers are shocking but no one reacts. They continue to believe that illegal immigrants are all refugees seeking a better future. Here are the results.

The pronoun *they* is recurrent in the anti-immigration comments in our data, but there is never any clarification of who these *they* are: the Swedish population, the Swedish authorities, the people who accept migrants? The article elicited a widespread and very strong anti-migrant and anti-Muslim reaction, for the most part extremely virulent, including illegal hate speech against Muslims.[15] Very few individuals responded with "civic monitoring" (Hatakka 2009, 10) – meaning rarely did anyone post contradictory evidence or condemn these racists posts. These few individuals include the author of one of the very first comments, strongly

12 https://xryshaygh.blogspot.com/2015/12/blog-post_52.html

13 http://www.iokh.gr/2016/01/biasmos-mexri-thanatou-sth-souhdia.html

14 livenews Friday, December 11, 2015 10:24:00 AM CET [en] [other]

15 Some of the comments were actually chosen in a perceptual experiment to test perception of hate speech among the young Greek Cypriot population (Assimakopoulos and Baider 2019).

condemning the newspaper and accusing it of purposefully republishing old news to incite fear and fan the fire of hatred:

> (2) this (newspaper) has become porn news. DESPICABLE. This news dates back to 2013. (….) you post it to terrorize the world and so that the haters can spit their bile. (….)[16]

Indeed, the news is false: the rapes and murder did not happen in 2015. Those who posted the article were though "spitting their bile and hatred", and encouraging others to become outraged too. Eighty-five per cent of the participants engaged in aggressive speech, whether in xenophobic speech or anti leftist statements. The comment quoted above, condemning the news story, is an example of a counter-argument, a concept which we explore below.

2.3 Countering hate speech online: creating a counter-narrative

The word *counter-narrative* is used to refer to any narrative (text, video, images, etc.) that aims to respond to the discourse deployed by extremists or malicious groups against certain targets. We cannot know the reasons why participants decide to write counter-narratives, and whether they have an explicit agenda when doing so or why they have the courage to do so. They can be activists or regular commenters, or both or none of these. In any case, they are among the few people trying to counter hate speech. However, counter-argumentation does not often lead to a thread, with counter narrators giving up quickly, most of the time because of ad hominem insults and strongly negative comments addressed to them. In the present thread only two people dared to provide counter arguments to the anti-migrant narrative.

Counter-narratives have been defined in several ways. Direct counter-narratives "deconstruct, discredit or demystify violent extremist messaging through ideology, logic, fact, or humour" (Warrington 2017). Their aim is

> to 'win the argument' by deconstructing and delegitimising extremist propaganda (…) making fun of, challenging and falsifying the extremist ideology's claims or demonstrating the contradiction between extremist utopias and their brutal realities. (Gemmerli 2015, 4)

16 Since the examples in the article have been translated, it is impossible to retrace the originals and this is in accordance with the GDPR rules. We cannot know the reasons why participants are embarking on writing counter-narratives, whether they have an explicit agenda when doing so and why they have the courage to do so. They can be activists or regular commenters or both or none of these categories. In any case they are few people trying to counter hate speech and the counter-argumentation does not often lead to a thread, counter narrators giving up quickly, most of the time because of direct insults and strong negative comments towards then. In that thread only two people dared to provide counter arguments.

Gemmerli (2015) qualifies such narratives as a short-term solution: "This approach *attempts to affect the behaviour* of those who sympathise with or take part in violent extremism *in the short term*" (our italics). They also encourage the passive reader to condemn hateful comments, while they also help to trigger positive feelings (such as empathy) for victims of hate speech. Some counter-tactics – unfortunately – include using discourteous means to oppose online racism; for example, sharing links with memes and parodies / use of ridicule (Hatakka 2019, 10).

Strategies called *alternative narratives* focus more on acknowledging the grievances and suggesting peaceful solutions; or they focus on the victims of hate speech and the consequences they suffer, rather than presenting direct counter-argumentation *per se* (Speckhard et al. 2018). They specifically aim to offer a (positive) alternative to the stereotyping contained in the messages, and try to listen to the ill-will and grievances that motivate such comments. Some studies claim these initiatives have improved the situation (Braddock and Horgan 2016, Warrington 2017, Speckhard et al. 2018). For instance, Speckhard et al. (2018) were able to observe that people reached by counter-narratives in Iraq, "have both thoughtfully engaged with the content of our videos and initiated online debates on ISIS and other contentious socio-political issues that often influence and drive violent extremism". This impact was measured by analysing Facebook reactions generated by four Facebook campaigns in Iraq. Braddock and Dillard (2016) performed a series of meta-analyses related to narratives and persuasion over a 30-year period, and concluded that "exposure to a narrative is *positively* related to the adoption of narrative-consistent viewpoints" (our italics). They argue that narratives have the potential to persuade, independent of context. However, many have expressed doubt about the efficiency of such counter strategies (de Graaf 2009, Gemmerli 2015, 2016, Woron 2018).

In any case, we will adopt the hypothesis that such narratives work, even if only in the short term, since even limited effectiveness is important as it can help prevent an *escalation* of violence (Baider 2020), thus working in the opposite way to radicalization, which has been shown to cause an escalation in violent actions such as revenge, retaliation, etc. (Roy 2008). A major value of counter-narratives is, therefore, their ability to prevent such developments in less radical commenters. The above-mentioned studies concluded that, overall, to be effective such narratives should target the underlying motivations that lead to violence, taking into account the argumentation, grievances and emotions that fuel hate speech. For this reason we believe that analysing data collected from websites may be a way to fulfil these aims and present more constructive counter-narratives.

In our data we examined the few exchanges in which there were both a racist argument and a counter-argument in order to see the kinds of strategies used

to heighten as well as to defuse the emotions described above (fear, anger and hatred).

Only five people engaged in counter argumentation, three very briefly with one or 2 exchanges, while two responded 10 times to other participants. One of those two engaged in exchanging poetry with another commenter during most of his or her exchanges, and the thread stopped very quickly after those messages. It is difficult to know whether they did this on purpose to stop the thread or whether it happened by chance. The second counter-narrator based his or her argumentation on facts (quoting Sharia law or referring to history).

4 ANALYSIS OF THE FRAMES USED IN HATE SPEECH AND COUNTER SPEECH

As mentioned earlier, most participants expressed anger in their statements, but only five people made more than six comments, with two being especially prolific (25 statements), and one of them sharing multiple videos about sexual violence and migrants.

The xenophobic exchanges were of the following kind:

- angry hate speech directed at Muslims with racist slurs and calls "to wipe out this race from this planet";
- reframing the assault as a class, political or religious issue;
- highlighting links to frightening videos, photographs, articles, etc., many of which were other old or fake news, spreading fear, anger and hatred.

4.1 Reframing as a cultural-religious issue

The sexual threat is most often framed as a cultural-religious issue, as in (3) below, or in reference to Sharia law as in (4). The first argument is that in "their" cultures it is acceptable to mistreat women since they consider women second-class citizens and mere objects. Therefore, as objects, they have no rights, rape is little more than "business as usual" so to speak, as explained in (3):

(3) For them this is not anything special, they have learned to be this way in their countries, to rape (…..) women are for them objects;

In the thread below, Sharia law, renamed Islamic law and falsely mentioned as being in the Koran (as hinted at in [2]: *All of this from the Koran)* is cited as the

reason that Muslim civilization is misogynist; it is used to argue that rape is not only normal given the lesser value of women, but it is also *demanded* by that law and by their God (*they think they obey ALLAH*). Such comments encouraged other participants to denounce *all* believers in Allah as monsters as in the subsequent comments:

> (4) (in English in the original) "Under Islamic law, rape can only be proven if the rapist confesses or if there are four male witnesses. Women who allege rape without the benefit of the act having been witnessed by four men who subsequently develop a conscience are actually confessing to having sex. If they or the accused happens to be married, then it is considered to be adultery" – WHEN Islamists rape, IF THERE are four pals to confirm the rape, the girl is accused of adultery!! *All of this from the Koran*!!! Afterwards we SPEAK of HUMAN RIGHTS for these animals (…) *they think they obey ALLAH*..... (our italics).

And while Sharia law is indeed unfair to women, it contains precepts that can be seen as unfair to all Muslims; in fact, Sharia has been denounced as violating the human rights of many sections of society and not only the rights of women.[17] Moreover, only a minority of Muslims believe in Sharia law, which is controversial within the community: more Muslim countries do not apply this law (or are opposed to its application) than those that do.[18] Further, the conclusion that rape is *demanded* by the religious text, i.e., Muslims not only rape women freely but they also have to do this to be a good "Muslim" and in the name of Allah, is a distortion of the quoted text. A graphic parallel made between words in capitals such as HUMAN RIGHTS and ALLAH supports this conclusion and implies a broad cognitive parallel between the lack of human rights and those who follow Allah in general.

Other comments conclude not only that "they" cannot be allowed in Europe, but that "they" have to be annihilated since they pose such a serious threat to women. The identity behind the pronoun *they* becomes blurred in these comments, since the word *Muslims* is also used (e.g., [10]) in links to videos or posts describing Muslims as rapists, thus confusing fundamentalists and non-fundamentalists, discrediting the whole Muslim population.

17 There is a lively debate in Europe regarding the issue of human rights and Sharia law, not only regarding womens' rights, including marriage with under-age girls: "It does not contain a right to freedom of religion, does not confirm the equality before the law of all men regardless of their religion" (cf. https://eclj.org/religious-freedom/pace/la-charia-est-elle-compatible-avec-les-droits-de-lhomme-).

18 See https://www.washingtonpost.com/opinions/five-myths-about-sharia, as well as a 2017 Pew research study showing that "in Eastern Europe and Central Asia – including Turkey (12%), Kazakhstan (10%) and Azerbaijan (8%) – relatively few favor the implementation of Sharia law" (https://www.pewresearch.org/fact-tank/2017/08/09/muslims-and-islam-key-findings-in-the-u-s-and-around-the-world/).

The thread following (4) – counter-arguments and support of the comment – is given below:

(5) Response to (4): Are we sure however that the "Islamic law" represents the Koran? Because in the Middle Ages we had "Christian laws" (i.e., forbidding a woman to have premarital relations, etc.) but *this does not mean that they represented the teachings of Jesus.*

(6) Response to (5): SORRY BUT WE DO NOT wage WAR for the holy VIRGIN and we do not burn people who have premarital relations.

(7) Response to (6): The Christians have committed horrific crimes. I refer not to the Crusaders 500 years ago, I say this for today. This has nothing to do with religion but with complete illiteracy and lack of education. Wars never, never (NEVER) happen because of religion. Religion is simply the excuse given to the masses.

(8) Response to (7): OK it happened with the banner and the cross CENTURIES ago, it is a disgrace, but we are talking about 2015!

(9) Response to (6): And us? Do we not have bizarre prophets?? "30 million Turks will die", etc ..; And if we do not practice what he [the bizarre prophet] said, it is because we cannot put his teachings into practice (we are a weak country).

The first (counter) argument in (5) hints at the fact that Sharia law does not reflect the way Muslims live, in the same way that Christian laws did not reflect how Christians lived. This argument was quickly refuted in (6) and (8), citing a false comparison between a recent event (a rape) and historical religious wars. Moreover, the explanation in (5) suggests that Muslims are dragging the West back to the Middle Ages, a common argument in anti-migration posts.

In (7) the counter-argumentation diverts the discussion to wars resulting from a lack of education, implying the same for sexual assaults (έλλειψη παιδείας και εκπαίδευσης); such a dubious link raises serious questions: does this post imply that educated people would be less violent? That education would stop violence? What kind of education would achieve this? The post is very vague.

Finally in (9) the participant refers to the Cypriot situation to perhaps foster self-criticism, quoting a local religious leader's call for violence and murder (*Do we not have bizarre prophets?? "30 million Turks will die"*). And although self-reflection can foster critical thinking (see section 5), this argument does not deny or question the generalization that Islam is by definition a violent religion, and that all Muslims are therefore dangerous.

These exchanges reveal that, first, counter-arguments engage on the terrain of the haters, disputing/explaining each point hoping to encourage self-reflection. They

do not use alternative narratives that would focus the debate on another dimension of the issue. None of the counter-attacks noted the generalizations and false reasoning used in the original post.

The second broad observation is that they work on logic and downplay any grievance. They do not acknowledge the misogyny of Sharia law, which might help to discredit the argument.

In sum, these examples reflect the general lines of argumentations on both sides:

- the anti-migration argumentation accentuates the difference, deepens the distance and thus fosters contempt /fear/ anger against "them";
- the counter-attackers emphasize similarities to foster understanding;
- neither side seems to convince the other; neither side is able to even raise doubts among those holding the opposite view.

4.3 Cultural reframing for political gains

Some comments reframe the issue of sexual violence by suggesting that Muslims have a political agenda. In (10) Muslims or/ and their leaders are described as having purposefully targeted the West:

(10) I do not want to engage in interminable discussions on these issues, but we should at least agree that the Muslims, *or if you prefer,* their leaders (…) had targeted, have targeted and will target the Western countries either from barbarism or jealousy (…) So let's imagine about what they think, for example, of women, of Christians, of the values that have defined the Greeks and therefore European culture, *and then the image will speak for itself*. Thanks.

1) You should shop in small GREEK-run shops
2) Those LISTENING TO NEWS FROM THE CHANNELS are causing serious harm to their health
3) ALL JOURNALISTS BELONG to a PARTY
4) DO NOT INVEST ONE EURO IN GERMAN PRODUCTS! (INVESTIEREN SIE NICHT 1 EYRO IN DEUTCHE PRODUKTE!) (our italics, capitals are used in the original quote)

The motives for targeting the West are described as "barbarism" (categorizing Muslims as dangerous) and jealousy (categorizing the Self as being in an enviable, favourable position). These attributed motives reveal the ideological square *us* vs *them*, demonizing Muslims and beatifying Christians (or Cypriots), while ignoring the reasons for the alleged "targeting" of the West, such as political turmoil or economic crisis, reasons possibly brought about by the West itself. "Their"

treatment of women and Christians, as well as 'their' opinion of the defining values of Greek culture and therefore European culture should be sufficient to end any such discussion, both now and in the future. The pronoun *they* is initially, and tentatively, applied to just the Muslim leaders, but seems to encompass all Muslims at the end of the quote, defining the entire population as incompatible with European culture and values, as mentioned earlier in our analysis.

Interestingly, the comment ends with some practical advice that *at first* seems to have little to do with the issue, a call to boycott German products. This is ironically written in German, and likely aims to punish Angela Merkel, the German Chancellor at the time, for her responsibility in encouraging the massive influx of migrants to the EU in 2015. It is also a reminder of the boycotts of Jewish-owned shops during the second world war in Germany. It is not only racist (against the Germans) but also at the same time nationalistic (buy in small Greek-owned shops) and xenophobic, with the call to boycott *all foreign-owned shops*. Thus the argument has expanded from a wish to exclude the misogynist Muslims from Cyprus to a desire to exclude all foreigners, a shift we found in an earlier work on extreme-right political discourse in Cyprus (Baider 2017). Another argument attacks the media, using another well-known theme in far-right discourse: all journalists defend the establishment agenda, i.e., the New World Order.

Therefore under the cover of protecting women's rights and denouncing misogyny we read a far-right agenda.

Other comments, more rarely, reframe sexual violence as a class issue, but the shift to hate speech, conspiracy theory included, seems almost unavoidable. In (11), it is stated that only the unimportant people will be affected by migrants' arrivals, not the ruling class who favours this influx of cheap labour:

(11) I would like to see what you would say *if they rape the daughter, the son of someone in the government*

(12) This not going to happen for two reasons (to the rulers, and to the propagandist monsters). First, because their children are guarded by 1200 cops each. And second, because the illegal reptiles have been ordered by their Imam not to harass them....

(13) There is only one solution ... to burn every Muslim. To eliminate them, not one should be left alive.

Immediately in (12) a conspiracy theory is implied, i.e., that immigration (and the ensuing violence) is organized on both sides-- the establishment, the pro-migration people called *propagandists,* and the Muslim countries. The extreme expression "propagandist monsters" (προπαγαδνιστικα εκτρωματα) to refer to liberal minded people and the metaphor "reptiles" (ερπετοειδης) to refer to migrants are in themselves hate speech. Such extremist language triggers or

encourages illegal hate speech in (13), where there is an explicit call for violence, and a virulent hatred expressed against a specific community (according to the Council of Europe 2008 definition of hate speech).

This thread illustrates how misinformation brings about the escalation of words and calls to punitive and even violent actions. The counter-argumentation that resisted and challenged misinformation from the radical right was unable to delegitimize the opposing commenters. In the next section we will offer suggestions for testing other means of counter-argumentation and organizing civic monitoring.

5 COUNTERING FRAMING MIGRANTS AS (SEXUALLY) VIOLENT

5.1 Creating counter-narratives

There are numerous reasons for the violent reactions read in comments (12) and (13), and we must understand them if we want to respond effectively.[19] Specialists in hate have tried to define hate speech by its motivations. Fischer et al. (2018), in an overview of the topic, summarised the variety of ways that hate speech can be motivated: by a kind of generalized anger (Frijda 1986); a generalized evaluation (Ben-Ze'ev 2000); a normative judgment (McDevitt and Levin 1993); a will to devalue others or to hurt the other person (Rempel and Burris 2005); a wish to take revenge or to be avenged (Baider 2013, 2014). It is unclear whether researchers / activists in counter-narratives and extremism have considered those possible motivators of hate speech to create more appropriate strategies.

As a matter of fact, specialists in counter-narrative argue that the most effective messages do not lecture the audience; rather, they offer something to think about and reflect on. They focus on argumentation strategies, recommending the following precepts for the content of counter-narratives (given in Tuck and Silverman 2016):[20]

- Be thought-provoking;
- Quote facts from credible sources to deconstruct, discredit and demystify haters' messages;
- Use emotional appeals to highlight the impact of hate speech;
- Use satire and humour to undermine the claims made online by haters;
- Spread positive stories and messages from people the audience admires.

19 We endorse the hypothesis that hate speech is indeed *discursive hatred,* although other researchers such as Brown (2015) dispute that position.

20 But see Gemmerli (2015, 2016), who disagrees with some of these recommendations.

These suggestions are *aimed at the audience* rather than the haters, and recommendations for writing alternative narratives are very similar.

Furthermore the option of alternative narratives has also been put forward by researchers such as Briggs and Feve (2013, 6). They argue that *alternative narratives* primarily comprise a "positive story about social values, tolerance, openness, freedom and democracy" – arguments that once again seem to target the audience and not the haters. As such, these narratives should take the form of a dialogue that will foster conversation online:

> Comments – and in particular, sustained engagements or conversations – provide a better indication for researchers about how individual users and target audiences overall may be reacting to the messages in counter-narrative videos. (Silverman et al. 2016)

The aim is to manage engagement, i.e., interactions, whether positive or negative, while also providing insight into reactions to a narrative. For instance, *Average Mohamed*[21] videos have inspired young Muslims to debate the role of gender in Islam and to deal with the struggle of having multiple identities. Fostering conversations online means facilitating exposure to alternative viewpoints, which can potentially foster critical thinking or plant a seed of doubt that later matures into a change in attitudes and behaviours (Gemmerli 2015, Silverman et al. 2016, 14):

> The effective narrative must, as a kind of Trojan Horse, sow a seed of doubt, which may be allowed to grow and break down the defence mechanisms from within. (Gemmerli 2015)

Our proposal is to build counter-frames with argumentation based as much on reasoning as on emotion. This will enable us to address the motivators of hate speech on both the cognitive and the affective levels.

5.2 Being prepared, knowing the facts to foster engagement

We have seen in section 4 that the counter-argumentation used by participants addresses neither the haters' grievances nor their "facts". Indeed only one commentator in our data, on the counter- argumentation side, acknowledges the situation in Sweden. Not knowing those facts and treating as propaganda the figures given by the haters without checking them will only feed distrust on their part and on the part of passive readers. For example, in (14) one participant attacks another because the source of the statistics (the Swedish government) he supplied is not acknowledged.

21 https://averagemohamed.com

> (14) *What specific event are you talking*??? Because I sent you specific events and Swedish government statistics !!!!! Or did you not read [my statistics] because in reality you are only interested in reading only your own opinion and propaganda !!!

Maybe a lack of knowledge explains the above lack of engagement with the "facts" given to reflect on; the haters interpret this as a lack of openness to dialogue (correctly, but also paradoxically). We would thus recommend knowing such facts that can be easily found in reports from GOs and NGOs. Some of them have indeed shown, for example, that males from cultures where male dominance and a macho attitude are accepted/expected are more likely to commit domestic or sexual violence. In Germany a government-sponsored study reported a rise in violent crime since 2015, and found that most (sexual) crimes are committed by a certain category of migrants, migrants from specific countries, or migrants with no hope of being allowed to stay:

> Afghans and Syrians were less likely to commit crimes than migrants from North Africa, who stand little chance of receiving permission to stay in Germany.[22]

Although it is only a small number of migrants responsible for such sexual assaults, generalizations are quick and feed into the local fear against all migrants. There are several criteria that could be used to build a counter-argumentation in this context:

- young males in general, *whether migrants or nationals*, are generally more likely to commit crimes, and more likely to become victims of violence;
- young males in specific categories within the local population, such as drug users, are also more likely to commit violent crimes than any other segment of the population;
- people are also more likely "to report crimes if they are committed by people who are different from them, causing some distortion in the crime statistics";[23]
- the media, as we have mentioned before, are more likely to report and give greater coverage to a crime committed by a foreigner.

Therefore knowing the data and how to contextualise the facts are important to foster critical thinking.

5.2 Strategies used to counter-frame the ideological square

As mentioned before, good counter or alternative narratives must foster critical thinking and spark reactions. We also noted earlier (section 4) several strategies

22 https://www.apnews.com/b5f9a0c0848b430c9cf6493d1d310c7b

23 https://nypost.com/2018/01/03/young-male-migrants-fuel-rise-in-violence-in-germany-study-says/

that the participants used included critical thinking and elaborate alternative/ counter- narratives. We will now examine these strategies and suggest ways to make them more effective.

5.2.1 Diverting the debate onto the self: misogyny at home

To respond to accusations that foreign cultures, i.e., those of many migrants, are misogynistic, a consideration of Cypriot attitudes towards women can offer food for thought. Cypriot women are often denied justice due to an imbalanced legal system (in terms of gender): until very recently, in fact, sexism was enshrined in the island's laws. For example, until the year 2000, children born of a *Cypriot mother* would not be recognized as Cypriot, while children with a Cypriot father were automatically awarded Cypriot citizenship. Children (of Cypriots) born abroad have the right to Cypriot citizenship upon application, submitted by one of the parents. However, prior to the 2000 amendment, the children of a Cypriot woman did not have this right, and could acquire citizenship *only with the approval of the Minister of Interior and provided that they fulfilled certain strict criteria*[24] (our italics).

Children of *women* with internally displaced status (i.e., who were from the northern part of the island which is now under Turkish Cypriot rule) have also been denied the status and benefits awarded to children of internally displaced *men (*until 2013 and sometimes later):[25]

> A member of the House of Parliament lodged a complaint in respect of the non-entitlement of persons whose mother, but not father, is a refugee, to acquire refugee status as opposed to persons who acquire refugee status because their father is a refugee and consequently become eligible for state benefits. It was propounded that this fact constitutes an adverse discrimination based on gender between the two categories of persons in breach of the principle of equality.[26]

Such unequal treatment on the basis of sex is also part of everyday life in Cyprus. A 2013 United Nations report[27] regretted "the persistence of patriarchal attitudes and deep-rooted stereotypes regarding the roles and responsibilities of women

24 http://www.mfa.gov.cy/mfa/mfa2016.nsf/5FC67546633D12DEC2257F95002BE630/$file/National%20Report%20of%20the%20Republic%20of%20Cyprus%20on%20the%20Implementation%20of%20the%20conclusions.pdf

25 http://evaw-global-database.unwomen.org/-/media/files/un%20women/vaw/country%20report/asia/cyprus/cyprus%20cedaw%20co.pdf?vs=3841

26 https://www.theioi.org/downloads/78t34/Europe_Cyprus_Commissioner%20for%20Administration_ANNUAL%20REPORT%202006%20Equality%20Authority%20EN.pdf

27 https://www.theioi.org/downloads/78t34/Europe_Cyprus_Commissioner%20for%20Administration_ANNUAL%20REPORT%202006%20Equality%20Authority%20EN.pdf

and men in the family, in the workplace, in political and public life and in society at large. In particular, the Committee notes with concern that the prevalence of such stereotypes contributes to women's disadvantaged position in the State party". Such a disparity in power and rights can manifest itself as sexual violence, reporting of which is often discouraged, if not downright forbidden/taboo. If it is reported then the victim becomes the guilty party.

When referring to sexual assault, the First Pancyprian Survey examining the interpersonal behaviour of the youth, entitled *Violent Behaviour in Interpersonal Relationships of Young Adults in Cyprus for people aged 18 – 2*5 (2012), reported that in Cyprus there is a widespread tendency to "blame the victim" in rape cases, as well as a widespread acceptance of this attitude:

> 70% of the participants… had opinions and attitudes that are conducive to violence such as 'victim blaming', and the belief that violence in relationships is a 'private' matter'.[28]

In fact, while the EU average for victim-blaming views in cases of rape is 22%, the percentage in Cyprus is 44%,[29] the highest in the EU. This figure means that almost half the respondents hold the view that "violence against women is often provoked by the victim".

The social situation described above could restrain commenters from taking the moral high ground in a debate over misogyny.

5.2.2 Diverting the debate onto the self: sexual violence at home

As noted earlier (section 4) some participants tried to counter the idea of Islam being a religion of violence by recalling earlier times (Crusades, wars) when the West was terrorizing the world and the very countries of today's migrants. One tried to diffuse the ideological square *us* (good) and *them* (bad) by recalling that Cyprus's history is not immune to calls to violence, e.g., made by "prophets" or politicians against "others". However, one need not recall the Middle Ages nor limit violence to politicians or priests; these attitudes are prevalent today. One could have focused the debate on the local situation of violence rather than the violence of others. It was an opportunity to publicize the statistics on sexual violence in Cyprus, e.g., widespread and unreported child abuse, with studies showing that one in four girls and one in six boys have been victims of some

28 2017 EIGE report: https://eurogender.eige.europa.eu/system/files/post-files/eshte_data_review_eige.pdf

29 https://eurogender.eige.europa.eu/system/files/post-files/eshte_data_review_eige.pdf

kind of sexual abuse at some stage of their lives – most perpetrated by someone they knew. The real problem is thus sexual violence locally; shifting this to the migrants only ignores the problem in Cyprus, where actions could be taken and be effective if the concern about sexual violence is genuine.[30]

Furthermore, an EU-wide survey carried out in 2015, at the time of the comments, revealed that 15% of women aged 18–74 years have experienced intimate partner physical and/or sexual violence at least once in their lifetime.[31] In the same year, 83.5 % of calls received by the national domestic violence helpline were from women and girls. International reports point out that Cyprus as a State reports very few investigations, prosecutions and/or convictions in cases of domestic violence, despite the high number of cases reported, including the mistreatment of domestic workers.[32] Reports on Cyprus, which were known at the time of the discussion, describe trafficking in human beings being rampant, owing to the shameful practice of offering young women (often minors) "artists" visas' (abolished in 2014, ten years after Cyprus's EU accession in 2004)[33] only to exploit them as prostitutes in the country. What do the outraged commentators say about this situation in Cyprus? This could be seen as a convenient "forgetfulness".

As a matter of fact, one might wonder, given the facts set out above (and in section 5.2.2), whether the online outrage regarding the sexual violence described in the article would have been the same, or even present, if both victim and perpetrator had been Swedish? If both had been migrants? If the woman had been a migrant and the man Swedish? The *fact that the man was a migrant* and violated a "white woman" seems at the heart of the outrage, not the violence *per se,* nor the act *per se*.

The rage seems more about race than about gender, morality or women's rights. At the outset of this article (section 1.1) we argued that migration is framed as a threat to both "national sovereignty and human security". The(sexual) violence perpetrated by migrants is an opportunity to build the threat on both fronts: the security of the nation and of individuals.

30 https://cyprus-mail.com/2018/11/18/one-in-four-children-experience-some-form-of-sexual-abuse-in-cyprus/

31 European Union Agency for Fundamental Rights, 2014. Violence against Women: An EU-wide Survey. Luxembourg: Publications Office of the European Union.

32 The prevalence of sexual violence and harassment experienced by women is very difficult to ascertain, since no administrative body follows the minimum standards as outlined in the Council of Europe guidance, (2016). When five foreign women and their two daughters disappeared between 2016 and 2018, and were found assassinated in 2019, the police were accused of gross negligence, not to mention racism, and the Minister of the Interior and the Chief of Police were forced to resign in 2019 because of public outrage at the lack of actions on the part of the authorities when they were warned at the time of their disappearance.

33 http://evaw-global-database.unwomen.org/-/media/files/un%20women/vaw/country%20report/asia/cyprus/cyprus%20upr%20wg%20report.pdf?vs=3906

5.3 Nationalism, border and gender violence

Research into the semantic frames used in hate speech directed at migrants, within the context of Greek Cypriot online discussions (Baider 2018), confirms the Foucauldian approach to Othering: any discussion of the Other will take place *within* networks of power and *for* the benefit of those who enjoy their economic, political, social and symbolic powers; in the same way that frames work within networks of powers, the system of concepts they employ is primarily created by the media and the public discourse in general.

Framing sexual violence within these power relations in counter-arguments could trigger a debate beyond sexual violence, a debate that considers the structure of social relationships based on violence in economic (exploitation) or political systems. Mardorossian (2002) explained that any analysis of sexual violence should not focus on individual actors but engage with the *social and structural theories* of sexual violence. Some feminist critics have claimed that rape is a way "to inscribe subordinate status on to an intimately known 'Other'" (Moffett 2006) and to maintain "patriarchal order", while Keskinen (2011), along with other analysts (Gingrich 2006), has noted the instrumentalization of gender and sexuality by right-wing parties. In fact, as we have witnessed in the analysis of the comments, several instances of sexual assaults have been used to push a far-right agenda and call for stronger borders within the EU. The most blatant example was the front cover of the Polish magazine *wSieci* ("The Network"), a mass-market politically conservative magazine, which published a highly inflammatory article after the 2015 sexual assaults in Germany. The cover showed a white-skinned, blond-haired woman draped in the European flag, screaming, while three pairs of dark-skinned male hands were groping her body. The headline read "The Islamic rape of Europe". It triggered outraged responses on social media and some Twitter users compared the image to Nazi propaganda. Inside the magazine, the article refers to the rape and sexual assault of hundreds of women in the German city of Cologne on New Year's Eve. "After the events of New Year's Eve in Cologne, the people of old Europe painfully realized the problems arising from the massive influx of immigrants," wrote the report's author, Aleksandra Rybińsk (wSieci.pl., 13 February 2016). Allowing migration into the EU has been metaphorized as allowing the rape and assault of European women.

Indeed, the idea of nation-building as intrinsically gendered and sexualized is not new. In nationalist discourse women appear as central figures in the reproduction of national collectivities and nation states (Anthias and Yuval-Davis 1992, Yuval-Davis 1997, Demetriou and Hadjipavlou 2018). This discourse also implies state control over marriage and reproduction, since the transfer of customs, language, traditions and religion is done usually through the family.

In this way, the woman is perceived as the best channel for nation-building; as a result women's bodies can symbolically and materially become battlefields on which group struggles are played out (Hadjipavlou 2010). Indeed, after analysing the exchanges which are the topic of this study, we concluded that the rage expressed was not so much about a rape, but about a non-European taking control of a European woman.

6 CONCLUSION

The mass media, whose cultural authority constructs or contributes to our understanding of the world, shapes the range of possibilities for understanding the story of migration (Newton 2008). Moreover, the media generally encourages punitive immigration policies and even violence against certain groups, such as those presented as perpetrators of sexual violence. Irregular migration is on the increase (Wohlfeld 2014, 67); consequently feelings of insecurity are also on the rise, feeding populist and far-right activism in EU countries, including Cyprus (Baider and Constantinou 2017). In this article we have analysed how online discussions escalate verbal violence, as they feed on and perpetuate this feeling of insecurity and threat, thus reinforcing stereotyping while we also examined the attempts to mitigate this escalation, which sadly were not very effective. Counter-strategies included among others: 1. debunking some statistics; despite the fact that these statistics were correct, to make an impact they should have been contextualized; 2. making historical comparisons with violence in the West – an argument that not only backfired but served to reinforce the cliché of Muslims being a backward civilization. Our data and analysis indicate a major difference in argumentative power of the two sides: the counter-attackers have neither the means nor the knowledge to make an impact, not even on their own audience (since sometimes facts they disputed were actually true). The analysis showed that arguing with haters or trolls online is very difficult; it requires training and preparation. Responders must have sufficient and appropriate knowledge (about statistics, laws, etc.) to clearly and effectively counter the arguments, while they must also be equipped with certain skills to guide the debate. We suggested an "inward looking" strategy would be effective to point out inconsistencies in behaviour and reasoning regarding the phenomenon of sexual violence, which is instrumentalized for political gains. Such training would require, according to Johnson et al. (2019), "good psychology, social psychology, and know[ing] some history that could actually engage", as well as working as a team.[34] Understanding the argumentation of hate speech narratives and devising effective counter-narratives

34 https://www.theguardian.com/science/2019/aug/22/online-hate-extremism-physics-science

are the first steps. There must be testing and evaluation of a range of counter-narratives in order to understand the impact of such tactics in the specific socio-cultural context under scrutiny.[35] Last, but not least, there must be some sort of constructive *measurement* of the impact – for example, a measurable change in behaviour. Identification of the factors that are responsible for an effective counter-narrative is the final aim, while recognizing that such counter-strategy must always be ready to adapt to new strategies of the other side (Braddock and Dillard 2016). For more long-term actions, it is important to educate the public and help them develop critical thinking to decipher malicious content; sensitizing young people is also key (Dilmaç and Kocadal 2019).[36] Indeed authors such as Ender (2019) envision counter-narratives being used as a pedagogical tool, especially in the communities targeted by hate speech, and a subject that could be included in the field of media literacy. This suggestions finds us in agreement and we would advocate to collaborate and build alliances with local NGOs that deal with migrants' issues so as to ensure an impact within the community the most affected by such hate speech.

References

Allport, G. (1954): *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley Publishing.

Anthias, F., and Yuval-Davis, N. (1992): *Racialized Boundaries*. London, Routledge.

Assimakopoulos, S., Baider F., and Millar S. (2017): *Online Hate Speech in the European Union – A Discourse-Analytic Perspective*. Springer Online. Available at: https://doi.org/10.1007/978-3-319-72604-5.

Assimakopoulos, S., and Baider, F. (2019): Hate speech in online reactions to news articles in Cyprus and Greece. In: *ICGL*, *Proceedings of the Greek linguistics Association*. Available at: http://icgl13.westminster.ac.uk/wp-content/uploads/sites/55/2020/01/ICGL2013_Proceedings_final_optimised.pdf (27 February 2020).

Baider, F. (2017): Thinking globally, acting locally: mainstream supremacist concepts within a local socio-historical context. *Journal of Aggression Language and Conflict*, 5 (2): 178–204.

Baider, F. (2019): Double Speech Act: Negotiating Inter-cultural Beliefs and Intra-cultural Hate Speech among the Youth. *Journal of Pragmatics*, 151: 155–166.

35Impact-of-Counter-Narratives_ONLINE_1 (1) Tanya Silverman Christopher J. Stewart Insights from a year-long cross-platform pilot study of counter-narrative curation, targeting, evaluation and impact www.againstviolentextremism.org Jonathan Birdwell Zahed Amanullah.

36 The authors propose two levels of intervention: a primary level which focuses on media literacy and prevention; a secondary level which focuses on reducing the damage done by cyber-humiliation.

Baider, F. (2019): Le discours de haine dissimulée: le mépris pour humilier. *Déviance et société*, 3 (43): 359–387.

Baider, F. (2020): Pragmatics lost? Reviewing Hate speech Approaches. *Pragmatics and Society*, 2 (11). Forthcoming.

Baider, F., and Constantinou, M. (2017): We'll come at night and find you, traitor. Cybercommunication in the Greek-Cypriot ultra-nationalist space. In O. Hatzidaki and D. Goutsos (eds.): *Crisis and Conflict in Contemporary Greece*: 411–452. Amsterdam: John Benjamins.

Baider, F., and Constantinou, M. (2018): Negotiating empathy in French and Greek Cypriot press: Christian values or social justice in migration discourse? *Studii de lingvistica,* 8: 71–83

Baider, F., and Kariolemou, M. (2014): Normes systémiques et imaginaire linguistique. In M. Karyolemou and P. Pavlou (eds.): *Language Policy and Planning in the Mediterranean World*: 130–151. Cambridge: Cambridge Scholars.

Barsalou, L. (1992): Frames, Concepts, And Conceptual Fields. In A. Lehrer and E. F. Kittay (eds.): *Frames*, *fields, and contrasts*: *New essays in semantic and lexical organization*: 21–74. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bauder, H. (2011): *Immigration Dialectic: Imagining Community, Economy, and Nation*. Toronto: University of Toronto Press.

Ben-Ze'ev, A. (2000): *The Subtlety of Emotions*. Cambridge, MA: The MIT Press.

Braddock, K, and Dillard, J. P. (2016): Meta-analytic evidence for the persuasive effect of narratives on beliefs, attitudes, intentions, and behaviors. *Communication Monographs*, 83 (4): 446–467.

Braddock, K., and Horgan, J. (2016): Towards a Guide for Constructing and Disseminating Counternarratives to Reduce Support for Terrorism. *Studies in Conflict & Terrorism*, 39 (5): 381–404.

Briggs, R., and Feve, S. (2013): *Review of programs to counter Narratives of violent extremism. What works and what are the implications for government*? London: Institute for Strategic Dialogue. Available at: https://www.publicsafety.gc.ca/cnt/ntnl-scrt/cntr-trrrsm/r-nd-flght-182/knshk/ctlg/dtls-en.aspx?i=127 (27 January 2020).

Cap, P. (2008): Towards the proximization model of the analysis of legitimization in political discourse. *Journal of Pragmatics,* 40 (1):17–41.

Cap, P. (2010): *Legitimisation in Political Discourse: A Cross-Disciplinary Perspective on the Modern US War Rhetoric Second Edition*. Cambridge Scholars Publishing.

Cap, P. (2017): *The Language of Fear: Communicating Threat in Public Discourse*. London: Palgrave Macmillan.

de Graaf, B. (2009): Counter-Narratives and the Unrehearsed Stories Counter-Terrorists Unwittingly Produce, *Perspectives on Terrorism*, 3 (2): 5–11.

Demetriou, C. (2014): *Report on measures to combat discrimination, Directives 2000/43/EC and 2000/78/EC, Country Report 2013 – Cyprus, State of affairs up to 1 January 2014*. European network of legal experts in the non-discrimination field. Available at: https://www.refworld.org/pdfid/541aa9e64.pdf (27 January 2020).

Demetriou, O., and Hadjipavlou, M. (2018): Engendering the Post-Liberal Peace in Cyprus: UNSC Resolution 1325 as a Tool. In O. Richmond and S. Pogodda (eds): *Post-Liberal Peace Transitions: Between Peace Formation and State Formation*: 83–104. Edinburgh: Edinburgh University Press

Dilmaç, J.A., and Kocadal, Ö. (2019): Prévenir le cyberharcèlement en France et au Royaume-Uni : une tâche impossible ? *Déviance et Société*, 3 (43): 380–419.

ECRI 2016 report. Available at: http://hudoc.ecri.coe.int/eng?i=CYP-CbC-V-2016-018-ENG (27 January 2020).

Ender, T. (2019): Counter-narratives as resistance: Creating critical social studies spaces with communities. *The Journal of Social Studies Research*, 43: 133–143.

Fillmore, C. J. (1982): *Frame Semantics*. *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing Co.

Frijda, N. H. (1986): *The Emotions: Studies in Emotions and Social Interaction*. Cambridge: Cambridge University Press.

Fairclough, N. (2015): *Language and Power* (3rd edition). New York: Routledge.

Fischer, A., Halperin, E., Canetti, D., and Jasini, A. (2018): *Why We Hate. Emotion Review*, 10 (4): 309–320.

Gemmerli, T. (2015): *Extremism on the internet. The fight against online radicalisation starts offlin*e. Danish Institute for International Studies. Available at: https://www.files.ethz.ch/isn/191363/pb_online_radicalisation_eng_web.pdf (27 January 2020).

Gemmerli, T. (2016): *Avoid the pitfalls of counter-narrative.* Danish Institute for International Studies. Available at: https://www.jstor.org/stable/resrep13218.

Gingrich, A. (2006): Nation, status and gender in trouble? Exploring some contexts and characteristics of neo-nationalism in Western Europe. In A. Gingrich and M. Banks (eds.): *Neo-nationalism in Europe & Beyond*: 29–49. New York: Berghahn Books.

Hadjipavlou, M. (2010): *Women and Change in Cyprus Feminisms, Gender in Conflict*. London: I. B. Tauris

Hatakka, N. (2019): Expose, debunk, ridicule, resist! Networked civic monitoring of populist radical right online action in Finland. *Information, Communication & Society*. Available at: https://doi.org/10.1080/1369118X.2019.1566392.

Johnson, N. F., Leahy, R., Johnson Restrepo, N., Velasquez, N., Zheng, M., Manrique, P., Devkota, P., and Wuchty, S. (2019): Hidden resilience and adaptive dynamics of the global online hate ecology. *Nature*, 573: 261–265.

Karydis, V. (2004 [1998]): Criminality or criminalization of migrants in Greece. An attempt at synthesis. In R. Vincenzo, N. South, and I. Tailor (eds.): *The New European Criminology: Crime and Social Order in Europe*: 350–368. Abingdon: Routledge.

Kecskes, I. (2013): *Intercultural pragmatics.* Oxford: Oxford University Press.

Keskinen, S. (2011): Borders of the Finnish nation – media, politics and rape by 'foreign' perpetrators. In E. Eide and K. Nikunen (eds.): *Media in Motion. Cultural Complexity and Migration in the Nordic Region*: 107–124. Abingdon: Routledge.

Kopytowska, M., and Baider, F. (2017): From stereotypes and prejudice to verbal and physical violence: Hate speech in context. *Lodz Papers in Pragmatics*, 13 (2): 133–145

Langton, R. (2017): The Authority of Hate Speech. *Oxford Studies in Philosophy of Law,* 3: 123–152.

Lippman, W. (1922): *Public Opinion*. New York: Harcourt, Brace and Co.

Mardorossian, C. (2002): Toward a New Feminist Theory of Rape. *Signs: Journal of Women in Culture and Society*, 27 (3): 743–775.

McDevitt, J., Levin, J., and Bennett, S. (1993): Hate Crime Offenders: An Expanded Typology. *Journal of Social Issues,* 58 (2): 303–317.

Moffett, H. (2006): 'These Women, They Force Us to Rape Them': Rape as Narrative of Social Control in Post-Apartheid South Africa. *Journal of Southern African Studies*, 32 (1): 129–144.

Musolff, A. (2015): Dehumanizing metaphors in UK immigrant debates in press and online media. *Journal of Language Aggression and Conflict*, 3 (1): 41–56.

Newton, L. (2008): *Illegal, Alien, or Immigrant: The Politics of Immigration Reform*. New York: NYU Press.

Rempel, J. K., and Burris, C. T. (2005): Let me count the ways: An integrative theory of love and hate. *Personal Relationships*, 12 (2): 297–313.

Roy, O. (2008): Radicalisation and De-radicalisation. *Perspectives on Radicalisation and Political Violence*, 11: 8–14.

Silverman, T., Stewart, C. J., Amanullah, Z., and Birdwell, J. (2016): *The impact of counter-narratives. Insights from a year-long cross-platform pilot study of counter-narrative curation, targeting, evaluation and impact*. London: Institute for Strategic Dialogue. Available at: https://www.isdglobal.org/wp-content/uploads/2016/08/Impact-of-Counter-Narratives_ONLINE_1.pdf (28 January 2020).

Speckhard, A., Shajkovci, A., and Izadi, N. (2018): Using Counter Narrative Campaigns on Facebook to "Break the ISIS Brand" in Iraq. *Homeland Security Today.* Available at: https://www.icsve.org/using-counter-narrative-campaigns-on-facebook-to-break-the-isis-brand-in-iraq/ (28 January 2020).

Stephan, W., Stephan, C., and Gudykunst, W. (1999): Anxiety in intergroup relations: a comparison of anxiety/uncertainty management theory and integrated threat theory. *International Journal of Intercultural Relations*, 23 (4): 613–628.

Thompson, C. (2013): Frontiers and Threats: Should Transnational Migration Be Considered a Security Issue? *Global Policy Journal*, 4 (4). Available at: https://nottspolitics.org/2013/11/28/frontiers-and-threats-should-transnational-migration-be-considered-a-security-issue/ (28 January 2020).

Tuck, H., and Silverman, T. (2016): *The Counter-Narrative Handbook*. London: Institute for Strategic Dialogue. Available at: https://www.isdglobal.org/wp-content/uploads/2016/06/Counter-narrative-Handbook_1.pdf (28 January 2020).

van Dijk, T. A. (1995): Discourse Semantics and Ideology. *Discourse & Society*, 6 (2): 243–289.

van Dijk, T. (1998): *Ideology: A Multidisciplinary Approach*. London: Sage.

Voulgaris, J., Dodos, D., Kafetzis, P., Lyrintzis, C., Michalopoulou, K., Nikolakopoulos E., Spourdalakis, M., and Tsoukalas, C. (1995): The Perception and Treatment of the Other in Contemporary Greece. Findings of an Empirical Research. *Greek Review of Political Science*, 5: 81–100.

Warrington, A. (2017): Countering Violent Extremism via de-Securitisation on Twitter. *Journal of Deradicalization*, 11: 258–280.

Wohlfeld, M. (2014): Is migration a security issue? In O. Grech and M. Wohlfeld (eds.): *Migration in the Mediterranean: human rights, security and development perspectives*: 61–77. Msida: Mediterranean Academy of Diplomatic Studies.

Woron, F. (2018): Dubious claims of counter-narrative videos, Tech and Terrorism. Available at: https://www.counterextremism.com/blog/dubious-claims-counter-narrative-videos (28 January 2020).

Yuval-Davis, N. (1997): *Gender and Nation*. London: Sage.

Social media as facilitator – a cultural and sociolinguistic analysis of hate speech discourse and performance in Japan

***Goran Vaage**, Kobe College, Japan*

Abstract

The last decade saw an increase in debate around hate speech in Japan and the introduction of new laws. In order to better understand hate speech discourse and performance in Japan this paper gives an overview of some historical and current socioeconomic trends in the country before providing a detailed analysis of the relationship between hate groups and target groups. As a case study, the right-wing organization *Zaitokukai* is introduced and its discourse investigated using data from a debate between its former leader and the former Mayor of Osaka. Through the application of performance theory it was found that social media and streaming sites such as YouTube play a role in staging the discussion of hate speech in the Japanese context, thus allowing it to reach a wider audience, and that social media also serve as a facilitator for hate speech on a street level.

Keywords: hate speech in Japan, *zainichi* Koreans, *Zaitokukai*, performance theory, YouTube

1 INTRODUCTION

This paper will give a cultural and sociolinguistic overview of hate speech in Japan, and subsequently position hate speech online and in social media in a broader context. Historical and current trends in discourse and the interaction between hate groups and target groups will be introduced to show how social media is used to organize and stage hate speech, thus serving as a facilitator.

If we accept that hate speech is communication, then we can easily use any elementary model of communication to identify the hate group, message, target group, context (such as medium, place, culture and religion), as well as potential feedback. This is a preliminary and valuable step towards better understanding the dynamics of instances of hate speech. Furthermore, once such variables have been established, it becomes possible to compare instances of hate speech. For example, Japanese instances of hate speech have different messages and target groups than hate speech in other countries, and moreover, the culture and context will be different.

Recent research on hate speech in Japan has mostly focused on the definition and framing of hate speech with the purpose of promoting understanding of this concept, one that remains a recent social and juridical term in the country (Norikoenet 2014, Ryang 2016). This chapter will apply performance theory, first advocated by Goffman (1956), to the analysis of the behaviour of the most central stakeholders in Japanese hate speech discourse, thus contributing to a better understanding of how stakeholders behave and why they do so.

2 CURRENT JAPANESE SOCIETY

Japan prides itself on being an island country (*shimaguni*), and while being an island country is not a good indicator of the amount of contact with other nation states, it is safe to characterize Japan as a developed country with a homogeneous population. Fearon (2003) puts Japan third to last on a world ranking of ethnic and cultural diversity by country, only surpassed by South Korea and North Korea. It is worth pointing out, however, that 16 years has passed since Fearon published his survey, and there has been a steep increase in immigration within this span. According to data from the *Statistical Handbook of Japan 2018* (Statistics Bureau Ministry of Internal Affairs and Communications Japan 2018), 1 in 30 marriages is now between a Japanese national and a non-Japanese national.

Furthermore, while Japan is known for its strict immigration laws these have been eased to allow for the immigration needed to compensate for labour shortages

due to a dwindling birth rate. The crude birth rate of Japan in 2018 was 7.50 per 1000 people, 223rd of 226 territories (CIA 2018). An OECD report from 2017 puts Japan at the top of countries with skills shortages. Over 80% of Japanese firms with over ten employees report having difficulties filling jobs, compared to 40% of firms in the United States and 20% of those in France (OECD 2017).

3 DISTINCTIVE CHARACTERISTICS OF HATE SPEECH IN JAPANESE CONTEXTS

A low birth rate, combined with strict immigration laws, have led to labour shortages in every sector. It is therefore possible to hypothesize that Japanese people are more likely to become engaged in hate speech in Japan because they have fallen out of society socially, and therefor seek somewhere to belong, rather than feeling a rage towards immigrants taking up jobs or using the welfare system. Ito (2014, 439) supports this possibility:

> Under decades of economic depression and neoliberal reform, people tend to feel that they are "vulnerable" in society, and that government would not care about individuals. More and more people would think that they are "abandoned" from society. Current anti-Korean sentiment coincides with such socioeconomic context where "human rights" sounds somewhat "empty".

Another distinctive characteristic of hate speech in Japan has to do with its peculiar and often misunderstood religious situation. The native religion of Japan, the all-pervasive Shintoism, coexists with other religions; two out of three weddings in Japan are Christian, and 85% of all funerals are Buddhist (Hendry 2013). Hate speech in Japan is therefore prone to deal with ethnicity rather than religion or dogmas (e.g. sexual orientation). The premise for hate speech in Japan is that it is primarily racial and politically motivated, catalysed by social changes.

4 DISCRIMINATION AND TARGET GROUPS IN MODERN JAPAN

Although the concept of hate speech and its Japanese form, *heito supiichi,* is a recent addition to the public sphere in Japan, certain groups have been marginalized over a long period of time. In particular, the discrimination of three different groups in Japan has been well documented; the *burakumin* outcast group,

foreigners (stereotypically "Westerners"), and *zainichi* Koreans and Chinese (special permanent residents).

Burakumin literally means "hamlet people", and refers to previous outcast communities in Japan. These communities have traditionally been linked to certain professions considered to be impure (Hankins 2014). Partially because of its taboo label, and partially because of a common belief that the problem of discrimination has already being resolved, the *burakumin* are hardly mentioned in the media anymore, and it is difficult to say anything about the extent of discrimination or hate speech towards them in contemporary Japan. However, the author has in fact come across several people in Japan mentioning the *burakumin* in relationship to certain areas or family names.

Secondly, incidents of foreign nationals reporting being treated in a discriminatory fashion in Japan have been, and still are written about in the news. The most thorough investigation in book form is probably Aradou (2004), with its in-depth look at the Otaru hot springs case. It was reported that several bath houses in the hot spring town of Otaru on the island of Hokkaido had denied foreigners access. Moreover, in 2014 the soccer club Urawa Reds was fined because one of its supporters' groups had denied foreigners access to the stands and put up discriminating banners (Orlowitz 2014). These and similar cases seem to be characterized by a fear or concern among some Japanese that foreigners will not be able to understand or respect Japanese customs or manners.

The third group, the so-called *zainichi* Koreans and *zainichi* Chinese form a sharp contrast to the potentially rule-breaking Western looking foreigners, as they are practically Japanese in appearance, name and culture, yet have become the prime target for hate speech in contemporary Japan. The word *zainichi* is made up of the two characters meaning 'stay' and 'Japan', and refers to residents in Japan with 'special permanent resident' status (*tokubetsueijuusha*). The Japanese Ministry of Justice also publishes detailed statistics on foreign nationals and their status. In 2015 there were 311,463 Koreans in Japan with special permanent resident status, down from 385,232 in 2010. Lee (2012, 1) explains the *zainichi* Korean's special position in Japan as follows:

> [They are] descendants of colonial-era migrants from the southern Korean peninsula during the first half of the twentieth century. It is in fact not always obvious who belongs to the zainichi Korean collective. They appear indistinguishable from the Japanese, and their cultural literacy, use of Japanese pass names, and native fluency in Japanese allows "passing" as a way of life, making them an invisible postcolonial community. Despite the community's high degree of social and cultural assimilation to Japanese society, zainichi Koreans are legally marginalized and treated as foreign residents. Although over 80 percent of zainichi Koreans were born in Japan,

> and the current demographics include highly assimilated second, third and fourth generations, they are categorized as foreign residents unless they go through the strict process of naturalization.

A substantial body of fieldwork and biographies gives a thorough picture of the lives of many *zainichi* Koreans (Lee 2012, Brown 2015, Cho 2016) who, even though they have lived and worked in Japan all their lives, choose not to become fully naturalized Japanese nationals.

For a thorough historical account of discrimination, hate speech and sentiment toward marginalized groups in Japan, including the three groups mentioned in this section, from a journalist/writer's perspective in Japanese, please refer to Yasuda (2015).

5 HATE SPEECH IN CURRENT JAPANESE SOCIETY

Although hate speech can certainly be covert, it is natural to start the investigation of hate speech within a confined culture or country with the kind that is visible and overt, and receives attention through outlets like the media or courts. According to Martin (2018) strong appeals for hate speech laws in Japan began in 2012 due to an increase in anti-Korean rallies and demonstrations, especially by protest groups such as the *Zaitokukai*, short for *Zainichi Tokken o Yurusanai Shimin no Kai,* literally meaning "citizens' group that does not forgive special rights for Korean residents of Japan". On the motivations behind such rallies, Martin (2018, 460-461) notes:

> A combination of North Korean nuclear weapons ambitions, territorial disputes over uninhabited islands with South Korea, and the ongoing friction over how to resolve the Japanese wartime sex-slave issue (euphemistically referred to in Japan as the "Comfort Women" issue), increased tensions between Japan and the Koreas. This, in turn, inflamed attitudes towards the Korean-Japanese community.

In other words, frustration over political tensions between Japan and Korea seems to have been taken out on the *zainichi* Korean group.

This political activity led to the first ordinance against hate speech being passed in the Osaka Assembly January 2016 (the Japan Times 2016). Osaka is the centre of the Kansai region, historically housing many *burakumin* as well as *zainichi* Koreans and Chinese, as well as being a stronghold of the *Zaitokukai*. This was later followed up by the "Act on the Promotion of Efforts to Eliminate Unfair Discriminatory Speech and Behavior against Persons Originating from

Outside Japan" in June the same year, with the first court order of compensation to a *zainichi* Korean September 2016 over defamation by the *Zaitokukai* (the Mainichi 2016).

6 ORGANIZATION AND ACTIVITIES OF THE *ZAITOKUKAI* ONLINE AND OFFLINE

We have thus established a primary target group for hate speech in Japan, the *zainichi* Koreans, and we have an organized group punished for hate speech, the *Zaitokukai*, that we now examine more in detail. The *Zaitokukai* is a nationalist, far-right political organization established and previously led by Makoto Sakurai, with the claim of 16,399 members on their website zaitokukai.info in June 2017. That website has now been offline for some time, but a new, smaller website, containing the rules and stipulations of the *Zaitokukai* and a blog by the new leader, was accessible at the time this chapter was submitted in August 2019. Makoto Sakurai is currently running another site called *Koudou Suru Hoshuu Undou*, which literally translates into "Act – Conservation Movement", advertising much of the same activities as the original *Zaitokukai* site.

The original *Zaitokukai* website in June 2017 featured on its front page a picture of a cartoon character wearing a white head band with a red sun, options for signing up and giving donations, a news feed, an explanation of the purpose the *Zaitokukai*, a link to a ready to print poster of the "four pillars" of the *Zaitokukai*, and an event calendar. These four pillars are the "special permanent resident permit" (*tokubetsu eijuu shikaku*), "subsidization of (North) Korean schools in Japan" (*Chosen gakkou hojou koufu*), "welfare benefit preferential treatment" (*seikatu hogo yuuguu*), and "alias dispensation" (*tsuumei seido*).

The event calendar lists upcoming rallies and demonstrations. Clicking on an entry leads to detailed information on the topic, time, meeting place, contact details (including the name of the person in charge), requests to bring flags and nationalistic symbols (e.g. the rising sun flag, used by the Imperial Japanese Navy), calls to avoid wearing national costumes of other countries than Japan, calls not to bring dangerous objects, and suggestions for paroles.

These paroles are primarily about political and economic policies, but at the street level and in social media calls for target groups to "return home" can be found, while derogatory comments are frequent. The visual aspects of such rallies and counter-demonstration by antifa-groups are well documented by Akiyama (2015). Speaking from the experience of the author, the antifa-groups appear

well organized. The same slogans printed in similar fonts appear repeatedly, including calls to stop racism, framing *Zaitokukai*-activities as hate speech, and even requests for racists to go home. The antifa-groups very often outnumber the original demonstrators.

Thus, the *Zaitokukai* webpage, and its social media functions, are not a medium of hate speech in itself, but rather a facilitator of hate speech, especially at street level, but also potentially for private individuals to use on social media on their own.

7 HATE SPEECH INTERACTION AND DISCOURSE

It is fortunate for the investigation presented in this chapter that footage is available from a debate about hate speech in front of the press in the Osaka City Hall from 10th of October 2014, between the leader of the *Zaitokukai*, Makoto Sakurai, and Touru Hashimoto, a lawyer and at the time Mayor of Osaka. This unlikely event came about during the peak of the clashes between the *Zaitokukai* and the antifa-groups, and it also received attention in the international media. An analysis of this incident, which a *Guardian* journalist described as a "show down" which descended into a "slanging match" (McCurry 2014), offers a unique chance to isolate aspects of the discourse on hate speech in Japan and its bipolarity. Figure 1 shows a summary of the course of events transcribed by the author. The timeline indicated is approximate, but is easily verifiable on streaming sites along with the contents. A complete transcription in Japanese of the debate and Sakurai's monologue to the press can be found in Logmi (2014).

To sum up the main events of the debate, the former Mayor of Osaka, Touru Hashimoto's main point is that the *Zaitokukai* should stop hate speech in Osaka, and that any dissatisfaction with the current system should be addressed through the appropriate democratic channels, like Congress or the court. On the other hand, Sakurai denies spreading hate speech and considers his own utterances to be unproblematic. He disagrees with the rights of the *zainichi* Koreans in Japan, and talks in negative terms about Korea because he believes that Koreans talk badly about Japan. He also claims that he is merely expressing his right to have an opinion, claiming that Hashimoto is trying to deny him his democratic right to freedom of speech.

Content wise there is little constructive dialogue, and no sign of agreement or solution. Rather, the debate functions as a stage for a duel with name-calling and attempts to frame the opponent as bad. This discourse structure is of a kind that is especially suitable for dissemination on social media. Altogether,

the various versions of videos of the debate had over 8 million views on streaming sites as of February 2020. A detailed linguistic analysis of the discourse is provided in the next section.

-12.00	Sakurai has arrived early and starts talking to the press alone.
-11.50	Sakurai says that he is critical of the way that the press work, and expresses dissatisfaction about being taped earlier that day without permission.
-11.00	Sakurai names several newspapers and television stations, and voices criticism about what they report, and asks them rhetorically why they have come.
-8.00	Sakurai criticizes the media for not reporting the truth about Koreans.
-5.00	The debate is about to start and the rules are being read. This is a one-on-one "exchange of opinions" (*iken koukan*). No questions are allowed.
-4.00	Sakurai leaves a copy of his book on Hashimoto's table, but it is returned. It is then shown to the press and displayed on Sakurai's table.
0.00	Hashimoto enters the room, and the debate starts. Hashimoto invites Sakurai to have the first word.
0.15	Sakurai says he wants to ask Hashimoto about hate speech. Hashimoto answers and a debate arises on what constitutes hate speech.
0.30	The debate moves on to the topic of what kind of statements about people of certain nationalities should be allowed.
0.50	An argument breaks out and they both stand up for a while.
1.30	The debate recommences, and Hashimoto says that if Sakurai has any issues with the current political and social system, he should take it to the Congress, instead of bullying minorities. Sakurai says he is not interested in politics.
4.00	Hashimoto urges Sakurai to use the appropriate channels of democracy. Sakurai answers that he is in fact obeying the rules of democracy, and claims that Hashimoto is trying to deny him his democratic rights.
4.4o	Hashimoto says that Osaka has no need for racists. Sakurai denies being a racist.
5.40	Hashimoto again urges Sakurai to use appropriate political channels, and to keep any statements within the law.
5.30	Sakurai says that the *Zaitokukai's* rallies are peaceful.
5.5o	Hashimoto gives examples of hate speech utterances directed to Koreans in the past, and asks Sakurai to stop making such utterances.
6.30	No new arguments are being made, and Hashimoto wants to end the debate.
7.00	Heated words are being exchanged and Hashimoto leaves the room.
7.10	Sakurai turns to the press, and says, as you can see Hashimoto ran away.
7.30	Sakurai wants to continue to talk to the press alone, but is asked to leave the room by the organizers.

Figure 1: Summary of events in the debate between former Mayor of Osaka, Touru Hashimoto and former *Zaitokukai* leader Makoto Sakurai.

8 METHODOLOGY: HATE SPEECH DISCOURSE AS PERFORMANCE AND SOCIAL MEDIA AS A STAGE

The various recordings of this debate available on streaming sites such as YouTube, with meta-information and commentary, demonstrate how social media play a role in staging discussions of hate speech in the Japanese context, and enabling them to reach a wider audience. This section will introduce methodology from performance theory, and analyse some concrete examples of how the two participants mentioned in the previous section present themselves and the other, and how they utilize the public stage given to them.

Hate speech inherently consists of words. But by saying something we are also doing something, as proposed in Austin's speech act theory (1962), and in this way the performance aspect is also a component of hate speech that should not be overlooked. This chapter will thus put special emphasis on this performance aspect in the analysis of hate speech discourse.

Firstly, it is necessary to establish what is meant by the term performance. Performance theory includes a wide array of potential study subjects. As Schechner (1988, xvii) elaborates,

> Performance is an inclusive term. Theater is only one node on a continuum that reaches from the ritualizations of animals (including humans) through performances in everyday life – greetings, displays of emotions, family scenes, and so on – through to play, sports, theater, dance, ceremonies, and performances of great magnitude.

The principal methodology of the sociologist Erving Goffman is to use every day concepts such as performance as a gateway to analysing and understanding human behaviour and interaction. He begins his chapter on performance in his seminal work *The Presentation of Self in Everyday Life* by stating, "When an individual plays a part he implicitly requests the observers to take seriously the impression that is fostered before them" (1956, 17). Hashimoto as a young, successful, but unorthodox mayor requests through his social media performance to be taken seriously as a politician. Sakurai, as an organizer and spokesperson of a right-wing group requests that his ideas and actions should be taken seriously, and that through social media they should reach a wide audience. Through the interaction one cannot help getting the impression that the they are talking across each other instead of together, and although they are performing before the same audience, they are appealing to completely different ones.

Hashimoto and Sakurai have conflicting opinions, and as such represent two different fronts. According to Goffman (1956, 22-30), the front of a performer can be

divided into three different parts, namely setting, appearance and manner. It is worth noting that Sakurai utilizes the setting to give his opinion on the media before the debate, as seen in the quotes below, and to discredit his opponent after the debate.

(1) "You should be ashamed. What about you, Kyodo News? You write bad things about the *Zaitokukai* all the time. You can criticize us, but in that case, what are you doing here? Go home!" (-9.40)

(2) "Where is Mainichi *Shimbun* (newspaper)? Is the perverted Mainichi *Shinbun* here? Do you remember the *waiwai*-incident? You wrote that Japanese are sex-maniacs in your English edition. Stop publishing hate articles!" (-9.20)

Sakurai voices his dissatisfaction with the mainstream media, and offers an alternative view on what hate speech is. These divergent attitudes are also fronted in the actual debate:

(3) S: "You brought up the issue of hate speech."

H: "In Osaka. I have told you to stop making those statements in Osaka."

S: "I'm asking you, what kind of statements?"

H: "Statements bracketing certain people or nationalities together, and then making value judgments."

S: "So, you're saying I'm not allowed to criticize Koreans?

H: "You." (*omae*)

S: "Don't call me *omae!*" (0.20)

Sakurai in his appearance does not see anything wrong with making these kinds of statements, and later in the debate reinforces the argument that he is allowed to have an opinion:

(4) S: "Saying, Koreans go back to the Korean Peninsula, is one opinion, isn't it?"

H: "Stop it, bracketing people together!"

S: "You (*Omae*), Stop renouncing democracy! Stop renouncing freedom of speech!" (6.00)

Another central component of Goffman's discussion of performance is "the popular view that the individual offers his performance and puts on his show 'for the benefit of other people'" (1954, 17). Sakurai believes that he is acting in the best interests of Japan, as evident from the following passage:

(5) "My wish is only to make Japan better, also, if someone uses abusive language about Japan or does something disrespectful, isn't it only natural that I get upset?" (6.30)

Hashimoto, as well, also claims that he is acting in the best interests of people in Osaka and Japan:

> (6) "At any hand, we don't need activities like yours (*omae*) in Osaka, so put forward your claims to the Diet…" (5.20).

Thus, for the benefit of people, both participants claim they are acting to protect the people and treasure democracy, although but with completely different meanings behind these acts.

In terms of the last component of the two men's respective fronts and manner, the use of the second person pronoun *omae* causes friction between the participants as seen in excerpts (3), (4) and (6) above. *Omae* is an informal pronoun, and second person pronouns are generally avoided with strangers, people of higher status, and people of out-groups (*soto*), and replaced with honorifics. *Omae* thus possibly takes on an emphatic, threatening or disrespectful function in this exchange. It is also striking and unique for an event with public officials that plain forms are used throughout the debate instead of polite ones, which would be conventional between two people who do not know each other well. Such diverging representations of the self and other are undeniably performances that will gain support for each participant's views in their respective camps, but not very constructive when it comes to finding solutions to the underlining issues.

9 THE JAPANESE CONTEXT

The label "hate speech" in Japan has until now mostly been used for anti-*zainichi* Korean discourse, possibly only encompassing a part of all potential hate speech. Martin (2018, 461-462) states the following on the extent of staged instances of hate speech in Japan:

> The first ever government study of the issue in 2015 found that there were 347 protests and demonstrations in 2013, and a total of close to 1,200 between April 2012 and September 2015. This is likely a conservative estimate, and the number of instances of lower levels or more individual forms of hate speech is likely several multiples of this number.

The reason why hate speech targeted at *zainichi* Koreans is particularly visible in Japan is most likely that Korea has territorial and historical disputes with Japan (e.g. the so-called "comfort women" issue), and that Japan is frequently targeted in staged demonstrations in Korea. In terms of hate speech in Japan however, thematically it seems to avoid touching upon these disputes, and instead focusing

on minorities' performativity in Japan, that is to say the minorities' political and economic rights, and their clashes with Japanese law.

10 CONCLUDING REMARKS

Hate speech comes in many forms, both in terms of linguistic content and context. These two correlate in that hate speech from political organizations and other organized groups is well-organized and tend to centre around publicized covert racism, whereas at the other end of the spectrum hate speech in computer mediated communication (CMC) is often anonymous, disarrayed, derogatory and direct. A scale going from overtness to covertness is shown in Figure 2 below.

Figure 2: Representation of scale showing the relationship between hate speech sender and hate speech content.

Organized hate speech will always try to balance itself on the line of what is considered to be legal and what falls under freedom of speech.

In a Japanese context, it was shown in this chapter that hate groups in Japan have heavily organized structures and that for these organizations online media in Japan mainly serve as a facilitator of hate speech rather than being a medium for hate speech in itself. Hate speech in Japan can obviously be defined by its content, but it is clear that certain types of activities and discourses also become hate speech through their performativity (Goffman 1956, Butler 1990). For most performances in Japan that can be considered hate speech, there are usually groups or individuals nearby stating that what is being staged is hate speech (Ito 2014, Akiyama 2015), and thus clearly marking it as such.

Hate speech is correlated to a chain reaction of historical and contemporary events, and it is not always the racism in the messages but rather the hate speech and politics as a performance, including the senders, receivers and the targets of the messages, which propels this chain reaction. Stopping such chain reactions is possible by first understanding the different variables included in communication, and based on this systematic knowledge, implementing the measures necessary to prevent them.

References

Akiyama, R. (2015): *Antifa*. Nishinomiya: Rokusaisha.

Aradou, D. (2004): *JAPANESE ONLY: The Otaru Hot Springs Case and Racial Discrimination in Japan*. Tokyo: Akashi.

Austin, J. (1962): *How to Do Things with Words.* Oxford: Clarendon Press.

Brown, L. (2015): Caught between two Countries: Zainichi Koreans in Japan. *Virginia Review of Asian Studies*, 17: 254–260.

Butler, J. (1990): *Gender Trouble: Feminism and the Subversion of Identity*. Abingdon-on-Thames: Routledge.

Cho, Y. (2016): Koreans in Japan: A Struggle for Acceptance. *Law School International Immersion Program Papers*, (2): 1–16.

CIA (2018): *The World Fact book.* Available at: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2054rank.html (25 July 2019).

Fearon, J. (2003): Ethnic and Cultural Diversity by Country. *Journal of Economic Growth*, 8 (2): 195–222.

Goffman, E. (1956): *The Presentation of Self in Everyday Life.* Edinburgh: University of Edinburgh Press.

Hendry, J. (2013): *Understanding Japanese Society.* Abingdon-on-Thames: Routledge.

Hankins, J. (2014): *Working Skin: Making Leather, Making a Multicultural Japan*. Oakland: University of California Press.

Ito, K. (2014): Anti-Korean Sentiment & Hate Speech in the Current Japan: A report from the Street. *Procedia Environmental Sciences*, 20: 434–443.

Koudou Suru Hoshuu Undou. Available at: http://www.koudouhosyu.info/ (25 July 2019).

Lee, S. (2012): Diversity of Zainichi Koreans and Their Ties to Japan and Korea. *Studies on Multicultural Societies*, (8): 1–28.

Logme (2014): *[Unabridged] Osaka Mayor Hashimoto Tohru VS Zaitokukai Leader Sakurai Makoto! 8 Minutes of Intense Debate, All of the Exchange.* Available at: https://logmi.jp/business/articles/25539 (7 January 2020).

Norikoenet (2014): *What is Hate Speech? What Kind of Thing is Racism?* Tokyo: Nanatsumori.

Martin, C. (2018): Striking the right balance: Hate speech laws in Japan, the United States, and Canada. *Hastings Constitutional Law Quarterly*, 45: 455–532.

McCurry, J. (2014): *Japanese 'hate speech' debate abandoned as insults fly.* Available at: https://www.theguardian.com/world/2014/oct/21/japan-hate-speech-debate-abandoned-toru-hashimoto-makoto-sakurai (25 July 2019).

OECD (2017): Getting Skills Right. Available at: https://www.oecd-ilibrary.org/employment/getting-skills-right_25206125 (25 July 2019).

Orlowitz, D. (2014): Twitter picture. Available at: https://twitter.com/aishiterutokyo/status/442286373806931970 (9 June 2017).

Ryang, S. (2016): *What is Japanese Style Hate Speech: Appearance of Racism that Leads to Distruction of Society.* Tokyo: Kage Shobo.

Schechner, R. (1988): *Performance Theory*. New York: Routledge.

Statistics Bureau Ministry of Internal Affairs and Communications Japan (2018): *Statistical Handbook of Japan 2018*. Available at: https://www.stat.go.jp/english/data/handbook/index.html (25 July 2019).

The Japan Times (2016): *Osaka Assembly Passes Nation's First Ordinance against Hate Speech.* Available at: https://www.japantimes.co.jp/news/2016/01/15/national/osaka-set-pass-japans-first-ordinance-hate-speech-will-name-shame-offenders/#.XTj_i-R7nmI (25 July 2019).

The Mainichi (2016): *Korean Resident of Japan's Legal Battle for Dignity Ends in her Favor but Problems Remain*. Available at: https://mainichi.jp/english/articles/20180320/p2a/00m/0na/003000c (25 July 2019).

Yasuda, K. (2015): *Hate Speech: The Hate and Violence of Patriots.* Tokyo: Bunshun.

Zainichi Tokken o Yurusanai Shimin no Kai. Available at: zaitokukai.info (9 June 2017).

Zainichi Tokken o Yurusanai Shimin no Kai. Available at: https://www.zaitoku814.com/ (25 July 2019).

About the authors

Fabienne H. Baider is an Associate Professor at the University of Cyprus and works on semantics and discourse from a socio-cognitivist and contrastive perspective. Her research interests include gender and language, conceptual metaphors and emotions in political discourse, and online communication with a focus on hate speech. Her methodology includes corpus linguistics and critical discourse analysis. She has published internationally in linguistic journals and edited volumes on emotions, on gender and on hate speech. She coordinates the EU Social Justice project C.O.N.T.A.C.T and the national project H.O.P.E. focused on counter narratives. She is a partner in the EU Social Justice projects SHELTER, focusing on hate crimes, and IMsyPP, focusing on the automatic detection of hate speech.

Anna Bobori holds a master's degree in business administration and in biotechnology. She is a PhD student at the Department of European Studies at the University of Cyprus. Her thesis focuses on the automatic detection of hate speech in online comments about migration on Greek portals from a discourse analysis perspective, using data analysis and machine learning tools. Before embarking on her PhD, she worked for the European Union Agency for Network and Information Security (ENISA) in Athens as Lead Auditor on Information Security Management Systems, for the European Institute of Innovation and Technology (EIT) in Budapest as project manager for Horizon 2020 projects and for the European Asylum Support Office (EASO) in Malta as a Quality Assurance and Evaluation Officer.

Zoran Fijavž holds a BA in interlingual communication from the Faculty of Arts, University of Ljubljana and is currently pursuing an MA in cognitive science at the Faculty of Education, University of Ljubljana. His main research interests are at the intersection of language, technology, cognition and sociality, in particular in technology-assisted critical discourse analysis.

Darja Fišer is an Associate Professor at the Department of Translation Studies at the Faculty of Arts, University of Ljubljana, and Senior Research Fellow at the Department of Knowledge Technologies at the Jožef Stefan Institute. She is also Vice Executive Director of CLARIN ERIC, a European Research Infrastructure for Language Resources and Language Technologies, and Chair of the Steering Committee of ESSLLI, the oldest and biggest European summer school on language, logic and information. As a researcher, she is currently active in the fields of computer-mediated communication and socially unacceptable online discourse practices using corpus-linguistics methods and natural language processing.

Vojko Gorjanc is a Full Professor at the Department of Translation Studies of the Faculty of Arts, University of Ljubljana, and Head of its Slovene Language Chair. He is Head of the central Slovene-language research program at the University of Ljubljana (*Slovene – Basic, Contrastive and Applied Research*). In the past few years he has been actively involved in the formation of the new university Center for Language Resources and Technologies. His teaching and research interests lie with sociolinguistics and critical discourse analysis, which he is combining with corpus-linguistic methodology. Presently his main focus rests on language and linguistic ideologies, especially the ideologies of standard language and standard-language culture, and power relations in discourse, including interpreter-mediated interaction.

Nikola Ljubešić is a Research Associate at the Department of Knowledge Technologies, Jožef Stefan Institute and the Faculty of Information and Communication Sciences, University of Ljubljana. His main research interests are hate and offensive speech detection based on social media data, user profiling, computational social science and social media data, natural language processing of non-canonical language and South Slavic languages, and representation learning for lexical semantics. He is coordinating the CLARIN Knowledge Centre for Language Resources and Tools for South Slavic Languages, and is a member of the Association for Computational Linguistics and the Slovene Society for Language Technologies.

Kristina Pahor de Maiti is a Research Assistant at the Department of Translation Studies of the Faculty of Arts, University of Ljubljana, and a PhD student at the Department of Comparative and General Linguistics at the same faculty. She earned her MA in interpretation studies at the University of Ljubljana and later on worked as a translator and interpreter in the private sector. Her research interests include spoken language and computer-mediated communication. She is currently focusing on the corpus-based analysis of socially unacceptable discourse online.

Philippa Smith is a senior lecturer in English and New Media in the School of Language and Culture at Auckland University of Technology in New Zealand. She teaches on the Bachelor's and Master's degrees in English and New Media, and supervises students on a range of topics from online friendships and satirical websites through to digital nomads. Her research interests involve critical discourse studies, language and communication, (new) media discourse, digital journalism and identity. In 2018 Philippa was a Visiting Fellow at the Oxford Internet Institute in the UK, where she researched negative online behaviours such as hate speech, trolling and online incivility.

Goran Vaage is an Associate Professor of Sociolinguistics at Kobe College, School of Letters. He earned his Master's degree from Osaka University of Foreign Studies in 2007, and his Doctoral degree from Osaka University, Graduate School of Language and Culture in 2010. His thesis topic was "Sociolinguistic Perspectives on Words for First Person in Japanese". From November 2010 to September 2012 he served as a JSPS postdoctoral research fellow at Osaka University, Center for Japanese Language and Culture. His research interests include cultural differences in language play and humour, Osaka studies, and stigma theory and attitudes towards language.

Name index