

# Exposure to Hate Speech in the Digital Age. Effects on Stereotypes About Roma People\*

Lecturer Dana Raluca BUTUROIU, PhD

Department of Communication  
College of Communication and Public Relations  
National University of Political Studies and Public Administration, Bucharest, Romania  
E-mail: raluca.buturoiu@comunicare.ro

Prof. Nicoleta CORBU, PhD

Department of Communication  
College of Communication and Public Relations  
National University of Political Studies and Public Administration, Bucharest, Romania  
E-mail: nicoleta.corbu@comunicare.ro

**Abstract.** *Hateful content directed against ethnic minority groups has gained momentum due to the open nature of social media platforms and the legal protection towards freedom of speech. Being the largest in the European Union, the Roma ethnic minority group faces generalized discrimination with regards to many vital areas such as access to employment, healthcare services, housing (including public housing), and education. In Romania, which is home for large numbers of Roma people (varying from 1,5 to 3*

- 
- \* This paper was financially supported by the **Human Capital Operational Program 2014-2020, co-financed by the European Social Fund, under the project POCU/380/6/13/124708 no. 37141/23.05.2019, with the title “Researcher-Entrepreneur on Labour Market in the Fields of Intelligent Specialization (CERT-ANTREP)”**, coordinated by the National University of Political Studies and Public Administration.

*million people), such discrimination is widely spread in common online settings. Based on a 4x2 between-subjects experiment (N=351), this study sheds light on the effects of exposure to hate speech directed against the Roma ethnic minority in the digital age, with particular reference to some dominant stereotypes about Roma people. Main results show that exposure to hateful content (i.e., neither mild/ derogatory, not extreme/ full hate) directed against Roma people might function in the sense of diminishing negative stereotypes rather than enhancing them. Such findings could be useful in future studies investigating hate speech effects and in the development of strategies and policies addressing discrimination against Roma ethnic minority.*

**Keywords:** Online hate speech; Roma minority; Stereotypes about Roma; Discrimination.

## Introduction

Recent years have witnessed important developments related to the digital domain. From a media and communication perspective, the unrestricted access to huge amounts of content and the proliferation and high use of social networking sites (SNS) are among the most important. Today's digital ecosystem allows people to freely find information and communicate at nearly marginal costs (Mondal *et al.*, 2017, 85). Within this virtual space, people are not alone, but grouped in more or less small networks, which favor the exchange of news, ideas, and opinions. Sometimes, the open nature of SNS and the legal protection of free speech (which is often mistaken for hate speech) posit a series of challenges. Particularly, SNS provide fertile ground for various discussions, usually including the "us" versus "them" polarization. This ideological divide fuels the emergence of hateful content (including more or less significant degrees of hate) directed against certain individuals and/ or groups of people, resulting in severe real-life consequences (Djuric *et al.*, 2015, 29), mainly due to the fact that it is disseminated in usual online settings (Keipi *et al.*, 2017).

Hateful content directed against ethnic minority groups is amongst the most common. With specific reference to Romania, ethnic groups such as the Roma are frequently targeted by both online and offline hate speech (Meza, 2016), taking various forms, ranging from offensive remarks to explicit incitement to violence (Scicluna, 2007). The Roma minority group is the biggest ethnic minority in Europe (Mills, 2019). Romania has the largest Roma population in absolute numbers (Dunlop, 2018), even though countries such as Bulgaria, Slovakia, Hungary, Serbia, Montenegro and other countries from Central and Eastern Europe (CEE) are home

for this minority group. According to the last Census in 2011<sup>1</sup>, the Roma population represents 3,3% (around 621,000 persons) of the total population in Romania, being uniformly distributed across the country. However, the total number of Roma people living in Romania varies between 1,5 million to 3 million people, being significantly higher than the official statistical data suggest, as many Roma do not self-declare as Roma (but as Romanians or Hungarians) in the Census<sup>2</sup>.

The first years of democracy in Romania were filled with increased ethnic conflicts (Rughiniş, 2010). From then on, the Roma ethnic minority community remained on a peripheral position in Romania as well as in the other European countries. According to some voices, in contemporary Romania, interethnic relations and policies are “sharply competitive and conflictual” when compared with those from other CEE countries (Magyari *et al.*, 2008, 6). This marginalized position was the engine for developing and perpetuating offensive language, aggressive attitudes, and ignorance behaviors regarding Roma people. The Romanian Roma minority faces discrimination with regards to access to employment, healthcare services, housing (including public housing), and education (Iordache, 2019). Furthermore, both the media and some public actors offer discriminatory public statements directed against the Roma minority group in Romania (e.g., the depiction of Roma as “crows” by Vladimir Tismăneanu<sup>3</sup>; the positions adopted by some representatives of the Conservative Party in the UK after the moment when the European Union eased restrictions in Romania and Bulgaria to access the UK labor market; the coverage of the Roma group in the Daily Mail, the most important media outlet that supports of the Conservative Party in the UK<sup>4</sup>; the discourse of some representatives of the extreme right movement like Nigel Farage<sup>5</sup>). Many of them are in connection with stereotypical representations portraying this specific group. These labelling shortcuts depicting main physical and personality traits of the group are deeply-rooted in the public imaginary and, thus, very hard to change. Main stereotypical representations of Roma people are linked to criminality, illiteracy, immorality, promiscuity, laziness, and resistance to integration into mainstream society (Breazu & Machin, 2019; Catalano, 2018; Creţu, 2014; Bernáth & Messing, 2013; Veselková, 2016).

---

1 [http://www.recensamantromania.ro/wp-content/uploads/2013/07/REZULTATE-DEFINITIVE-RPL\\_2011.pdf](http://www.recensamantromania.ro/wp-content/uploads/2013/07/REZULTATE-DEFINITIVE-RPL_2011.pdf).

2 <https://www.gandul.ro/stiri/recensamant-2011-doua-treimi-dintre-romi-se-declara-romani-700000-2000000-3000000-cati-romi-traiesc-in-romania-8883047>.

3 <https://www.rferl.org/a/depiction-of-roma-as-crows-exposes-deeper-racism-within-romania/30558933.html>.

4 <https://foreignpolicy.com/2019/12/12/gypsies-travellers-boris-roma-crackdown-conservative-tory/>.

5 <https://www.politics.co.uk/news/2013/09/06/nigel-farage-romanians-want-to-move-to-a-civilised-country>.

Against this background, this paper aims to analyze if and how the exposure to different degrees of online hateful content directed against the Roma ethnic minority group influences the main stereotypes about them. The originality of this approach resides in the fact that it sheds light on two interconnected elements referring to the Roma ethnic minority group in Romania (i.e., hate speech against this minority and stereotypes about Roma people) which could offer possible explanations of the way in which online hateful content transfers to offline settings (i.e., if it influences stereotypical perceptions of Roma people). Furthermore, another point of originality of the present research paper resides in the fact that it combines the two above-mentioned elements. To the best of our knowledge, there is no other similar research study performed in Romania so far. Therefore, it might be also insightful regarding the particularities of the Romanian context (i.e., as the European country being home for the largest number of Roma people).

### Exposure to online hate speech. Effects and implications

Researchers tend to agree that there is no universal definition of hate speech and that the existing definitions have been largely contested over time (e.g., Kimotho & Nyaga, 2016; Jubany & Roiha, 2016). In online settings, the situation is even more complex. On the one hand, the online medium is less regulated and generally perceived as a space where the freedom of speech should be above any restriction or censorship (Meza, 2016; West, 2012). On the other hand, there are some variables which might influence the exact illustration of what hate speech is or is not. Anonymity (i.e., the use of false identities and pseudonyms facilitates the free expression of hateful content), permanence (i.e., the hateful content is available for longer periods of time), itinerancy (i.e., even if the content is removed from a specific platform, it might be posted somewhere else on the same platform or on another one), and transnationality of the Internet (i.e., the hateful content can be accessed from different parts of the world simultaneously) are among the variables characterizing hate speech in online settings. The openness of social media adds more complexity to the whole phenomenon of hate speech since social media platforms allow and even encourage people to make their voice known by offering them the necessary ingredients for greater participation in the public sphere (Uzochukwu & Okafor, 2019, 146; Banks, 2010). The issue of social media openness is linked to what researchers depict as an “online disinhibition effect” (Enarsson & Lindgren, 2019, 4) which refers to the fact that people tend to be less restrained to openly express their views and more eager to self-disclosure in online settings than in offline ones (Rorke, 2015).

Furthermore, social media platforms also contribute to the expansion of personal communication networks, while exposing users to messages that are

not always ideologically congruent with their previous views (Barnidge *et al.*, 2019); thus contributing to creating a space with multiple ideologically-driven views which can further fuel the development of counter-ideologies expressed in more or less harmful ways. Therefore, when discussing any issue regarding hate speech and its effects or implications, one must first define the term (Meza *et al.*, 2018). For the purpose of our research, we will use the definition of hate speech as representing “all forms of expression which spread, incite, promote or justify racial hatred, xenophobia, anti-Semitism or other forms of hatred based on intolerance, discrimination, and hostility against minorities, migrants and people of immigrant origin” (Council of Europe, 2016, 76; Kimotho & Nyaga, 2016; Gitari *et al.*, 2015; Pálmadóttir & Kalenikova, 2014; Kaakinen *et al.*, 2018).

Accordingly, hate speech can take various forms of expression (i.e., not only verbal, but also non-verbal or symbolic) directed against an individual or a group of people considered to be inferior due to some constitutive and often innate characteristics (e.g., race, ethnicity, gender, sexual orientation, etc.). These forms of harmful positioning towards other individuals or groups are often intended to incite violence, to develop discriminatory treatment or to offend the human dignity of those individuals or groups (Kimotho & Nyaga, 2016; Stakić, 2011). Nevertheless, despite the fact that online hate speech can be encountered in different forms, our focus is on verbal forms of expression, specifically on language, as a means of “organizing people and directing their behavior” (Kimotho & Nyaga, 2016, 192; Ayeomoni & Akinkuolere, 2012). People use language for various reasons, from persuasion purposes to critically assessment of other individuals or groups; “words are powerful weapons, which used within a biased framework can further marginalize and demean groups which have already been rendered vulnerable” (Jubany & Roiha, 43). Against this backdrop, online hate speech represents a specific form of language developed and used in order to reach a specific purpose (Kimotho & Nyaga, 2016) or to induce certain effects.

In terms of effects of exposure to hateful content, studies document that online hate leads to both indirect and direct harm (Keipi *et al.*, 2017; Lee & Leets, 2002). Indirect social harm raises both ethical and legal questions referring to whether hateful content is natural and should be allowed or not. On the other hand, content that is built on direct harm can have both social and psychological negative effects at both individual and group levels. For example, exposure to hateful online content might result in reinforcing discriminatory behaviors against already vulnerable groups. In turn, targeted individuals which are the victims of such content might develop long lasting defensive and extremely vigilant behaviors. In other terms, while there are some people for whom the exposure to online hateful content remain unaffected and most hateful content is not directly linked to violent, discriminatory or harmful attitudes and behaviors, exposure to such content can cause short- and long-term effects for other people (Costello *et al.*, 2017; Leets, 2001). Immediate

effects include mood swings, anger, loneliness, and fear (Costello *et al.*, 2017, 589), while long-term effects include erosion of social trust, lead to the intergenerational transmission of extremist ideologies, and be an important source of radicalization.

On one hand, frequent exposure to hate speech results in decreasing affective responses when people are further facing hostile messages (i.e., through a desensitization process, those people who were previously exposed to hate speech content will exhibit decreased sensitivity to other hateful messages) (Soral *et al.*, 2017, 2). On the other hand, there is ground to believe that positive outgroup attitudes might function in the opposite direction, namely in the sense of sensitizing people to hate speech (i.e., making them more aware of the possibility of encountering hate speech and sometimes more compassionate for the victims of this type of speech) (Soral *et al.*, 2017, 4).

When these facets of hate speech are not taken into consideration in due time, they might lead to the escalation of the online abuse, and, in turn to the normalization of such practices in offline settings (Awan, 2014, 134; Del Vinga *et al.*, 2017). Notably, the effects of exposure to hateful content should be interpreted and analyzed having in mind the idea that hateful messages, irrespective of their focus, are disseminated within the normal, routine-based online communication, where users might not be expecting violent, discriminatory, or aggressive intentions. Therefore, while being less aware of encountering hateful content, people might be more vulnerable (Keipi *et al.*, 2017).

Within this context, one should acknowledge that there are different degrees of hateful content. Therefore, their effects might be different, in the sense that the more hateful the content, the more visible the negative effects of that message. The literature to date is rather scarce in revealing and explicitly characterizing different degrees of hateful verbal content. Nevertheless, there are studies documenting a spectrum of classes of hate speech (e.g., Sharma *et al.*, 2018). For example, taking into account the hateful intent of the harmful speech, the three possible classes of hate speech are as follows: class I – violence, extremism (racist, sexist, and other groups), propaganda; class II – intimidation, trespassing, accuse, threat/ fear, enmity, and class III – sarcasm, irony, trolling, bullying (Sharma *et al.*, 2018). The three classes of possible online hate speech describe decreasing degrees of hateful intent embedded in online harmful content.

A closer look into these three different classes of hateful content reveals the fact that class I hate speech incites violent actions mainly with the intention of hurting “sentiments of certain isms (extremism) for a violent response to be possible in return” (Sharma *et al.*, 2018, 109). Class II hate speech is also related to hurting sentiments, but not in such a provocative way to generate violent response in return; however, it can also be very provocative as it is based on accusing, threatening and aggressive language used in order to disagree or provoke a verbal duel against an individual or a group. Class III hate speech is mildly provocative as the content is

not necessarily targeting a specific group, while revolving around an ironic and/or sarcastic tone of voice. Class III hate speech can be used as partially synonym to derogatory language, which is, for some authors the first and least harmful type of “toxicity” in discourse (Tirrell, 2017). With specific reference to hate speech directed against ethnic groups, derogatory language includes ethnophaulisms and other linguistic factors (like derogatory outgroup jokes) intended to prejudice other individuals or groups (Mullen & Leader, 2008, 192). In the ethnic sphere, plain labels depicting certain ethnic minorities (such as Negro, Italian, Jew, Catholic, etc.) might have an emotional tone. Their higher key equivalents (such as nigger, wop, kike, papist, harp, etc.) are even more provocative, in the sense that, whenever used, the speaker’s intention is not only to characterize an individual, but also to denigrate and dismiss (Mullen & Leader, 2008, 193; Hornsby, 2001).

Based on the relatively scarce literature showing different degrees of hateful content and their different potential effects, we have developed our own categorization regarding the degrees of online hate speech targeting ethnic minorities, with specific focus on the Roma ethnic minority group. We believe that, in online settings, people are prone to encounter the following types of content referring to Roma people: neutral content (i.e., content referring to Roma people as to any other ethnic minority group, in a neutral tone of voice), derogatory content (i.e., content causing relative mild offence towards the Roma ethnic minority group, by using specific labels or specific words related to mild intolerance against the group), hateful and extremely hateful content (i.e., content which is explicitly provocative based on accusing, threatening and aggressive language). We assume that these delineations should not be seen as separate, but rather as different points on a continuum, varying from no hateful content to extremely hateful one. We expect that the exposure to different degrees of online hate speech to elicit different effects, in the sense that either the more prominent the hateful intent of the message, the more negative people’s attitudes towards targeted individuals or groups, either we might find a boomerang effect.

### **Stereotypical representations of the Roma ethnic minority group in the context of hate speech**

The Roma ethnic minority group, known in Romania by the ethnonyms “Romi” or “Țigani” is one of the largest ethnic minorities in Europe. Studies document that Romania has probably the biggest Roma population in absolute numbers, while large numbers of Roma people are also living in Bulgaria, Hungary, Serbia and Montenegro, Slovakia, Turkey, Macedonia, Spain, France, Italy, and Germany (Rughiniş, 2010; Cernat, 2011; Mădroane, 2012). The documented history of the Roma ethnic minority group goes back to the 14th century, when the Roma were recorded as slaves. Even though the slavery was abolished in 1856, the Roma population did not own their

lands. As a result, from then onwards, they have been occupying a marginal social position both in Romania and in Europe, being referred to as “others” most of the time (Kligman, 2001; Petranova, 2018), being treated as “objects, rarely subjects” or being depicted as a “problem” (Sigona, 2005; Sonneman, 1999; Okely, 2014; Gross, 2006; Rughiniș, 2010; Breazu & Machin, 2019; Vrăbiescu & Kalir, 2018; Fekete, 2014; McGarry, 2011; Langaas, 2017; Belavusau, 2013).

This peripheral position of the Roma minority group has permanently fed the development and wide use of stereotypical representations of the group. Studies show that in all the countries from CEE, the Roma are the most rejected ethnic minority group and that this rejection is mainly based on views such as: “The Roma are genetically inclined to commit crimes”, “You can’t trust or rely on Gypsies”, and “The Gypsies are lazy and irresponsible” (i.e., discrimination and stigmatization based on stereotypical representations) (Csepeli & Simon, 2004, 133).

Stereotypical representations refer to standardized and simplified depictions about certain individuals or groups. In the case of intergroup stereotypes, they are generally based on personality characteristics (either positive or negative) and their main effects are linked to the formation of the general relation and positioning towards the targeted group (Jozef & Juraj, 2018). There is consensus among researchers interested in the study of stereotypical perceptions of Roma people that this particular ethnic group is mainly perceived and referred to in negative rather than positive terms mainly because of the emotions associated with those specific representations (Klíma, 2018). Another factor contributing to the mainly negative perceptions about Roma people is linked to the belief that negative portrayals of a specific individual or group function as elements which authorize “the exclusion of members of that group from the scope of justice” (Hadarics & Kende, 2019, 258). In other terms, one main reason why the positive stereotypes seem to be missing when portraying the Roma ethnic minority group refers to the scarcity in everyday images circulating in the public sphere. This, in turn, determines people to avoid or even cannot think about the Roma people in “ordinary terms”, as if they are “normal people” (Tremlett, 2017, 736).

Main media and public perceptions suggest that Roma ethnic group is linked to criminality, illiteracy, immorality, promiscuity, laziness and resistance to integration into mainstream society (Breazu & Machin, 2019; Catalano, 2018; Crețu, 2014; Bernáth & Messing, 2013; Veselková, 2016). Roma are depicted as people with constrained access in vital areas such as education (Roma pupils are discriminated by both colleagues and teachers, they are known for low educational results and early school leaving), labor market (low employment rate), and health (they live in poor conditions and, thus, they cannot see the doctor regularly, their health being seriously affected) (Lášticová & Findor, 2016; Mauček, 2013). Specifically, they are depicted as asocial individuals; as criminals; as parasites; unwilling to work and to study; equated to animals (rats and pigs), insects (cockroaches) or things; as too



fertile and having too many children, and as destroyers of their dwellings (Miškolci *et al.*, 2020, 135; End, 2017).

Therefore, main depictions of Roma revolve around the stereotypes of Roma as deviants and criminals, people who are living in dirty and poor conditions, and who voluntarily choose to be dependent on social benefits (Miškolci *et al.*, 2020; Castañeda, 2014; Stăiculescu & Gherasim, 2013). Despite their deeply historical rooted nature, there are voices that suggest that all these stereotypes (both negative and positive ones) are inaccurate since they imply “a homogeneity, a coherence we would not expect from any group” (McGarry, 2014, 761; Maučec, 2013). Even though main stereotypes regarding Roma people are negative, some traditional, “romanticized” stereotypes do exist and are widely shared. For example, the romantic depiction of Roma people as having musical and dancing talent, as being capable of passionate love and strong emotions, as being spontaneous, free and spiritual, and as having an innate relatedness to nature as well as a tenacious ability to enjoy themselves (Villano *et al.*, 2017, 31).

All the above-mentioned negative depictions of Roma ethnic minority group are even more visible in online hateful contexts. Hate speech is one important source of attitudes such as evictions, school segregation, and physical aggression. Hatred and exclusion of the Roma minority group are often based on the depictions and representations provided by both the media and the officials (ministers, politicians, public authorities). Their interventions are seen as expressions of dominant social norms in CEE countries (Kende *et al.*, 2017), and often fuel people’s negative stereotypes about Roma people, thus indirectly contributing to the discrimination and stigmatization of the Roma minority group and, even more dramatically, to the proliferation of a “dangerous tide of hatred” which is not limited to online settings, but rather transfers to offline ones (Scicluna, 2007, 47).

In terms of counter hate speech interventions in online settings and their influence on either negative or positive stereotypes about Roma, studies are scarce. Specifically, there are contradictory findings about the real benefits of positive stereotypes. Even though positive stereotypes have proved efficient in minimizing the negative effects of being negatively-stereotyped and consequently in feeding prosocial and proactive behaviors (Lášticová & Findor, 2016), positive stereotypes can also be perceived as negative by their “recipients” (Dobai, 2018). Counter hate speech measures directed in the sense of decreasing the negative impact of hate speech on Facebook did not prove efficient; most often people maintained mainly hateful attitudes. However, there is evidence to believe that counter hate speech measures function in the sense of activating users to get engaged into the conversation in a greater capacity than they would do before. At the same time, positive comments about the Roma ethnic minority group seem to function as a barrier against further hate; when exposed to pro-Roma comments users stopped presenting any more hate (Miškolci *et al.*, 2020).

Against this background, this study aims at analyzing the main positive and negative stereotypes about Roma ethnic minority group in Romania in the context of online hate speech. As there is no clear delineation in the literature about various effects related to both degrees of hate speech and reactions to it, we guide this exploratory study by research questions (as opposed to hypotheses). Specifically, we formulate three research questions, as follows:

*RQ1. Are people’s stereotypes about the Roma people influenced in either a positive or a negative way, as a result of exposure to hate speech? (i.e., if there is any type of effects of exposure to hateful content)*

*RQ2. Under what conditions people’s stereotypes about Roma people could be influenced in either a positive or a negative way, as a result of exposure to hate speech? (i.e., if the effects are either positive or negative in a hateful context)*

*RQ3. What kind of stereotypes are more prone to be influenced in either a positive or a negative way, as a result of exposure to hate speech? (i.e., if the positive or negative effects are related to the content/ degree of hateful speech).*

**Method**

In this study we constructed a 4x2 between-subjects experiment (see Table 1), plus a control condition. We made up a story based on verifiable information about statistics of the European Union related to the penetration rate of illiteracy among Roma people (9% of Roma people are enrolled in secondary education, while only 2% of them go to University) and state aids (almost 40% of the Roma population rely on this type of financial aid as the sole source of revenue for their families). We manipulated the story, keeping the information correct, but framing it using different degrees of hate content (neutral, derogatory, hate and full hate) and using either positive or negative comments to each type of framing. The positive ones were meant to counter the hateful content, whereas the negative ones added more hate speech as a form of approval and support (see Table 1 for N for each condition). All stimuli were constructed as Facebook post (suggesting high engagement). The control condition was framed neutrally, with no visible comments.

	<b>Neutral</b>	<b>Derogatory</b>	<b>Hate</b>	<b>Full hate</b>
Positive comments	Neutral with positive comments (N=40)	Derogatory speech with positive comments (N=38)	Hate speech with positive comments (N=42)	Full hate speech with positive comments (N=40)
Negative comments	Neutral with negative comments (N=39)	Derogatory speech with negative comments (N=36)	Hate speech with negative comments (N=39)	Full hate speech with negative comments (N=38)

+ control (N=40)

**Table 1.** Overview of the experimental conditions + control (neutral, no comments)

## Sample and procedure

We used students enrolled in social sciences Higher Education studies in Romania (N=351) as participants in the experiment, which was conducted on February 17-28, 2020, based on pen and paper self-administered questionnaires. The questionnaires were randomized, randomization being successful with regards to gender ( $F_{8,333}=1.89$ ,  $p>.05$ ) and opinion about Roma people ( $F_{8,343}=1.88$ ,  $p>.05$ ). At the end of the survey, participants were debriefed about the fictional stimulus and thanked.

## Stimuli

The layout of the stimuli mimicked a Facebook post, suggesting high level of engagement (309 comments, 21 shares, 467 reactions), which was kept constant in all conditions, including control. Of the 309 comments, three were visible in the 8 non-neutral conditions, being either positive, counter-arguing the initial post, or negative, in line with the initial post.

## Measures

*Stereotypes about Roma people* were measured using 7-point semantic differentiator scales, measuring people's estimations that the adjectives generally apply to Roma people: trustworthy – untrustworthy; hardworking – lazy; honest – dishonest; sympathetic – unsympathetic; intelligent – limited; educated – uneducated. The scale grouped in one factor, with loadings ranging from .696 to .832 ( $\alpha=.846$ ,  $M=3.20$ ,  $SD=1.05$ ).

## Manipulation checks

To make sure people perceived the stimuli as intended, we used three manipulation checks, proven successful. People in the control, neutral, and derogatory with positive comments conditions perceived the Roma people as being portrayed in a negative way to a lesser extent than people in the other conditions ( $F_{8,341}=16.74$ ,  $p<.01$ ). Positive comments were perceived as positive only by the people in the conditions that used positive comments ( $F_{7,302}=16.98$ ,  $p<.01$ ), while people in the conditions containing negative comments perceived the comments to be derogatory toward the Roma people ( $F_{7,300}=13.06$ ,  $p<.01$ ).

## Findings

Results show that people in both negative ( $\beta=.181$ ,  $p=.011$ ) and positive ( $\beta=.175$ ,  $p=.015$ ) hate speech conditions hold significantly more positive stereotypes than people in the control group, while people in all other conditions (neutral, derogatory or full hate speech) were not influenced in any way. Controlling for own opinion

about Roma people in general does not influence results. This shows that mild hate speech content, as well as very hateful and prejudiced content do not affect people’s representation of the Roma. However, the right “dosage” of hate speech lead to more positive stereotypes, acting as a boomerang. None of the negative content lead to more negative stereotypes (see Table 2 for Descriptives by condition). Moreover, given that stereotypes are relatively stable representations, difficult to change in a one-shot exposure to hateful content, it might be very useful to look at the trend of the results. Repeated exposures, in time, might reinforce a type of effects that now only shows as potentiality. The unexpected result is that none of the combination of content and comments lead to more negative portrayals than the control, which suggest that hate speech toward Roma people might in fact have a general boomerang effect.

Condition	N	Mean	SD
hate speech + negative comments	39	<b>4.24</b>	0.80
hate speech + positive comments	41	<b>4.21</b>	1.03
neutral + negative comments	39	4.00	1.13
derogatory + negative comments	36	3.94	1.12
full hate speech + positive comments	40	3.87	1.12
neutral + positive comments	39	3.80	1.02
derogatory + positive comments	38	3.73	1.13
full hate speech + negative comments	36	3.70	1.18
control	40	3.64	0.82

Table 2. Descriptives of stereotypes of Roma people, by condition

Looking at the type of stereotypes we used to measure people’s portrayals of Roma, there are not many differences from the general results (see Table 3 and Table 4), but exceptions are interesting to look at: other than the level of sympathy and appreciation about education who yielded no significant results (however, descriptives show similar trends with the aggregate level of stereotyping), all other stereotypes were significantly more positive than those hold by people in the control group only in the “hate” condition (that is harsh language, not mild, not really extreme). Comments seem to matter to some extent only for trustworthiness (positive comments reduce consistently the effect) and even less for portrayals of Roma people as either hardworking or lazy (negative comments somewhat reduce the effect). Honesty is affected a little differently, in the sense that the significantly influential conditions seem to be either hate speech with positive comments or neutral with negative comments.

Condition	stereotypes overall		trustworthy	hardworking	honest	sympathetic	intelligent	educated
	Mean	SD	untrustworthy	lazy	dishonest	unsympathetic	limited	uneducated
hate speech + negative comments	4.24	4.26	4.21	4.05	4.21	4.47	4.26	
	39	39	39	39	39	38	38	
	0.80	1.16	0.98	1.10	1.30	0.98	1.20	
hate speech + positive comments	4.21	3.78	4.39	4.22	4.37	4.46	4.02	
	41	41	41	41	41	41	41	
	1.03	1.46	1.28	1.41	1.51	1.27	1.65	
neutral + negative comments	4.00	3.97	3.87	4.18	4.13	4.00	3.82	
	39	39	39	39	39	39	39	
	1.13	1.42	1.47	1.37	1.36	1.40	1.65	
derogatory + negative comments	3.94	3.56	3.97	3.78	4.39	3.97	3.92	
	36	36	35	36	36	36	36	
	1.12	1.71	1.36	1.33	1.50	1.52	1.61	
full hate speech + positive comments	3.87	3.50	3.85	3.75	4.20	4.10	3.83	
	40	40	40	40	40	40	40	
	1.12	1.66	1.35	1.43	1.45	1.15	1.69	
neutral + positive comments	3.80	3.54	3.82	3.74	4.00	4.05	3.64	
	39	39	39	39	39	39	39	
	1.02	1.48	1.52	1.46	1.30	1.23	1.61	
derogatory + positive comments	3.73	3.37	4.00	3.76	3.97	3.79	3.51	
	38	38	37	38	38	38	37	
	1.13	1.44	1.47	1.46	1.24	1.47	1.57	
full hate speech + negative comments	3.70	3.36	3.81	3.56	3.83	3.92	3.72	
	36	36	36	36	36	36	36	
	1.18	1.42	1.55	1.48	1.58	1.25	1.61	
control	3.64	3.53	3.50	3.50	3.85	3.65	3.80	
	40	40	40	40	40	40	40	
	0.82	1.18	1.28	1.18	1.05	1.17	1.20	

Table 3. Descriptives of individual stereotypes, by condition (results highlighted significantly higher, if compared with the control group)

	stereotypes overall (β)	trustworthy – untrustworthy (β)	hardworking – lazy (β)	honest – dishonest (β)	sympathetic – unsympathetic (β)	intelligent – limited (β)	educated – uneducated (β)
hate speech + negative comments	.181*	.159*	.162*	.127	.082	.200**	.094
hate speech + positive comments	.175*	.057	.210**	.170*	.122	.204**	.047
neutral + negative comments	.108	.097	.086	.157*	.064	.086	.004
derogatory + negative comments	.086	.006	.104	.062	.120	.076	.023
full hate speech + positive comments	.071	-.005	.081	.058	.082	.112	.005
neutral + positive comments	.049	.003	.074	.056	.035	.098	-.033
derogatory + positive comments	.029	-.034	.113	.060	.028	.034	-.058
full hate speech + negative comments	.018	-.034	.068	.012	-.004	.063	-.015

Table 4. OLS Regression results for main effects, by type of stereotype

It should be noted that not only stereotypes that were the object of the stimuli were influenced (hardworking – lazy; trustworthy – untrustworthy; educated – uneducated), but all others equally or even more. This shows that, regardless of the topic circulated on social media in a hateful way, many stereotypes about Roma might get positively influenced. If we look at general trends, there are two interesting results: in the overwhelming majority of cases stereotypes are only made more positive, by any of the combinations we tested in this study. The exceptions, however, are interesting too, as they concerned two of the stereotypes referred to in various ways in both the content and the comments. The stereotype regarding trustworthiness seems to have the potential to be affected in a negative way by the full hate speech content (regardless of the valence of the comments), as well as by mild derogatory portrayals with positive comments. Stereotypes about Roma being educated/ uneducated, which was the main topic of the stimuli, seem to be potentially affected in a negative way by neutral or derogatory language, with positive comments or by full hate speech with negative ones. Overall, it seems that stereotypes directly referred to in the stimuli (if more powerful or maybe repeated) could have negative effects in two type of settings: neutral (or almost language) with positive comments, which might seem too positive to the Roma people and could lead to the opposite effect, or by the most extreme combination: full hate speech with negative comments. The differences are, however, very small.

## Discussion

Overall, this study suggests that effects of hate speech in social media might not have so powerful negative effects on stereotypes as previously thought, at least not among young educated people; if anything, it backfires. We found significant results of both positively and negatively commented posts constructed using what we called “hate speech” (which is neither mild, nor extreme hateful language against Roma people), in the sense of attenuating the negative stereotypes people hold of this population rather than enhancing them. All other combinations were not significant. However, provided that stereotypes are very stable structures (Esses & Zanna, 1995, 1052; Poppe, 2001), looking at trends of non-significant results is worth reporting, as one-time exposure could give researchers hints into the direction and size of repeated exposure effects. The trends are consistent across discrete stereotypes and overall effects: all types of combinations lead to less negative portrayals of the Roma, with very few exceptions. The exceptions are either full hate speech regardless of the valence of the comments or derogatory with positive comments, which are some extreme of the spectrum: portraying the Roma either too positively (given the context, derogatory content with positive comments could be considered as positive) or too negatively might lead to negative effects or, at least, to no effect at all, but this is only true for the already very

deeply rooted and rather general stereotypes, those related to trustworthiness and education (with education being also the main topic used in the stimuli).

When looking at the way other people's reaction might add to or diminish the effect, we see that they seem to count less than the original post; it could be the case that people are less inclined to judge reactions to posts on SNS and more prone to react to the posts themselves. Whether or not such comments might influence people's willingness to speak on behalf of the Roma people is a different story, and not the focus of this study. However, as far as stereotypes are concerned, it seems that they do not matter much.

As any study, this too comes with limitations. The most important one is related to the sample, and results should be interpreted in the light of the sample used: the results are applicable to young educated people from urban areas, rather than to a general population. The size or even direction of the results might be different when re-tested for the general population. However, it should be noted that this study could also offer a key to act against hate speech, in the sense of trying to limit exposure to hate speech to more vulnerable populations, while feeling safer when discussing impact on educated people. At the same time, this study is conducted in a country with a consistent Roma minority, which has been repeatedly subject to negative portrayals (Meza, 2016) and discrimination (Iordache, 2019); therefore, results should be cautiously generalized to different contexts.

## Conclusions

Negative stereotypes against Roma people have raised concerns about perpetuating integration problems in the last decades (Andor, 2018; Filipescu, 2009). The SNS have brought about new concerns about the proliferation of hate speech against various minorities, Roma included, as they offer both a spotlight or a more participatory space (Uzochukwu & Okafor, 2019, 146; Banks, 2010) and a safer environment for people to express themselves (Enarsson & Lindgren, 2019, 4; Rorke, 2015). In this context, effects of hate speech directed against Roma people becomes a point of inquiry worth investigating. In this study, we look at effects of various degrees of hate speech negatively portraying Roma people among educated people, testing at the same time whether or not approval or disapproval of these types of speech could influence results.

Generally speaking, we found that negative portrayals of Roma people on social media rather helps diminish negative stereotypes than enhancing them, and that the effects are powerful especially for the "right amount" of hate speech, not mild, nor extreme. Comments are less important than content itself, at least at a first glance; it is difficult to assess what might be the effects, should people in fact take the time to read extensively other people's opinions about the matter. However, the social media platforms have transformed people, arguably, in even more "cognitive



misers” (Fiske & Taylor, 1984) than before. Therefore, this study suggests that information snacking (rather than consuming) in the social media era favor full post content rather than reactions to it in this specific context.

Findings could have implications into policies regarding Roma people integration and discrimination. Rather than generally fighting hate speech on every platform and all media, it could be more effective to care about non-educated communities, as it seems that in educated ones hate speech manifestation backfires. As a future line of investigation, we urge scholars to investigate effects based on categories of people since there might be significant differences across age, education, residence, etc., which could help in a tailored effort of addressing discrimination or integration.

## References

1. Andor, L. (2018). EU policy and Roma integration (2010–14). *Journal of Poverty and Social Justice*, 26(1), 113-126.
2. Awan, I. (2014). Islamophobia and Twitter: A Typology of Online Hate Against Muslims on Social Media. *Policy and Internet*, 6(2), 133-150.
3. Ayeomoni, O. M., & Akinkuolere, O. S. (2012). A Pragmatic Analysis of Victory and Inaugural Speeches of President Umaru Musa Yar’ Adua. *Theory and Practice in Language Studies*, 2(3), 461-468.
4. Banks, J. (2010). Regulating hate speech online. *International Review of Law, Computers & Technology*, 24(3), 233-239.
5. Barnidge, M., Kim, B., Sherrill, L. A., Luknar, Z., & Zhang, J. (2019). Perceived exposure to and avoidance of hate speech in various communication settings. *Telematics and Informatics*, 44.
6. Belavusau, U. (2013). Anti-Roma Hate Speech in the Czech Republic, Hungary, and Poland. In M. Goodwin & P. De Hert (eds.) *European Roma Integration Efforts – A Snapshot* (pp. 141-181). Brussels University Press.
7. Bernáth, G., & Messing, V. (2013). Pushed to the Edge Research Report on the Representation of Roma Communities in the Hungarian Mainstream Media. CEU. Retrieved on April 16, 2020, from <https://cps.ceu.edu/publications/working-papers/pushed-to-the-edge>.
8. Bienkov, A. (2013). *Nigel Farage: Romanians want to move to a ‘civilised country’*. Retrieved on April 14, 2020, from <https://www.politics.co.uk/news/2013/09/06/nigel-farage-romanians-want-to-move-to-a-civilised-country>.
9. Breazu, P., & Machin, D. (2019). Racism toward the Roma through the affordances of Facebook: bonding, laughter and spite. *Discourse & Society*, 30(4), 376-394.
10. Castañeda, H. (2014). European mobilities or poverty migration? Discourses on Roma in Germany. *International Migration*, 53, 87-99.
11. Catalano, T. (2018). The use of zingari/ nomadi/ rom in Italian crime discourse. *Nomadic Peoples*, 22(1), 123-142.

12. Cernat, V. (2011). Extended Contact Effects: Is Exposure to Positive Outgroup Exemplars Sufficient or Is Interaction With Ingroup Members Necessary?. *The Journal of Social Psychology, 151*(6), 737-753.
13. Crețu, G. (2014). Roma minority in Romania and its media representation. *Sfera Politicii, 22*, 112-120.
14. Csepeli, G., & Simon, D. (2004). Construction of Roma identity in Eastern and Central Europe: perception and self-identification. *Journal of Ethnic and Migration Studies, 30*(1), 129-150.
15. Costello, M., Hawdon, J., & Ratliff, T. (2017). Confronting Online Extremism: The Effect of Self-Help, Collective Efficacy, and Guardianship on Being a Target for Hate Speech. *Social Science Computer Review, 35*(5), 587-605.
16. Council of Europe (2016). Recommendation No. R (97) 20 of the Committee of Ministers to member states on “hate speech”. Recommendations and declarations of the Committee of Ministers of the Council of Europe in the field of media and information society. Retrieved on April 16, 2020, from <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680645b44>.
17. Del Vigna, F., Cimino, A., Dell’Orletta, F., Petrocchi, M., & Tesconi, M. (2017). Hate me, hate me not: Hate speech detection on Facebook. Proceedings of the First Italian Conference on Cybersecurity (ITASEC17), Venice, Italy.
18. Djuric, N., Zhou, J., Morris, R., Grbovic, M., Radosavljevic, V., & Bhamidipati, N. (2015). Hate Speech Detection with Comment Embeddings. Proceedings of the 24th International Conference on World Wide Web, May, 29-30.
19. Dobai, A. (2018). Everyday Discrimination: The Experiences of Hungarian Roma. Doctoral Thesis.
20. Dunlop, T. (2018, June 25). No homeland, no hope – Europe’s Roma are back in the firing line. Retrieved on April 14, 2020, from The Guardian. <https://www.theguardian.com/commentisfree/2018/jun/25/europe-roma-matteo-salvini-ethnic-minority>.
21. Enarsson, T., & Lindgren, S. (2019) Free speech or hate speech? A legal analysis of the discourse about Roma on Twitter. *Information & Communications Technology Law, 28*(1), 1-18.
22. End, M. (2017). Subtle images of antigypsyism: An analysis of the visual perception of “Roma”. *Identities, 24*(6), 668-683.
23. Esses, V. M., & Zanna, M. P. (1995). Mood and the expression of ethnic stereotypes. *Journal of Personality and Social Psychology, 69*(6), 1052-1068.
24. Fekete, L. (2014). Europe against the Roma. *Race & Class, 55*(3), 60-70.
25. Filipescu, C. (2009). Revisiting minority integration in Eastern Europe: Examining the case of Roma integration in Romania. *Debate: Journal of Contemporary Central and Eastern Europe, 17*(3), 297-314.
26. Fiske, S.T., & Taylor, S.E. (1984). *Social cognition*. McGraw-Hill Ryerson.
27. Gitari, N. D., Zuping Z., Damien, H., & Long, J. (2015). A Lexicon-based Approach for Hate Speech Detection. *International Journal of Multimedia and Ubiquitous Engineering, 10*(4), 215-230.

28. Gross, P. (2006). A Prolegomena to the Study of the Romani Media in Eastern Europe. *European Journal of Communication*, 21(4), 477–497.
29. Hadarics, M., & Kende, A. (2019). Negative stereotypes as motivated justifications for moral exclusion. *The Journal of Social Psychology*, 159(3), 257-269.
30. Hornsby, J. (2001). Meaning and Uselessness: How to Think about Derogatory Words. *Midwest Studies in Philosophy*, XXV, 128-141.
31. Institutul Național de Statistică. (2011). *Rezultate definitive ale Recensământului Populației și al Locuințelor – 2011 (caracteristici demografice ale populației)*. Retrieved on April 15, 2020, from [http://www.recensamantromania.ro/wp-content/uploads/2013/07/REZULTATE-DEFINITIVE-RPL\\_2011.pdf](http://www.recensamantromania.ro/wp-content/uploads/2013/07/REZULTATE-DEFINITIVE-RPL_2011.pdf).
32. Iordache, R. (2019). Country report. Non-discrimination. European Commission. Retrieved on April 14, 2020, from <https://op.europa.eu/en/publication-detail/-/publication/1ea6a084-f49d-11e9-8c1f-01aa75ed71a1>.
33. Jozef, V., & Juraj, M. (2018). Hetero- and Auto-Stereotypes of Ethnic Minorities in Slovakia. *Individual & Society*, 21(4), 3-15.
34. Jubany, O., & Roiha, M. (2016). Backgrounds, Experiences and Responses to Online Hate Speech: A Comparative Cross-Country Analysis. Retrieved on April 12, 2020, from [http://www.unicri.it/special\\_topics/hate\\_crimes/Backgrounds\\_Experiences\\_and\\_Responses\\_to\\_Online\\_Hate\\_Speech\\_A\\_Comparative\\_Cross-Country\\_Analysis.pdf](http://www.unicri.it/special_topics/hate_crimes/Backgrounds_Experiences_and_Responses_to_Online_Hate_Speech_A_Comparative_Cross-Country_Analysis.pdf).
35. Kaakinen, M., Oksanen, A. & Räsänen, P. (2018). Did the risk of exposure to online hate increase after the November 2015 Paris attacks? A group relations approach. *Computers in Human Behavior*, 78, 90-97.
36. Keipi, T., Näsi, M., Oksanen, A., & Räsänen, P. (2017). *Online Hate and Harmful Content. Cross-National Perspectives*. Routledge.
37. Kende, A., Hadarics, M., & Lášticová, B. Anti-Roma attitudes as expressions of dominant social norms in Eastern Europe (2017). *International Journal of Intercultural Relations*, 60, 12-27.
38. Kimotho, S. G., & Nyaga, R. N. (2016). Digitized Ethnic Hate Speech: Understanding Effects of Digital Media Hate Speech on Citizen Journalism in Kenya. *Advances in Language and Literary Studies*, 7(3), 189-200.
39. Kligman, G. (2001). On the social construction of “otherness”: identifying “the Roma” in post-socialist communities. *Review of Sociology*, 7, 61-78.
40. Klíma, O. (2018). A Deep Story of the Czech: The Social Meanings of Roma in Czech Society. Bachelor thesis. Retrieved on April 19, 2020, from [https://is.muni.cz/th/v20l7/A\\_Deep\\_Story\\_of\\_the\\_Czech.pdf](https://is.muni.cz/th/v20l7/A_Deep_Story_of_the_Czech.pdf).
41. Langaas, R. F. (2017). New Policies, Old Attitudes? Discrimination against Roma in Norway. Master thesis. Retrieved on April 17, 2020, from <http://bora.uib.no/handle/1956/16014>.
42. Lášticová, B., & Findor, A. (2016). Developing explicit measures of stereotypes and anti-Roma prejudice in Slovakia: Conceptual and methodological challenges. *Human Affairs*, 26(3), 233-252.
43. Lee, E., & Leets, L. (2002). Persuasive storytelling by hate groups online: Examining its effects on adolescents. *American Behavioral Scientist*, 45(6), 927-957.

44. Leets, L. (2001). Response to Internet hate sites: Is speech too free in cyberspace? *Communication and Law Policy*, 6, 287-317.
45. Magyari, N., Mark, L., Harbula, H., & Magyari-Vincze, E. (2008). Country Report on Ethnic Relations: Romania. Central European University. Budapest, Hungary.
46. Maučec, G. (2013). Identifying and Changing Stereotypes Between Roma and Non-Roma: From Theory to Practice. *Innovative Issues and Approaches in Social Sciences*, 6(3), 181-203.
47. Mădroane, I. D. (2012). Roma, Romanian, European: A Media Framed Battle over Identity. *Critical Approaches to Discourse Analysis across Disciplines*, 5(2), 102-119.
48. McGarry, A. (2011). The Roma Voice in the European Union: Between National Belonging and Transnational Identity. *Social Movement Studies: Journal of Social, Cultural and Political Protest*, 10(3), 283-297.
49. Meza, R. (2016). Hate-Speech in the Romanian Online Media. *Journal of Media Research*, 9(3), 55-77.
50. Meza, R., Vincze, H. O., & Mogoş, A. Targets of Online Hate Speech in Context. A Comparative Digital Social Science Analysis of Comments on Public Facebook Pages from Romania and Hungary. *Intersections. EEJSP*, 4(4), 26-50.
51. Mills, Y. (2019, April 8). Roma people: 10 ways Europe's biggest minority faces discrimination. Reuters.com. Retrieved on April 14, 2020, from <https://www.reuters.com/article/us-global-roma-rights/roma-people-10-ways-europes-biggest-minority-faces-discrimination-idUSKCN1RK01Y>.
52. Miškolci, J., Kováčová, L., & Rigová, E. (2020). Countering Hate Speech on Facebook: The Case of the Roma Minority in Slovakia. *Social Science Computer Review*, 38(2), 128-146.
53. Mondal, M., Silva, L. A., & Benevenuto, F. (2017). A Measurement Study of Hate Speech in Social Media. In Proceedings of HT '17, Prague, Czech Republic, July 04-07.
54. Mullen, B., & Leader, T. (2008). Linguistic Factors: Antilocutions, Ethnoyms, Ethnoplaulisms, and Other Varieties of Hate Speech. In J. F. Dovidio, P. Glick, & L. A. Rudman (eds.) *On the Nature of Prejudice: Fifty Years after Allport* (pp. 192-208). John Wiley & Sons.
55. Mutler, A. (2020). *Depiction of Roma As Crows Exposes Deeper Racism Within Romania*. Retrieved on April 13, 2020, from <https://www.rferl.org/a/depiction-of-roma-as-crows-exposes-deeper-racism-within-romania/30558933.html>.
56. Okely, J. (2014). Recycled (mis)representations: Gypsies, Travellers or Roma treated as objects, rarely subjects. *People, Place and Policy*, 8(1), 65-85.
57. Pálmadóttir, J. A., & Kalenikova, I. (2014). Hate speech; an overview and recommendations for combating it. Retrieved on April 12, 2020, from <http://www.humanrights.is/static/files/Skyrslur/Hatursraeda/hatursraeda-utdrattur.pdf>.
58. Petranova, D. (2018). Othering, Stereotyping and Dehumanization – Online Discourse on Minority Groups in Slovakia. Master thesis.
59. Poppe, E. (2001). Effects of changes in GNP and perceived group characteristics on national and ethnic stereotypes in Central and Eastern Europe 1. *Journal of Applied Social Psychology*, 31(8), 1689-1708.

60. Rorke, B. (2015, December 21). FREE TO HATE? ANTI-GYPSYISM IN 21ST CENTURY EUROPE. European Roma Rights Centre. Retrieved on April 18, 2020, from <http://www.errc.org/news/free-to-hate-anti-gypsyism-in-21st-century-europe>.
61. Rughiniș, C. (2010). The forest behind the bar charts: bridging quantitative and qualitative research on Roma/ Țigani in contemporary Romania. *Patterns of Prejudice*, 44(4), 337-367.
62. Scicluna, H. (2007). Anti-Romani Speech in Europe's Public Space – The Mechanism of Hate Speech. European Roman Rights Centre. Retrieved on April 17, 2020, from <http://www.errc.org/roma-rights-journal/anti-romani-speech-in-europes-public-space--the-mechanism-of-hate-speech>.
63. Sharma, S., Agrawal, S., & Shrivastava, M. (2018). Degree based Classification of Harmful Speech using Twitter Data. Proceedings of the First Workshop on Trolling, Aggression and Cyberbullying, Santa Fe, USA, 106-112.
64. Sigona, N. (2005). Locating 'The Gypsy Problem'. The Roma in Italy: Stereotyping, Labelling and 'Nomad Camps', *Journal of Ethnic and Migration Studies*, 31(4), 741-756.
65. Sonneman, T. F. (1999). Dark Mysterious Wanderers: The Migrating Metaphor of the Gypsy. *Journal of Popular Culture*, 119-139.
66. Soral, W., Bilewicz, M., & Winiewski, M. (2017). Exposure to hate speech increases prejudice through desensitization. *Aggressive Behavior*, 1-11.
67. Stakić, I. (2011). Homophobia and hate speech in Serbian public discourse: how nationalist myths and stereotypes influence prejudices against the LGBT minority. *The Equal Rights Review*, 7, 44-65.
68. Stăiculescu, A. R., & Gherasim, M. (2013). Roma representation in the media. *Contemporary Readings in Law and Social Justice*, 5, 947-955.
69. Sultănoiu, M. (2012). RECENSĂMÂNT 2011. Două treimi dintre romi se declară români. 700.000, 2.000.000, 3.000.000 ... câți romi trăiesc în România?. Retrieved on April 13, 2020, from <https://www.gandul.ro/stiri/recensamant-2011-doua-treimi-dintre-romi-se-declara-romani-700000-2000000-3000000-cati-romi-traiesc-in-romania-8883047>.
70. Tirrell, L. (2017). Toxic Speech: Toward an Epidemiology of Discursive Harm. *Philosophy of Language*, 45(2), 139-162.
71. Tremlett, A. (2017). Visualising everyday ethnicity: moving beyond stereotypes of Roma minorities. *Identities*, 24(6), 720-740.
72. Uzochukwu, C. E., & Okafor, E. G. (2019). Social media, hate speech and conflict: Interplay of influences. *International Journal of Social Sciences and Humanities Reviews*, 9(1), 144-158.
73. Veselková, Z. (2016). Where others fear to tread: the role of journalists in the creation and spreading of negative stereotypes about Roma people. *Srodkowoeuropejskie Studia Polityczne*, 2, 141-154.
74. Villano, P., Fontanella, L., Fontanella, S., & Di Donato, M. (2017). Stereotyping Roma people in Italy: IRT models for ambivalent prejudice measurement. *International Journal of Intercultural Relations*, 57, 30-41.
75. Vrăbiescu, I., & Kalir, B. (2018). Care-full failure: how auxiliary assistance to poor Roma migrant women in Spain compounds marginalization. *Social Identities*, 24(4), 520-532.

76. Wenham-Ross, K. (2019). *Britain's Conservatives Pledge to Target Roma*. Retrieved on April 17, 2020, from <https://foreignpolicy.com/2019/12/12/gypsies-travellers-boris-roma-crack-down-conservative-tory/>.
77. West, C. (2012). Words that Silence? Freedom of Expression and Racist Hate Speech. In I. Maitra & M. K. McGowan (eds.) *Speech and Harm: Controversies Over Free Speech* (pp. 222-248). Oxford University Press.