The first results of the content analysis of the media in the course of migration crisis in Hungary

Endre Sik, Bori Simonovits

Chemnitz, July 2019

CEASEVAL
RESEARCH ON THE COMMON EUROPEAN ASYLUM SYSTEM; Nr. 35

CEASEVAL received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No 770037.

Herausgeberschaft:

Prof. Birgit Glorius and Dr. Melanie Kintz
Technische Universität Chemnitz
Institut für Europäische Studien
Humangeographie mit Schwerpunkt Europäische Migrationsforschung
09107 Chemnitz

http://www.tu-chemnitz.de/phil/europastudien/geographie

Alle Rechte vorbehalten, Chemnitz 2019

ISSN 2627-339X
The first results of the content analysis of the media in the course of migration crisis in Hungary

Abstract:
After a brief and critical overview of the methodological basis of the research on the role of media in the politicization process of the migration crisis, we came to the conclusion that our knowledge is often based on analyses using rather unreliable methodology. As an illustration of a better solution we created a large dataset from online articles published in Hungary between January 2015 and April 2018, and carried out an analysis of prevalence and some aspects of the content of the politicization process. With this exercise we intended to set an example for EU authorities why to create a proper dataset to follow the mediatized form of the politicization processes would be useful. At the end of the case study, to demonstrate the essence of our approach for more methodology-sensitive Readers we briefly defined the basic terms we used, and in the Appendices we summarized the main characteristics of the research design and the dataset we constructed.

Keywords: Hungary, politicization, content analysis, migration discourse, moral panic button

Please cite as:

---

1 We are very grateful for the comments of Márton Gerő, Gábor Hajdu, Eszter Katona, Árpád Knap, Fanni Máté and Zsolt Szabolcsi.

2 The paper was written as part of Task 3 (WP5) for the CEASEVAL project. The corpus building was funded by Peripato (ELTE University) and carried out by Precognox and ELTE PostDoctorate Research Fund provided by the National Research Development and Innovation Fund (PD-121095).
Table of Content

Introduction ................................................................................................................................. 3

Chapter 1 A methodology-sensitive treatise of the literature and the research design ............. 4

1.1 The Migrant Crisis... (2017) ............................................................................................................ 7
1.2 The Special Issue of the Journal of in 2018 and Refugee Studies ................................................. 9
1.3 The basic characteristics of the research design ......................................................................... 11

Chapter 2 The politicization of migration crisis in contemporary Hungary ................................. 16

2.1 Prevalence of the relevant articles .............................................................................................. 16
2.2 Labelling ....................................................................................................................................... 21
2.3 Actors ........................................................................................................................................... 26
2.4 Organizations and locations ........................................................................................................ 29
2.5 Sentiments ................................................................................................................................... 30
2.6 Some preliminary tests of the MPB ............................................................................................. 33
   2.6.1 The case of Origo .................................................................................................................. 33
   2.6.2 The pre- and post-fence experience .................................................................................... 36

Chapter 3 Summary and further possibilities ............................................................................... 39

   The politicization of the migration crisis in contemporary Hungary ................................................. 41
   Further possibilities of using the dataset .......................................................................................... 43

References ..................................................................................................................................... 46

Glossary ..................................................................................................................................... 50

Appendix 1 The steps of building the dataset (András Hering and Endre Sik) ............................ 54

   A1.1 The sources of the research ...................................................................................................... 54
   A1.2 The selection of articles ............................................................................................................ 57
   A1.3 Variables .................................................................................................................................... 61
   A1.4 Data output ............................................................................................................................... 63

Appendix 2 Milestones and basic statistics of the migration crisis in Hungary in 2015 and 2016 .... 64

Appendix 3 Pictures of the subsequent pressings of the MPB ...................................................... 71

Appendix 4 The brief story of Origo ............................................................................................ 75
**Introduction**

In the proposal of CEASEVAL we argued that in order to analyze the development of migration discourses and politicization we should analyze

“... the importance attributed by the media to the arrival of refugees and EU and Member States’ responses (prevalence), the object or objects of politicization (issues discussed) and the actors involved in the media discourse (expansion of actors and audiences). CEASEVAL will establish a common research design and methodology in order to obtain data, guide the analysis at the national level and facilitate cross-national comparison.” (Proposal, 2017, p.66)

Later on when defining the role of the various teams as part of Task 3 in WP5 we decided that Tárki will lead a research team to

“... establish a common research design and methodology in order to obtain data, guide the analysis at the national level and facilitate cross-national comparison.” (ibid, p. 66)

The first part of these efforts were used in preparing the quantitative analysis of salience of the relocation quota discourse in the national reports (Patterns..., 2018) and in the comparative analysis (Pasetti and Garcés-Mascareñas, 2018). In this paper we deepen our quantitative analysis of the migration crisis (see Glossary) focusing on a single country (Hungary) and by developing a special dataset to illustrate the potential of such approach for future research and policy making.

In Chapter 1 we offer a brief and critical overview of the methodological basis of the research on the role of media in the politicization process of the migration crisis, to point out that (1) our knowledge is often based on rather shaky empirical ground, and (2) what were the basic methodological ideas we used to build our research design to minimalize these shortcomings. Chapter 2 contains a case study of Hungary between January 2015 and April 2018 about the prevalence and some aspects of the content of the politicization process. This and the following working papers are intended to serve for EU authorities as a methodological experiment to outline the necessary (though insufficient) basis of an all-European discourse monitoring system. The development and regular updating of such a large scale and properly designed corpus could be the basis (1) to analyze the discursive trends of the media (both prevalence- and content-wise), (2) prepare simple (descriptive) comparative analyses of the media content on more focused topics, and (3) create a corpus which can be used to make valid and reliable in-depth discourse analyses to uncover the hidden content of the media. Chapter 3 concludes and outlines further possible uses of such a dataset. To demonstrate the essence of our approach for more methodology-sensitive Readers, in the Glossary we briefly define the basic terms we used, and in the Appendices we summarize the main characteristics of the research design and the dataset we constructed.

We assume that with the help of our case study we can illustrate the usefulness of developing a reliable and large dataset and applying content analysis. We think this is important if we accept that (1) the media has a strong impetus on public opinion, (2) and this will even more be the case in the future when (3) issues which are morally sensitive (such as migration is, a topic in which case the simplest news has the tendency of becoming heavily loaded emotionally) will be more and more the target of fake news and other forms of “alternative reality”. And (4) it will be especially the case if governments start to use organized forms of propaganda to manipulate “the man of the street”, using migration and refugee issues as a source of scapegoating, hatemongering and generating moral panic (see Glossary) for political advantages, especially (5) in its most developed form (the moral panic button, MPB, see
Glossary). Finally, (6) this database and analysis would help the Commission to understand the triangle of actual migration discourses, voting behavior in the Council and in the European Court of Justice concerning asylum and migration issues of the Member States.

However, before we start our analysis it is unavoidable to explain why we use the term “migration crisis” throughout the paper instead of other options and most of all “migration crisis”. The core of the problem is that whichever term we decide to use we immediately enter into the labelling debate, as our use “refugee” instead of “migrant” (or vice versa) will be interpreted by many readers as making a political statement.

Writing the proposal of CEASEVAL we used both terms, and while the title of the project refers implicitly (asylum seekers) to a refugee system, the first sentence on CEASEVAL home page reads as follows:

“Since 2015, migration to and within Europe has challenged the adequacy of the Common European Asylum System (CEAS).”

As the title of the paper indicates, in this paper we opted for using the term migration crisis. We would therefore like to stress that our use of the term “migration crisis” should not be seen as a political statement (i.e. we do not mean to imply that those who came were NOT refugees). We are using both the terms “migrant” and “migration” in a purely technical, neutral sense.

Chapter 1 A methodology-sensitive treatise of the literature and the research design

In a concise meta-analysis of the state-of-art of the role of media in the politicization process of the European migration crisis in and about 2015 (State..., 2018), the author not only gives a useful overview in regard with the role of media’s framing (see Glossary) of immigration (often using securitization or threat frames), and often causing public anxieties but also concludes that historical legacies and attitude towards the EU are the key drivers to account for cross-national variation of these alternative discourses. We also learn that the

“... media does not operate in isolation to the wider politics of the issue, and the majority of research has found that political discourse and media discourse are highly interlinked and presenting a “causality dilemma” between the two spheres...” (p. 11)

The paper refers to a comparative analysis in eight countries prepared for the Council for Europe in 2017 and summarizes the main findings as follows:

---

3 And avoiding to use the vague term „mixed flow” designed by IOM (IOM (2009), IOM (2015).
4 While outlining our research design (part B) we used the terms refugee 218 times (and as a procedural subcategory of it, asylum seeker 64 times) and migrant 170 times, i.e. we mixed the two terms but tended to favour the former (49+14% versus 37%).
5 Of course whether the terms “migrant” and “migration” are indeed neutral terms depends on the context. In case of Hungary, for example, these terms are NOT neutral but have negative connotations of “being foreign” and are often used in contrast to “refugee” which has positive connotations (i.e. people who need and deserve charity and help).
Firstly, the European press played a central role in framing refugee arrivals in 2015 as a crisis for Europe – new arrivals were seen as outsiders and different from Europeans. Secondly, the authors found significant differences across regions – there was a stark contrast between media coverage on the West and the East and especially between media in receiving and non-receiving countries. Thirdly, ... found significant temporal trends, with sympathetic and empathetic frames in the early stages of the crisis gradually being replaced with suspicion and in some cases hostility towards refugees. Fourthly, the authors found systematic and persistent hate speech and hostility towards refugees, especially in Eastern Europe (and Hungary specifically). Fifthly, refugees and migrants had limited agency and no voice in the press ... Sixthly, the media played little attention to the context of the refugee plight – there were limited connections between stories on new arrivals and war reporting for example.” (p.15)

Another exercise compared the framing techniques of the media in five countries in (Spain, Italy, Germany, the UK and Sweden) in 2014 and 2015 (Berry et al, 2015) The main findings were:

“The report found major differences in terms of the sources journalists used (domestic politicians, foreign politicians, citizens, or NGOs), the language they employed, the reasons they gave for the rise in refugee flows, and the solutions they suggested. Germany and Sweden, for example, overwhelmingly used the terms ‘refugee’ or ‘asylum seeker’, while Italy and the UK press preferred the word ‘migrant’. In Spain, the dominant term was ‘immigrant’. ... there were wide differences between the framing and themes. For example, humanitarian themes were more common in Italian coverage than in British, German or Spanish press. Threat themes (such as to the welfare system, or cultural threats) were the most prevalent in Italy, Spain and Britain. ... the Swedish press was the most positive towards refugees and migrants in contrast to the UK which was the most negative and polarized. ... Britain’s right-wing media was uniquely aggressively in its campaigns against refugees and migrants’. ... governing parties or coalitions tended to dominate political sourcing ... the key challenger or challengers coming from the anti-immigration right’.” (p.16)

In the concluding analysis of the Special Issue of the Journal of Immigrant and Refugee Studies (Triandafyllidou, 2018), the author introduced two competing frames: the moralizing frame and the threat frame. The moralizing frame placed the responsibility of the refugee flows on wars, conflict and violence in the countries of origin and presents refugees as victims. The threat frame perceived the movement of people as a natural disaster. According to the author, these discourses have to be understood in relation to: 1) the positioning of each country as a “frontline or final destination”; 2) past experiences of seeking or offering refuge and hosting migrants (or lack thereof); and 3) local politics.

Finally, in a book on national discourses on the refugee crisis in 2015 (The Migration ..., 2017) thirty-one authors gave an overview of the main characteristics of the discursive processes and the role of media in eighteen European countries. In the introductory chapter the editors wrote that

„distinctive European perspective on the migration crisis does not exist but rather is, parallel to the diverse range of European countries with different cultural and political traditions, characterized by multifaceted national views.” (Ritter et al, 2017, page 14)
If we accept the claims of Czymara and Dochow (2018)⁶, that

“...comparing periods of vivid discussions with times where the issue was hardly discussed in the German media ... the predicted probability of being very concerned increased by about 13 percentage points. Deeper investigations reveal that the media effect is most potent for individuals living in areas with lower share of ethnic minorities and for those with lower education or conservative ideology...”

(abstract)

than the conclusions of the previously mentioned comparative analyses should be treated as the solid basis of hypothesis to analyze the role of media in the politicization process of the migration crisis in contemporary Europe.

The problem, however, is that often the aforementioned analyses do not have solid methodological background enough to serve as the basis of comparative conclusions. To illustrate the validity of such criticism we start with the excellent literature review of the European media research (Eberl et al, 2018) based on the meta-analysis of 89 articles. The authors summarized the characteristics of these papers from the point of view prevalence as follows:

“...Quantitative content analyses gather textual data mainly via available online archives ... To identify relevant stories, most studies rely on extensive Boolean search strings referring, for instance, to events or actors associated with the discourse on immigration ... Most commonly, prevalence is measured based on the absolute number or relative share of news stories referring to immigration or to the presence of certain migrant groups ... online media, especially social media sites including user-generated content, are largely neglected ... there is little comparative research on the prevalence of immigration-related issues or actors in the media across different European countries... research often focuses on the terms of ‘(im)migrants’ (e.g. ‘asylum seekers’) ... there is only a low number of studies analyzing different migration groups comparatively or using longitudinal approaches. (p. 3-4)

The authors’ conclusion from methodological point of view is that

“...the bulk of research in this field has focused on traditional news media while news ... through social media has largely stayed on the side-lines. ... This literature review also showed that this field is restricted to a handful of European countries, oftentimes with similar media systems. Research thus far does not allow for a proper mapping of a European discourse on immigration in Europe. (p. 12-13)

While we agree with the critical remarks of the previous analysis, we missed those critical points which would point out the limitations the media research as far as sampling, representativity, i.e. „standard” statistical elements of a quantitative analysis would require to be able to make to reliable and valid generalizations.

For example the previously mentioned and otherwise elaborated analysis of the media in five countries (Berry et al, 2015) followed the „standard path” of media research, i.e. covering only some artificially selected newspapers from quality and popular press, and the samples were also rather limited (300 news per country) to cover a whole year⁷. The multiple sampling and the manual coding procedures were very careful (see in detail the descriptions on pages 25-28) as well as their efforts to adjust the

---

⁶ Their analysis is based on a longitudinal panel data (N=25,000) of persons who were interviewed twice between 2001 and 2015 in Germany.
⁷ Especially in light of their own decision to cover a single story for one week by 100 articles per country (p 27).
search strings to the local needs, i.e. adding special terms to the translated general ones (see the Spanish example in footnote 13).8 The research ignored electronic media9 and websites, local media, extreme sources, i.e. those forms of the media which reaches the largest audience.10

Another comparative analysis of media representation of the migration crisis in eight countries (Chouliaraki et al, 2017) used the same standard methodological solutions, i.e. focusing on a small and artificially selected sample of sources (two (in the U.K. four) „key newspapers” one from the „left”, one from the „right”), and in each newspaper 60 articles (20 articles focusing on the three pre-defined major events)11 were selected during the ten working days (?) following the event12.

To illustrate the merits and the limits of creating proper methodological tools to analyze the politicization processes of the migration crisis, we focus on two – earlier already mentioned - exercises which claim comparativity in European level: a book (The Migration… 2017) and a Special Issue of the Journal of Immigration and Refugee Studies. We selected these two cases partly because they have very different ideas concerning the methodology needed to analyze the processes of politicization, partly because they have the largest European coverage.

The ultimate limit of these two – otherwise excellent - selections, however, is that the sophisticated linguistic discursive analytical techniques, including the most often used discursive historical analysis, can provide very high level of internal validity (see in Glossary) but hardly any external one. In other words, we get a very detailed description of the framing processes of the media but we do not know whether these cases were the exceptions or the typical techniques, i.e. we do not know how to generalize our findings.

1.1 The Migrant Crisis... (2017)

In this book (The Migrant …, 2017) the eighteen national case studies focused on the same issue and used more or less identical time frame, i.e. the period of the heights of the migration crisis in 2015. The authors of were encouraged to use any methodological techniques to prepare their case studies. As far as the empirical design of the analysis was concerned, the result (including the Hungarian one) was rather disappointing: mostly excerpts from a few articles, political speeches mixed with basic statistics and results from public opinion polls. In cases there were more serious attempts to analyze the prevalence and content of the media on a representative sample, the data used and the methods applied were mostly simple ones. We identified seven case studies in which the author(s) made efforts to do more systematic analysis of the media discourse and provided the Reader with more or less proper documentation13:

8 Their efforts to come up with an approximately identical number of articles in all countries, is however, debatable, especially since there is no information concerning the technology of „down-sampling” of the German articles (p. 27).
9 The additional small sample of the two main evening television programs in the U.K. obviously does not mend this neglect.
10 Just as was done earlier for Fundamental Rights Agency (2012).
11 Erecting the wall on the Hungarian-Serbian border (13 July), the drowning of Alan Kurdi (3 September), and the terror attack in Paris (13 November).
12 This pattern could be flexibly changed if the selection did not reach the expected 20 cases (see footnote 5).
13 The German case study (Lichtenstein et al, 2017) used qualitative content analysis of the five most prominent television political talk shows. They identified four framing elements: 1) problem definition; 2) the
The **Albanian** case study (Pasha-Niemetz, 2017) took internal and external validity (see in Glossary) seriously by analyzing 327 articles, selected by one keyword (refugee) in three online sources (though in fact 74% came from one source which has the largest audience). Developing the coding system, they used the concept of grounded theory, i.e. starting with some basic codes and after several iterations built more and more complex coding trees.

The **Greek** case study (Kaitatzi-Whitlock – Kenterelidou, 2017) based its analysis on a larger size (N=994) of articles during three locally define „crisis periods” from various print, online and television sources. Since, however, the selection process and keyword(s?) were not described properly, it is difficult to evaluate their research from a methodological point of view.

The sample of the **Serbian** case study (Jevtovic-Aracki, 2017) was even bigger (N=3583) containing the cover pages of the seven printed newspapers between 1 January 2015 and 30 June 2016 with the terms refugee, asylum seekers, migration crisis, migrants appearing on the front page. In the course of the analysis they covered not only the title but the subtitles, abstracts pictures and cartoons as well.

The **Spanish** case study (Perez, 2017) covered the period between September 2015 (the peak of the Syrian migration crisis) and April 2016 (the visit of Pope Francis to the refugee camps in Greece) They selected articles (of any genre) using the keyword ‘refugees’ (refugiados) from a database (Proquest Newstand) which files a plain text copy of the articles published by El Mundo and El País. But they considered the dataset too big and selected only 50 articles from each newspaper randomly for the analysis. They coded some characteristics of the articles as well such as dateline (Paris, London, etc.), their section (local, national or international news, etc.), genre (news, background analysis, editorials, opinion columns, etc.).

The **Swiss** case study (Bonfadelli, 2017) used as the basis of the analysis previous researches with the coverage of immigration issues in radio and television programs from 2007. The data covered the Spring 2015, and the selection of articles was restricted to news and political program if related to Switzerland. The broadcast data was complemented by a quantitative analysis of press articles dealing with the refugee issue of the three most important newspapers in the German part of Switzerland.

The case study of the **U.K.** (Griessler, 2017) based its analysis on an excellent report that covered the national print editions of all national UK publications (171,401 migration related articles from 19 newspapers) from January 2006 to May 2015 (Allen, 2016). The report uses three techniques: a frequency analysis of migration-related words and articles, a collocational analysis (see in Glossary) to analyze how strongly one word is linked with a target word, and content analysis where certain narratives (in this case EU/European and illegal immigration) were selected manually. Despite having such an excellent benchmark report, the paper in the book focuses only two newspapers and analyze only, and rather superficially, migration articles published in August and September 2015.

---

structural cause of the problem; 3) the attribution of blame; and 4) the proposed solution in coding the discussions.
1.2 The Special Issue of the Journal of in 2018 and Refugee Studies

One of the editors summarized the basic characteristics of the research designs in a Table (Triandafyllidou, 2018 p. 208) In the Table the author compiled the general features of the ten case studies which shows a high variety of foci (the widest is the German case study which covers the refugee flow related discourse from Spring 2015 to early 2016, the most focused is the Serbian/Croatian one covering one event for one month (September 2015)).

From methodological point of view, this collection of case studies was different from the previous one in two respects: (1) effort was made to use more or less similar analytical techniques, and (2) the data contained not only online news but also special media (tweets and blogs). But while some case studies were aiming to produce results with higher external validity (Greek, Austrian), most of the other papers maximized the internal validity but focused only on very special segments of the media (Polish, Swedish, Serbian/Croatian, UK, German).

To introduce the Reader with some of the methodological solutions (both excellent and questionable) the author(s) applied, we briefly summarize the most relevant features of the empirical analysis of the case studies.

In the Greek case study, the authors based their large-scale quantitative content analysis on the results of a small-scale pilot study. The quantitative analysis focused on the „politicians’ tweets and the published articles on the refugee crisis in three Greek newspapers, the populist right-wing Proto Thema, the highbrow conservative Kathimerini, and the left-wing Efimerida ton Syntakton (all national newspapers). ... Our newspapers’ sample comprises articles that are available on the online version ... and contain the key words “refugee crisis,” “refugees,” “migration crisis,” “migrants,” “Europe,” “Greece,” “Schengen.” (Boukala and Dimitrakopoulou, 2018 pp. 185)

As to the selection of articles, they included articles (of all genres, i.e. editorials, columns, reports) within the category Politics. While the number of selected sources was small (and artificial in the sense that it did not take into consideration the size of their audience), they set the starting and ending dates of their analysis very narrowly (from January to March 2016) as well. Still their database that was large (N=2,991 articles). They did content analysis using NVivo10 using the concepts that were identified in the pilot study.

In a separate step they carried out critical discourse analysis (using the Discourse Historical Approach – see in Glossary) on a database containing all official tweets of two politicians: the Greek prime minister, Alexis Tsipras and the president of the main opposition party, Kyriakos Mitsotakis. The time span of this analysis was the same as the previous. Their dataset contains 283 and 290 tweets, respectively.

The Austrian case study (Rheindorf and Wodak, 2018) used a very sophisticated method to analyze to changing discourse in regard with border and border management. They conceptualized the linguistic term “debate” to make it more appropriate for their analysis for several reasons first of all having defined beginning and ending of the discourse. They defined the new concept the so called “discourse strand” (see in Glossary) and analyzed the discourse on refugees in Austrian media throughout 2015. Their data contained a large corpus (N=6,701 texts) compiled from fifteen newspapers/magazines with wide circulation (?). The ultimate focus of their research were two discourse strands: building fence (“Grenzzaun”) and setting a maximum limit for asylum seekers (“Obergrenze”). These two
(overlapping) subcorpora contained 1,697 texts about building a border fence and 372 texts for setting a maximum limit (between 4 April, 2015 and 24 February 24, 2016).

The Polish qualitative case study focused on the discourse of right-wing populism. In the course of the research the author started with an “entry-level” (topic, theme, and content oriented) followed by an in-depth (uncovering the strategies, especially to that of argumentations) analysis. The author lucidly describes the “technology” he applied:

“At first, the analysis of topics encompasses close reading of each of the texts forming a respective discourse (e.g., speeches, tweets). The reading results in the production of a general list of themes for each of the texts and, in a cumulative manner, for the entire discourse, thus forming an overall map of its contents. The thematic analysis eventually also serves as the basis for later, argumentation-oriented (i.e., topos) analysis. There, the aim is to identify the key argumentative frames in the analyzed discourse as those that structured key arguments and expressions. The analysis, thereby, points to how the key themes (topics) were deployed in formation of different arguments…” (Krzyżanowski, 2018a, pp. 82-84)

The period the case study covers is only thirty days in 2015 (between 15 September and 15 October 15). The data contains the parliamentary speech of J. Kaczynski on the migration crisis and immigration in the Polish Sejm, the website summary of the speech on the PiS website, the tweets using PiS’s official Twitter profile, some additional web material (?), and statements on the issue made by J. Kaczynski at a pre-election rally.

The U.K. analysis had a similar structure to that of the Polish. The data covered three periods in 2015. In each period the author selected some “key events” as the foci of the discourse historical approach. The primary-level texts (tweets) were retrieved for five political parties/politicians. These tweets were coded for relevance. The final number of cases was 180 tweets. The author applied a version of the discourse historical approach (see in Glossary)

The focus of the Serbian and Croatian case was the reactions of public intellectuals (?) to the escalation of the migration crisis in September 2015, i.e. when Hungary completed the fence on the Serbian border, and thus diverting the flow of refugees into Croatia. The author “apologizes” and convincingly argues why he decided to build his data from an “outdated” type of source.

“In order to capture Croatian and Serbian intellectuals’ viewpoints on the significance and implications of wall building as a bordering practice in the context of the refugee crisis, I shall focus on a specific genre: the opinion piece. The choice of focusing on this more “traditional” genre rather than on digital discourse is motivated by the notion,... that the digitally mediated public sphere, despite its abundance of voices, discussion, and access, is often devoid of the familiar characteristics of intellectual and expert discourse, ... Therefore, although their position of supremacy in orienting public opinion might have been partly undermined by talk shows, blogs, and social media at large, editorials and opinion pieces, I maintain, remain a privileged site wherein certain representations (or truths) about society are articulated, promoted, and legitimized.” (Sicurella, 2018 p. 62)

---

14 Two from the Conservative Party (the then prime minister, David Cameron and the official account of the Prime Minister’s Office; two from the Labour Party (Jeremy Corbyn and the official account of the Labour Party; and Nigel Farage, the then leader of UKIP and a Member of the European Parliament (MEP).
The data consisted of a sample of political genres such as editorials, op-eds, commentaries, and interviews, which were published from late August through September 2015. To ensure the balance (sic!) and representativeness of the sample, the author selected fourteen sources “with different cultural-political leanings and which appeal to various audiences.” (ibid p. 63). From the selected 115 opinion pieces only 50 were finally selected (after going through again some sort of “balancing”).

The very focused Swedish qualitative analysis used material collected between 1 November and 15 December 2015, when (1) the number of refugees arriving in Sweden rose sharply, resulting some change of refugee and asylum policy, and (2) there were terrorist attacks in Paris (Nov 13, 2015). The analysis focused

“.. on the SAP (Social Democratic Party) Twitter discourse and examined, at first by means of topic-oriented analysis, the overall development of the key themes and frames of the SAP’s mediated political discourse in the studied period of 6 weeks. On the other hand, a closer look is also taken on SAP social media discourse of one selected day, Nov 24, 2015, when the party announced the tightening of asylum-seeker regulations. Within that second step of analysis, the focus is on how policy and related political action is discursively legitimized by the SAP as a political actor behind the policy changes. The focus is on how social phenomena—in our case those related to immigration and specifically the “refugee crisis”—are constructed and represented linguistically, in this case in political discourse. This allows seeing discourse as a carrier of different forms of legitimation of discursively constructed forms of political identity and agency.” (Krzyżanowski, 2018 b p 105)

The German analysis focused on “leading” German newspapers and magazines (Leitmedien), i.e. opinion leaders, “agenda (or trend) setters” for other media and for the political elite. The time span of their analysis was a whole year (between March 2015 and 2016). They were

“… applying critical discourse analysis (CDA) … However, considering the growing significance of images and visuals in ways of communication and the media...., we find it important not to limit our analysis to content, expressed in textual form but to include the visual dimension...We used a sample of visualizations, which had an impact on the public discourse. As we will show in the exemplary case of the image of the dead body of Aylan Kurdi, images have the capacity to shift the focus in discourses and possibly lead to a shift of framing.” (Vollmer and Karakayali, 2018, pp 119-120)

1.3 The basic characteristics of the research design

Since our aim was to reach the highest possible level of external validity (while retaining the possible highest level of internal validity as well) the first priority was to define the limits of our sampling design and within it to maximize coverage. As to the limits of our targeting strategy, we sampled only online sources and ignored (1) the printed editions of the same media and the weekly or monthly magazines, (2) the TV and the radio, (3) the social media, and (4) all forms of the local media.

In the following we briefly summarize the main characteristics of the selected sources of our analysis. We selected ten sources to represent the Hungarian online media taking into consideration three dimensions:

(1) the type of source (online version of printed newspaper versus online site),
(2) the size of the online audience, and
(3) their relation to the government (governmental and non-governmental) as well as (among the non-governmental sources) their political attitude (right-wing/radical versus moderate/left-wing).
Figure 1 shows the estimated size of audience of online media in 2018 (Bognár, 2018), five of our selected sources are among those with the largest audience (red bars).

1. **Figure** The estimated size of the audience of the online media, multiple response (%)\(^{15}\)

![Bar Chart: Estimated Size of Audience of Online Media](image)

Source: Bognár, 2018

As to their relation to the government and political attitude, we have to explain briefly (1) why do we consider the first dimension as a high priority one, and (2) why do we classify only the non-governmental media according their political attitude.

The reason why the distinction between governmental and non-governmental source has the highest priority is that over the past years, the government reached hegemonic position in the media through a wide range of means such as: (1) initiation restrictive media laws, (2) restructuring the public service media to make sure that it acts as a government mouthpiece, (3) through buy-outs and takeovers created a situation whereby most of the mainstream media are in the hands of government allies.

---

\(^{15}\) The proportion of the population used the source last week
(Mertek, 2017), (4) with the strategic placing of state advertising made sure that the pro-government media are flourishing, while media that are critical of the government struggle to survive.

As to the relevance of the political attitudes of non-governmental sources, though they differ from each other along a traditional right/left dimension but their attitude towards the discourse on migration crisis could be strongly influenced by other consideration such as (1) the composition of their potential voters (e.g. the strong xenophobic attitude of Jobbik voters is likely to have an impact on the operation of alfahir.hu), (2) their marketing strategy (e.g. atv.hu is the online addition of a television channel), etc.

The five media with large audience consists unsurprisingly (they have to reach large audience to maximize their profit) four non-governmental media and one governmental source (origo.hu) which used to be also for-profit before its nationalization (see Chapter 2.6.1 and Appendix 4). The other five sources consist of three right-wing and two leftist online media outlets with a smaller but very active audience (alfahir.hu, pestisracok.hu, magyaridok.hu and nepszava.hu, and nepszabadsag.hu). The last medium, however, does not exist anymore but used to have the largest audience during most of the time of our research (until 8 October 2016 when it was forced to go out of business).

Figure 2 shows the typology of the sources we will use as a standard in the course of analysis. On the left side there are the three governmental sources (origo.hu, magyaridok.hu and pestisracok.hu) and one right-wing non-governmental site (alfahir.hu). On the right side the non-governmental sources (the first four of them having high ranking value on the Reuters’ scale (Bognár, 2018), the last two the leftist ones.

The reason to construct this oversimplified dichotomous typology was that we wanted to focus on the MPB, and assumed that from this point of view governmental and right-wing non-governmental sources will show similar behavior.

---

16 Table A1.1. in Appendix 1.1 summarizes the general characteristics of the selected sources.
17 The dark grey color of Nol indicates that since this media does not exist anymore, we used its archive only.
18 Obviously in the future other typologies of the sources might be much better as will be illustrated at the end of Chapter 2.2 where we used a pairwise comparative analysis. Moreover, the size of the sample is large enough to analyze all sources separately as well.
Since we wanted to describe the main characteristics of the media discourse in Hungary during the migration crisis we had to define the starting and (if there was) the ending dates of the time-span of the migration crisis.

As to the former, we started our analysis on 1 January 2015. Not because it was the beginning of the calendar year but because according to the results of a case study (Bernáth-Messing, 2015) the Hungarian Prime Minister gave several interviews between January 7 and 11 by which time the discourse on the migration crisis had started. Moreover, we assumed that the use of this occasion as the first pressing of the MPB was a reaction on behalf the government to increase their popularity which has been rapidly decreasing since late 2014 (Sik, 2016a, Barlai-Sik, 2017).

The latter decision, i.e. how long should the research period last, has been derived from our assumption that MPB is a long-term form of manipulation, consequently, the discourse on migration crisis did not end with the disappearance of refugees. Since several qualitative studies illustrated the pressing of the MPB between 2015 and 2018, we decided to extend our research period to cover the Parliamentary election of 2018.20

Finally, within this time-span we had to define shorter periods which – according to the available qualitative observations – could be identified as separate instances of pressings of the MPB within the time-span (Table 1). The idea was to set the exact date of campaign events on the basis of identifying their (1) starting and the (2) culminating points when the MPB was initiated and pressed. The general idea was to start the scraping (see on the term Glossary, on the technique and results Appendix 1.1) of the relevant articles some time before and finish some time after the culmination point.

### Table 1. The focused observation periods

<table>
<thead>
<tr>
<th>Periods</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 January - 4 February 2015</td>
<td>Orbán's first anti-immigration</td>
</tr>
<tr>
<td>15 April - 31 November 2015</td>
<td>The peak period of the relocation</td>
</tr>
<tr>
<td>11 September - 9 October 2016</td>
<td>National consultation „Stop“</td>
</tr>
<tr>
<td>28 March - 1 June 2017</td>
<td>National consultation „Stop“</td>
</tr>
<tr>
<td>10 November - 8 December 2017</td>
<td>The „Stop Soros!“ campaign</td>
</tr>
<tr>
<td>17 January - 20 February 2018</td>
<td>Parliamentary election</td>
</tr>
<tr>
<td>201 March - 15 April 2018</td>
<td>(cca. 6 weeks)</td>
</tr>
</tbody>
</table>

19 Following the terrorist attacks against Charlie Hebdo on January 11, 2015 the Hungarian Prime Minister expressed his interpretation of the events in the following way: „We should not look at economic migration as if it had any use, because it only brings trouble and threats to European people. . . . Therefore, immigration must be stopped. . . . We will not allow it, at least as long as I am prime minister and as long as this government is in power. ("Hungary PM Orbán Says," 2015, para. 4-6) According to Dessewfy and Nagy (2016) immediately a number of political and communication tools were applied to reinforce this message. (Dessewfy-Nagy, 2016)

20 To illustrate the message of the various pressings of the MPB between 2015 and 2018 we selected representative pictures that was created by the framing experts to manipulate the population (Appendix 3): No.1 on national consultation on terrorism and immigration, No 2. on relocation quota referendum, No 3. “Stop Brussels”, No 4. “Stop Soros”, No 5a. anti-UN campaign and 5b Parliamentary election. Since then there have been several occasions in which we could again and again identify the various new versions of the previous propaganda oeuvre (e.g. anti UN migration regulation), the newest being an anti-EU campaign (See the picture No 6. in Appendix 3).
To find the articles that were likely to contain relevant information in relation with the migration crisis, we defined four keywords („menekül…” „bevándor…” „migrá…” „betelepít…””) and selected articles that contained at least one of these keywords. The keywords were constructed to identify the core part of the Hungarian equivalents of the following terms: refugee, immigrant, migrant and relocation. In the course of identifying a keyword some words were stemmed. These selections words that were needed stemming was made manually and later was validated in the entire corpus. To avoid false positive matches, words with lower than five characters were forced to match exactly, longer words were allowed to have affixes and suffixes.\(^26\) (For more detail see Appendix A1.2 and A1.3.)

\(^{21}\) See footnote 17.

\(^{22}\) This period actually covers several acute crisis periods, see details in Appendix 2.

\(^{23}\) The date when the government expressed its intention of having a referendum. The campaign itself started in July.

\(^{24}\) The questionnaires were posted since 31 March. This period overlaps with the first phase of the anti-CEU and anti-NGO campaigns.

\(^{25}\) This campaign had another peak during the next period (Parliamentary election), and continued beyond the time-span of our research until June 2018, when two new anti-NGO laws (on foreign-funded organizations and on „organizations aiding migration”) were passed by the Parliament.

\(^{26}\) Case differences were ignored.
Chapter 2 The politicization of migration crisis in contemporary Hungary

The following analysis is based on a corpus of media articles scraped between January 2015 and April 2018. In the following, we intend to introduce the Reader of the results of a descriptive analysis of prevalence and some aspects of the content of the politicization process.

2.1 Prevalence of the relevant articles

We analysed the prevalence of relevant articles in two dimensions: time-span and source, i.e. we were interested whether there are significant differences along the time-span and the among the different types of sources focusing on the migration crisis.

There are three aggregated approaches we use to analyze the trend of prevalence:

- The first level of aggregation is when we computed the weekly average of articles publish (the number of selected weeks was 65 (see Figure 3)). It does not have any theoretical relevance but smooths out the daily oscillation of the data.27

- The second variable (period, N=29, see Figure 4) splits the research period into shorter and more or less equally long periods. The allocation of the observations follows the logic of the MPB, therefore, we consider this proxy as an event-specific type of periodization.

- Finally, we split the research period into ten units (deciles) containing equal number of observations (peridec, N=10). This proxy is not sensitive to concrete events of the migration crisis but offers a general design to trend analysis.

The most detailed trend of prevalence can be seen in Figure 3 where we aggregated the data on a weekly basis.

27 The lowest daily number of articles (12/day) was on April 18 2015. The peak season (fourteen days with at least 300 articles/day) was between August 31 and September 25 2015 (with the highest value (393/day) on September 1 2015), when large masses of asylum seekers were marching through Hungary (see the daily numbers of border crossings in Appendix 2 Figure A2.2 and the event-history in Table A2.1).
3. **Figure** The weekly number of articles (N=45,875)*

*Numbers on the horizontal axis indicate the number of week between 7 January 2015 and 14 April 2018. The red lines indicate a split of the timeline (between weeks 38/39, 42/43, 51/52 and 55/56). In some cases, (due to the split of the timeline, i.e. not to aggregate days from another month and at the end of the time-span) the number of days per week slightly deviate from seven: weeks 37, 50, 59 and 65 contain 5, 9, 6 and 9 days, respectively.

The first salience of the refugee discourse occurred during the time when masses of refugees stayed temporarily in Hungary (Appendix 2 Figure A2.1) and when the Hungarian/Serbian border was sealed (and initiated violence at the border, see Appendix 2 Table A2.1, Figure A2.2). The second salience occurred almost a year later when there was hardly any refugee invasion anymore (see the last row of Table A2.2 in Appendix 2) but the relocation quota debate was still raging (Table 1).

As to the event-specific prevalence of the articles (Figure 4), the trend is unsurprisingly the same as was the weekly distribution (Figure 3). The first conclusion, however, should be not the event-specific increases but the fact that while the number of refugees became minimal after October 2015, the number of refugee related articles almost never has fallen below a thousand cases per period, and between the sharp increasing period there were long period of stability of the presence of the discourse.
4. Figure The number of articles by event-specific periods (N=45875) *

*The red lines indicate splits of the timeline

We identified the following periods in which, compared to the preceding and following periods, there were smaller or bigger (milder or steeper) saliences:

in 2015

- January: The more than 1000 articles is likely associated with the first pressing of the MPB (the post-Hebdo Orbán speech), see also the Precognox data in Figure 5.
- May: The time of national consultation on terrorism and immigration.
- second part of June: The on-going anti-immigration “information” campaign.
- between early August to early October: The “Hot Autumn” of the migration crisis.

in 2016 from mid-September to early October: The campaign in relation to the quota referendum.

in 2017 from mid-November to early December: The national consultation of the “Soros Plan”.

This result is validated by Balogh et al (2016), who used slightly different keywords (migrant, immigrant, refugee) and covered a partially overlapping period (October 1 2014 to June 11 2016, N=42845). The data (Figure 5) shows that „migration crisis” related articles were almost nonexistent before 2015 – despite the fact that in late 2014 the number of refugees has started to increase (Appendix 2 Figure A2.1).28 The first, though relatively small raise of the number of articles occurred after the first pressing of the MPB (January 2015). The salience of mass migration related topic occurred at the same time as in our model (between the end of August to mid-September 2015). After October 2015 the number of articles has dropped somewhat but remained constantly on the level of just preceding the peak period and started to raise again in June 2016 as a result of the first stage of campaign (see footnote 23) of the relocation quota referendum.

28 This was due to the temporary salience of illegal migration from Kosovo to Austria and Germany (Appendix 2, Table A2.1 and Figure A2.3).
5. **Figure** Weekly frequency of migration-related articles between 1 October 2014 and 8 June 2016

**Notes:** x-axis indicates the first day of the week. Key events: (1) 7 June 2015: Fence along the Serbian border announced. (2) 27 August 2015: Dead migrants found in a truck; 29 August 2015: Preliminary fence at the Serbian border ready; 4 September 2015: Migrants start to walk to Vienna; 16 September 2015: Refugees clash with police at Rőszke. (3) 7 June 2016: Parliament votes on terror law; 13 June 2016: Parliament votes on new refugee law.

**Source:** Bognár et al (2019) on the basis of Precognox dashboard, number of weeks = 88; weekly average (red line) = 486.

The salience of the migration crisis related articles in early August-September 2015 is validated also by the cumulative trend in Figure 6 which shows an inflection point of the prevalence of the terms refugee and immigrant in this period (Bodor-Fokasz, no date).29

---

29 This database consists of 2236 articles related to refugee/migrant topics published in the electronic version of two Hungarian dailies: Magyar Nemzet (MNO) and Népszabadság (NOL).
6. Figure The cumulative trend of the terms refugee and immigrant in 2015

Source: Bodor-Fokasz, no date

The analysis of two TV channels (the governmental MTV1 and the non-governmental RTLK) showed (Szabolcsi 2017) that the topic of migration was high in 2015/2016 but only on MTV1.

„Only some special events, such as the Euro Football Cup or the Oscar Gala, interrupted the lengthy coverage of migration. These events grabbed public awareness and dominated the news agenda in early February, June and mid-August 2016. The length of migration-related news was significantly higher on November 17, 2015, July 25 and October 3, 2016. ... In each case an exceptional event, such as terrorist attacks or the quota referendum increased news coverage... RTLK news, on the other hand, seems to show the inverse of the tendency in the presentation of migration-related topics. Presentations on migration are rather short, unless news on legislative changes, international events, war or terrorist attacks increase the length of presentations. ... In 2016 the length of these topics significantly dropped or, in many instances, completely disappeared. By the end of the examined period in August 2016, migration-related topics also increased in RTLK news due to the upcoming quota referendum. However, the length of these news blocks was still significantly lower than in MTV1 news. (Szabolcsi 2017 p 86)

To sum up, migration crisis related news have the largest prevalence in the news oriented world where non-governmental media dominates. In Hungary, however, due to the repeated pressings of the MPB the prevalence of the migration issues remained high. In other words, while the high level of prevalence during the peak of the migration crisis of 2015 can be considered as the normal reaction of a news-oriented media, the high level of prevalence (especially in governmental sources and with higher-than-average level of negative sentiment) without any migration-related event occurring (except those artificially created by the government), is a sign of the operation of the MPB.
2.2 Labelling

The first – though (since the propaganda may intentionally conceal its underlying message) often unrecognizable – component of every framing effort is the term the article uses to label the people crossing the border. As Figure 7 shows the dominant term was the *menekült* (refugee), followed by *migráns* (migrant) and *bevándorlás* (immigration).

7. Figure The prevalence of the labels (% in the articles, N=45875)

If, however, we aggregate the labels related to the term refugee and those of (im)migration, we find that – if we assume that the overlapping among the items do not differ significantly - the aggregation of (im)migration is larger than that of refugee.

If we focus on the general trend of the labels, the lifecycle of the most often used five labels shows significant changes (Figure 8). The term *refugee* dominated the scene in 2015 and 2016, it was losing strength by 2018. The most salient period of this term was the culmination of the migration crisis in September 2015.

---

30 Refugee system (menekültügyi rendszer), asylum seeker (menedékkérő), refugee crisis (menekültválság) and immigrant (bevándorló), migration (migráció).

31 Since due to the size of the sample almost all statistical analysis was significant on level p=0.0000, we omit to refer to them.
The various versions of migration have different lifecycles:

- the prevalence of migrant increases fast in the beginning but stay unchanged until 2018 when it becomes more prevalent,
- the terms immigration and immigrant have U-shape trend and in 2017 their prevalence start increasing
- the term migration shows a steady increasing trend during the entire period.

The aggregated result of these changes was that by the end of the period the prevalence of all four terms of migration exceeded the prevalence of the term refugee.

If we shift our focus from the general toward a more event-specific trend, we get a somewhat deeper understanding of the association between event and labelling (Figure 9).

Before reaching its peak, the term refugee lost its strength twice, first between May and early June 2015 (i.e. during the first national consultation), secondly in July. Later on there was a drop after the
peak in last two weeks of September 2015 (when the border between Hungary and Serbia was closed). The major decrease of the strength of the term *refugee* was in mid-September 2016 and continued until early 2017 (the quota referendum and its aftermath). Finally - in the beginning of 2018 (during the anti-UN, anti-CEU campaigns).

The first loss of the term *refugee* in 2015 occurred with a sharp increase of the prevalence of the terms *immigration* and *immigrant* simultaneously. When the second loss of the term *refugee* (in July 2015) has occurred, the migration term has decreased as well. While in 2016 and 2017 *immigration*, *immigrant* and *migration* started to replace the term *refugee*, the term *migrant* has increased significantly.

The fact that by the end of the period various forms of *migration* replaced *refugee* can be interpreted as the result of the operation of the MPB, i.e. while there were masses of refugees (and migrants and asylum seekers) marching through Hungary the media used the term which seemed to fit best to label the mixed population of foreigners. After this short period, however, the term migration became more popular in the increasingly government dominated media.

According to the MPB approach we assumed that compared to the non-governmental sources the state controlled ones (see Figure 2) are more likely to use the term migration than the term refugee.

**10. Figure** The prevalence of the most often used labels by sources (%)

![Graph showing the prevalence of the most often used labels by sources](image)

Comparing the two groups of sources by the prevalence of using the various labels (Figure 10), we concluded that the term *refugee* was indeed significantly less used in governmental than in non-governmental sources, and the opposite was true to all four versions of the term *migrant*. Three governmental sources (pestisracok.hu, magyaridok.hu and alfahir.hu) used the terms *migrant* as often as the term *refugee*, and in the latter the term immigration was also relatively frequently used. Among the non-governmental sources, index.hu, 24.hu and 444.hu were less, atv.hu, népszava.hu and nol.hu were a little bit more likely to use the term *migrant*.

To analyze the use of the two most often use labels by sources during the time-span of the research, we selected pairs of sources relatively similar as far as their type (online version of printed newspaper
versus online site) and the size of their audience were concerned but differed in their relation to the government and political attitudes.

First we compared a governmental and a non-govermental online website having large audience (Figure 11).

11. Figure The prevalence of the labels refugee and migrant in index.hu and origo.hu by period deciles (%)

In case of the non-governmental source (index.hu) a bell curve fits to the refugee and (following a sharp rise after the first period) there is not much change of using the term migrant. The prevalence of the former always exceeds the latter. As to origo.hu (which became governmental source sometime in 2016 (see Appendix 3)) the general trend of using the term refugee is a bell curve as well but with a fast drop in 2016, and the term migrant is characterized by a rising trend which became fast in 2017.

Comparing the online version of a governmental (magyaridok.hu) and a non-governmental (nepszava.hu) daily (Figure 12), we find that in case of the governmental source the decreasing trend of using the term refugee and the opposite trend of using the term migrant explains why by the mid-period deciles the latter reached the prevalence of the former. In 2018 the opposite trends became even sharper resulting an increasing gap using the migrant more frequently than the term refugee.
12. **Figure** The prevalence of the labels *refugee* and *migrant* in nepszava.hu and magyaridok.hu by period deciles (%)

Comparing three non-governmental sources with liberal (444.hu), pro-governmental (pestisracok.hu) and radical right-wing (alfahir.hu) attitudes (Figure 13), we find more similarities than differences. As to 444.hu, there is a combination of a bell curve (*refugee*) and an increase (*migrant*) trend which resulted equal frequency of the two terms by the end of the time-span (while using the term migrant remained below the level of the two right-wing sources).

13. **Figure** The prevalence of the labels *refugee* and *migrant* in 444.hu, pestisracok.hu and alfahir.hu by period deciles (%)

The pro-govermental and radical right-wing sources have more or less identical use of the term *refugee* (though while pestisracok.hu follows the familiar bell curve, alfahir.hu shows a decreasing linear trend), and the prevalence of using the term *migrant* is completely similar in the two sources.

Finally, we compared the prevalence of using the term *quota* (Figure 14) in case of all sources. The result is a strikingly similar trend in almost all cases. There is a mild increase in the fourth decile which disappears by the sixth decile followed by a significantly faster increase in decile seven (during the
campaigns preceding and following the referendum on the relocation quota, see Table 1) and a slow
decrease after then.

14. Figure The prevalence of the label *quota* by sources and period deciles (%)

The two outliers differ from this general trend only having a faster increase in decile seven which lasted
in case of nol.hu longer (and remained unfinished due to its closing down).

2.3 Actors

By analyzing the prevalence of and the context in which the actors appear in the media can be
considered as the first level analysis of the role a person has (or more precisely as it is mirrored in the
media) in the politicization process of the migration crisis. In the dataset we have a large number of
actors identified individually. Appendix 1, section A1.3 shows those who appear in at least 2% of the articles.
The dominant position of Orbán and the presence of two more government officials among the most often mentioned actors proves that the government dominates the discourse of migration crisis in Hungary, which again can be a sign of the operation of the MPB.

The general trend of prevalence of the two main actors (Figure 16) shows that while Orbán’s presence is relatively stable (except during the peak of the migration crisis), Soros appeared and had risen fast only when the government propaganda pushed the MPB, i.e. decided to make him the ultimate scapegoat of the migration crisis.

Orbán is the Prime Minister of Hungary, Soros is treated as his arch-enemy, the world-wide fighter of liberal values, Merkel is the Germany Chancellor, Szijjártó is the Minister of Foreign Affairs of Hungary, Lázár was the head of the prime Minister’s Office, Vona was the president of Jobbik (radical right-wing opposition party) between 2006 and 2018.
As to the prevalence of Orbán and Soros by sources (Figure 17), we found that their presence in governmental and non-governmental source is different.

17. Figure The proportion of articles mentioning Orbán and Soros by source (%)

Soros was more visible in the governmental media (except in alfahir) than in the non-governmental sources.34 The case of Orbán needs further analysis because there was as higher intra- than extra-group deviation in both the governmental and among the non-governmental groups. In average, however, Orban is likely to be mentioned more often in the non-governmental than in the governmental media.

34 The extremely low visibility in nol.hu is due to the fact that this newspaper was closed before the Soros campaign entered into full force.
2.4 Organizations and locations

The three dominant types of organizations in the corpus were Hungarian media organizations, political parties and international organizations. Figure 18 contains the most often mentioned organizations. 35

18. Figure The most often mentioned organizations (% proportion of articles mentioning the organization)

MTI is the Hungarian news agency, Fidesz is the governing political party, Brussels is the pseudo-term of the EU, Jobbik was (at least at the beginning of the time-span) a radical right-wing opposition party, Schengen is the pseudo-term of the Treaty, M1 is the government controlled Hungarian television channel, MSZP is the Hungarian Socialist Party.

As a first step towards analyzing the locations mentioned in the corpus Figure 19 contains the most often mentioned locations. 36

---

35 Appendix 1 section A1.3 contains those organizations which appear at least once in more than 3% of the articles.

36 Appendix 1 section A1.3 contains those locations which appear at least once in more than 4% of the articles.
19. Figure The most often mentioned locations (% proportion of articles mentioning the location)

The most often mentioned locations outline the main directions of the Balkan route of the migration crisis.37

2.5 Sentiments

In the future the addition of sentiments will make the content analysis much more robust since we would be able to differentiate between cases when an actor, an organization or a location is referred in the media positively or negatively, e.g. whether Orbán (or Soros) is a “hero” or a “culprit”. At this stage we only introduce the Reader to the most often used positive and negative sentiments in the selected articles (Figure 20).38

37 The presence of the USA can be explained on the hand by its world-wide influence, on the other hand because our dataset contains migration related news which are unrelated to the Hungarian processes.
38 Appendix 1 section A1.3 contains those sentiments which were mentioned at least 3% in the articles.
20. Figure The most often mentioned five positive and negative sentiments (% proportion of articles mentioning the sentiment)

The trend of positive sentiments does not show significant changes (Figure 21) their general shape is a mild U-curve which suggests that during the peak of the migration crisis positive sentiments have somewhat - and temporarily - lost their prevalence.

21. Figure The positive sentiments by period (% the proportion of articles having the sentiment mentioned)

As to the trend of negative sentiments (Figure 22), the only one with significantly changing prevalence is the term *illegal*. This term has been frequently used during the peak of the migration crisis of 2015 and after 2017 became the epithet ornament of the term *migrant* and therefore an oft-used metaphor of MBP.
22. **Figure** The negative sentiments by period (%, the proportion of articles having the sentiment mentioned)

Though it gives only an imprecise picture\(^{39}\), positive sentiments (with some exception such as atv.hu and 24.hu) are more frequent in the non-governmental sources than in the governmental ones (Figure 23).

23. **Figure** The positive sentiments by source (% the proportion of articles having the sentiment mentioned)

As to negative sentiments (Figure 24), the term illegal is used more often in the governmental media than in the non-governmental ones (except atv.hu and nepszava.hu) – again a possible sign of the operation of the MPB - but there is not much difference between the two groups in regard with the prevalence of the other negative sentiments.

\(^{39}\) While a xenophobe regulation can be judged *good* in the governmental media and a civic action against is *bad*, in the non-governmental media the exact opposite might occur.
2.6 Some preliminary tests of the MPB

The following two exercises go beyond the previous general descriptions in one sense: they give more in-depth knowledge concerning our MPB related hypotheses. In the course of these experiments we re-arranged the dataset into focused sub-samples, and in this way created quasi-experimental desings to check whether the differences (if any) would fall into the direction expected on the basis of the MPB approach.

2.6.1 The case of Origo

The case of Origo can be considered as a natural experiment (see Glossary), since while during the beginning of the time-span this website (origo.hu) was a non-governmental source, by the end of the time-span it became a state controlled one.40

We assume that as the ownership of Origo has moved from non-governmental (i.e. market-oriented consequently news-oriented) towards state-controlled owners (who based their profit on state allocated propaganda funding), so had the content of the articles changed into more and more serving the manipulative purposes of the MPB.

The analysis of labelling shows these expected changes very clearly (Figure 25).

---

40 A brief summary of this history between 2015 and early 2018 can be read in Appendix 4.
25. **Figure** The prevalence of labels in the three selected periods of Origo (percentage of articles the label was used)\(^{41}\)

As expected, there was a significant decrease of using *refugee*, *asylum seeker*, *refugee crisis* and *refugee system*, and an increasing use of *migrant* and *migration*, and less visibly to *immigration*.

Using the trend of the two most prevalent labels in the entire corpus as a benchmark (Figure 26), we can see the impact of origo.hu becoming a governmental propaganda source, i.e. there was a general decline of using the term *refugee* but in case of origo.hu this was significantly stronger than in the corpus. The trend of using the term *migrant* also fits to our expectations: compared to the general trend (a mild increase) origo.hu started to use this term significantly more often after being subservient to MPB.

26. **Figure** The prevalence of *refugee* and *migrant* labels in the three selected periods of Origo and of all articles (percentage of articles the label was used)

---

\(^{41}\) The total number of articles was 4069, and in 2015, 2016/7 and 2018 were 2022, 792, and 1455, respectively.
Since we did not use a sophisticated classification of positive and negative sentiments, we did not raise any assumption and were only interested to see if there were any changes in the course of the shift of Origo from independence to subserviency (Figure 27).

27. Figure The prevalence of positive and negative sentiments in the three selected periods of Origo (percentage of articles the sentiments were used)

The only significant change occurred in the use of the term illegal and only in the last period. This trend deviates from the general trend (Figure 22) where the peak of using the term illegal was in the beginning of the time-span. This deviation from the general trend, however, fits very well to the MPB concept since this was the period when the Parliamentary election campaign focused on illegal migrants (brought by Soros - in collaboration of civil organizations, the EU and the UN - to Europe) heavily.

Last but not least, we assumed that pressing the MPB would result in case of both two main actors to become more prominent (Figure 28).
Contrary to our expectation (and the trend in the corpus) the presence of Orbán has decreased but exactly as was assumed the visibility of Soros fastly (and earlier than in the corpus) increased. While the explanation of the former needs further and more careful analysis, the latter case is clear since Soros was the main scapegoat of migration related threats against Hungary and Europe (and the World).

2.6.2 The pre- and post-fence experience

The second example is based on the fact that during the time-span of the research two significantly different periods can be defined: the one before and the one after the building of the fence (September 2015). The explanation is that while in the pre-fence period there were large numbers of refugees in Hungary and they - by the then-existing more independent sources - were made visible for the public, in the post-fence period neither refugees not equally strong non-governmental media existed anymore. In other words, in the pre-fence period (1) because of the visibility of refugees (Boda-Simonovits, 2016) and (2) because of the existence of more polemic and informative media, the operation of the MPB was somewhat tempered but these balancing effect disappeared (or were significantly weaker) in the post-fence period. Consequently, we assume that the previously identified signs of the effect of the MPB will be more visible in the latter period compared to the former one. To illustrate the relevance of these assumptions, we defined two periods: an early one (between 7 January 2015 to 16 September 2015, -4 indicating the first four deciles), and a late one (after 2016, 7- indicating the last four deciles).

In the next two figures (Figure 29 and 30) we compare the two most often used labels, the refugee and the migrant. The proportion of articles using the two terms changed in the opposite direction in all sources, and in cases they did not (e.g. migrants in some non-governmental sources) there was no significant change between the two periods.

42 During these weeks the level of xenophobia, which has been increasing fast since 2014, has decreased (Sik, 2016b).
29. **Figure** The proportion (left figure) and rate (right figure) of using the label *refugee* in the two periods by sources (percentage of articles the label was used, rate: proportion of articles after 2016 divided by the proportion of articles before October 2015)

![Graph showing the proportion and rate of using the label refugee](image)

30. **Figure** The proportion (left figure) and rate (right figure) of using the label *migrant* in the two periods by sources (percentage of articles the label was used, rate: proportion of articles after 2016 divided by the proportion of articles before October 2015)

![Graph showing the proportion and rate of using the label migrant](image)

Therefore, we concluded that – unless we assume that the strength of the MPB was so strong that all sources were dominated by it - since the trend is similar among both governmental and non-governmental sources, the operation of the government-induced MPB cannot be justified.

If, however, we compare the rates of change of prevalence of the two labels (the right hand blocks in the two Figures), we find that the decrease of using the *refugee* label is much stronger among the governmental sources than the average which is a clear indication of their stronger readiness to meet with the expectations of the state authorities.43

43 We carried out the same analysis in case of all other labels and found that the trend in the two groups of sources was rather similar: increasing rate of *migration* and decreasing rate for all other labels (*immigrant, immigration, asylum seeker, refugee crisis and refugee system*).
We assumed that due to the pressing of the MPB, the governing party (Fidesz) will be more in focus in the second than it was in the first period. As Figure 31 shows this indeed was the case but since this trend was again similar in all sources, irrespective their political status, and the fact that in both groups there were outliers, i.e. sources with extremely high increase rate, the MPB assumption cannot be proved in this dimension either.

31. **Figure** The proportion (left figure) and rate (right figure) of mentioning *Fidesz* in the two periods by sources (percentage of articles the label was used, rate: proportion of articles after 2016 divided by the proportion of articles before October 2015)

As to the trend of using the most often applied negative sentiment (illegal) we did not find any systematic difference between governmental and non-governmental sources (Figure 32). In both groups there were sources hardly changing their practice and sources which used this term significantly less in the second period. Consequently, we should reject the hypothesis of the influence of MPB in this respect.

---

44 We found increasing proportion of mentioning the two larger opposition parties (Jobbik and MSZP) as well. There was some difference in the proportion and rate of referring to them by the sources which are known to be “close” to them (in case of alfahir.hu Jobbik, and in case of népszava.hu MSZP).
32. Figure The proportion (left figure) and rate (right figure) of mentioning illegal in the two periods by sources (percentage of articles the label was used, rate: proportion of articles after 2016 divided by the proportion of articles before October 2015)

Finally, we assumed that Orbán Viktor will be more in focus in the second period compared to the first. As the following figures (Figure 33) shows (except in one source) this indeed was the case. However, we again did not find clear signs of the operation of the MPB – the trend was more or less similar in both groups (with one deviation – alfahir became extremely Orbán-focused by the second period).

33. Figure The proportion (left figure) and rate (right figure) of mentioning Orbán in the two periods by sources (percentage of articles the label was used, rate: proportion of articles after 2016 divided by the proportion of articles before October 2015)

Chapter 3 Summary and further possibilities

The aim of this paper is twofold: first - as a methodological experiment - illustrate the usefulness of content analysis for EU authorities. With this case study we intend to outline a technique which is a necessary (though insufficient) basis of an all-European discourse monitoring system. The development and regular updating of such a large scale and properly designed corpus could be the basis (1) to analyze the discursive trends of the media (both prevalence- and content-wise), (2) prepare simple (descriptive) comparative analyses of the media content on more focused topics, and (3) create a corpus which can be used to make valid and reliable in-depth discourse analyses to uncover the hidden content of the media.
We think such exercise are important if we accept that (1) the media has a strong impetus on public opinion, (2) and this will even more be the case in the future when (3) issues which are morally sensitive (such as migration) will be more and more the target of fake news and other forms of “alternative reality”. And (4) it will be especially the case if governments start to use organized forms of propaganda to manipulate “the man of the street”, using migration and refugee issues as a source of scapegoating, hatemongering and generating moral panic for political advantages, especially (5) in its most developed form (the moral panic button, MPB. Finally, (6) this database and analysis would help the Commission to understand the triangle of actual migration discourses, voting behavior in the Council and in the European Court of Justice concerning asylum and migration issues of the Member States.

Secondly, we intended to continue and deepen our analysis of the politicization process of the migration crisis in Hungary, we did in the comparative analysis of salience in WP5 (Patterns..., 2018, Pasetti and Garcés-Mascareñas, 2018). Moreover, unlike in the papers we have written on the theoretical approach moral panic button (MPB) earlier where we focused on the outcome, in this paper we analyze the process of the pressing of the MPB itself. In other words, this analysis, though it is hardly more than a descriptive overview of these processes can be considered as a first attempt to illustrate the operation of the MPB.

First (Chapter 1.1 and 1.2), we summarized the limits of the content and discourse analyses literature of the migration crisis in 2015 as far as its methodological basis. With several examples we illustrated that even the best opi have rather simple and/or biased empirical basis, consequently they can hardly be considered as proper starting point for comparative analysis. The usual limit of these analyses was that even the most sophisticated linguistic discursive analytical techniques (having very high level of internal validity) have hardly any external validity. In other words, we get a very detailed description of the framing processes of the media but we do not know whether these cases were the exceptions or the rules, i.e. we do not know how to generalize our findings.

As to our research design (Chapter 1.3), since our aim was to reach the highest possible level of external validity (while retaining the possible highest level of internal validity as well) the first priority was to define the limits of our sampling design and within it to maximize coverage. As to the limits of our targeting strategy, we sampled only online sources and ignored (1) the printed editions of the same media and the weekly or monthly magazines, (2) the TV and the radio, (3) the social media, and (4) all forms of the local media. We selected ten sources to represent the Hungarian online media taking into consideration three dimensions: (1) the type of source (online version of printed newspaper versus online site), (2) the size of the audience, and (3) their relation to the government (governmental and non-governmental) as well as (among the non-governmental sources) their political attitude (right-wing/radical versus moderate/left-wing). We selected five media with large audience four non-governmental (index.hu, 24.hu, atv.hu, and 444.hu) and one governmental source (origo.hu) which used to be also for-profit before its nationalization. The other five sources consist of three right-wing and two leftist online media with smaller but very active audience (alfahir.hu, pestisracok.hu, magyaridok.hu and nepszava.hu, nepszabadsag.hu).

To describe the main characteristics of the media discourse in Hungary during the migration crisis we had to define the starting and (if there was) the ending dates of the time-span of the migration crisis.

---

45 The increase of popularity of the ruling government (and the political party (Fidesz) behind it) and the increasing level of xenophobia (Sik, 2016a, Barlai-Sik, 2017).
As to the former, we started our analysis on 1 January 2015 because in early January 2015 the Hungarian Prime Minister (Viktor Orbán) gave several interviews between which could be seen as the first pressing of the MPB. As to the end of the time-span, we assumed that one of the main characteristics of MPB is being a long-term form of manipulation, consequently, the discourse on migration crisis did not end by the disappearance of refugees. Since several qualitative studies illustrated the pressing of the MPB between 2015 and 2018, we decided to extend our research period to cover the Parliamentary election in 2018. Finally, we defined shorter periods which were identified as separate pressings of the MPB within the time-span such as the pre- and post-campaign weeks of the national consultations.

To find those articles which were likely to contain relevant information in relation with the migration crisis, we defined four keywords („menekül...” „bévándor...” „migrá...” "betelepít...") and selected articles which contained at least one of these keywords.

The politicization of the migration crisis in contemporary Hungary

We start the analysis of the politicization of the migration crisis in contemporary Hungary focusing on prevalence (Chapter 2.1). We analysed the prevalence of relevant articles in two dimensions: time-span and source, i.e. we were interested whether there are significant differences along the time-span and the among the different types of sources focusing on the migration crisis.

We identified the following periods in which, compared to the preceding and following periods, there were smaller or bigger changes in the number of articles: in 2015 the early January period (the first pressing of the MPB, i.e. the post-Hebdo Orbán speech), May (the national consultation on terrorism and immigration). second part of June (the on-going anti-immigration “information” campaign, and between early August to early October (the “Hot Autumn” of the migration crisis. In 2016 from mid-September to early October the national consultation of the “Soros Plan”, in 2017 from mid-November to early December the campaign in relation to the quota referendum created a high-selience periods.

To sum up, migration crisis related news have the largest prevalence in the news oriented world where non-governmental media dominates. In Hungary, however, due to the repeated pressings of the MPB the prevalence of the migration issues remained high. In other words, while the high level of prevalence during the peak of the migration crisis of 2015 can be considered as the normal reaction of a news-oriented media, the high level of prevalence (especially in governmental sources and with higher-than-average level of negative sentiment) without any migration-related event occurring (except those artificially created by the government), is a sign of the operation of the MPB.

In the following part of the paper we analysed the (un)changes of the content of the Hungarian along the time-span and in the different sources.

Tha analysis of labelling showed that the dominant term was the menekült (refugee), followed by migráns (migrant) and bevándorlás (immigration). The term refugee dominated the scene in 2015 and 2016 and was losing strength by 2017/2018 (during the quota referendum and its aftermath, and the anti-UN, anti-CEU campaigns. The most salient period of this term was the culmination of the migration crisis in September 2015 but by 2018 the prevalence of all four terms of migration exceeded the prevalence of the term refugee. The fact that by the end of the period various forms of migration replaced refugee can be interpreted as the result of the operation of the MPB, i.e. while there were

---

46 The Hungarian equivalents of the terms refugee, immgrant, migrant and relocation.
masses of refugees (and migrants and asylum seekers) marching through Hungary the media used the term which seemed to fit best to label the mixed population of foreigners. After this short period, however, the term migration became more popular in the increasingly government dominated media.

As the MPB concept predicted, the term *refugee* was indeed significantly less used in governmental than in non-governmental sources, and the opposite was true to all four versions of the term *migrant*.

To analyze the use of the two most often use labels by sources during the time-span of the research, we selected pairs of sources relatively similar as far as their type and the size of their audience were concerned but differed in their relation to the government and political attitudes.

In case of the two websites having the highest number of readers, the non-governmental source (index.hu) the prevalence of *refugee* shows a bell curve and (following a sharp rise after the first period) there is not much change of using the term *migrant*. The prevalence of the former always exceeds the latter. As to origo.hu (which became governmental source in 2016) the general trend of using the term *refugee* is a bell curve as well but with a fast drop in 2016, and the term *migrant* is characterized by a rising trend which became fast in 2017.

Comparing the online version of a governmental (magyaridok.hu) and a non-governmental (nepszava.hu) daily (Figure 12), we find that in case of the governmental source the decreasing trend of using the term *refugee* and the opposite trend of using the term *migrant* explains why by the mid-period deciles the latter reached the prevalence of the former.

We only very briefly illustrated the main characteristics of the prevalence of actors, organizations, localities and sentiments. The dominant position of Orbán and the presence of two more government officials among the most often mentioned actors proves that the government dominates the discourse of migration crisis in Hungary. The trend of prevalence of Orbán and Soros showed that while Orbán’s presence is relatively stable (except during the peak of the migration crisis), Soros appeared and had risen fast only when the government propaganda pushed the MPB, i.e. decided to make him the ultimate scapegoat of the migration crisis. The three dominant types of organizations in the corpus were Hungarian media organizations, political parties and international organizations. The most often mentioned locations outline the main directions of the Balkan route of the migration crisis. The trend of positive sentiments has a mild U-curve which suggests that during the peak of the migration crisis positive sentiments have somewhat - and temporarily - lost their prevalence. As to the trend of negative sentiments, the term *illegal* has been frequently used during the peak of the migration crisis of 2015 and after 2017 became the epitheton ornans of the term *migrant* and consequently an oft-used metaphor of MBP especially in governmental sources.

In the course of analysis we re-arranged the dataset into focused sub-samples and in this way created quasi-experimental designs to check whether the differences (if any) would fall into the direction expected on the basis of the MPB approach.

The case of Origo can be considered as a natural experiment since while during the beginning of the time-span this website (origo.hu) was a non-governmental source, by the end of the time-span it became a state controlled one. We assumed that as the ownership of Origo has moved from non-governmental (i.e. market-oriented consequently news-oriented) towards state-controlled owners (who based their profit on state allocated propaganda funding), so had the content of the articles changed into more and more serving the manipulative purposes of the MPB.
As expected, there was a significant decrease of using *refugee, asylum seeker, refugee crisis* and *refugee system*, and an increasing use of *migrant and migration*, and less visibly to *immigration*.

Using the trend of the two most prevalent labels in the entire corpus as a benchmark, we could show the role of MPB quite clearly, i.e. in the corpus there was a general decline of using the term *refugee* but in case of origo.hu this was significantly stronger. The trend of using the term *migrant* also fits to our expectations: compared to the general trend (a mild increase) origo.hu started to use this term significantly more often after being subservient to MPB.

The use of the term *illegal* deviated from the general trend as was expected: while in the corpus the peak of using the term *illegal* was in the beginning of the time-span, origo.hu applied this term in the last period when the Parliamentary election campaign focused on illegal migrants (brought by Soros in collaboration with civil organizations, the EU and the UN - to Europe) heavily.

The second example is based on the fact that during the time-span of the research two significantly different periods can be defined: the one before and the one after the building of the fence (September 2015). We assumed that in the **pre-fence period** (1) because of the visibility of refugees and (2) because of the existence of more polemic and informative media, the operation of the MPB was somewhat tempered but these balancing effect disappeared (or significantly weaked) in the **post-fence period**. Consequently, we assume that the previously identified signs of the effect of the MPB will be more visible in the latter period compared to the former one.

Comparing the rates of change of prevalence of the *refugee* and *migrant* labels we found that the decrease of using the *refugee* label is much stronger among the governmental sources than the average – a clear indication of their stronger readiness to meet with the expectations of MPB.

**Further possibilities of using the dataset**

In the future there are two ways to use the dataset to improve our knowledge on the politicization of the migration crisis in Hungary: (1) fine-tune the content analysis and (2) develop various discourse analyses. In the following we offer a list of further content and discourse analyses which can be envisioned using the dataset we developed:

**as to content analysis:**

- We can apply on the article-level dataset more sophisticated analytical techniques such as principal component and cluster analysis to identify groups of articles containing particular types of words significantly more frequently (for example key terms of MPB (illegal and migrant simultaneously or the cumulaation of “scapegoats” (such as Soros, Brussels, civil society, etc.) in the same article.
- If we can identify certain elements of the general discourse on migration crisis as particularly relevant, we could select subsamples containing such terms and test whether our main findings are present in these focused analyses and what other associations (if any) are identifiable. We envision such focused analysis for example in relation to the relocation quota which was a core topic of MPB in Hungary.
- If we assume that MPB operates with short, simple, stereotypical messages, it might make sense to select a subsample of containing only short articles, and test whether on this special segment of articles our original findings are present more lucidly.
- If we assume that it is not the main text but headlines, leads, and have particularly strong impact on the public that it makes sense to repeat our analysis on these elements of the articles.
- If we assume that pictures have a particularly significant effect on the public, we should analyse the pictures belonging to the articles (by using the analytical technics of the excellent paper on the typology of pictures (The faces of migration, xxx)) separately.
- Using principal component and cluster analysis we could describe the structure of the positive and negative sentiments in a systematic way. Once we created the clusters that contain sentiments strongly belonging to each other (and differ from those in other clusters) we could differentiate among the articles whether (1) they are focusing on a particular sentiment or contains a mixture of them and (2) what are the relations of these sentiment types to the other elements of the message the article contains (e.g. what sentiment structure is more or less likely to be attached to the appearance of Merkel, etc.).
- The dataset contains articles which are not related to Hungary directly such as terror cases or the USA/Mexico migration crisis, etc. If we assume that the use of these type of news can also be part of the MPB propaganda than we should identify those articles which contain both domestic and international elements of MPB as a special technique of manipulation.

As to discourse analysis:

- Considering the results of our qualitative analysis of the media and Parliamentary speeches (Bognár et al, 2018) as a first level of a code tree, we can start to develop the basic typology of frames (such as “power struggle”, sovereignty”, “effectiveness”, “democratic”, and “humanistic” frames). Moreover, this qualitative analysis could also serve as a first step towards identifying the typical argumentation discoursive elements and the diverse approaches towards solidarity and responsibility.
- Analysing the characteristics actors are present in the discourse, there are several dimensions to analyse such as (1) we should differentiate between those occasions when the actors act and when they are only referred to, (2) we could identify metaphors associated closely to a certain actor, (3) we could analyse the context of sentiments in which an actor appears (such as a culprit or a hero, a leader or a follower, a representative of an organization or of a nation (or idea), etc.
- Focusing on the localities, we can draw the mental map which is communicated to the Readers by the media, and analyse what type of discourse is associated to these localities (such as “flooding”, “chanelling”, “obstacle”, etc.
- Mutatis mutandis, similar exercise can be envisioned in case of actors and/or organizations (i.e. what role is associated to the particular person and organization), and how are these roles associated to each other and to other characteristics of the actors/organizations (such as sentiments).
- Networks of actors and/organizations can be relevant dimensions in understanding the deeper relations within the discourse. Once such structures are identified their (in)stability, the sentiments they are associated with and metaphores they hold, etc. can useful to reveal the less visible aspects of the discourse.
- There is a unique opportunity to compare the overlap of the frames in the media and public opinion. In 2016 and 2018 a representative survey used a questionnaire in which a block focused on migration related fears and those who are responsible for them (scapegoats). We
could identify the words (sentences) used in the survey in the media and analyze their association, i.e. to find identical structures (if there are any) in the media and the public opinion.
References


Fokasz, N. and Bodor, P. (2015) The construction of “crises” From the “Debt crisis” to the “refugee problem” in the Greek and Hungarian dailies (Manuscript)


Proposal of CEASEVAL RESEARCH ON THE COMMON EUROPEAN ASYLUM SYSTEM (2017) (Manuscript)


Glossary


Cluster analysis ... is the task of grouping a set of objects in such a way that objects in the same group (called a cluster) are more similar (in some sense) to each other than to those in other groups (clusters). ... Cluster analysis itself is not one specific algorithm, but the general task to be solved. It can be achieved by various algorithms that differ significantly in their understanding of what constitutes a cluster and how to efficiently find them. Popular notions of clusters include groups with small distances between cluster members, dense areas of the data space, intervals or particular statistical distributions.

Corpus (based on Allen, 2016 p. 23)

Corpus linguistics is an approach to studying language that analyses collections of texts called ‘corpora’ (singular: corpus) to look for usage patterns, among other aspects, that aren’t necessarily apparent when researchers read a small set of articles. ... despite computerized assistance, researchers still can—and arguably must—view how the quantitative patterns exist in context and interpret the significance of them.


Critical discourse analysis (CDA) is an interdisciplinary approach to the study of discourse, or put simply talk and text, that views language as a form of social practice. Scholars working in the tradition of CDA generally argue that (non-linguistic) social practice and linguistic practice constitute one another and focus on investigating how societal power relations are established and reinforced through language use. In this sense, it differs from discourse analysis in that it highlights issues of power asymmetries, manipulation, exploitation, and structural inequities in domains such as education.

Discoursive historical analysis

„The analysis will follow a multilevel pattern of analysis ... It distinguishes between the entry-level (thematic, topic oriented, or content oriented) and in-depth (strategy oriented including argumentation oriented) analysis. Within the former, the focus is mainly on discourse topics that allow one to map the key tendencies in the contents of analyzed mediated texts and to sketch the tendencies in overall framing of the discussed issues. The in-depth analysis, on the other hand, focuses on the key patterns of argumentation. It will rely on the notion of discursive strategy ... and, thereby, on the category of topoi (or argumentation schemes ...” (Kryzanocsky (2018b p. 105)
“...In the first stage, texts are carefully read, and discourse topics are identified ..., and at the second level, an in-depth analysis of texts is undertaken to ascertain how intertextuality and interdiscursivity is achieved (through recontextualization) and performed through certain linguistics strategies at the text or sentence level. In this case, the focus will be on just two related elements of the Discourse Historical approach: topoi and legitimation.” (Bennett, 2018, p 145-6)

**Discourse strand**

"topical threads within discourses ... In summary, the arguments for implementing the notion of discourse strand in such a fashion are

- topical continuity
- strong intertextual links (often explicit) among its elements
- relative temporal proximity
- an often limited group of social actors (focused social field)
- an initiating event or events
- topically and temporally bounded
- high keyness values (of subcorpora vis-a-vis corpora).” (Rheindorf and Wodak, 2018 page 22-23)


Frames in communication consist of the communication of frames between different actors. Framing involves social construction of a social phenomenon – by mass media sources, political or social movements, political leaders, or other actors and organizations. Participation in a language community necessarily influences an individual’s perception of the meanings attributed to words or phrases. Politically, the language communities of advertising, religion, and mass media are highly contested,


In computing, JavaScript Object Notation (JSON) is an open-standard file format that uses human-readable text to transmit data objects consisting of attribute–value pairs and array data types (or any other serializable value).

**Moral panic**

A moral panic is the process of arousing social concern over an issue, usually through the work of “moral entrepreneurs” (people initiating the panic create a clear message and set the agenda) and the
mass media. This process can be described as intensive fear and a high level of anxiety raised by news transmitted by the mass media. As Cohen defined it the moral panic is a threat to societal values and interests” (Cohen 1972).

**MPB (Barlai-Sik, 2016a)**

The MPB is a complex set of state-propaganda techniques. It includes various forms of manipulation such as ‘national consultation’, the ‘quota referendum’, Parliamentary election and overlapping waves of billboard, TV, radio spots campaigns. The aim of using the MPB is win popularity by framing and manipulating the political discourse in the public sphere. This is a long-term, high-cost manipulation technique which monopolizes all forms of media and reduces the room for maneuver of non-governmental actors. Through this mechanism, citizens can be reached with messages that are tailored by professional framing experts to influence the way they think, feel and behave.

In our case the message is always a migration crisis related issue (i.e. anti-Soros, anti-UN, anti-EU, anti-civil society, etc.) but with different scapegoats. These targets, however, are framed in a way to form a conspiracy against the Hungarian nation, against which only the government can save the people: „the government is using the MPB to build national unity by „creating an enemy image as subject of the moral panic which is perceived as a threat to social values and national interests, thus to moral order.” (Barlai-Sik, ibid p. 153)

**Natural experiment** ([https://en.wikipedia.org/wiki/Natural_experiment](https://en.wikipedia.org/wiki/Natural_experiment))

A natural experiment is an empirical study in which individuals (or clusters of individuals) are exposed to the experimental and control conditions that are determined by nature or by other factors outside the control of the investigators. The process governing the exposures arguably resembles random assignment. Thus, natural experiments are observational studies and are not controlled in the traditional sense of a randomized experiment. Natural experiments are most useful when there has been a clearly defined exposure involving a well-defined subpopulation (and the absence of exposure in a similar subpopulation) such that changes in outcomes may be plausibly attributed to the exposure

**Parsing** ([https://en.wikipedia.org/wiki/Parsing](https://en.wikipedia.org/wiki/Parsing))

Within computational linguistics the term is used to refer to the formal analysis by a computer of a sentence or other string of words into its constituents, resulting in a parse tree showing their syntactic relation to each other, which may also contain semantic and other information.

Robust statistics are statistics with good performance for data drawn from a wide range of probability distributions, especially for distributions that are not normal. One motivation is to produce statistical methods that are not unduly affected by outliers. Another motivation is to provide methods with good performance when there are small departures from parametric distributions.


Web scraping, web harvesting, or web data extraction is data scraping used for extracting data from websites. While web scraping can be done manually by a software user, the term typically refers to automated processes implemented using a bot or web crawler. It is a form of copying, in which specific data is gathered and copied from the web, typically into a central local database or spreadsheet, for later retrieval or analysis.


In linguistic morphology and information retrieval, stemming is the process of reducing words to their word stem, base or root form. The stem need not be identical to the morphological root of the word; it is usually sufficient that related words map to the same stem, even if this stem is not in itself a valid root.


Internal validity is determined by how well a study can rule out alternative explanations for its findings. External validity concerns the extent to which the (internally valid) results of a study can be held to be true for other cases, for example to different people, places or times. In other words, it is about whether findings can be validly generalized.
Appendix 1 The steps of building the dataset (András Hering and Endre Sik)

In Appendix 1 we add some details to Chapter 1.3 where we introduced the Reader to the steps we made in the course of building our datasets. In Chapter A1.1 we focus on two aspects of coverage: sources and period. In Chapter A1.2 we focus on the techniques we used in the course of selecting the articles in general and in particular those which (from our point of view) had relevant content. Finally, in Chapter A1.3 and A1.4 we introduce the Reader to the variables the original database consisted and the variables we constructed later as well as the formats of the output (SPSS and JSON). This section offers the Reader a chance to check the internal validity of our analysis and can be used as a manual for further research as well.

A1.1 The sources of the research47

Table A 1.1.1 The general characteristics of the selected sources

**Origo** ([www.origo.hu](http://www.origo.hu))


Searching by the site’s own search engine covering the whole period. All relevant articles are selected in the domestic and foreign politics sections.

**Pesti Srácok** ([www.pestitracok.hu](http://www.pestitracok.hu))


Searching by the site’s own search engine covering the whole period. All relevant articles are selected in the following sections: Magyar ugar, Világ ugar, Forró drót

**Alfahír** ([www.alfahir.hu](http://www.alfahir.hu))


Searching by the site’s own search engine covering the whole period. All relevant articles are selected of the entire site since there are no separate sections.

**Magyar Idők** ([www.magyaridok.hu](http://www.magyaridok.hu))


Searching by the site’s own search engine covering the whole period. All relevant articles are selected in the domestic and foreign politics, Opinion and the Situation sections.

**Index** ([www.index.hu](http://www.index.hu))


Searching by the site’s own search engine covering the whole period. All relevant articles are selected in the domestic and foreign politics sections.

**24** ([www.24.hu](http://www.24.hu))

Searching by the site’s own search engine covering the whole period. All relevant articles are selected in the domestic and foreign politics sections.

444 (www.444.hu)
Searching by the site’s own search engine covering the whole period. All relevant articles are selected in the domestic and foreign politics sections.

Atv (www.atv.hu)
Established in 1989. Owned by Hungarian Faith Church. Orientation: ???
We create a list of all articles during the entire period and a crawler select the relevant articles backward until January 1 2015 in the domestic and international sections.

Népszava (www.nepszava.hu)
Searching by the site’s own search engine covering the whole period. All relevant articles are selected in the domestic and foreign politics sections.

Nol.hu
Existed between 1956 and 2016. In 2015 bought by Mediawork from a Foundation owned by the Hungarian Socialist Party only to close down next year. Orientation: left, social democratic.
Searching by the site’s own search engine covering the period between January 1 2015 and October 8 2016 (the last day the newspaper was published). All relevant articles are selected in the following sections: domestic and international politics, Opinions sections.

Table A1.1.2 The results of scraping in the domestic and international sections by sources (number of articles) 48

<table>
<thead>
<tr>
<th>Source</th>
<th>Total</th>
<th>Selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domestic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>origo.hu</td>
<td>33021</td>
<td>2271</td>
</tr>
<tr>
<td>pestisracok.hu</td>
<td>4111</td>
<td>351</td>
</tr>
<tr>
<td>alfahir.hu</td>
<td>50628</td>
<td>3986</td>
</tr>
<tr>
<td>magyaridok.hu</td>
<td>44907</td>
<td>4513</td>
</tr>
<tr>
<td>index.hu</td>
<td>26919</td>
<td>1716</td>
</tr>
<tr>
<td>444.hu</td>
<td>77204</td>
<td>4181</td>
</tr>
<tr>
<td>atv.hu</td>
<td>42340</td>
<td>2408</td>
</tr>
</tbody>
</table>

48 In case of alfahir.hu, 444.hu and nepszava.hu there were no international sections but there were special sections (See Table A1.1.1). Articles in these sections were considered as part of the domestic scraping.
<table>
<thead>
<tr>
<th>Source</th>
<th>Selected articles</th>
<th>All articles</th>
<th>Proportion (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.hu</td>
<td>45933</td>
<td>2306</td>
<td></td>
</tr>
<tr>
<td>nepszava.hu</td>
<td>38932</td>
<td>4314</td>
<td></td>
</tr>
<tr>
<td>nol.hu</td>
<td>12159</td>
<td>1682</td>
<td></td>
</tr>
<tr>
<td><strong>International</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>origo.hu</td>
<td>24746</td>
<td>2012</td>
<td></td>
</tr>
<tr>
<td>pestisracok.hu</td>
<td>26968</td>
<td>2969</td>
<td></td>
</tr>
<tr>
<td>magyaridok.hu</td>
<td>48551</td>
<td>4988</td>
<td></td>
</tr>
<tr>
<td>index.hu</td>
<td>25590</td>
<td>2227</td>
<td></td>
</tr>
<tr>
<td>atv.hu</td>
<td>30820</td>
<td>1742</td>
<td></td>
</tr>
<tr>
<td>24.hu</td>
<td>32744</td>
<td>2568</td>
<td></td>
</tr>
<tr>
<td>nol.hu</td>
<td>12810</td>
<td>2193</td>
<td></td>
</tr>
</tbody>
</table>

**Table A1.1.3** The proportion of the selected articles of all articles by sources
### Table A1.1.4 The allocation of observations between 1 January 2015 and 18 April 2018

<table>
<thead>
<tr>
<th></th>
<th>7 January - 4 February 2015 (4 weeks)</th>
<th>15 April - 31 November 2015 (cca. 30 weeks)</th>
<th>11 September - 9 October 2016 (cca. 30 weeks)</th>
<th>28 March - 1 June 2017 (cca. 9 weeks)</th>
<th>10 November - 8 December 2017 (4 weeks)</th>
<th>17 January - 20 February 2018 (cca. 5 weeks)</th>
<th>201 March - 15 April 2018 (cca. 6 weeks)</th>
</tr>
</thead>
<tbody>
<tr>
<td>origo.hu</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>pestisraco.k.hu</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>alfahir</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>magyarido.k.hu</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>index.hu</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>24.hu</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>444.hu</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>atv.hu</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>nepszava.hu</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>nol.hu</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

### A1.2 The selection of articles

As a result of **scraping**, every article contained at least one matching keywords. If an article contained any of these given keywords (in the title, subtitle, abstract, body or tag), the matched word were logged in an array form with the field it was found in. (Picture A1.2.1)
Two examples of the selection of the suffix (“menekül”) of the Hungarian word refugee in the corpus.

```json
"matches": [ {
  "match": {
    "field": "body",
    "word": "menekülőben"
  }
}, {
  "match": {
    "field": "body",
    "word": "bevándorlóként"
  }
},
"matches": []
]
```

The observation unit of parsing was the article. The original dataset (already filtered by Precognox) was formatted as follows:

- source
- sourceUrl (used as unique identifier for an article)
- publication time
- headline of the article (if any)
- subtitle (if any)
- authors (can be one person, multiple people, anonymous, or news org)
- lead (if any)
- body (the content of the article)
- tags (defined by the author, if any)
- images (array of links to pictures and descriptions)
- category (the basic genre of the article)

To illustrate the necessary cautiousness to work with non-mandatory elements (subtitle, lead, tag) in the corpus, we show the results of identifying these elements (Table A1.2.1).

**Table A1.2.1** The number of having tags, subtitles and abstracts by source

<table>
<thead>
<tr>
<th>Source</th>
<th>Tag present</th>
<th>No tag</th>
<th>Subtitle present</th>
<th>No subtitle</th>
<th>Lead present</th>
<th>No lead</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.hu</td>
<td>4751</td>
<td>123</td>
<td>0</td>
<td>4874</td>
<td>3443</td>
<td>1431</td>
</tr>
<tr>
<td>444.hu</td>
<td>4096</td>
<td>85</td>
<td>4132</td>
<td>49</td>
<td>0</td>
<td>4181</td>
</tr>
<tr>
<td>alfahir.hu</td>
<td>3947</td>
<td>39</td>
<td>38</td>
<td>3948</td>
<td>3829</td>
<td>157</td>
</tr>
<tr>
<td>atv.hu</td>
<td>0</td>
<td>4150</td>
<td>4150</td>
<td>0</td>
<td>4149</td>
<td>1</td>
</tr>
<tr>
<td>index.hu</td>
<td>3943</td>
<td>0</td>
<td>3678</td>
<td>265</td>
<td>956</td>
<td>2987</td>
</tr>
<tr>
<td>magyaridok.hu</td>
<td>7833</td>
<td>1668</td>
<td>8356</td>
<td>1145</td>
<td>8279</td>
<td>1222</td>
</tr>
<tr>
<td>nepszava.hu</td>
<td>4337</td>
<td>3</td>
<td>1879</td>
<td>2461</td>
<td>3841</td>
<td>499</td>
</tr>
</tbody>
</table>

49 The approximate translation of the selected two words in English are “in the course of fleeing” (menekülőben) and “as immigrant” bevándorlóként).
Table A1.2.1 shows that the differences of the design of a source can significantly bias the weight of the source when we analyze the prevalence or content of these non-mandatory elements. Moreover, some sources might be completely missing or seriously underrepresented (e.g. in case of tags atv.hu and magyaridok.hu).

As to the limit of using the term **category** (i.e. from which section of the source the article was scraped), as Table A1.2.2 shows, while the basic genres (domestic or international news) work well in most cases, there are sources where further analysis is needed to decide about the handling of special sections.

<table>
<thead>
<tr>
<th>Source</th>
<th>Domestic</th>
<th>International</th>
<th>The situation</th>
<th>Hotline</th>
<th>Opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>24.hu</td>
<td>2306</td>
<td>2568</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>444.hu</td>
<td>4181</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>alfahir.hu</td>
<td>3986</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>atv.hu</td>
<td>2408</td>
<td>1742</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>index.hu</td>
<td>1716</td>
<td>2227</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>magyaridok.hu</td>
<td>4513</td>
<td>4301</td>
<td>146</td>
<td>0</td>
<td>541</td>
</tr>
<tr>
<td>nepszava.hu</td>
<td>4340</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>nol.hu</td>
<td>1682</td>
<td>1606</td>
<td>0</td>
<td>0</td>
<td>587</td>
</tr>
<tr>
<td>origo.hu</td>
<td>2271</td>
<td>2012</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>pestisracok.hu</td>
<td>351</td>
<td>66</td>
<td>0</td>
<td>2903</td>
<td>0</td>
</tr>
</tbody>
</table>

To avoid **duplication**, the scraping was organized in a way which identified articles with the same URL and ignored them. Picture A1.2.2 illustrates the format of an article in the corpus looked like.
The JSON files were parsed in program R, with the rjson package. The arrays in the observations were flattened and concatenated when needed.

Some articles had zero body length (after rounded in hundreds). Most of the time these were either photo galleries or videos or had non-standard formatting resulting in a stub article. These zero-length articles were omitted from the file. This filtering again affected some sources more than others (e.g. atv.hu frequently uploaded segments from their TV programs) (Table A1.2.3).

**Table A1.2.3** The number of filtered and parsed articles by sources

<table>
<thead>
<tr>
<th>Source</th>
<th>Filtered</th>
<th>Parsed</th>
</tr>
</thead>
<tbody>
<tr>
<td>origo.hu</td>
<td>14</td>
<td>4283</td>
</tr>
<tr>
<td>pestisracok.hu</td>
<td>11</td>
<td>3320</td>
</tr>
<tr>
<td>alfahir.hu</td>
<td>182</td>
<td>3986</td>
</tr>
<tr>
<td>magyaridok.hu</td>
<td>0</td>
<td>9501</td>
</tr>
<tr>
<td>index.hu</td>
<td>2</td>
<td>3943</td>
</tr>
<tr>
<td>24.hu</td>
<td>109</td>
<td>4874</td>
</tr>
</tbody>
</table>
A1.3 Variables

We started to build the SPSS file, we used the Precognox dashboard as a benchmark and added new items when their frequency reached 2-3% in the corpus. In a separate phase we checked the proper names to filter out “deviant cases” and pool those which had the same stem. Just to illustrate the difficulty to identify a proper name we attached the following list in the raw file which contains the most frequent proper names of the total list (N=63562).

Table A1.2.4 The list of the most frequent thirty proper names in the raw file

<table>
<thead>
<tr>
<th>Name</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Magyarország</td>
<td>14593</td>
</tr>
<tr>
<td>Fidesz</td>
<td>14221</td>
</tr>
<tr>
<td>Orbán Viktor</td>
<td>8564</td>
</tr>
<tr>
<td>Jobbik</td>
<td>8283</td>
</tr>
<tr>
<td>Magyarországon</td>
<td>7391</td>
</tr>
<tr>
<td>Európa</td>
<td>6386</td>
</tr>
<tr>
<td>Orbán</td>
<td>6334</td>
</tr>
<tr>
<td>Magyarországra</td>
<td>5365</td>
</tr>
<tr>
<td>Európai Unió</td>
<td>4231</td>
</tr>
<tr>
<td>Viktor</td>
<td>4097</td>
</tr>
<tr>
<td>Soros</td>
<td>3972</td>
</tr>
<tr>
<td>Magyarországot</td>
<td>3388</td>
</tr>
<tr>
<td>Az</td>
<td>2971</td>
</tr>
<tr>
<td>Európában</td>
<td>2864</td>
</tr>
<tr>
<td>Európai Bizottság</td>
<td>2617</td>
</tr>
<tr>
<td>Soros György</td>
<td>2614</td>
</tr>
<tr>
<td>Budapesten</td>
<td>2481</td>
</tr>
<tr>
<td>Országgyűlés</td>
<td>2389</td>
</tr>
</tbody>
</table>
The lists of proper names in the SPSS file (those mentioned most frequently in the articles)\(^{50}\) were as follows:

**Actors:** Orbán Viktor (23\%), Soros (6\%), Merkel (6\%), Szijjártó (5\%), Lázár (4\%), Vona (3\%), Rogán (2\%), Juncker (2\%), Pintér (2\%), Obama (2\%), Putyin (2\%), Kósa (2\%), Gyurcsány (2\%), Trump (2\%).

**Organizations:** MTI (20\%), Fidesz (19\%), Brussels (15\%), ENSZ (8\%), Jobbik (8\%), Schengen (7\%), Isis (6\%). m1 (6\%), MSZP (6\%), Facebook (6\%), ellenzéki part (4\%), DK (4\%), KDNP (4\%), Amnesty Int. (3\%), V4 (3\%), LMP (3\%), NATO (3\%).

**Locations\(^{52}\):** EU, Európai Unió (50\%), Magyarország (48\%), Németország (17\%), Budapest (16\%), USA, Egyesült Államok (12\%), Ausztria (11\%), Törökország (8\%), Görögország (8\%), Olaszország (6\%), Franciaország (6\%), Balkán (5\%), Berlin (5\%), Szerbia (4\%), Horvátország (5\%), Rózske (4\%), Nagy-Britannia (4\%), Szlovákia (4\%).

A special program (http://opendata.hu/hu/dataset/hungarian-sentiment-lexicon) was used to identify the sentiments. This program identified 5939 negative and 1748 positive sentiment words. We searched only for exact matches, because the lexicon already contained suffixed formats. In our current workfile both for the negative and positive keywords only the most often occurring 50 negative and positive terms were kept. These were as follows:

**Positive sentiments:** jó-good (14\%), fontos-important (11\%), jól-well (10\%), szabad-free (9\%), sikerült-was successful (9\%), megfelelő-proper (6\%), támogatja-supports (5\%), megoldás-solution (5\%), támogatás-supporting (4\%), egyértelmű-unequivocally (4\%), megfelelően-properly (4\%), biztonságos-safe (4\%), valódi-true (4\%), egyértelmű-unequivocal (4\%), érdemes-worthy (4\%), hajlandó-ready (3\%), lehetőség-possibility (3\%), központi-central (3\%), alkalmas-suitable (3\%), sikerül-succeed (3\%), sikeres-successful (3\%), támogat-support (3\%), szakértő-expert (3\%).

---

\(^{50}\) The percentages indicate the proportion of articles an item was mentioned at least ones.

\(^{51}\) Further analysis might change this value since the terms „magyar miniszterelnök“, „miniszterelnök“ and „Orbán“ were mentioned 27\% and 4\%, respectively.

\(^{52}\) Two quasi-locations for obvious reasons were coded as „disguised organization“: Brussels and Schengen.
Negative sentiments: illegális-illegal (19%), félszorok-scarred (8%), válság-crisis (7%), probléma-problem (6%), rossz-bad (6%), háború-war (6%), illegálisan-illegally (5%), támadás-attack (4%), terrorizmus-terrorism (3%), veszélyes-dangerous (3%), háborús-warlike (3%), szélsőséges-extreme (3%), tartózkodó-reserved (3%), erőszak-violence (3%), baj-trouble (3%), terrorista-terrorist (3%), konfliktus-conflict (3%), menekülő-escaping (3%), bűncselekmény-crime (3%), halál-harmful (3%), tiltott-forbidden (3%).

We created a new variable to measure the size of the article (bodylength) by the number of characters (space included). The average size of an article was 2527 characters (standard deviation 2612) with a significantly lower median and mode (both 1000 characters) which indicates a large number of longer articles. The length of the bodies was rounded to hundreds.

A1.4 Data output

The corpus itself was archived in JSON format but for the current work file the results were exported to SPSS format. The exports have been truncated, because most formats (xlsx, sps) cannot handle long strings. These exports do not contain the body part either. The zero body length was filtered out, and the body length was rounded to hundreds.

In the original JSON file, to save space the export was formatted differently: instead of nesting one observation as seen in the example, we nested by the fields themselves. This way we avoid printing out the field’s names 309 times the number of observations. This can be transformed to the original form with the following syntax in R:

```r
> library(rjson)
> files_list <- read_json("./out/json/sample.json")
> files_df <- data.frame(sapply(files_list,c))
```

The total number of words in the corpus was 9 583 225. This number, however contains the so called “stop-words” (the, and, but...). If we select only the body text and apply a rudimentary stemming we end up with 8 179 984 words.
Appendix 2 Milestones and basic statistics of the migration crisis in Hungary in 2015 and 2016

Table A2.1 The event-history of the Hungarian migration crisis in 2015

<table>
<thead>
<tr>
<th>Date</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>11 January</td>
<td>The first relevant official statement related to immigration into Hungary: Prime Minister Orbán’s speech in Paris after the commemoration ceremony for the victims of the Charlie Hebdo terror attack: economic migration is bad, Hungary will therefore not provide asylum for economic migrants.</td>
</tr>
<tr>
<td>April</td>
<td>Preparations begin for the so-called ‘national consultation’ on immigration initiated by the government (mailing a questionnaire to all Hungarian adults to canvass their opinion on immigration). Increasing number of anti-immigration communiques by the government.</td>
</tr>
<tr>
<td>May</td>
<td>‘National consultation’ on immigration and terrorism</td>
</tr>
<tr>
<td>June</td>
<td>Government-sponsored anti-immigration billboard campaign nationwide and a counter campaign is organized by a fringe political party with pro-immigration messages.</td>
</tr>
<tr>
<td>25–30 June</td>
<td>The formation of pro-migrant grassroots organizations.</td>
</tr>
<tr>
<td>13 July</td>
<td>The Hungarian government starts building a fence along the Hungarian–Serbian border.</td>
</tr>
<tr>
<td>5–8 August</td>
<td>The opening of transit zones at a railway stations in Budapest with the volunteers and grassroots providing street social work and aid for asylum seekers.</td>
</tr>
<tr>
<td>8 August</td>
<td>71 dead migrants are found in a van in Austria close to the Hungarian border, obviously en route from Hungary.</td>
</tr>
<tr>
<td>25–29 August</td>
<td>Negotiations are under way about a central transit zone in Budapest, controlled by the Municipality of Budapest.</td>
</tr>
<tr>
<td>end of August – early September</td>
<td>Increasing tensions at the Budapest railway stations, where thousands of asylum seekers are waiting for the opportunity to travel on to Germany. Tensions are increased by the erratic reactions of the Hungarian authorities and the state railway company: at one point no asylum seekers are allowed on the trains leaving for Germany – even those with valid tickets; later it is again possible to get on the trains for a few hours, but then all international trains leaving for Germany are cancelled for a few days; on 3 September a train leaves Budapest Keleti railway station with asylum seekers who were informed that they are heading for Germany; the train, however, is stopped at the Bicske reception camp (Hungary), with asylum seekers feeling cheated and trapped.</td>
</tr>
<tr>
<td>3–6 September</td>
<td>Right after the train incident at Bicske on 3 September, asylum seekers at Budapest Keleti railway station set out to walk to Austria along the M1 motorway. In response, the government provides buses to transport the asylum seekers direct to the Austrian border from the motorway and from</td>
</tr>
</tbody>
</table>
Keleti station. A statement is made by the head of the Catholic Church in Hungary, Cardinal Peter Erdős, justifying the limited involvement of the church in the crisis, in stark contrast to statements made by the Pope.

8 September
Petra László, a camerawoman for a right-wing Hungarian TV channel, trips refugees running from the police at Rőszke (Serbian border).

15–16 September
‘The battle of Rőszke’ takes place between police and asylum seekers after the physical and legal closure of the Hungarian–Serbian border. The migration flow heads towards the Hungarian–Croatian border.

17 September
The Hungarian–Serbian border is closed.

21–23 September
Repercussions of a speech delivered by Prime Minister Orbán, who said ‘the government has given financial support to the NGOs’; the volunteers and grassroots protest, as they have in fact not received any state funds; the prime minister meant only those established charity organizations that were commissioned by the government with providing aid at the Croatian and Austrian borders after 15 September.

16 October
The fence along the Hungarian–Croatian border is completed, the border is closed: the end of the mass inflow of asylum seekers and migrants into Hungary.

Source: Juhász et al, 2016

Table A2.2 The number of asylum seekers by the citizenship 2010-2016

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghan</td>
<td>702</td>
<td>649</td>
<td>880</td>
<td>2 328</td>
<td>8 796</td>
<td>46 227</td>
<td>11 052</td>
<td>1 432</td>
</tr>
<tr>
<td>Algerian</td>
<td>35</td>
<td>56</td>
<td>59</td>
<td>1 116</td>
<td>98</td>
<td>599</td>
<td>710</td>
<td>62</td>
</tr>
<tr>
<td>Armenian</td>
<td>11</td>
<td>12</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>44</td>
<td>–</td>
</tr>
<tr>
<td>Bangladeshi</td>
<td>4</td>
<td>3</td>
<td>15</td>
<td>679</td>
<td>252</td>
<td>4 059</td>
<td>279</td>
<td>9</td>
</tr>
<tr>
<td>Cameroonian</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>54</td>
<td>94</td>
<td>642</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Chinese</td>
<td>12</td>
<td>10</td>
<td>6</td>
<td>5</td>
<td>11</td>
<td>8</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>Egyptian</td>
<td>14</td>
<td>20</td>
<td>6</td>
<td>5</td>
<td>13</td>
<td>23</td>
<td>218</td>
<td>6</td>
</tr>
<tr>
<td>Ethiopian</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>43</td>
<td>32</td>
<td>5</td>
</tr>
<tr>
<td>Georgian</td>
<td>68</td>
<td>21</td>
<td>12</td>
<td>41</td>
<td>40</td>
<td>30</td>
<td>13</td>
<td>6</td>
</tr>
<tr>
<td>Indian</td>
<td>3</td>
<td>11</td>
<td>12</td>
<td>84</td>
<td>11</td>
<td>345</td>
<td>123</td>
<td>6</td>
</tr>
<tr>
<td>Iraqi</td>
<td>48</td>
<td>54</td>
<td>28</td>
<td>63</td>
<td>497</td>
<td>9 279</td>
<td>3 452</td>
<td>812</td>
</tr>
<tr>
<td>Iranian</td>
<td>62</td>
<td>33</td>
<td>45</td>
<td>61</td>
<td>268</td>
<td>1 792</td>
<td>1 286</td>
<td>109</td>
</tr>
<tr>
<td>Country</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
<td>--------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>Kosovo</td>
<td>379</td>
<td>211</td>
<td>226</td>
<td>6 212</td>
<td>21 453</td>
<td>24 454</td>
<td>135</td>
<td>4</td>
</tr>
<tr>
<td>Mali</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>305</td>
<td>117</td>
<td>291</td>
<td>14</td>
<td>–</td>
</tr>
<tr>
<td>Macedonian</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>11</td>
<td>11</td>
<td>18</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>Moroccan</td>
<td>14</td>
<td>30</td>
<td>47</td>
<td>496</td>
<td>49</td>
<td>267</td>
<td>1 033</td>
<td>24</td>
</tr>
<tr>
<td>Moldova</td>
<td>14</td>
<td>9</td>
<td>7</td>
<td>10</td>
<td>5</td>
<td>8</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Pakistani</td>
<td>41</td>
<td>121</td>
<td>327</td>
<td>3 081</td>
<td>401</td>
<td>15 157</td>
<td>3 873</td>
<td>163</td>
</tr>
<tr>
<td>Palestinian</td>
<td>225</td>
<td>29</td>
<td>19</td>
<td>136</td>
<td>875</td>
<td>1 036</td>
<td>206</td>
<td>17</td>
</tr>
<tr>
<td>Russian</td>
<td>23</td>
<td>11</td>
<td>4</td>
<td>11</td>
<td>19</td>
<td>17</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Somali</td>
<td>51</td>
<td>61</td>
<td>69</td>
<td>191</td>
<td>194</td>
<td>352</td>
<td>331</td>
<td>9</td>
</tr>
<tr>
<td>Sudanese</td>
<td>14</td>
<td>4</td>
<td>2</td>
<td>104</td>
<td>71</td>
<td>278</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>Turkish</td>
<td>59</td>
<td>25</td>
<td>30</td>
<td>86</td>
<td>116</td>
<td>292</td>
<td>425</td>
<td>29</td>
</tr>
<tr>
<td>Tunisian</td>
<td>10</td>
<td>30</td>
<td>21</td>
<td>234</td>
<td>44</td>
<td>77</td>
<td>67</td>
<td>4</td>
</tr>
<tr>
<td>Ukrainian</td>
<td>9</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>37</td>
<td>28</td>
<td>23</td>
<td>7</td>
</tr>
<tr>
<td>Vietnamese</td>
<td>37</td>
<td>11</td>
<td>3</td>
<td>8</td>
<td>28</td>
<td>33</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Other</td>
<td>117</td>
<td>118</td>
<td>109</td>
<td>1 881</td>
<td>1 916</td>
<td>5 819</td>
<td>732</td>
<td>84</td>
</tr>
<tr>
<td>Total</td>
<td>2 104</td>
<td>1 693</td>
<td>2 157</td>
<td>18 900</td>
<td>42 777</td>
<td>177 135</td>
<td>29 432</td>
<td>3 395</td>
</tr>
</tbody>
</table>

Source: Hungarian Statistical Office

http://www.ksh.hu/docs/eng/xstadat/xstadat_annual/i_wvn0o2b.html
Figure A2.1 The number asylum seekers and refugees in Hungary (January/March 2013 – January/March 2016)

Figure A2.2 The daily number of illegal border crossings and the most relevant migration crisis related events (1 January 2015 – 31 November 2015)

Source: Juhász-Molnár, 2016
Figure A2.3 The daily number of illegal border crossings by the country of origin (1 January 2015 – 31 August 2015)

Source: Hungarian ... (2015)
**Figure A2.4** The daily number of illegal border crossings and the most relevant migration crisis related events (1 January 2016 – 1 July 2016)

Source: Juhász-Molnár 2016
Appendix 3 Pictures of the subsequent pressings of the MPB

No. 1 We do not want illegal immigrants! The Hungarian reforms work!

No. 2. “We send the message to Brussels so they will understand it”
No 3. „Stop Brussels!”

No 4. The Soros plan endangers the Christian culture in Europe
5a “The UN wants us to accept migrants continuously. Hungary decides, not the UN.”

5b. Soros hugging the leaders of the four parties running in the election (holding wire cutters in their hands): „They would dismantle the fence”.
“You have the right to know want Brussels is prepared to do: they want to introduce the mandatory relocation quota! To weaken the border defense of the member states! To make migration easier with a migrant-visa!”
Appendix 4 The brief story of Origo

Origo was established in 1998. The title of the portal expressed the aim of the owners – to establish a portal which is widely used as the very basic source of news. By 2000 Origo became the most popular webpage in Hungary.

According to rumours, after the parliamentary elections in 2014, Origo became a target of the government. Background agreements between certain government actors and the at-that-time owner (the Hungarian affiliate of Deutsche Telekom) were made that in return to some favours to increase their profit the owner will make sure that the site will stop criticizing the government.

In December 2015 Origo was sold to an oligarch well-known of his good connections to the government. In early 2017 most of the staff was either fired or left expressing their solidarity with those fired.

By the following brief quotations we on the one intend to substantiate some elements of the previous story and add some evaluation of Origo’s changing status in the context of the Hungarian online market, on the other hand we want to justify how did we define the two cezures we use in the empirical analysis to split the three periods.

As to 2016 (Bognár, 2017):

„The Hungarian online sphere is still vibrant and plural, with a wide range of political and news portals/blogs operating. Index and Origo are the two biggest online news portals in Hungary, founded in the late 1990s after legacy media struggled to adjust to the online environment. Although Origo can no longer be considered a space for independent journalism under its new ownership, critical opinion and investigative articles are produced by a number of online sites including Index, 444, and Mandiner. Partly crowdfunded watchdog NGOs and investigative journalism centres, Atlatszo and Direkt36, specialise in high-quality, labour-intensive investigative journalism. However, because of the nature of the media environment, these stories rarely get featured in the governmental or mainstream media. This raises serious concerns over Hungarian citizens’ right to be informed. Research shows that the amount of state advertising to government-friendly media outlets (such as Origo and TV2) has significantly increased in recent years“ (page 74).

„In Hungary the online market is the most balanced among the media markets, but providers with close ties to the government have significantly proliferated in this realm, too. The portal with the greatest reach, origo.hu, was acquired by the New Wave Media Ltd. in February 2016. Previously, origo.hu was owned by the telecommunications company Magyar Telekom, Hungary’s leading telecom corporation. Magyar Telekom in turn is a subsidiary of Deutsche Telekom, and the German telecommunications decided to sell off the online news provider. Now New Wave Media is owned by the businessman István Száraz, whose name became widely known in April 2016. At that time a major scandal erupted because another portal owned by Száraz, vs.hu (which is operated by New Wave Productions Ltd.) was awarded over half a billion forints from foundations affiliated with the Hungarian National Bank (MNB). The circumstances of the subsidy were totally inscrutable to the public. Before the change in ownership, origo.hu had a reputation for investigative reporting, but since then it has changed substantially and is now discernibly more loyal toward the government. (Recent..., page 12)
The research project CEASEVAL ("Evaluation of the Common European Asylum System under Pressure and Recommendations for Further Development") is an interdisciplinary research project led by the Institute for European studies at Chemnitz University of Technology (TU Chemnitz), funded by the European Union’s Horizon 2020 research and innovation program under grant agreement No 770037.) It brings together 14 partners from European countries aiming to carry out a comprehensive evaluation of the CEAS in terms of its framework and practice and to elaborate new policies by constructing different alternatives of implementing a common European asylum system. On this basis, CEASEVAL will determine which kind of harmonisation (legislative, implementation, etc.) and solidarity is possible and necessary.