www.reminder-project.eu

REMINDER

ROLE OF EUROPEAN MOBILITY AND ITS IMPACTS
IN NARRATIVES, DEBATES AND EU REFORMS

Multilingual Dictionary Construction: A Roadmap to Measuring Migration Frames in European Media Discourse

WORKING PAPER

Authors: Fabienne Lind
Jakob-Moritz Eberl
Tobias Heidenreich
Hajo G. Boomgaarden
Eva Luisa Gómez Montero
Beatriz Herrero
Rosa Berganza
Will Allen
Peter Bajomi-Lazar

Published: November 2019

universität wien

Universidad Rey Juan Carlos

1857 BGE

COMPAS CENTRE ON MIGRATION, POLICY & SOCIETY

REMINDER

ROLE OF EUROPEAN MOBILITY AND ITS IMPACTS
IN NARRATIVES, DEBATES AND EU REFORMS

Multilingual Dictionary Construction: A Roadmap to Measuring Migration Frames in European Media Discourse

Authors: Fabienne Lind[1], Jakob-Moritz Eberl[1], Tobias Heidenreich[1], Hajo G. Boomgaarden[1], Eva Luisa Gómez Montero[2], Beatriz Herrero[2], Rosa Berganza[2], Will Allen[3] and Peter Bajomi-Lazar[4]

[1] University of Vienna
[2] Universidad Rey Juan Carlos
[3] University of Oxford
[4] Budapest Business School

Submitted: March 2018

Working paper prepared as part of the REMINDER project

www.reminder-project.eu

Correspondence address:

Fabienne Lind, University of Vienna, Department of Communication
Rathausstraße 19, 1010 Vienna, Austria, Mail: fabienne.lind@univie.ac.at

Abstract

Since the so-called “refugee crisis” in 2015, the concept of free movement has been put under pressure by several EU member states. Still, EU citizens’ attitudes toward free movement are very different from one member state to the other. In order to understand possible sources of such attitudes and trace back changes in discourses on free movement to the early 2000’s, we argue that one has to take a closer look at media coverage on migration in Europe.

Since the analysis of such large corpora of texts is very resource intensive both in terms of time as well as money, one needs to rely on computer assisted methods of analysis, such as so-called dictionary approaches. However, research using automated procedures of multi-lingual text analyses are still sparse and methods are still only in early stages of their development and tend to focus on West-European languages. This is why, in this working paper, we provide a comprehensive literature review of the state-of-the-art of research in dictionary construction with a focus on multilinguality. Furthermore, we outline strategies to construct and validate such a multilingual dictionary. Finally, we give an overview of the implementation of such strategy for the measurement of migration frames in different European languages in the context of the REMINDER project. Eventually, this approach will allow the mapping of the migration discourse in seven European member states across time, providing possible insights and differences as well as shifts in attitudes related to free movement.

Introduction

Freedom of movement is one of the fundamental rights of citizens of the European Union and one of the Union's core principles. However, with the so-called "refugee crisis" in 2015 and the 2016 Brexit referendum, intra-EU mobility and non-EU immigration into Europe are under debate, with several EU members now seriously questioning the future of the free movement of people (Hobolt, 2016). Mass media play a crucial role as a link between politics and citizens. If one wants to understand citizens' attitudes toward free movement mass media may be an important place to start. In fact, when there is a lack of personal experience of mobility within the EU, media coverage may actually be the most important source of information; frankly, it may even be the only source citizens rely on when forming an opinion on free movement. An increasing number of studies provide evidence that media affect migration and EU related political attitudes and it has been shown that to most people, media is the most important source of information about EU topics (e.g., Boomgaarden & Vliegenthart, 2009; Boomgaarden, Vliegenthart, de Vreese, & Schuck, 2010; Vliegenthart, Schuck, Boomgaarden, & de Vreese, 2008). In order to understand public opinion on free movement, a systematic overview of media coverage on this issue in EU member states is needed.

The main goal of the Work Package 8 (WP8) project is to map discourses on migration and intra-EU mobility over time and across EU countries. Is a change in media coverage the precursor to changes in public opinion? More specifically, the aim is to measure the salience of frames, actors, and tone in large-scale multilingual text corpora consisting of texts from traditional mass media and social media by means of computer-assisted content analysis methods. One of these methods is the usage of a dictionary approach; an approach using the rate at which specific key words appear in a text to classify documents into substantive categories (e.g., frames). The text corpora analysed in WP8 originate from Spanish, English, German, Swedish, Polish, Hungarian, and Romanian text sources, and are thus multilingual. This multilinguality is a challenging and thus far rarely examined issue in computer-assisted content analysis.

Accordingly, this deliverable includes the following two parts:

- Part 1: a comprehensive literature review, which looks at multilingual dictionary construction.
- Part 2: a roadmap, which outlines strategies to construct such a multilingual dictionary for the measurement of migration frames (i.e., migration related subtopics that are based on varying problem definitions or causal interpretations) in Spanish, English, German, Swedish, Polish, Hungarian, and Romanian print and online news articles.

While we focus in this second part on the measurement of migration frames in traditional mass media (e.g., printed and online), the introduced strategies and techniques go beyond this particular case and can potentially be applied to manifold types of text data. At later stages of this project, the best practices and techniques identified will be applied to the measurement of other concepts (i.e., actors, tone) and will be adjusted in order to analyse other text types (i.e., social media), and thus contribute greatly to the overarching goal of this project, which is the mapping of media discourses on intra-EU mobility over time and across countries.

Part 1: Multilingual Dictionary Construction: A Review

With the increasing use of computer-assisted content analysis methods, dictionaries have become a decisive tool for concept measurement in digital texts. As a toolkit, "a dictionary is a set of words, phrases, parts of speech, or other word-based indicators (e.g., word length, number of syllables) that is used as the basis for a search of texts" (Neuendorf, 2002, p. 126-127). Researchers have several possibilities to select the best performing, most valid or most appropriate dictionaries to perform the measurement task. They apply dictionaries that are available open-source or commercially[1], modify and adjust these dictionaries to their own needs, or construct them from scratch, and apply strategies that include

[1] Widely-used dictionary examples – all in English - include the "General Inquirer" (Stone, Dunphy, & Smith, 1966) with its wide range of targeted concepts (i.e. sentiment, affect, cognition), the "DICTION" dictionary (Hart, 1984) originally developed to analyse rhetoric and political speech, and the "LIWC dictionary" (Linguistic Inquiry and Word Count) (Pennebaker, Booth, & Francis, 2007) to measure psychological concepts (i.e., emotions).

deductive or inductive approaches, or a mix of both. When researchers choose a dictionary approach and are interested in concepts that go beyond those that can be measured with available dictionaries, they will have to construct new customized dictionaries (Loughran & McDonald, 2011; Grimmer & Stewart, 2013). The necessity of this task becomes particularly clear in the attempt to conduct an analysis of a multilingual text corpus. Dictionaries in languages other than English are rare (Boumans & Trilling, 2016; Pang & Lee, 2008) and finding dictionaries that match specific research interests in these languages, and which are also comparable across languages, is close to impossible.

Generally, most of the literature on computer-assisted content analysis has dealt with English language texts (Pang & Lee, 2008). Obviously, this is not due to a lack of research questions that would greatly benefit from analyses drawing on multilingual dictionaries. There are, or at least to some extent were, manifold reasons for the delayed attention to languages other than English, and for the lack of simultaneous work with multiple languages. Among these reasons are: the tremendous effort needed to build the necessary resources for “only” one language (Laver, Benoit, & Garry, 2003; Young & Soroka, 2012), arguably, the dominance of English as the language of science (Ammon, 2001), the scarcity of multilingual resources and tools that assist the text analysis, and the high costs to hire professional translators, and the immaturity of machine translation (MT) technology.

Today, many of these factors are in a time of upheaval and changing. For example, multilingual text corpora, multilingual text analysis resources (e.g., dictionaries, annotated corpora), as well as other helpful tools such as parallel corpora or multilingual thesauri, are becoming increasingly available. As a consequence, work on “analysis strategies”, which look at how to ideally use and combine analysis resources and tools, has become an active research field. The first social science research teams constructed and applied dictionaries for multilingual text analysis (e.g., Baden & Stalpouskaya, 2015; Benoit, Schwarz, & Traber, 2012). Currently missing, however, is a systematic review and empirical comparison of different multilingual dictionary construction strategies and analysis approaches. This makes the roadmap outlined in part 2 of this deliverable not only highly relevant for the achievement of the goals set in this project, the study of European media discourses, but also particularly timely for the research field of comparative communication science in general.

Working with Dictionaries

Researchers using a dictionary for the analysis of a text corpus employ a top-down approach. This means that the corpus (consisting of documents, text entities such as articles or sentences) is searched with the help of a predefined list of words/word stems and phrases that represent the concept(s) of interest. The analysis strategy assumes that these target concepts are reflected through dictionary features (i.e., keywords). Feature counts (i.e., frequency of features per text) offer a reliable analysis of a text.

The construction of a new dictionary is relatively straightforward for clearly defined target concepts such as the occurrence of a specific actor in the news. For this purpose, a dictionary could simply include variations of the person's name or position. However, such a strategy reaches its limits quickly when it comes to the dictionary construction for less clear-cut concepts, such as frames or sentiment. The level of difficulty of the two central construction steps, feature pre-selection (i.e., identification of potentially relevant keywords) and feature evaluation (i.e., re-assessing the appropriateness of these keywords and their usefulness for the final feature selection), increases considerably.

Advantages, challenges and in particular limitations of computer-assisted content analysis methods for social sciences, especially in contrast to traditional manual content analysis, have been comprehensively discussed elsewhere (e.g., Bouman & Trilling, 2016; Grimmer & Stewart, 2013; Riffe, Lacy, & Fico, 2014; Young & Soroka, 2012). In short, the main benefits of text analysis using a dictionary approach are the perfect reliability (i.e., the ability to produce consistent measures after repeated trials) and capability to process large quantities of text. The main challenge is validity; meaning whether we are actually measuring the concepts, which we are interested in. "Bag-of-words" text analysis approaches, such as the dictionary approach, process individual words regardless of order and context, and as a consequence wrongly assume "semantic independence" (Young & Soroka, 2012, p. 209). It is clear that computer-assisted methods, which classify and process text, are not equivalent to manual coding or to human understanding of text. Quoting Grimmer and Stewart (2013), “all quantitative models of language are wrong—but some are useful.” (p. 269). Overall, the construction of good-quality dictionaries has been described as “very difficult” (Young & Soroka, 2012) and an “extensive effort” (Laver, Benoit, & Garry, 2003, p. 312) that requires time, money and strong collaboration with human coders (i.e., experts who help in the validation process). Both, feature pre-selection and feature evaluation, are strongly impaired by one’s “subjective conception” and “limited domain knowledge” (Burscher, Odijk, Vliegenthart, De Rijke, & De Vreese, 2014, p. 192).

The outlined advantages, challenges and weaknesses apply not only to monolingual dictionaries. We argue that some of the issues are even more critical in a comparative multilingual context (e.g., naturally limited domain knowledge of multiple national contexts). Some additional concerns (e.g., machine translation errors) will become clear within the next paragraphs where we outline the steps of multilingual dictionary construction.

Multilingual Dictionaries

Social scientists have started to integrate computer-assisted text analytic methods into their work with multilingual text corpora. When using dictionaries, they usually decide for one of the following two approaches, thus not being able to compare them to each other: (approach A) the application of a multilingual dictionary to a multilingual text corpus, and (approach B) the translation of the multilingual text corpus into a target language and the application of a dictionary in this language.

With regards to approach (A), the construction of a multilingual dictionary receives most attention. Such a dictionary aims to hold feature lists in different languages but for the measurement of the same concept. The interrelatedness of the language-specific feature lists depends on the goal of the analysis:

On the one hand, in a non-comparative framework, a feature list in one language often supports the creation of a feature list in another language, as long as the studied concept is not understood diametrically differently in the different cultural contexts. This new and translated feature list is subsequently perfectly adaptable to the respective context.

Ultimately, the two multilingual feature lists can be detached from each other and used separately in the different (country-) contexts. Up to now, this has been the predominate direction of development and usage of multilingual feature lists in the social sciences (e.g., Duval & Pétry, 2016; Sevenans, Albaugh, Shahaf, Soroka, & Walgrave, 2014).

Applying a comparative perspective, the aim is often to identify cross-country similarities and differences with regard to one concept. For this purpose, a multilingual dictionary includes several language-specific feature lists that take context (i.e., country-/regional) into account, and at the same time map comparable concepts that describe the general discourse. Such endeavours, although highly beneficial for the studying of cross-national media discourses on topics such as climate change, or migration, are still scarce. An exception is the INFOCORE project (Baden & Stalpouskaya, 2015), where the research group used a multi-step mixed-method strategy to construct concept keyword lists in nine languages. First, native speakers constructed language-specific lists with relevant concepts and related indicators (words) that were based on their work with monolingual text samples. Concepts, and later indicators, were then compared, integrated, and revised across languages. Cross-checking across languages, thesauri, word frequency analysis, as well as disambiguation strategies further assisted to homogenize and to improve the feature lists. This approach included many feedback loops and a strong collaboration with native speakers.

A very different approach (approach B) to obtaining an idea about the content of a multilingual text corpus is to first translate the entire corpus into one language, and to subsequently, apply analysis resources designed for a single language. Lucas et al. (2015), for example, translated an originally Chinese-Arabic document-term matrix into English and applied the Structural Topic Model, a topic model that can control for the original language. Working with a dictionary using this approach fully relies on monolingual, mostly English dictionaries, which are either adapted and refined or, in a "minimalist" approach, applied as they are**.** Benoit, Schwarz, and Traber (2012) used this method in their analysis of policy positions in legislative speeches (originally German, French and Italian) in Switzerland. They contrasted German and English as target language of the initial translation step (see other examples from political science [Pennings, 2011] or computer science [Denecke, 2008]).

While it is rather likely that approach B misses out on important context specific features, it is also an open question as to how important such context-specific features (e.g., “mojados” in Mexico; “boat people” in Spain or Australia) actually are in a comparative analysis as compared to non-context-specific features (e.g., “refugee” or “migrant”). The great advantage of approach B certainly is that it is a short-cut, which bypasses the labor-intensive effort of selecting features for multiple languages (approach A). But at what price? At this point, we know nothing about how well both approaches perform in direct comparison to each other.

Steps of Multilingual Dictionary Construction

The first important task for the construction of a multilingual dictionary is the definition of the target concept. After specifying what the dictionary should actually be measuring, the following steps are the pre-selection of dictionary features, the pre-processing of both the features and the text corpus, and finally the evaluation of different construction steps. All these steps are performed in an iterative manner (i.e., insufficient evaluation results may lead to a further specification of features, which in turn would have to be evaluated again).

Feature pre-selection

When it comes to the selection of features, researchers may choose among different techniques.

Within monolingual dictionary construction, social scientists have applied several strategies, which include among others: Extracting seemingly relevant sentences, words and phrases from text corpus samples (e.g., Vliegenthart & Roggeband, 2007), combining available dictionaries (e.g., Lawlor, 2015; Young & Soroka, 2012), consultation with human experts (e.g., Bengston & Xu, 1995) or, making use of the “wisdom of the crowd” (e.g., Haselmayer & Jenny, 2017). Researchers also make use of available resources in other languages. When they do so, they start off with a monolingual, mostly English-language template, that is first translated word by word into the target language, and in a second step enriched by working with various language-specific tools. Duval and Pétry (2016), for example, selected this strategy for their creation of the French Lexicoder Sentiment Dictionary (LSDFr). After a manual translation of the source dictionary, the English Lexicoder Sentiment Dictionary

(LSD), they applied stemming, eliminated duplicates, added synonyms, and worked with KWIC (keyword in context), as well as stop word lists (Duval & Pétry, 2016). Following a similar approach but with regard to topics, the Dutch-language Lexicoder Topic Dictionary was constructed (Sevenans, Albaugh, Shahaf, Soroka, & Walgrave, 2014; see also Gao, Hao, Li, Gao, & Zhu, 2013).

So far only applied in monolingual contexts, an alternative automated approach to creating a dictionary that fully emerges from the text corpus alone, is based on principle component analysis (i.e., frequently co-occurring words that form word clusters) (e.g., Greussing & Boomgaarden, 2017). Similarly, Lawlor and Tolley (2017) searched in their text corpus for the most frequently used words and phrases, then applied hierarchical clustering and examined whether terms that clustered together formed a logical frame. These terms were then used as features to construct English-language dictionaries for the automated measurement of frames (see also Balaban, Meza, & Vincze, forthcoming; McLaren, Boomgaarden, & Vliegenthart, 2017).

Great freely-available resources to extract further multilingual dictionary features are large-scale multilingual parallel language resources, such as those made available by the European Commission's Joint Research Centre (JRC) (Steinberger et al., 2014). Examples include the Digital Corpus of the European Parliament (DCEP) (Hajlaoui, Kolovratnik, Väyrynen, Steinberger, & Varga, 2014), which social science projects are starting to use to correct dictionaries (e.g., Proksch, Lowe, & Soroka, in press). Other useful resources are the "JRC-EuroVoc" (Steinberger et al., 2014), a multilingual thesaurus, or "JRC-NAMES" (Ehrmann, Jacquet, & Steinberger, 2017), which includes variants of names (persons, organizations, events) extracted from hundreds of millions of news articles since 2004 in 21 languages.

Following this review of recent multilingual dictionary construction projects, one needs to conclude that feature selection strategies are manifold and versatile. Nevertheless, the field lacks a systematic comparison of these different strategies.

Pre-processing of features and the target text corpus

Prior to analysis, dictionary features and words in the target text corpus have to be pre-processed in order to achieve greatest and most accurate matching. Commonly applied techniques are stemming (i.e., reducing words to their *word stem*) or lemmatizing (i.e., determining the base form of a word following its intended *meaning*), conversion to lower case and – exclusively for the target text corpus – the removal of stop words (i.e., very common and short function words), and punctuations.

Translation is the central pre-processing step for multilingual text analysis. It is usually applied before the stemming or lemmatizing of words. Both human translation and machine translation (MT) have played a major role for multilingual dictionary construction. It is argued that MT technology, evolving from phrase-based to neural machine translation models, has matured. Though it may not outperform manual translation, it can complement the work with multilingual text in meaningful ways (e.g., Aslerasouli & Abbasian, 2015; Balahur & Turchi, 2014; Lotz & Van Rensburg, 2014).

Given the optimization through automated translation procedures, it is an open question as to how beneficial the often costly (both in terms of time and financial resources) collaboration with native speakers still is. Comparative research projects with a large number of countries and languages may thus be especially interested in empirical assessment of human vs. machine feature translation.

Evaluation approaches

There are different techniques and criteria to consider for the evaluation of a multilingual dictionary. A frequently applied technique to assess the quality of a dictionary is to compare dictionary coding decisions to manual coding decision (e.g., Young & Soroka, 2012). This technique contrasts the final output of different coding processes (dictionary vs. manual coding) and the application of different dictionaries, as well as quantifies their performance. Similarly, to conduct an inter-coder reliability test for two manual coders, one would compare the manual with the dictionary based coding decisions, for example, using Krippendorff’s alpha (i.e., a procedure by which the agreement between codings is compared and evaluated, Krippendorff, 2004). After all, manual classifications are still referred to as the “gold standard”, or as the “most reasonable benchmark” (Rauh, 2017, p.

9), acknowledging the obvious differences between machines and humans in text processing approaches.

Other means to evaluate multilingual dictionaries focus on specific construction steps. Individual features are, for example, selected and evaluated based on their ability to, on the one hand represent the target concept, and, on the other hand to match the vocabulary of the target text corpus. Related to these evaluation objectives is the dilemma of generalizability vs. domain-specific knowledge discovery (demonstrated here: Loughran & McDonald, 2011). Should the dictionary rather aim at measuring the concept in wide applications, or be customized for the application to one specific text corpus? While an answer to this question depends certainly on the individual intended purpose and target concept, researchers should adapt dictionary features to the text domain in order to obtain meaningful results. The domain refers to the text type (e.g., language in legislative texts, differs from news texts; language in print articles differs from social media posts) and in a multilingual framework it also refers to the respective language-, and country-specific context. Different languages have for example a different diversity of words to express the same or a similar meaning (richness of a language). As another example, languages differ in terms of their morphologic complexity. For example, Hungarian, a Finno-Ugric language, is a highly inflective and agglutinative language which requires special efforts (Pajzs et al., 2014) such as the application of customized pre-processing tools (e.g., lemmatizing). Features have to be assessed in terms of their consideration of country-specific contexts. It is clear that a dictionary constructed to measure sentiment is less affected by country contexts than a dictionary designed to measure the occurrence of relevant political actors. Designed for the measurement of topics or frames from a comparative perspective across countries, a multilingual dictionary is evaluated by its ability to first, account for the individual national discourses, and second, to include elements that are part of a supranational discourse (i.e., general components likely to occur in any national context).

Machine translation (MT) is a key pre-processing step of multilingual text analysis and, therefore, also a central subject of evaluation. MT quality is evaluated based on human assessment (e.g., accuracy and fluency) and, with increasing popularity, automatic metrics such as BLEU, NIST, METEOR scores, which compare a machine translation output with human translation (Banerjee & Lavie, 2005; Doddington, 2002; Papineni, Roukos, Ward, & Zhu, 2002). Both methods serve different purposes, have different strengths and are ideally applied in combination. Regardless of the chosen method, MT quality is assessed for language pairs (e.g., English <-> German). MT quality is different for each language pair, as well as for each individual text and depends on the direction of translation, the chosen MT technology, and changes with further improvement of these technologies (Koehn, 2009). As general rules, translation quality is likely to be related to 1) the similarity of languages (e.g., English and Spanish are more similar than English and Hungarian) and 2) the spread of a language (Spanish is widely spoken, Catalan is not widely spoken). MT quality is often higher if English is chosen as the target language (e.g., translation from German into English), and lower if English is used as the source language (e.g., translation from English into German) (e.g., Durrani, Haddow, Heafield, & Koehn, 2013).

The strategies to construct a multilingual dictionary, which have been analysed here contribute to a roadmap for the construction of a multilingual migration frame dictionary, which will be presented next.

Part 2: The Roadmap to Constructing a Multilingual Migration Frames Dictionary

We now present a roadmap for the design of a multilingual dictionary to measure migration frames in European media discourse. We chose frames relating to the concept of migration as central target concepts for this project; lessons learnt here can also be used for the comparatively simpler construction of tone and actor dictionaries. From the different text corpora (i.e., traditional mass media, social media) that we collect and map within this project, we select the following corpus for the construction and evaluation of the multilingual migration frame dictionary.

Text Corpus

The selected text corpus is the largest text corpus mapped by this project (about *N* = 1.5 mil articles), and is used to provide a historical analysis of European medial discourse on migration and intra-EU mobility in order to better understand public opinion about free movement. This text corpus is multilingual, and consists of print and online news articles dealing with emigration and immigration that were published between January 2000 and December 2017 in Spain, the UK, Germany, Sweden, Poland, Hungary and Romania. Table 1 (Appendix) includes a list of media outlets for which this project collected news articles. Further, it lists the language-specific search strings together with their recall and precision measures (i.e., measures, which assess the quality of these search strings). The languages are all written in Latin script and belong to different language families and subfamilies, namely the Uralic and Indo-European language family, with the main subfamilies being Germanic languages (English, German, Swedish), Romance languages (Spanish, Romanian), and Slavic languages (Polish) (Beekes, 2011). These seven languages are the native languages for over half of the people in the European Union (Special Eurobarometer 386, 2012).

Concept Definition

Studying media coverage on migration with computer-assisted methods, we look at frames as topics formed through re-occurring patterns of specific words that help us categorize documents (Jacobi, van Atteveldt, & Welbers, 2016). *Frames* can be seen as schemes of interpretation that promote a particular problem definition or causal interpretation of an issue (see Entman, 1993; Goffman, 1974); in this case – among others – a causal interpretation of emigration and immigration in Europe. Through framing, media provide a context for individual audiences' understanding and public discussion of a policy issue. Media frames on migration have been examined – mainly using manual content analysis – in numerous national contexts.[2] As the economy, welfare, social action, security, and/or culture frames have been identified to be the most relevant in relation to migration coverage (e.g., Eberl et al., 2017), they will also be the focus of this project. Table 2 presents the five frames and shows how each of them is conceptualized.

[2] e.g., Spain (Checa & Arjona, 2011), the UK (Caviedes, 2015), Germany (Helbling, 2014), Sweden (Horsti, 2008), Poland (Galasińska, 2010), Hungary (Vicsek, Keszi, & Márkus, 2008), Romania (Light & Young, 2009).

Table 2: Manual Codebook Excerpt: The Economy, Welfare, Social Action, Security and Culture Frame

Frame	Items (Manual Content Analysis)	Text Example (Article Excerpt)
Economy	Does the article refer to economy/budget-related aspects of migration? Does the article refer to labour-related aspects of migration?	“A drive to encourage more foreign workers to move to Scotland was launched yesterday, in an attempt to reverse the country’s declining population and flagging economic growth. The population of Scotland is in decline, and ministers believe they must attract more immigrant workers to stave off long-term problems.” (Source: The Guardian)
Welfare	Does the article refer to welfare-related aspects of migration? (By welfare, we refer to the areas public education, public healthcare, public housing, public family support, unemployment support, state subsidies (food, electricity, etc.), pension/retirement or state/public services in general.)	“Related: David Miliband: failure to take in refugees an abandonment of UK’s humanitarian traditions “The situation on the islands is dramatic in terms of the sheer numbers flowing in, lack of shelter and ever worsening hygiene conditions,” Local NGO’s and volunteers, working around-the-clock to support insufficient state services now stretched to breaking point, described the situation as “utterly overwhelming.” (Source: The Guardian)
Social Action	Does the article refer to social action for/in support of migrants/migration? charity, donations, volunteering, solidarity, “welcome culture”, fundraising events, holding awareness events, sponsoring, etc.	“For the 20 or so refugees gathered in the warmth of Abigail Housing's base, up some back steps on an industrial estate just outside Bradford city centre, this is their last hope. It is a source of profound anger among the volunteers who provide a room, food parcel and £15 a week to save these people from homelessness and poverty that British law condemns them to destitution.” (Source: The Guardian)
Security	Does the article refer to security and/or crime-related aspects of migrants/migration? (Those who are held responsible for crime are migrants. For violent crime, those whose security is threatened can be both non-migrants and/or migrants.)	“According to his father, the refugee Hussein K., who was accused in the Freiburg murder trial before the Youth Chamber, is considerably older. An official document shall indicate 29 January 1984 as the date of birth. The defendant's father, who lives in Iran, told the court that in a telephone call. Hussein K. would have been 33 years old, at the time of the crime almost 14 months ago he would have been 32 years old. He himself had stated 17 years. He is accused of murder and particularly serious rape.” (Source: Die Zeit, machine translated, Google Translate)
Culture	Does the article refer to the promotion of a culturally diverse society? Does the article refer to the preservation of a culturally homogeneous society?	“Of course there's a floor where we all stand. Only that is much more diverse than we all do. Germany was never as homogeneous as after the Nazi dictatorship and World War II. Thank God, we have become more diverse again. There is also a variety in Christianity. The same must be said of Islamic fellow citizens.“ (Source: Die Zeit, machine translated, Google Translate)

Feature Pre-Selection Strategies

We strive towards a multilingual dictionary that contains feature lists for the economy, welfare, social action, security, and culture frames in seven languages. In fact, we construct many different feature lists per language and frame, to be able to systematically contrast their classification performance. We follow two paths in building a first basic stock of features for each migration frame.

Path 1. Exploiting the available monolingual dictionaries, we search for dictionaries with closely related categories and identified multiple relevant feature lists (Albaugh, Sevenans, & Soroka, 2013; Balaban, Meza, & Vincze, forthcoming; Greussing & Boomgaarden, 2017; Lawlor & Tolley, 2017; Vliegenthart & Roggeband, 2007). The features are either included in the published articles or were kindly provided by the authors in response to our request.

Path 2. Extracting the most frequent words from multilingual annotated sentences, we select – individually for each language – all sentences from the "The Comparative Manifesto Project" (Volkens et al., 2015a) database that are annotated with codes relating to our frames of interest. Based on these sentences, we are able to extract the most frequent words to be used as features for each frame. The CMP database contains parties' electoral manifestos from over 50 countries since 1945, annotated by native speakers. The CMP has been recommended and tested for the development of issue-specific dictionaries in multiple languages (Merz, Regel, & Lewandowski, 2016). Notably, the CMP has also been used for comparative research about immigration (e.g., Alonso & Fonseca, 2012; Burgoon, 2012; Lehmann & Zobel, 2018).

Table 3: Migration Frames: Dictionary Subcategories (Path 1) CMP Categories (Path 2)

Path 1						Path 2	
Dictionaries						Annotated Sentences	
Frame	Albaugh, Sevenans, & Soroka, (2013)	Balaban, Meza, & Vincze, (forthcoming)	Greussing & Boomgaarden (2017)[a]	Lawlor & Tolley (2017)	Vliegenthart & Roggeband (2007)	CMP Category	Code
Economy	Economy; Finance; Macro-economics; Labour	National costs	Economisation; Labour market integration	Economy	n/a	Economy related categories; Labour Groups: Positive; Labour Groups: Negative	401-416; 701; 702
Welfare	Services; Healthcare; Education; Social Welfare	n/a	n/a	n/a	n/a	Welfare State Expansion; Welfare State Limitation	504; 505
Social Action	Advocacy	Humanitarian / International	Humani-tarianism	n/a	n/a	Civic Mindedness: positive	606
Security	Security; Defence	Danger/ criminality; Securitisation	Criminality; Securitisation	Security; Crime	n/a	Law and Order: Positive	605
Culture	n/a	n/a	n/a	Ethnicity	Multicultural	Multiculturalism: Positive/ Negative; National Way of Life: Positive/ Negative	607; 608; 601; 602

Table 3 shows how we link the subcategories of other dictionaries (path 1) and sentence codes from the CMP codebook (Volkens et al., 2015b) (path 2) to the five target frames (i.e., the economy, labour market, welfare, social action, security, and culture frame).

Enrichment of Basic Feature Stock

Working with this basic stock of features, we apply two techniques to collect additional features. First, we make use of "JRC-EuroVoc" (Steinberger et al., 2014), the multilingual thesaurus of the European Union, to identify additional original language keywords. Second, we collaborate with native speakers (one per language), who have excellent knowledge of the individual national contexts. All have a graduate degree or are currently graduate students, either in communication, translation, or linguistics. They review the so far collected feature lists, propose refinements and additional words to be added.

Pre-Processing

All features, compiled up to this point, are combined to one list of keywords per frame. For each frame, we then translate all features into multiple languages, and thus create feature lists for seven languages. English is used as pivot (bridge) language. Hence, non-English features are, for example, first translated into English and then into other languages. Translation is performed by native speakers and in a second version using MT technology. MT is automated using Google Translate and the R package translateR (see Lucas et al., 2015). After machine translation, native speakers review the lists to identify MT issues, language-, as well as country-specific challenges. Finally, native speakers and researchers collaboratively evaluate the entire feature lists for each frame and language, and decide whether to include or exclude them. Evaluation criteria are, for example, the "concept fit" (e.g., the English feature "tax" fits to the "economy" frame, the feature "bring" does not) and the "ambiguity of a feature" (e.g., the German feature "betrieb" has several meanings; it refers to an enterprise, which would be relevant for the "economy frame", but it also refers to the grammatical past tense form of "to practise", which is too ambiguous clearly match any of the frames).

To keep track, every individual feature is categorized. The categories refer to different feature characteristics such as its current language (necessary since we the researchers do not understand all languages), source of origin (e.g., CMP, added by native speaker), its original language (the language of the text where it was extracted from), the type of translation (manual vs. MT), and its current status in the project (e.g., excluded because of MT translation issues). Filter variables based on these characteristics are the basis for the systematic comparison of different feature lists, and thus the contrast of different feature analysis approaches, selection strategies, and translation types. Table 4 depicts examples for labelled features per frame.

Table 4: Features (Examples) and Feature Characteristics (Selection)

Feature	Frame	Current Language	Source	Translation	Manual evaluation
growth	Economy	English	CMP	Not translated	Pending
munkanélküliség	Economy	Hungarian	Other Dictionary	Google Translate	Pending
acord colectiv	Economy	Romanian	EuroVoc Multilingual Thesaurus	Google Translate	Pending
housing	Welfare	English	CMP	Not translated	Pending
sjukhus	Welfare	Romanian	Other Dictionary	Google Translate	Pending
krankheit	Welfare	German	Manually added	Manual Translation	Pending
solidarity	Social Action	English	CMP	Not translated	Pending
darowizna	Social Action	Polish	Other Dictionary	Google Translate	Pending
hjälpa	Social Action	Swedish	CMP	Google Translate	Pending
terrorist	Security	English	Other Dictionary	Not translated	Pending
przestępstwo	Security	Polish	Other Dictionary	Google Translate	Pending
grenzübergang	Security	German	Other Dictionary	Manual Translation	Pending
diversity	Culture	English	Other Dictionary	Not translated	Pending
church	Culture	English	Other Dictionary	Not translated	Pending
respect	Culture	Romanian	Other Dictionary	Google Translate	Pending

Features are then stemmed with the Python-based application txtorg (see, Lucas et a., 2015) and converted to lowercase. In order to find the stemmed features in the text corpus, and subsequently count their occurrence per article, the text corpus is also pre-processed using txtorg. Stop words and punctuations are removed, words are stemmed, and converted to lowercase. The entire corpus is transformed into a document-feature matrix (one article per row, one unique feature per column).

In Part 1, we introduced an alternative approach (B) to analysing a multilingual text corpus with dictionaries. Here, the multilingual text corpus is translated into just one target language and a dictionary in this language is used for the search of features. To take this approach into account, we translate a randomly selected subsample (*N* = 1,000 articles per language) of the Spanish, German, Swedish, Polish, Hungarian, and Romanian article corpus into English with Google Translate.

Frame Evaluation and Measurement

We use the frame specific feature lists to classify the articles in the text corpus. Following Lawlor (2015), the frame salience per article depends on the number of frame-specific features. Thus, one article can relate more or less strongly to each specific frame as well as include multiple frames at the same time.

Based on the categorization of the features we can select different feature lists (e.g., just the Polish features for the measurement economy frame which originate from other previously used dictionaries and which were manually translated by native speakers) and track their usefulness for the classification of news articles. As demonstrated by Rauh (2017) in an application for monolingual text, the coding decisions resulting from the application of different feature lists are systematically compared with each other as well as to the human annotated results. To create such a manual benchmark, seven native speakers manually code 1,000 articles (those 1,000 articles that are translated for a test of approach B) in their respective native language. The native speakers assert the presence of each frame (yes, no) through the answer of frame-specific questions to the text (Table 2). The inter-coder reliability test for the manual content analysis includes two parts. First a test, where all seven coders classify 50 English language articles, is conducted and then a second one is carried out, where each native speaker codes 50 articles in his or her respective native language. These results are then compared with the coding decisions of an English native speaker, who codes the translated version of each of these 50 articles.

The classification decisions of different feature lists and decisions by human coders are compared through different quality measures (i.e., recall and Krippendorff’s alpha [Krippendorff, 2004]). Working with subsamples of the classified articles, native speakers then assess the chosen given classifications as well as deviations.

Summary, Next Steps and Embedding in the Framework of REMINDER

In this working paper, we provided a comprehensive review of the state-of-the-art in dictionary construction for automated text analysis. We focused specifically on multilingual dictionary construction, where research is still scarce since most studies tend to focus on

English language countries or refrain from comparative analyses altogether. After that, we presented a roadmap outlining key steps of automated dictionary construction, beginning with the corpus selection and concept definition. As our study focusses on the migration discourse in different EU countries, we outlined strategies of feature selection, refinement as well as validation to eventually be able to measure migration related frames in a multilingual European context using a dictionary approach.

Currently ongoing is the manual coding of 1,000 news articles per language and the enrichment of the collected basic feature stock. The next main tasks are feature translation, pre-processing steps and evaluation. Based on the comparison of the coding performance of the different dictionary construction approaches, we will identify the best performing dictionary. This will result in the provision of a new, validated, multilingual dictionary for measurement of migration frames in seven languages. We will also know how a multilingual dictionary applied to multilingual untranslated text (approach A) performs in contrast to English feature lists applied to a translated text corpus (approach B), which is a methodological achievement crucial for the remainder of the project. Following these comparisons, we will select the best performing dictionary and analysis approach to classify the previously described text corpus and track the salience of frames in media coverage across the seven European member countries between 2000 and 2017.

The analysis of text corpus of about 1.5 million articles, which consists of original mass media news articles published between 2000 and 2017, is central for the historical mapping of the emigration immigration discourse in Europe. Mapping this discourse will eventually allow us to understand citizens' attitudes toward free movement and may give us insights into possible sources of shifts in these attitudes. Also, the multilingual dictionary and analysis approaches are ready to be applied to two other text corpora collected for REMINDER. The first one is the text corpus, which will be used as input for the public opinion survey that is being conducted in cooperation with Work Package 9 (WP9).[3] The second text corpus is social media, which is currently being gathered in a multiple step

[3] It consists of online news articles, which are downloaded daily from media outlets' websites and print news articles, collected from online archives, for the same time period that the survey from WP9 is in the field. Data collection started at the same time as the first wave of the panel survey in December 2017 and is ongoing.

process.[4] The techniques and lessons learnt in the process of constructing a multilingual dictionary for measurement of migration frames are subsequently copied and adopted to measure the salience of other concepts (i.e., tone and actors) analyzed in the framework of REMINDER.

[4] We first identify the Facebook and Twitter accounts of identified important stakeholders (e.g., political or civil society actors) in every country of our sample and on the European level and then download posts, Tweets, comments and answers from the respective application programming interface (API) for every account. Moreover, we collect additional (meta-)data on user interactions and shared content to get a deeper understanding of the actual message.

Literature

Albaugh, Q., Sevenans, J., Soroka, S., & Loewen, P. J. (2013, June). *Lexicoder Topic Dictionaries, June 2013 versions, McGill University, Montreal, Canada.* Retrieved from www.lexicoder.com

Alonso, S., & Fonseca, S. C. da F. (2012). Immigration, left and right. *Party Politics, 18*(6), 865–884.

Ammon, U. (2001). *The dominance of English as a language of science: Effects on other languages and language communities*. Berlin: Walter de Gruyter.

Aslerasouli, P., & Abbasian, G. R. (2015). Comparison of Google Online translation and human translation with regard to soft vs. hard science texts. *Journal of Applied Linguistics and Language Research, 2*(3), 169-184.

Baden, C., & Stalpouskaya, K. (2015). *Common methodological framework: Content Analysis. A mixed-methods strategy for comparatively, diachronically analyzing conflict discourse.* INFOCORE Working Paper 2015/10. Retrieved from http://www.infocore.eu/wp-content/uploads/2016/02/Methodological-Paper-MWG-CA_final.pdf

Balaban, D., Meza, R. & Vincze, H. O. (forthcoming). *The role of religion in Romanian news of the refugee crisis. A clusters-based frame analysis.* Working Paper.

Banerjee, S., & Lavie, A. (2005, June). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. In *Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization, Ann Arbor, United States* (pp. 65-72). Retrieved from http://www.aclweb.org/anthology/W05-0909

Beekes, R. S. (2011). *Comparative Indo-European linguistics: an introduction*. Amsterdam: John Benjamins Publishing.

Bengston, D. N., & Xu, Z. (1995). *Changing National Forest Values: a content analysis*. Research Paper NC-323. United States Department of Agriculture, Forest Service, North Central Forest Experiment Station. St. Paul, MN, US.

Benoit, K., Schwarz, D., & Traber, D. (2012, June). *The sincerity of political speech in parliamentary systems: A comparison of ideal points scaling using legislative speech and votes.* Paper presented at the 2nd Annual Conference of European Political Science Association (EPSA), Berlin, Germany.

Boomgaarden, H. G., and Vliegenthart, R. (2009). How news content influences anti-immigration attitudes: Germany, 1993-2005. *European Journal of Political Research, 48*(4), 516–542.

Boomgaarden, H. G., Vliegenthart, R., De Vreese, C. H., and Schuck, A. R. (2010). News on the move: Exogenous events and news coverage of the European Union. *Journal of European Public Policy*, *17*(4), 506-526.

Boumans, J. W., & Trilling, D. (2016). Taking stock of the toolkit: An overview of relevant automated content analysis approaches and techniques for digital journalism scholars. *Digital Journalism*, *4*(1), 8-23.

Burgoon, B. (2012). Partisan embedding of liberalism: How trade, investment, and immigration affect party support for the welfare state. *Comparative Political Studies, 45*(5), 606–635

Burscher, B., Odijk, D., Vliegenthart, R., De Rijke, M., & De Vreese, C. H. (2014). Teaching the computer to code frames in news: Comparing two supervised machine learning approaches to frame analysis. *Communication Methods and Measures, 8*(3), 190-206.

Caviedes, A. (2015). An emerging 'European' news portrayal of immigration?. *Journal of Ethnic and Migration Studies, 41*(6), 897-917.

Checa, J. C., & Arjona, Á. (2011). Españoles ante la inmigración: el papel de los medios de comunicación [Spaniards' Perspective of Immigration. The Role of the Media]. *Comunicar, 19*(37), 141-149.

Denecke, K. (2008, April). Using sentiwordnet for multilingual sentiment analysis. In *Data Engineering Workshop, 2008. ICDEW 2008. IEEE 24th International Conference on* (pp. 507-512). IEEE.

Doddington, G. (2002, March). Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In *Proceedings of the second international conference on Human Language Technology Research, San Diego, United States* (pp. 138-145). Retrieved from http://delivery.acm.org/10.1145/1290000/1289273/p138-doddington.pdf?ip=131.130.173.239&id=1289273&acc=PUBLIC&key=9074CF143665B1C6%2E828B6F6B55044937%2E4D4702B0C3E38B35%2E4D4702B0C3E38B35&__acm__=1518001924_320706ce742748d425009b8785e76b5d

Durrani, N., Haddow, B., Heafield, K., & Koehn, P. (2013, August). Edinburgh's machine translation systems for European language pairs. In *Proceedings of the Eighth Workshop on Statistical Machine Translation Association for Computational Linguistics, Sofia, Bulgaria* (pp. 114-121). Retrieved from http://www.aclweb.org/anthology/W13-2212

Duval, D., & Pétry, F. (2016). L'analyse automatisée du ton médiatique: construction et utilisation de la version française du Lexicoder Sentiment Dictionary [Automated media tone analysis: construction and use of the French version of Lexicoder Sentiment Dictionary]. *Revue Canadienne de Science*, *49*(2), 197-220.

Eberl, J.-M., Heidenreich, T., Boomgaarden, H. G., Herrero B., Berganza, R., Allen, W., & Bajomi-Lazar, P. (2017). Discourses on intra-EU mobility and non-EU migration in european media coverage: A comprehensive literature review. Paper prepared as part of the REMINDER project. Retrieved from project website: http://www.reminder-project.eu/wp-content/uploads/2017/05/REMINDER-D8-1_Discourse_mobility_European_media_web.pdf

Ehrmann, M., Jacquet, G., & Steinberger, R. (2017). JRC-Names: Multilingual entity name variants and titles as Linked Data. *Semantic Web*, *8*(2), 283-295.

Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4), 51–58.

Galasińska, A. (2010). Leavers and stayers discuss returning home: Internet discourses on migration in the context of the post-communist transformation. *Social Identities, 16*(3), 309-324.

Gao, R., Hao, B., Li, H., Gao, Y., & Zhu, T. (2013). Developing simplified Chinese psychological linguistic analysis dictionary for microblog. In K. Imamura, S. Usui, T. Shirao, T. Kasamatsu, L. Schwabe, N. Zhong (Eds.) *Lecture Notes in Computer Science: Vol. 8211. Brain and Health Informatics BHI 2013* (pp. 359-368). Cham: Springer.

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. Cambridge: Harvard University Press.

Greussing, E., & Boomgaarden, H. G. (2017). Shifting the refugee narrative? An automated frame analysis of Europe's 2015 refugee crisis. *Journal of Ethnic and Migration Studies, 43*(11), 1749-1774.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis, 21*(3), 267-297.

Hajlaoui, N., Kolovratnik, D., Väyrynen, J., Steinberger, R., & Varga, D. (2014, May). DCEP-Digital Corpus of the European Parliament. In *Proceedings of Language Resources and Evaluation Conference, Reykjavik, Iceland.* (pp. 3164-3171). Retrieved from http://www.lrec-conf.org/proceedings/lrec2014/pdf/943_Paper.pdf

Hart, R. (1984). *Verbal Style and the Presidency: A Computer-Based Analysis.* Orlando, FL: Academic Press.

Haselmayer, M., & Jenny, M. (2017). Sentiment analysis of political communication: combining a dictionary approach with crowdcoding. *Quality & Quantity*, *51*(6), 1-24.

Helbling, M. (2014). Framing immigration in Western Europe. *Journal of Ethnic and Migration Studies, 40*(1), 21-41.

Hobolt, S. B. (2016). The Brexit vote: a divided nation, a divided continent. *Journal of European Public Policy*, 23(9), 1259-1277.

Horsti, K. (2008). Europeanisation of public debate: Swedish and Finnish news on African migration to Spain. *Javnost-the public, 15*(4), 41-53.

Jacobi, C., van Atteveldt, W., & Welbers, K. (2016). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, *4*(1), 89-106.

Koehn, P. (2009). *Statistical machine translation*. Cambridge: Cambridge University Press.

Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research*, *30*(3), 411-433.

Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *American Political Science Review*, *97*(2), 311-331.

Lawlor, A. (2015). Local and national accounts of immigration framing in a cross-national perspective. *Journal of Ethnic and Migration Studies*, *41*(6), 918-941.

Lawlor, A., & Tolley, E. (2017). Deciding who's legitimate: News media framing of immigrants and refugees. *International Journal of Communication*, *11*, 967-991.

Lehmann, P. & Zobel, M. (2018). Positions and saliency of immigration in party manifestos: A novel dataset using crowd coding. *European Journal of Political Research.* Online published first. doi:10.1111/1475-6765.12266

Light, D., & Young, C. (2009). European Union enlargement, post-accession migration and imaginative geographies of the 'New Europe': Media discourses in Romania and the United Kingdom. *Journal of Cultural Geography, 26*(3), 281-303.

Lotz, S., & Van Rensburg, A. (2014). Translation technology explored: Has a three-year maturation period done Google Translate any good?. *Stellenbosch Papers in Linguistics Plus*, *43*(1), 235-259.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance*, *66*(1), 35-65.

Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, A., & Tingley, D. (2015). Computer-assisted text analysis for comparative politics. *Political Analysis*, *23*(2), 254-277.

McLaren, L., Boomgaarden, H., & Vliegenthart, R. (2017). News Coverage and Public Concern About Immigration in Britain. *International Journal of Public Opinion Research*. Advanced online publication. doi: 10.1093/ijpor/edw033.

Merz, N., Regel, S., & Lewandowski, J. (2016). The Manifesto Corpus: A new resource for research on political parties and quantitative text analysis. *Research & Politics*, *3*(2), 1-8.

Neuendorf, K. A. (2002). *The content analysis guidebook*. Thousand Oaks, CA: Sage Publications.

Pajzs, J., Steinberger, R., Ehrmann, M., Ebrahim, M., Della Rocca, L., Simon, E., & Váradi, T. (2014, May). Media monitoring and information extraction for the highly inflected agglutinative language Hungarian. *Proceedings of the 9th edition of the Language Resouces and Evaluation Conference (LREC), Reykavik, Iceland*, p. 2049-2056.

Pang, B. & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval, 2*(1–2), 1-135.

Papineni, K., Roukos, S., Ward, T., & Zhu, W. J. (2002, July). BLEU: a method for automatic evaluation of machine translation. *In Proceedings of the 40th annual meeting on association for computational linguistics, Philadelphia, United States* (pp. 311-318). Retrieved from http://delivery.acm.org/10.1145/1080000/1073135/p311-papineni.pdf?ip=131.130.173.239&id=1073135&acc=OPEN&key=9074CF143665B1C

6%2E828B6F6B55044937%2E4D4702B0C3E38B35%2E6D218144511F3437&__acm__=1518001841_5b96bf6b2dbb720fe20118e3d3d293de

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2001). *Linguistic Inquiry and Word Count: LIWC.* Austin; TX: LIWC.net.

Pennings, P. (2011). Assessing the 'Gold Standard' of party policy placements: Is computerized replication possible?. *Electoral Studies*, *30*(3), 561-570.

Proksch, S-O, Lowe, W., & Soroka, S. (in press). *Multilingual sentiment analysis: A new approach to measuring conflict in parliamentary speeches.*

Rauh, C. (2017). Validating a sentiment dictionary for German political language. *Working paper.* Retrieved from https://www.researchgate.net/profile/Christian_Rauh2/publication/321488707_Validating_a_sentiment_dictionary_for_German_political_language/links/5a2509dbaca2727dd87e73e7/Validating-a-sentiment-dictionary-for-German-political-language.pdf

Riffe, D., Lacy, S., & Fico, F. (2014). *Analyzing media messages: Using quantitative content analysis in research.* New York, NY: Routledge.

Sevenans, J., Albaugh, Q., Shahaf, T., Soroka, S., & Walgrave, S. (2014, June). *The automated coding of policy agendas: A dictionary-based approach (v. 2.0).* Paper presented at the 7th Annual Conference of the Comparative Agendas Project (CAP), Konstanz, Germany.

Special Eurobarometer 386 (2012*). Europeans and their Languages.* Conducted by TNS Opinion & Social at the request of. Directorate-General Education and Culture, Directorate-General for Translation and Directorate-General for Interpretation. Survey co-ordinated by the European Commission.

Steinberger, R., Ebrahim, M., Poulis, A., Carrasco-Benitez, M., Schlüter, P., Przybyszewski, M., & Gilbro, S. (2014). An overview of the European Union's highly multilingual parallel corpora. *Language Resources and Evaluation, 48*(4), 679-707.

Stone, P. J., Dunphy, D. C., & Smith, M. S. (1966). *General Inquirer: A Computer Approach to Content Analysis.* Boston, MA: The MIT Press.

Vicsek, L., Keszi, R., & Márkus, M. (2008). Representation of refugees, asylum-seekers and refugee affairs in Hungarian dailies. *Journal of Identity and Migration Studies, 2*(2), 87-107.

Vliegenthart, R., & Roggeband, C. (2007). Framing immigration and integration: Relationships between press and parliament in the Netherlands. *International Communication Gazette*, *69*(3), 295-319.

Vliegenthart, R., Schuck, A. R. T., Boomgaarden, H. G., & de Vreese, C. H. (2008). News Coverage and Support for European Integration, 1990-2006. *International Journal of Public Opinion Research*, *20*(4), 415–439.

Volkens, A., Lehmann, P., Matthieß, T., Merz, N., Regel, S., & Werner, A. (2015a). *The Manifesto Data Collection. Manifesto Project (MRG /CMP / MARPOR). Version 2015a. Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB). Version 2015a.* Berlin: Wissenschaftszentrum Berlin für Sozial-forschung (WZB).

Volkens, A., Lehmann, P., Matthieß, T., Merz, N., Regel, S., & Werner, A. (2015b). *The Manifesto Project Dataset - Codebook. Manifesto Project (MRG / CMP / MARPOR). Version 2015a.* Berlin: Wissenschaftszentrum Berlin für Sozial-forschung (WZB). Retrieved from https://manifestoproject.wzb.eu/down/documentation/codebook_MPDataset_MPDS2015a.pdf

Young, L., & Soroka, S. (2012). Affective news: The automated coding of sentiment in political texts. *Political Communication*, *29*(2), 205-231.

Appendix

Table 1: Media Outlets per Country, Search Strings, Recall and Precision for Subsample

			Search string validation [c]				
Country	Media Outlets [a]	Search String [b]	N Coded articles	*N* Relevant Articles	Recall	*N* Retrieved Articles	Precision
Spain	Print (ABC, El Mundo, El País); Online (abc.es, elpais.com, elmundo.es, lavanguardia.com)	asilo! OR inmigra! OR refugiad! OR migrante! OR migratori! OR "sin papeles" OR "campo de desplazados" OR patera! OR emigra! OR "libre circulación" OR "fuga de cerebros"	2113	104	0.92	105	0.93
UK	Print (Daily Mail, Daily Mirror, Metro, The Daily Telegraph, The Guardian, The Sun, The Times); Online (dailymail.co.uk, express.co.uk, telegraph.co.uk, thesun.co.uk, thetimes.co.uk)	asyl! OR immigrant! OR immigrat! OR migrant! OR migrat! OR refugee! OR foreigner! OR "undocumented worker!" OR "guest worker!" OR "foreign worker!" OR emigrat! OR "freedom of movement" OR "free movement"	3418	170	0.88	165	0.90
Germany	Print (Bild, Die Tageszeitung / taz, Die Welt, Frankfurter Allgemeine Zeitung, Frankfurter Rundschau, Süddeutsche Zeitung); Online (spiegel.de, sueddeutsche.de, taz.de, welt.de, zeit.de)	asyl! OR immigrant! OR immigriert! OR immigrat! OR migrant! OR migrat! OR flüchtling! OR ausländer! OR zuwander! OR zugewander! OR einwander! OR eingewander! OR gastarbeiter! OR "ausländische arbeitnehmer!" OR emigr! OR auswander! OR ausgewander! OR personenfreizügigkeit! OR arbeitnehmerfreizügigkeit! OR "freier personenverkehr!"	1203	119	0.89	111	0.94
Sweden	Print (Aftonbladet, Dagens Industri, Dagens Nyheter, Expressen/ GT/KvP vardag, Svenska Dagbladet, Metro); Online (dn.se)	asyl! OR invandr! OR migrat! OR migrant! OR flykting! OR utlänning! OR immigrant! OR ensamkommande! OR EU-migrant! OR "utländsk bakgrund" OR gästarbetar! OR "utländsk! arbet!" OR papperslös! OR emigr! OR utvandr! OR "fri rörlighet"	1244	85	0.67	60	0.93
Poland	Print (Gazeta Wyborcza, Rzeczpospolita, Dziennik Gazeta Prawna), Online (gazeta.pl, wyborcza.pl)	azyl! OR migr! OR imigr! OR uchodźca OR uchodźcy OR uchodźcę OR uchodźcą OR uchodźco OR uchodźców OR uchodźcom OR uchodźcami OR uchodźcach OR cudzoziem! OR obcokrajow! OR "robotni! z zagranicy" OR "pracowni! z zagranicy" OR gastarbeiter! OR "nielegaln! pracowni!" OR emigr! OR "swobodny przepływ"	1391	63	0.77	63	0.76
Hungary	Print (Blikk, Magyar Hirlap, Metropol, Magyar Nemzet, Napi Gazdaság, Népszabadság, Nepszava); Online (blikk.hu, magyarhirlap.hu, nepszava.hu, hir24.hu, napi.hu, Index.hu, hvg.hu)	menedék! OR bevándor! OR immigrá! OR migrá! OR menekült! OR vendégmunk! OR elvándor! OR emmigrá! OR mozgásszabadság!	1200	102	0.83	101	0.81
Romania	Print (Adevarul, Evenimentul Zilei, Jurnalul National, Romania Liberia, Ziarul Financiar,)	azil! OR imigra! OR migra! OR emigra! OR refugiat! OR "muncitor străin" OR "muncitori străini" OR "muncitorii străni" OR "muncitorilor străni" OR "lucrător străin" OR "lucrători străini" OR "lucrătorii străini" OR "lucrătorilor străini" OR "libera circulație a persoanelor" OR "libertatea de circulație a persoanelor" OR "libera circulație a lucrătorilor" OR "libertatea de circulație a lucrătorilor"	1415	63	0.71	61	0.71

Note. [a] Time periods for which we have access to archives differs across outlets. For each country, we have data from 2000-2017 for at least one media outlet; [b] The search strings, and correspondingly the news articles are in the most-widely spoken language for each country (e.g., nor Catalan, or Basque but Spanish for Spain); [c] For each country, a native speakers read and coded between *N* = 1200 and *N* = 3418 news articles. The randomly selected articles were published in 1-2 media outlets per country. The native speakers coded the relevance (yes/no) and the retrieval (yes/no) when applying the search strings. Both types of information were used to calculate recall and precision.

REMINDER

ROLE OF EUROPEAN MOBILITY AND ITS IMPACTS IN NARRATIVES, DEBATES AND EU REFORMS

The REMINDER project is exploring the economic, social, institutional and policy factors that have shaped the impacts of free movement in the EU and public debates about it.

The project is coordinated from COMPAS and includes participation from 14 consortium partners in 9 countries across Europe

UNIVERSITY OF OXFORD

This project has received funding from the European Union’s Horizon 2020 research & innovation programme under grant agreement no 727072

COMPAS

CENTRE ON MIGRATION · POLICY & SOCIETY