

MIGRATION DISCOURSE PRECEDING THE GREAT MIGRANT CRISIS IN THE CROATIAN AND ENGLISH LANGUAGE – A CORPUS-BASED CRITICAL DISCOURSE ANALYSIS

JELENA GUGIĆ

Juraj Dobrila University of Pula, Faculty of Educational Sciences, Pula, Croatia
jelena.gugic@unipu.hr

Abstract In this paper, a corpus-based critical discourse analysis of adjectival and verbal collocates of the words *refugee*, *asylum seeker*, *immigrant* and *migrant* are presented. The research was conducted using two corpora available on the Sketch Engine interface - hrWaC and enTenTen2015. The aim of the research was to list the 15 most typical adjectival and verbal collocates of the aforementioned nodes in the Croatian and English language and compare them. The critical discourse analysis included in-depth reading of concordance lines including the extracted collocations in the two languages, as well as defining the discourse surrounding them in a more qualitative way. The findings suggest that prior to the outbreak of the 2015 migrant crisis the discourse surrounding the four groups of people was judgemental and prejudiced, unwelcoming, and sometimes even extremist, but that the Croatian language discourse was less xenophobic than the English language discourse.

Keywords:

corpus,
critical discourse
analysis,
migrant crisis,
Croatian,
English

1 Introduction

Croatia is, from the historical perspective, a country from which people have mostly emigrated. According to Jerić (2019) the emigration of Croatian people was caused by different economic or political reasons, in eight different historical periods. The first emigration wave occurred in the 15th century due to the onslaught of the Turks, followed by the second at the end of the 19th and the beginning of the 20th century when people emigrated because of the disease which affected grapes preventing people from earning their daily bread by producing wine, and then the third due to the Great Depression and its consequences between the two world wars. The fourth emigration wave occurred due to the founding of a communist regime, and it was followed by the emigration of the 1960s when the former Yugoslavia opened its borders to western countries. The sixth emigration was caused by the Croatian Homeland War (the Croatian War of Independence 1991 – 1995), and that was followed by the seventh wave linked to the financial crisis of 2008. The eighth wave started in 2013 when Croatia became a member of the European Union, and it is still going on. Immigration also occurred throughout history in similar periods, and the immigrants coming to Croatian territory mostly originated from its neighbouring countries (Hungarians, Italians from the Venetian Republic, Austrians, Germans, Croats from Bosnia, Slovaks, Czechs, Jews, and Poles) (Zlatković Winter, 1993). The European migrant crisis of 2015 was marked by a huge inflow of refugees trying to escape the turmoil occurring in the countries of the Middle East, especially in the decade preceding this year. Croatia was a country which migrants only wanted to pass through, but when its neighbouring countries stopped accepting them, many were forced, or even decided themselves, to stay and seek asylum.

The interest of this paper lies in the representation of *refugees*, *asylum seekers*, *immigrants* and *migrants* in a Croatian and English language corpus (hrWaC and enTenTen2015) consisting of texts from the web from the period prior to the European migrant crisis, and in obtaining an insight into how the speakers of the two languages depicted them at the time when the huge migration wave had not yet occurred. After presenting similar research conducted about this topic before the present one, the paper presents the research aims which it wanted to achieve and the adopted methodology. The research included the analysis of the most typical adjectival and verbal collocates of the four nodes (*refugees*, *asylum seekers*, *immigrants* and *migrants*) in the two languages and after defining them quantitatively, they were also analysed in

a qualitative fashion with the intention to confirm or reject the findings reported by the collocation analysis.

2 Theoretical background

Since the European migrant crisis started in 2015, numerous studies about the topic and discourse surrounding migrants have been conducted. Thus, Barlai, Fähnrich, Griessler and Rhomberg (2017) edited a book about the migrant crisis in which they included articles written by authors from almost all European countries tackling some aspect of the migrant crisis, mostly from the socio-political and public discourse perspectives. In 2016 Lendaro addressed the migrant crisis from the perspective of border policies and discussed their employment or lack of employment in different European countries which should apply the regulations thus guaranteeing the same human rights to all people, no matter what their country of origin. There were other numerous authors addressing the same topic from the national security perspective (Adamson, 2006; Lalić et al., 2015; Metelev, 2016; Ivanova, 2017; Vulević, 2018; Bježančević, 2019; Gryshova et al., 2019). Goodman et al. (2017) analysed the discourse surrounding the migrant crisis and found out that it changed according to the geographical position of the most serious event in the denoted period (i.e. from the “Mediterranean migrant crisis” to “Calais migrant crisis” and to the “European migrant crisis”, then turning into the “refugee crisis”). In 2015 Berry et al. prepared a report about the press coverage of the refugee and migrant crisis in the EU for the United Nations High Commission for Refugees. They analysed five European countries (UK, Sweden, Germany, Spain, and Italy). Knowing that the use of particular linguistic features and word combinations can affect public opinion, the migrant crisis covered by the media, especially by the press, was of huge interest to linguists as well. Most of them analysed the words *migrants*, *immigrants*, *refugees*, *asylum seekers*, *Islam* and *Muslims* using the corpus method and discourse analysis (Baker et al., 2008; Crymble, 2010; Khorsavinik, 2010; Baker et al., 2013; Blinder & Allen, 2016; Salashour, 2016; Al Fajri, 2017; Sakellariou, 2017; O’Regan & Riordan, 2018; Serafis et al., 2020). What seems to be the common ground of all these studies is that the arrival of the aforementioned groups of people was not seen as a positive occurrence, and attitudes toward them have been mostly negative.

In Croatia the number of studies addressing the migration crisis from a linguistic point of view is smaller. The most recently published one is by Dobrić Basanež and Ostojić (2021) who investigated the migration discourse in the Croatian news media employing the corpus linguistics method and critical discourse analysis. After conducting a detailed analysis, the authors concluded that “RASIM¹ are often discriminated against not only when they are physically pushed back at the borders, but also in the words the media use to report on these incidents.” (Dobrić Basanež & Ostojić, 2021, p. 22). They also claimed that although the migration discourse in Croatian news was less discriminatory than in the UK press, it still created xenophobic attitudes (ibid., p. 22–23). Bezić and Petrović (2019) analysed the discourse strategies in the media representation of migrants on a corpus of Italian and British daily newspapers. They paid special attention to the construction of the opposite personal pronouns we/they (representing the European Union/migrants) and analysed the nomination strategies used which were more frequent in the Italian press due to Italy being the country from which the wave of migration starts and the UK being at the edge of this process. Finally, Podboj (2019) described the discourse construction of identity in narratives about personal migration experiences. She conducted a semi-structured interview with 10 participants who emigrated from Croatia in 2010. Although her dissertation did not include refugees, asylum seekers, immigrants and migrants in the analysis, her work expressed the need for “new, discourse- and identity-oriented criteria” approaches to linguistic analysis (ibid, p. 13).

3 Empirical research

3.1 Research aims

The first aim of the present research was to compare the most typical collocates of the four terms usually analysed as part of the migrant crisis discourse (namely *refugees*, *asylum seekers*, *immigrants* and *migrants*) in the Croatian and English language using the hrWaC corpus (Ljubičić & Klubička, 2014) and the enTenTen2015 corpus (Jakubiček et al., 2013). The second aim was to conduct a critical discourse analysis and compare the representation of these four groups in the two corpora and understand how they were pictured on the web in the period preceding the outbreak

¹ Abbreviation for Refugees, Asylum Seekers, Immigrants and Migrants (Gabrielatos & Baker, 2008; Baker et al. 2008)

of the migrant crisis in 2015. The assumptions behind these aims were that both languages would have similar collocates of the words examined and that, as former studies proved, they would carry negative connotations.

3.2 Methodology

Studies interested in the linguistic aspect of the migrant crisis discourse usually use a combination of the Corpus Linguistics method (CL) and Critical Discourse Analysis (CDA). McEnery and Hardie (2012, p. 1) define CL as “an area which focuses upon a set of procedures, or methods, for studying language.” The most commonly used CL techniques in linguistic analysis are key word, collocation and concordance. CDA, on the other hand, is defined by Huckin (1997, p. 87) as “a highly context-sensitive, democratic approach which takes an ethical stance on social issues with the aim of improving society.” Taking into consideration discourse or “language above the sentence or above the clause” (Stubbs, 1983, p. 1), CDA looks at language in a critical manner and is interested in how ideologies and power relations are expressed in it (Baker, 2008). Since to apply the strategies of CDA it is important to have larger amounts of text, it does not come as a surprise that CDA found solid grounds for its application in CL. As already mentioned, one of the most exploited CL techniques is collocation. Collocations are defined as “strings of words that seem to have certain ‘mutual expectancy’, or a greater-than-chance likelihood that they will co-occur in any text.” (Nattinger and DeCarrico (1992, p. 21, as cited in Moehkardi, 2002). In their cooccurrence, words can carry certain connotations (Stubbs, 1996) and are therefore used as markers signalling the mental image people can make by seeing those words together. However, as Stubbs (1996, p. 195) suggests “if collocations and fixed phrases are repeatedly used as unanalysed units in media discussion and elsewhere, then it is very plausible that people will come to think about things in such terms.” Considering that Stubbs mentioned ‘unanalysed units in media discussion’, it is of the utmost importance not to leave the language we study unexplained. The quantitative nature of CL thus becomes empowered by the qualitative nature of CDA. The same two methods were applied in this research as well.

To achieve the aims of this research two corpora available on the online concordancer Sketch Engine (Kilgariff et al., 2004; Kilgariff et al., 2014) were used, namely the hrWaC 2.2 and the enTenTen2015. The web “contains a massive collection of data that is ever-growing” (McEnery & Hardy, 2012, p. 7) and “contains genres that are not found in traditional written corpora, such as blogs and online discussion forums” (Al Fajri, 2017, p. 383). Therefore, corpora of texts crawled from the web were seen as a suitable ground for analysis by which to achieve the aforementioned aims. In Sketch Engine the hrWaC corpus consists of 1,211,328,660 words. It was created in 2014 and is composed of texts taken from the web (the .hr domain). To analyse the situation in the English language the enTenTen2015 corpus was analysed. It consists of 13,190,556,334 English words extracted from the internet. Although it would have been better to have two corpora belonging to the same corpus family (hrWaC and ukWaC), the enTenTen2015 was chosen in order to have a matching time period of texts crawled from the web in the two languages since the ukWaC corpus was compiled in 2007 (it is important to mention that the enTenTen2015 version is not the last one in the TenTen family corpus; it was followed by enTenTen2018 and enTenTen2020).

First, collocations were extracted from both corpora using the Word Sketch tool. Only lexical or content words appearing as collocates were considered since they bear semantic content and contribute to the meaning of the sentence or clause they appear in offering various types of information. According to Petrović (2007, as cited in Košuta, 2012) the most common types of collocations in the Croatian language are adjective + noun, verb + noun and adverb + verb. Since the four nodes observed in this research are nouns, the collocates analysed, at the same time content words, were adjectives and verbs. The same procedure was applied in the study of collocates for the English nodes. The number of possible collocates for each of the four nodes was high in both languages (for instance, the word *izbjeglica* has 88 adjectival collocates, while *refugee* attracts 100 adjectival collocates). Therefore, only the top 15 collocates with the highest logDice score were considered. “The logDice score has a reasonable interpretation, scales well on a different corpus size, is stable on subcorpora, and the values are in reasonable range.” (Rychlý, 2008, p. 9). Its comparability among different corpora sizes makes it suitable for this research as well. LogDice is fixed at its maximum value of 14 and is therefore easy to interpret, because the closer the value to 14, the more typical the collocation. However, in this respect, Rychlý (2008) states that the logDice value is usually less than 10. To reach

the second research aim, the Concordance tool was employed. It enables researchers to look at words or collocations of interest in context which is crucial to conduct discourse analysis.

3.3 Results and discussion

Since the four words to be analysed in this research (*izbjeglice*, *azilanti*, *imigranti* i *migranti* - *refugees*, *asylum seekers*, *immigrants* and *migrants*) had been set prior to the commencement of the research, the first step to be taken was to search these four words using the Word Sketch tool offered in Sketch Engine. This was done in both the hrWaC and the enTenTen2015 corpus. The collocations were then sorted in order to extract those whose collocate was a content word (adjective or verb), and among them only the top 15 according to their logDice score.

3.3.1 Adjectival and verbal collocates of *izbjeglica* and *refugee*

Table 1: Adjectival and verbal collocates of *izbjeglica*

Adjectival collocate	logDice	Verbal collocate	logDice
palestinski	8.7	zbrinjavati	8.5
sirijski	7.9	zbrinuti	7.9
muslimanski	7.7	naseliti	7.6
afganistanski	7.4	boraviti	7.5
tibetanski	7.1	naseljavati	7.4
bošnjački	6.9	doći	6.8
irački	6.7	useliti	6.7
čečenski	6.6	doseljavati	6.7
kosovski	6.6	bježati	6.6
somalijski	6.6	doseliti	6.5
kubanski	6.6	pristizati	6.4
albanski	6.6	smjestiti	5.7
srpski	6.4	živjeti	5.6
sudanski	6.4	biti	5.5
židovski	6.4	prevoziti	5.3

Source: Sketch Engine

Table 2: Adjectival and verbal collocates of *refugee*

Adjectival collocate	logDice	Verbal collocate	logDice
Syrian	11.0	resettle	10.0
Palestinian	9.6	arrive	7.9
Afghan	8.8	displace	7.3
Somali	8.7	repatriate	7.2
Iraqi	8.3	relocate	6.9
Sudanese	8.3	deport	6.9
Palestine	8.2	welcome	6.7
Burmese	7.6	admit	6.6
Eritrean	7.5	accept	6.5
Tibetan	7.4	assist	6.5
Rohingya	7.3	detain	6.5
Jewish	7.2	settle	6.5
Vietnamese	7.1	strand	6.5
(North) Korean	7.0	house	6.5
Bhutanese	6.9	shelter	6.5

Source: Sketch Engine

The first most noticeable information given in Table 1 and Table 2 is that both the Croatian and English language associate the word *refugee* with their nationality. In fact, by checking the list, of the top 15 adjectival collocates, one sees that all are nationalities. *Palestinian* and *Syrian refugees* are the most typical collocations in both corpora (in reversed order). Besides them, nationalities mentioned in both corpora are *Afghan*, *Tibetan*, *Iraqi*, *Somali*, *Sudanese* and *Jewish*. When it comes to verbal collocates, it is extremely interesting to notice that the most typical verbs associated with *izbjeglica* in the Croatian language are *zbrinjavati* and *zbrinuti* ('to take care of, to give shelter to'), whereas the verb *shelter* is the least typical collocate in the English language. A more detailed analysis shows that all the Croatian verbs have a positive or neutral connotation. The only one which could be considered negative is the verb *bježati* ('to run away from'), but a closer look at its concordances shows that the verb only describes the situation when the refugees had to flee from their countries:

1. *državu dobrovoljno, u potrazi za boljim životom i ako se odluči vratiti kući nastavit će uživati zaštitu svoje vlade. Izbjeglice bježe zbog straha od proganjanja i ne mogu se u datim okolnostima sigurno vratiti svojoj kući Osobe za koje se putem*
2. *komentare, pozzzzz. Ipak znam o čemu da pišem. O strašnom ratu na Bliskom istoku. Izrael i dalje bombadira Libanon, a izbjeglice bježe u Siriju, Cipar i druge zemlje. STRAŠNO Gadno je dolje, ali mijenjam temu I onda kažu da je zrak onečišćen, pa kako*

On the other hand, there are more verbs in the English language which could be classified as bearing a negative connotation (*repatriate, detain, strand*). In the end, it can be said that the verbal collocates in both languages are mainly concerned with giving refugees a place to live (*zbrinuti, naseliti, useliti, doseliti, smjestiti* and *resettle, relocate, settle, house, shelter*).

3.3.2 Adjectival and verbal collocates of *azilant* and *asylum seeker*

Table 3: Adjectival and verbal collocates of *azilant*

Adjectival collocate	logDice	Verbal collocate	logDice
politički	3.7	odobriti	2.7
afganistanski	3.7	primati	1.9
lažan	3.6	priznati	1.6
potencijalan	3.2	primiti	1.4
smjestiti	3.0		
ruski	2.5		
ilegalan	2.4		
takozvani	2.4		
mnogobrojan	1.9		
stran	0.2		

Source: Sketch Engine

Table 4: Adjectival and verbal collocates of *asylum seeker*

Adjectival collocate	logDice	Verbal collocate	logDice
destitute	6.7	deport	8.2
Eritrean	6.5	detain	8.1
Lankan	5.9	refuse	6.8
unaccompanied	5.9	fail	6.8
failed	5.7	resettle	6.2
Tamil	5.5	demonise	6.2
Afghan	5.5	relocate	6.1
LGBT	5.3	intercept	6.0
Rohingya	5.2	deter	5.9
Sudanese	5.0	repatriate	5.6
bogus	4.9	reject	5.5
Montagnard	4.8	imprison	5.4
Hazara	4.6	arrive	5.3
Syrian	4.5	house	5.2
Somali	4.5	disperse	4.9

Source: Sketch Engine

A first glance at Table 3 shows that the word *azilant* is seldomly used as a collocation node in the Croatian language corpus. It has only 10 adjectival and four verbal collocates. On the other hand, in the English language corpus there was a sufficient number of both types of collocates for the node *asylum seeker*. It can be discussed that if the Croatian node was *tražitelj azila* instead of *azilant* (both have the same meaning – *asylum seeker*) there would be a larger number of collocates. However, the assumption proved to be wrong. The search resulted in only four adjectival and five verbal collocates. Besides the small number of collocates for *azilant*, their typicality is low. According to the *logDice* score, the closer the score value to 14, the more typical the collocate-node association. The highest score is 3.7 for the collocate *politički* and *afganistanski* so it seems that Croatians were mostly, but not too seriously, concerned about the number of asylum seekers coming from Afghanistan, and the fact that they fled their country due to political reasons. It is also interesting to see that the Croatian corpus ranked third, fourth and seventh the collocates *lažan*, *potencijalan* and *takozvani* (*fake*, *potential* and *so-called*) which indicates a lack of trust toward the true reason of people wanting asylum. We can argue that this is a xenophobic stand toward the RASIM group in view of the crisis that was about to happen. Similarly, Maričić et al. (2014) found out that in 2013 the attitude toward asylum seekers was negative, but that there was also a lack of information among the public linked to their rights and the barriers they had to face in Croatia. We expected to see Afghan people depicted as *fake asylum seekers* (since they had a higher *logDice* score), but surprisingly, the topic was about asylum seekers from the territories of Serbia, Macedonia and Bosnia and Herzegovina seeking asylum in Sweden. Regarding verbal collocates, all the four verbs had a positive connotation (*odobriti*, *primati*, *priznati*, *primiti* – *approve*, *accept*, *recognise*, *admit*).

When it comes to the adjectival collocates of *asylum seeker*, the English language corpus still revolves around their nationality or sexual orientation (LGBT), their poor financial means being the main concern (*destitute asylum seekers*). In the top 15 adjectival collocates there is one mention of asylum seekers being fake - *bogus*. *Bogus* has a higher *logDice* score (4.9) than the three adjectives of the same discourse in the Croatian language, so it can be argued that although there are less words depicting asylum seekers as fake, English language speakers make a stronger association between this group of people and their honesty in seeking a shelter and new place to live. Regarding the verbal collocates of *asylum seeker*, they have a much

higher typicality score than verbs in the Croatian corpus, and all except one (*house*) bear a negative connotation.

1. *allegedly to the reason why he sought asylum. IHR once again urges the European countries not to **deport** Iranian asylum seekers back to Iran. Mahmood Amiry-Moghaddam, the spokesperson of IHR said: The new charges against him are due to his*
2. *in its "Centers for Illegal Foreigners," in breach of its Law on Foreigners and Stateless Persons. **Rejected** asylum seekers are **detained** in extremely poor conditions pending their removal. Seven percent of asylum seekers live in substandard*
3. *State practice with regard to detention. While some EU Member States such as Germany and Italy rarely **detain** asylum seekers, more than 13,000 asylum seekers entered detention in the UK in 2012 and Malta continues to detain for months the vast*

Most of the verbal collocates in the English language corpus bear the meaning of ‘get rid of something’ (*deport, refuse, resettle, relocate, repatriate, reject*), and this is how *asylum seekers* were depicted on the English web.

3.3.3 Adjectival and verbal collocates of *imigrant* and *immigrant*

Table 5: Adjectival and verbal collocates of *imigrant*

Adjectival collocate	logDice	Verbal collocate	logDice
ilegalan	10.7	deportirati	8.3
afrički	7.1	protjerati	6.3
muslimanski	6.7	zapošljavati	6.2
legalan	6.3	prevoziti	6.2
meksički	6.3	prebacivati	5.4
alžirski	6.2	uvoziti	5.2
sjevnoafrički	6.2	napadati	4.2
friski	6.1	mrziti	3.6
azijski	6.0	izbaciti	3.3
kubanski	5.8	privlačiti	3.0
nelegalan	5.7	privući	3.0
tuniski	5.6	primati	3.0
irski	5.5	spasiti	2.6
pakistanski	5.5	ubiti	2.3
bangladeški	5.5	spriječiti	2.1

Source: Sketch Engine

Table 6: Adjectival and verbal collocates of *immigrant*

Adjectival collocate	logDice	Verbal collocate	logDice
undocumented	11.0	deport	8.9
illegal	10.6	detain	8.3
Mexican	8.1	arrive	7.9
Latino	7.7	criminalize	6.4
Muslim	7.4	legalize	6.3
unauthorized	7.4	assimilate	6.3
Irish	7.2	settle	6.2
Italian	7.0	blame	6.0
Jewish	6.9	naturalize	6.0
German	6.8	orphan	5.9
second-generation	6.8	smuggle	5.9
first-generation	6.7	arrest	5.8
Polish	6.7	apprehend	5.7
Chinese	6.7	welcome	5.6
Hispanic	6.7	bar	5.5

Source: Sketch Engine

With the node *immigrant* in the Croatian language corpus the situation changes again. It is again concerned with the origin of the immigrants coming to the territory of Europe and Croatia (*afrički, muslimanski, alžirski, sjevernoafrički*, etc.), and again its speakers seem to be bothered by the legal status of the immigrants. The most typical collocate, with a score higher than 10 (the *logDice* value is usually less than 10, as already mentioned) is *ilegalan* (10.7), with *nelegalan* (5.7) also being given importance. Since *ilegalan* usually “collocates with words such as *drugs, activities, trade, abortion, logging, rave* and *drug*” (Al Fajri, 2017, p. 384), its association with the word *immigrant* bears a strong negative connotation. “It seems to more criminalise the people rather than their actions, and it can also dehumanise them.” (ibid, p. 384). With verbal collocations the situation gets serious for the first time. Of the top 15 collocates, 11 have a strong negative collocation, from *deportation, banishment, attacking* and *bating* to even *killing!*

u Švedskoj. Upravo u toj državi izbili su neredi početkom ovog mjeseca, nakon što je policija ubila 69 - godišnjeg imigranta . Ukrajinska atletičarka Natalija Dobrinska osvojila je zlatnu medalju u petoboju na Sjjetskom dvoranskom

It seems encouraging that the third most typical collocate is *zapošljavati* (*employ*). It is an indication of a strong sense of humaneness that people felt toward the poor immigrants fleeing from their countries. However, a look at the concordances showed a completely different situation – immigrants were usually employed as a cheap labour force or to work as unregistered employees who are easy to exploit.

1. predstavila je prijedlog direktive prema kojoj bi se strogo sankcionirali svi poslodavci koji svjesno **zapošljavaju** imigrante **za rad na crno**. Europski povjerenik za pravosuđe, slobodu i sigurnost Franco Frattini izjavio je kako Europa neće više
2. da rade za minimalnu nadnicu, a istodobno otpuštali stare radnike koji bi se mogli buniti **i zapošljavali** novodošle imigrante koji **se nikada nisu bunili** - Ne. Pozivanje na neku specifičnost Hrvatske nije opravdano. Slični problemi postoje i u
3. tradicionalno gostoljubiva zemlja ", objašnjava Vergara. Stručnjaci se slažu da neki farmeri **zapošljavaju** imigrante **jer ih je lako izrabljivati**. Sindikalni vođa Canamero kaže da se svake godine prijavi od 15 do 20 slučajeva

Besides again being concerned with the nationality of the immigrants, the English language corpus showed high similarity to the Croatian. *Undocumented*, *illegal* and *unauthorised* score very high among adjectival collocates, showing the same concerns as Croatians had. *Undocumented* and *illegal* have a *logDice* score higher than 10, which makes them extremely typical for that discourse. It is encouraging that *undocumented* is more typical than *illegal*. Being milder in its negative connotation than *illegal*, *undocumented* is also less dehumanising because it shows people who need to get documents in the country of their arrival. The verbal collocates of the English language corpus are typically negative. The immigrants' deportation and detention is still something that concerns the English most, but one of the worst verbs which shows immigrants as objects or goods for which it is only important to be transported from one point to another – for good money, of course – is the verb *smuggle*.

1. risks that some immigrants are prepared to take to get into Canada. There is a growing business of attempting **to smuggle** immigrants into the country, many of which are hoping to acquire political asylum. However, the fate of most of these illegal
2. Police] has questioned 212 migrant traffickers in the Calais area and dismantled 5 international rings **smuggling** immigrants into the United Kingdom, organized from several European countries. Eric Besson and Damian Green welcomed the fact

It needs to be said that there are, of course, verbs which show a positive attitude towards immigrants and depict them as people in need of help whose arrival in the country is seen as something inevitable and it is therefore better to make it easier for them to assimilate. These are: *legalize, settle, naturalise, welcome*.

3.3.4 Adjectival and verbal collocates of *migrant* (Croatian and English word)

Table 7: Adjectival and verbal collocates of *migrant* – Croatian

Adjectival collocate	logDice		Verbal collocate	logDice
ilegalan	8.4			
skilled	7.5			
neregularan	7.5			
neregistriran	6.1			
nezakonit	6.1			
dnevni	5.5			
afrički	4.1			
legalan	3.5			
ekonomski	3.3			
kineski	2.5			
potencijalan	2.5			
muslimanski	2.4			
siromašan	2.3			
svakodnevnan	0.6			
ratni	0.6			

Source: Sketch Engine

Table 8: Adjectival and verbal collocates of *migrant* – English

Adjectival collocate	logDice		Verbal collocate	logDice
undocumented	9.8		deport	8.1
irregular	9.3		detain	7.6
illegal	8.1		smuggle	7.5
skilled	7.5		strand	7.5
neotropical	7.2		rescue	7.4
labour	7.0		arrive	7.2
long-distance	6.9		drown	6.2
Bangladeshi	6.9		apprehend	6.2
Burmese	6.9		intercept	6.1
forced	6.6		repatriate	6.0
Filipino	6.6		force	5.8
would-be	6.5		deter	5.7
African	6.4		return	5.6
passage	6.3		criminalize	5.5
vulnerable	6.3		traffic	5.5

Source: Sketch Engine

The first thing to notice for the Croatian node *migrant* is that the *Word Sketch* tool did not produce any results which would be considered verbal collocates. Since this was considered improbable, the word *migrant* was checked using the Concordance tool. As many as 93 concordances were produced and a deeper analysis of the concordance lines found out that there were no verbs associated with the word *migrant* in the hrWaC corpus. It was also found that many concordance lines were in English, which explained why the second most typical adjectival collocate of *migrant* in the Croatian language was an English word. By examining the wider context, it was clear that the word *migrant* was usually associated with the adjective *skilled* because used by Croatian speakers who explained the migration procedure to their interlocutors, future emigrants:

1. iz Red Dot-a (§ 2 shop-a) Znam inženjere koji zaradjuju nesto vise nego cistaci Znam i jednu Dalmatinku **skilled migrant** koja ima 2 (prirodnjackska) fakulteta od doma. Tu je vec petnaestak godina i nezaposlena je. Nedavno je ovdje
2. zadnji je za lokalnu upravu. Ovisno koji sektor te zanima. Jel ovo dovoljno ili želiš još Meni su ovi iz Highly **Skilled Migrant** Programe-a javili da imam dobre izgleda za dobivanje radne dozvole ali bi sada trebala platiti nekaj da detaljne
3. skupiti AUD20k koliko treba za karte, papire i prvih mjesec-dva života ne treba nigdje ni ici. Takav niti nije **skilled migrant** Aha. Nivo produktivnosti agrarnog sektora je statican. Dosli smo do kraja razvoja agrarnog sektora i to je to. Zamisli

As earlier, the adjectival collocations of the word *migrant* in the Croatian language corpus also showed that according to the texts found on the web, Croatians were mostly preoccupied with the legal status of the migrants. Hence the collocates *ilegalan*, *neregularan*, *neregistriran*, *nezakonit* (*illegal*, *irregular*, *unregistered*, *unlawful*). Regarding their typicality, all have a logDice score above 5 which confirms that the discourse about the lack of a legally approved status of migrants was prevalent even in the years preceding the 2015 migration crisis.

Regarding the English node *migrant*, both adjectival and verbal collocates were found in the enTenTen2015 corpus. The most typical adjectival collocates reflect the same concern among the British as among Croatians. The texts on the web mostly referred to *undocumented*, *irregular* and *illegal migrants*. It is interesting to note that the fourth most typical collocate is *skilled* which can be correlated to the results of the Croatian language corpus concordance analysis where the word *skilled* was also highly

associated with *migrant*. In terms of the 15 most typical adjectival collocates of *migrant*, an adjective describing them as fake also appears: the adjective *would-be*.

1. *and asylum seekers show the reality behind the European Union's bogus internationalism. The situation for **would-be migrants** has deteriorated significantly since these photos were taken, and in particular since uprisings and*
2. *, Texas, to see the pope. With mild weather ahead, southern Europe is once again bracing for new boatloads of **would-be migrants** and asylum seekers from North Africa. Italy has borne the brunt of this migrant flow for two decades, and it has*
3. *for 2015 : "And struggling economies on the continent will mean there will be more desperate young **would-be African migrants** washing up on the shores of Lampedusa in Italy. "The African Union will hold regular summits and our leaders will make*

This is in line with the findings by Al Fajri (2017). He analysed the ukWaC corpus and found that a very strong collocate of immigrant was *would-be*. The ukWaC corpus he analysed showed *would-be immigrants* as *unwelcome*, *illegal* and *detained*, as well as *dying* at sea, which was also confirmed in the enTenTen2015 corpus.

The verbal collocates, again, bear a negative meaning when it comes to *migrants*, and regarding their typicality, all can be considered highly typical (logDice score above 5). The largest number of them describe ways of physically preventing the migrants from moving (*detain*, *strand*, *apprehend*, *deter*, even *force*), there are verbs describing the need to send the migrants back to their country of origin (*deport*, *repatriate*, *return*), and sadly, there are also two verbs which indicate that the activities around the migrants are illegal (*smuggle*, *traffick*). Luckily, among them there is a positive verbal collocate depicting the efforts made by local people and authorities trying to save lives while not caring about their status.

1. *are responsible for saving tens of thousands of more lives. In 2014 seafarers aboard 800 merchant ships **rescued** 40,000 migrants . Their role in the large scale rescue of migrants should be recognised and commended. However, EU governments are*

2. *champions in Kenyans Sharon Cherop and Caroline Continue reading → Gallery: The Daily Edit 04.21.15 A man **rescues** a migrant from the Aegean sea, within the japanese² island of Rhodes, Monday, April 20, 2016. Greek authorities stated that no*
3. *for NGOs to move away from our comfort zone, strengthen our advocacy and bring about change." Local people **rescued** 700 migrants from a sinking ship last week The UN has condemned the refusal of South-East Asian countries to rescue thousands of*

4 Conclusion and further implications

The RASIM are an extremely sensitive group of people who are in need, due to various circumstances, of equal rights and opportunities as the citizens of the country they wish to live in. However, they are often seen as a threat to the local community, and the media have greatly contributed to that. "...by the choice of topic and the way of reporting on the phenomenon of migrations and migrants, the media shape attitudes about them" (Bezić & Petrović, 2019, p. 82). However, this is done through language, so it is very important to understand the role that language plays in depicting and shaping reality.

The results obtained in this paper demonstrate that the public opinion about the RASIM group presented on the Internet in the period prior to the 2015 migrant crisis was mostly negative. One of the most prominent concerns speakers of the two languages had was the nationality of those coming to their country. In his book entitled "The Nature of Prejudice," Gordon Allport defines prejudice as "an antipathy based upon a faulty and inflexible generalisation" (Allport, 1954, p. 13, as cited in Killan, 1955). It can be argued that the concern of people in finding out about the nationality of others is linked to the image, or maybe prejudice, they have about a certain nationality. Furthermore, people were interested in the reasons why those people fled from their countries and if their intentions were trustworthy. Hence, they were *fake, potential, so-called, would-be, bogus*. However, when it comes to their honesty, Croatians, who do not have a tradition of immigration as the English do, seem to be more gullible in this respect. Not only was public opinion concerned about the RASIM people's good intentions, but even more with their legal status. Both were most afraid of them being *illegal, undocumented, irregular, unregistered, unlawful*. When it comes to verbs used to describe the actions undertaken around the RASIM

² The island of Rhodes is a Greek island, but the mistake was not corrected in order to retain the original text found on the web.

group, the similarity between the two languages is smaller with verbs than with adjectives. Throughout the whole analysis, it is obvious that the Croatian language corpus uses more positive-connotation verbs associated with RASIM. It is concerned with giving RASIM a place to live (*resettle, relocate, settle, house, shelter*), they wish to *approve* their status, *accept, recognise* and *admit* them, but there is a sub-group, namely immigrants, towards which the attitude is mostly negative - from *deportation, banishment, attacking and hating* to even *killing!* There are fewer positive verbs used in the English language corpus linked to the RASIM group. The UK, as a country used to, and perhaps tired of, immigration is concerned with finding a way to *repatriate* RASIM. Very seldomly do they use verbs such as *house* or *shelter* thus proving them unwelcome. Sadly, the English web gives examples of the RASIM dehumanisation when describing activities of *smuggling* and *trafficking*. This is extremely important if correlated with the strong association made between the RASIM group and their image of illegal and undocumented people.

To conclude, in the period preceding the 2015 migrant crisis, the RASIM were described not too negatively except for the suspicion of their legal status in both languages. However, the actions undertaken around these people were less positive in the English than Croatian corpus. Considering the different traditions of the two countries – the UK being used to immigrations; Croatia being a country of emigrants – the milder, more positive verbs used on the Croatian web were expected. However, considering the research as a whole, the attitude toward RASIM proved to be judgemental and prejudiced (nationality), unwelcoming (repatriation) and sometimes even extreme (smuggling, trafficking, killing). Therefore, both the assumptions behind the research aims could be considered valid. Hope in the human nature of men must, however, not be renounced. The few examples depicting the RASIM group, and the actions linked to them positively represent the tiny, but arguably strong, glimmer of hope that the 21st century society is able to strive for and reach the values set by the modern humanities – a healthy democracy which leads to human participation in solving modern concerns and maintaining human values.

References

- Adamson, F.B. (2006). Crossing borders: International migration and national security. *International Security*, 31(1), 165–199. doi: <https://doi.org/10.1162/isec.2006.31.1.165>
- Al Fajri, M.S. (2017). Hegemonic and minority discourses around immigrants: A corpus-based critical discourse analysis. *Indonesian Journal of Applied Linguistics*, 7(2), 381–390.

- Allport, G.W. (1954). *The Nature of Prejudice*. Doubleday Anchor.
- Baker, P., Gabrielatos, C., KhorsaviNik, M., Krzyżanowski, M., McEnery, T. & 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. <https://doi.org/10.1177/0957926508088962>
- Baker, P., Gabrielatos, C. & McEnery, T. (2013). *Discourse Analysis and Media Attitudes. The Representation of Islam in the British Press*. CUP.
<https://doi.org/10.1017/CBO9780511920103>
- Barlai, M., Fähnrich, B., Griessler, C., Rhomberg, M. & Filzmaier P. (2017). *The migrant crisis : european perspectives and national discourses*. Lit Verlag.
- Berry, M., Garcia-Blanco, I. & Moore, K. (2015). *Press Coverage of the Refugee and Migrant Crisis in the EU: A Content Analysis of Five European Countries*. Cardiff School of Journalism, Media and Cultural Studies.
- Bezić, M., & Petrović, I. (2019). Diskursne strategije u medijskom prikazu migranata. *Zbornik radova Filozofskog fakulteta u Splitu*, (12), 81–97.
- Bježančević, S. (2019). Migrantska kriza u Europskoj uniji – izazovi, europska sigurnost i zaštita temeljnih prava. *Zbornik Pravnog fakulteta Sveučilišta u Rijeci*, 40(3), 1231–1248. <https://doi.org/10.30925/zpfsr.40.3.11>
- Blinder, S., & Allen, W. L. (2016). Constructing immigrants: Portrayals of migrant groups in British national newspapers, 2010–2012. *International Migration Review*, 50, 3–40. <https://doi.org/10.1111/imre.12206>.
- Crymble, L. (2010). *Textual Representations of Migrants and the Process of Migration in Selected South African Media: A Combined Critical Discourse Analysis and Corpus Linguistics Study*. Master of Arts Thesis. Rhodes University.
- Dobrić Basanež, K., & Ostojić, P. (2021). Migration discourse in Croatian news media. *Medijaska istraživanja*, 27(1), 5–27.
- Gabrielatos, C., & 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.
- Goodman, S., Sirriyeh, A. & McMahon, S. (2017). The evolving (re)categorisations of refugees throughout the “refugee/migrant crisis.” *Journal of Community & Applied Social Psychology*, 27(2), 105–114. doi: <https://doi.org/10.1002/casp.2302>
- Gryshova, I., Kofman, B. & Petrenko, O. (2019). Migration cultures and their outcomes for national security. *Journal of Security and Sustainability Issues*, 8(3), 521–530. [https://doi.org/10.9770/jssi.2019.8.3\(18\)](https://doi.org/10.9770/jssi.2019.8.3(18))
- Huckin, T. N. (1997). Critical discourse analysis. In T. Miller (Ed.), *Functional approaches to written text: Classroom applications*, 87–92. United States Information Agency.
- Ivanova, D. (2017). Hungarian security policy and the migrant crisis (2015–2017). *International conference Knowledge-Based Organisation*, 23(1), 166–171.
- Jakubiček, M., Kilgarriff, A., Kovář, V., Rychlý, P., & Suchomel, V. (2013). The TenTen corpus family. In *7th International Corpus Linguistics Conference CL*, 125–127.
- Jerić, M. (2019). Suvremeno iseljavanje Hrvata: kakva je budućnost Republike Hrvatske. *Oeconomica Jadertina*, 2, 21–31.
- Khorsavinik, M. (2010). The representation of refugees, asylum seekers and immigrants in British newspapers: A critical discourse analysis. *Journal of Language and Politics*, 9(1), 1–28. doi: <https://doi.org/10.1075/jlp.9.1.01kho>
- Kilgarriff, A., Baisa, V. Bušta, J., Jakubiček, M., Kovář, V., Michelfeit, J., Rychlý, P. & Suchomel, V. (2014). The Sketch Engine: ten years on. *Lexicography*, 1, 7–36.
- Kilgarriff, A., Rychlý, P., Smrž, P., Tugwell, D. (2004). The sketch engine. *Proceedings of the 11th EURALEX International Congress*: 105–116.
- Killan, L.M. (1955). The Nature of Prejudice. By Gordon W. Allport. *American Journal of Sociology*, 61(3), 267–268.

- Košuta, N. (2012). O pitanju stjecanja kolokacijske kompetencije u stranome jeziku. *Strani jezici*, 41(3), 263–276.
- Lalić, V. & Đurić, S. (2015). Izbjeglička ili migrantska kriza u Evropi – bezbjednosne implikacije i izazovi. In B. Kovačević (Ed.), *Ržjici i bezbjednosne prijetnje*, (pp.53–64). Udruženje nastavnika i saradnika Univerziteta u Banjoj Luci.
- Lendaro, A. (2016). A ‘European migrant crisis?’ Some thoughts on Mediterranean Borders. *Studies in Ethnicity and Nationalism*, 16(1), 148–157.
- Ljubešić, N. & Klubička, F. (2014). {bs,hr,sr}WaC– Web corpora of Bosnian, Croatian and Serbian. In Felix Bildhauer & Roland Schäfer (Eds.), *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*, 29–35.
- Maričić I., Pezerović, A., Pezerović, S. & Radalj, J. (2014). Jesmo li spremni za promjene? Prevencija diskriminacije azilanata kroz prijedlog mjera za odvjješćivanje javnosti. *Socijalna politika I socijalni rad*, 2(1), 128–145.
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: Methods, theory and practice*. Cambridge University Press.
- Metelev, S. (2016). Migration as a threat to national security. *Indian Journal of Science and Technology*, 9(14), 1–6. <https://doi.org/10.17485/ijst/2016/v9i14/91086>
- Moehkardi, R.R.D. (2002). Grammatical and lexical English collocations: some possible problems to Indonesian learners of English. *Humaniora*, 14(1), 53–62.
- Nattinger, J.R., & DeCarrico, J.S. (1992). *Lexical Phrases and Language Teaching*. Longman.
- O'Regan, V., & Riordan, E. (2018). Comparing the representation of refugees, asylum seekers and migrants in the Irish and UK press. A corpus-based critical discourse analysis. *Journal of Language and Politics*, 17(6), 744–768. doi: <https://doi.org/10.1075/jlp.17043.ore>
- Petrović, B. (2007). Razvijanje kolokacijske kompetencije u hrvatskome kao stranom i drugom jeziku. *Strani jezici*, 36(1), 31–38.
- Podboj, M. (2019). *Diskursna konstrukcija identiteta u narativima o migrantskom iskustvu*. Dissertation. Filozofski fakultet u Zagrebu.
- Rychlý, P. (2008). A Lexicographer-Friendly Association Score. *Proc. 2nd Workshop on Recent Advances in Slavonic Natural Languages Processing, RASLAN*, 2, 6–9.
- Sakellariou, A. (2017). The representation of the refugees and the immigrants in the Greek press: A corpus linguistics and critical discourse analysis approach. In J. House, & T. Kanikliodu (Eds.), *Conference proceedings of the 1st International Conference Eurpoe in Discourse. Identity, Diversity, Borders* (pp. 397–407). Hellenic American University.
- 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(B), 73–81. doi: <https://doi.org/10.1016/j.dcm.2016.07.002>.
- Serafis, D., Greco, S., Pollaroli, C., & Jermini-Martinez Soria, C. (2020). Towards an integrated argumentative approach to multimodal critical discourse analysis: evidence from the portrayal of refugees and immigrants in Greek newspapers. *Critical Discourse Studies*, 17(5), 545–565. doi: 10.1080/17405904.2019.1701509
- Sketch Engine (2022), Retrieved 20 April 2022, from <https://www.sketchengine.eu>
- Stubbs, M. (1983). *Discourse analysis: The sociolinguistic analysis of natural language*. University of Chicago Press.
- Stubbs, M. (1996). *Text and corpus analysis: Computer-assisted studies of language and culture*. Blackwell.
- Vulević, S. M. (2018). Migrantska kriza kao izazov socijetalnoj bezbednosti u Evropskoj uniji. *Vojno delo*, 70(3), 55–74. <https://doi.org/10.5937/vojdela1803055V>
- Zlatković Winter, J. (1993). Imigracije u Hrvatskoj: skica povijesnog toka. *Migracijske teme*, 9(3-4), 303–323.