

‘This is England, speak English!’: A corpus-assisted critical study of language ideologies in the right-leaning British press

David Wright

School of Arts and Humanities, Nottingham Trent University, Nottingham, UK

Gavin Brookes

Department of Linguistics and English Language, Lancaster University, Lancaster, UK

Abstract

This article examines right-leaning press representations of people living in the UK who can't speak English, or at least speak English well, following the 2011 Census, which was the first to ask respondents about their main language and proficiency in English. The analysis takes a corpus-assisted approach to critical discourse analysis, based on a 1.8 million-word corpus of right-leaning newspaper articles about ‘speak(ing) English’ in the years following this historic Census (2011 to 2016). The analysis reveals the tendency for the press to focus on immigrants – particularly in the contexts of education and health – who are represented with recourse to a series of argumentation strategies, or ‘topoi’. Over the course of this paper, we argue that these topoi are problematic, as they present paradoxes, obscure the role of the Government in ensuring integration, overlook the difficulties of language learning and cultural assimilation, and generally contribute to a broader anti-immigrant UK media narrative which serves to legitimise exclusionary and discriminatory practices against people from minority linguistic and ethnic backgrounds.

Keywords

Speak English, 2011 UK Census, Multilingualism, Language Ideology, Media, Immigrant Representation, Corpus-Assisted Critical Discourse Analysis

Introduction

In 2011, for the first time, the Census of the United Kingdom (UK) asked respondents to identify their ‘main language’ and, if the answer was something other than English, to indicate how well they could speak English. The results revealed that, of the 4.2 million usual UK residents whose main language was not English, 79 per cent (3.3 million people) said they could speak English ‘Well’ or ‘Very well’, while 17 per cent (726,000 people) said they could speak English but ‘Not well’, and 3 per cent (138,000 people) said they could not speak English at all (ONS, 2013). The result of the Census therefore showed remarkable linguistic homogeneity, with 99.74 per cent of UK residents (53.9 million people) having some command of English either as a main or additional language.

These historic Census items and their results were widely reported across the press (particularly by right-leaning newspapers) at the time of the Census, on the publication of the

resulting data in 2013, and have been drawn upon in a wide range of stories published in the years since. However, this press coverage has not been studied from a linguistic or critical discourse perspective. Yet, studying newspaper data is important in this context because the media has the potential not only to reflect but also powerfully influence understandings of and attitudes toward the social world for large numbers of people (Fowler, 1991). This is certainly the case for debates surrounding issues like language, immigration and citizenship since, as Blackledge (2005, p. 67) points out, ‘representations of minority ... groups in news media can validate diversity and solidarity, or contestation and differentiation’. This paper aims to address this gap in research by examining the UK print media’s response to the Census result. Based on a 1.8 million-word corpus of right-leaning UK national newspaper articles published between 2011 and 2016, this study takes a corpus-assisted approach to critical discourse analysis to identify the argumentation strategies, or ‘topoi’, on which the right-leaning press draw to represent people who can’t speak English in the years following this historic Census. We then interpret the observed topoi in terms of their underlying language ideologies which, following Vessey (2017, p.278), we understand to be ‘beliefs about languages (or a particular language) that are shared and that become so well-established that their origin is often forgotten by speakers; the beliefs accordingly become naturalized, perceived as common sense, and are socially reproduced’.

The article begins by introducing the 2011 Census and the relevant aspects of the UK socio-political context in which it took place, outlines the connections – as we and others view them – between language, identity and nationalism, and then reviews relevant research concerned with language ideologies in the press. Following this, the methodology section describes the compilation of the corpus and introduces our corpus-assisted approach to critical discourse analysis. The analysis outlines the recurring topoi used in the articles to represent people who can’t speak English. The implications of these topoi are then considered in the discussion section, before we conclude the article by reviewing the main findings and offering suggestions for future research on this topic.

Background

The 2011 Census, language and politics in the UK

The Office for National Statistics (ONS), the recognised national statistical institute of the UK, holds a national ‘Census’ survey every ten years. The Census aims to collate ‘the most complete source of information about the population’ to allow ‘central and local Government, health authorities and many other organisations to target their resources more effectively’ (ONS, 2016). The 2011 Census was the first to include specific questions about language. Specifically, Question 18 asked respondents, ‘What is your main language?’ and required them to tick one of two options: ‘English’ or ‘Other’. If the language was ‘Other’, there was a space to specify which language. Respondents whose main language wasn’t English were then required to move to Question 19, which asked, ‘How well can you speak English?’. To answer this question, respondents were required to tick one of the following options: ‘Very well’, ‘Well’, ‘Not well’, or ‘Not at all’. These questions are problematic for a number of reasons, not least because the

very notion of a ‘main language’ is ambiguous and likely to cause problems for people who are competent in more than one language (see Sebba (2017) for a critique of these Census items). However, for the purpose of this discussion we set these difficulties aside to focus on the results of Census and the press response to them. As mentioned in the previous section, the results for Questions 18 and 19 revealed that English was the main language for 92 per cent (49.8 million) of usual UK residents aged three years and over. Of the remaining 8 per cent (4.2 million) who had a different main language, 79 per cent (3.3 million) stated that they could speak English ‘Well’ or ‘Very well’, 17 per cent (726,000) responded that they could speak English but ‘Not well’, and 3 per cent (138,000) indicated that they could not speak English at all (ONS, 2013). Despite the Census showing that 99.74 per cent of usual residents (53.9 million people) were able to speak English either as a main or additional language, the immediate media response to the results focussed on the relatively small portion of the respondents who indicated that they could not speak English at all. For example, the BBC ran an article with the headline ‘138,000 speak no English – Census’ (BBC, 2013). Such a headline is indicative of the wider media response to the Census results – a response which was at best sensationalist, and at worst served to ‘feed xenophobia’ (Sebba, 2017, p. 279).

As Sebba (2017, p. 271) notes, the decision taken in 2007 to include a question in the 2011 Census asking respondents about their language coincided with ‘a period when language was very much part of the public policy agenda’. Sebba (2017, pp. 269–71) positions this agenda as emerging from the emphasis placed by successive UK Governments on the integration and cohesion of incoming immigrants to UK communities. First, the Community Cohesion Review Team was established in 2001 and considered the English Language as a national value, setting out expectations that everyone entering the country should use English in the interests of community cohesion. The Commission on Integration and Cohesion was formed in 2007 which called for a reduction in the provision of translated documents by local Government authorities on the basis that such translations undermine efforts to integrate non-English speaking residents. Therefore, Sebba (2017, p. 274) argues that the addition of the language-related questions to the Census in 2011 occurred against a policy backdrop that could be characterised as ‘pro-integration’ and ‘anti-multilingualism’.

Language issues have remained subject to scrutiny and policy changes since the Census and throughout the timeframe which will be covered by our analysis (2011–2016). English language proficiency has been a legislative requirement for UK citizenship applications since 1981, and in more recent years, language requirements and proficiency testing have changed in a number of ways. The British Nationality Act of 1981 stated that applicants for citizenship must have ‘sufficient knowledge of the English, Welsh or Scottish Gaelic Language’. The Nationality, Immigration and Asylum Act of 2002 added the requirement that applicants demonstrate sufficient knowledge about life in the UK. In 2005, the Knowledge Of Language and Life (KOLL) requirement was introduced as a statutory condition for citizenship applications and, in 2007, became a mandatory part of applications for permanent settlement in the UK. By 2011, the KOLL requirement could be demonstrated either by passing the ‘Life in the UK’ test or by obtaining an English for Speakers of Other Languages (ESOL) qualification. In 2013, at the mid-point of the time period covered by our analysis, the requirements were made more stringent, as it became compulsory for applicants both to pass the Life in the UK test *and* obtain a speaking and listening qualification in English at B1 CEFR

or higher from an accredited institution. Furthermore, prior to that, in 2010, language requirements and pre-entry English tests were extended to partners and spouses of British citizens, a group who had previously been exempt from such measures. In 2015, the lawfulness of this change was upheld in two highly-publicised cases wherein the Supreme Court ruled that pre-entry English tests of spouses did not breach their right to a private and family life. In 2016, the required linguistic proficiency of all immigrants to the UK (not just those seeking citizenship) was written into legislation, as section 19 of the Immigration Act 2014 amended the Nationality, Immigration and Asylum Act 2002 by inserting Article 8 of the European Convention on Human Rights (public interest considerations), which legislates:

It is in the public interest, and in particular in the interests of the economic well-being of the United Kingdom, that persons who seek to enter or remain in the United Kingdom are able to speak English, because persons who can speak English—

- (a) are less of a burden on taxpayers, and
- (b) are better able to integrate into society.

Most recently, the Immigration Act 2016 enforced a manifesto pledge made by the Conservative Government to ensure that every person who works in the public sector (e.g. in schools, hospitals, law enforcement, social care) and who has a customer-facing role speaks ‘fluent English’.

There is, therefore, a rich legislative and policy context, characterised by these key changes, for the period leading up to and following the 2011 Census. These contextual factors are worth bearing in mind, as they are likely to play a role in shaping public discourse surrounding matters pertaining to language, nationality and immigration. Sebba (2017, p. 269) attributes some of the anti-multilingualism agenda enacted by the Labour Government in the early 2000s, particularly in relation to Asian languages, to increased and sustained public concern ‘about immigration and race issues’ in the wake of terrorist attacks in New York City on 9/11 2001 and London in 2005. If this connection between public opinion on immigration and language-related Government policy changes stands true, then it may be the case that the UK’s decision to leave the European Union following a referendum in June 2016 – approaching the end-point of our data – flavours some of these immigration-language public media debates. Indeed, a headline from the *Guardian* on 1st June 2016, a few weeks before the aforementioned referendum, reads ‘Migrants must have good English in post-Brexit UK, says leave campaign’. The various policy and legislative changes in relation to language requirements for immigrants, and public attitudes towards immigration more generally, provide important context for the intersection of discourses around immigration, identity and language in the British print media, both in the years preceding the 2011 Census, which is the stimulus for this study, and in the years following the Census, which provide the data for our analysis.

Language, identity and nationalism

In the context of language ideological discussions, notions of language and national identity are often couched in terms of Anderson's (2006) 'imagined communities' (Barbour, 2000; Pavlenko & Norton, 2007). Anderson (2006, p. 6) posits that 'nations' are socially constructed communities, 'imagined' because members will never meet or know their fellow members, 'yet in the minds of each lives the image of their communion'. It is commonly held that 'language is a prime determinant of nationalist identity' (Billig, 1995, p. 29), and Blackledge (2009, p. 69) notes that one obvious way 'to link language and national identity is through language policy, planning and standardization practices which legitimize particular language varieties and link them to specific identities'. Indeed, the testing of migrants' proficiency in the official language(s) of the country to which they are moving is increasingly dominating public debate surrounding integration and citizenship (Milani, 2015). As long as language is marked as central to the concept of 'the nation', then language and linguistic variation (or the lack of) will inform individuals' *imagined* nation. Ideologies and practices hinging on the desirability or necessity of some languages over others create a 'linguistic market, dominated by the official language' of a given nation (Bourdieu, 1991, p. 45). Under such conditions, language becomes a means by which to determine 'who is in' and 'who is out', with linguistic minorities commonly "invented" by, and in relation to, the linguistic majority' as an inferior out-group (Blackledge, 2009, p. 69).

Ideological systems such as this can contribute to the notion that the imagined nation is a linguistically homogenous and, in many cases, monolingual society. In countries such as the UK, this imagined 'one language, one nation' myth (Piller, 2001, p. 261) is at odds with the heterogeneous, multilingual reality, and serves to exclude native speakers of minority languages (in this case, languages other than English) who may be denied access to, or membership of, the linguistically dominant majority community, whether that is real or imagined (Blackledge, 2005, p. 46). Such ideologies of inclusivity *versus* exclusivity can give rise to interethnic conflict within a nation where attitudes favouring monolingualism interlink with ideals of a monoethnic, monoreligious, monoideological state, in which differences are viewed as being 'dangerous' to the 'purity' of society (Blommaert & Verschueren, 1998, p. 265). As Blackledge and Pavlenko (2001, p. 243) note, this leads not only to the exclusion of those who are unable or unwilling to fit 'the norm', but also to the active discrimination against such individuals or groups. This view aligns with Lippi-Green's (2012, p. 68) 'standard language ideology', which proposes that an idealized nation-state has one perfect homogenous language. This ideology, Lippi-Green (2012, p. 74) argues, props open a 'back door' to discrimination and prejudice that would normally be considered racism (e.g. discriminating against people on the basis of their race or ethnicity), but which is couched within, and concealed by, talk about language.

Language ideologies regarding homogeneity and the creation of the linguistic marketplace are upheld by powerful national institutions (Bourdieu, 1991, p. 46), the most obvious example being Government policies such as proficiency testing. However, Blackledge (2005) demonstrates how UK media and political discourse more generally has the potential to reinforce language ideologies and legitimize linguistic discrimination by consistently positioning English as 'above other languages spoken and written in multilingual England' (Blackledge, 2005, p. 225), all the while constructing negative associations with non-English, particularly Asian, languages. Building on the pioneering work by Reisigl and Wodak (2000,

2001), Blackledge (2005) draws on the concept of topoi, as we do in this study, to interpret the representational and argumentational strategies identified in his data. According to Blackledge (2005, p. 67), topoi can be defined as ‘parts of argumentation that belong to obligatory, either explicit or inferable, premises. They are the content-related warrants or “conclusion rules” that connect the argument with the conclusion of a claim’ (see also: Reisigl & Wodak, 2001, p. 75). Topoi can be used to frame social issues as problems and then provide (or at least imply) solutions to those problems. For Blackledge (2005), topoi resemble key discursive strategies in media texts surrounding language debates, while Reisigl & Wodak (2000) suggest that topoi are powerful and influential discourse strategies for the reproduction of discriminatory discourse. Taking various examples from local news reportage and political speeches, Blackledge (2005) demonstrates how the topos of advantage/usefulness is used to argue that if an action is useful then it should be done. In the context of a multicultural society, this topos can be used to argue that an action can be justified or rationalized on the grounds that it is of benefit to the minority (linguistic) group(s). At the same time, Blackledge (2005) explores how the topos of danger or threat is used to link speakers of Asian languages and their actions to specific dangerous or threatening consequences, while the topoi of finance and burdening/weighting down were used to represent the speakers of Asian languages as financially costly and otherwise burdensome for ‘the taxpayer’. These topoi are similar to those identified by Hart (2010) in his critical analysis of media representations of immigrants. However, Hart (2010, p. 67) identifies a number of topoi that are not discussed by Blackledge, but which can be used in the context of non-native English speaker representations. Namely, the topos of culture holds that the out-group have irreconcilable differences in norms and values to the in-group, and the topos of displacement reflects the belief that the in-group will eventually be outnumbered by the out-group and be usurped as the dominant group within a given population. The similarity and overlap between the topoi identified by Blackledge in relation to multilingualism and Hart in relation to immigration show that the discourses surrounding both of these issues are likely to draw upon and perpetuate similar beliefs and ideologies which group language, identity and belonging under the umbrella of nationality or nationalism. Connections drawn between language and nationality thus can be aligned with the ‘racialization’ of the non-native English speaker, as identified by Shuck (2006). In the context of the United States, Shuck (2006, pp. 259–60) describes the ways in which public discourse around non-English languages is expressed ‘using rhetorical patterns most often associated with discussions of race and ethnicity’. The racialization and the nationalization of public discourse about language and multilingualism serve the same purpose – to relate language to categories of identity, and to emphasize differences between dominant majorities and marginalised minorities.

Minority languages and negative representations in the press

Lippi-Green (2012, p.144) argues that the power of the media in the political process of language subordination ‘can hardly be underestimated’. Blackledge (2002) undertakes a detailed qualitative analysis of a short local news item from the *Birmingham Evening Mail* and concludes that the piece is ‘monolingualizing’, in that it expresses the implicit ideology that monolingualism is representative of ‘Britishness’, while multilingualism is ‘retrograde and

tribal' (Blackledge 2002, p. 84). Similarly, Blackledge (2005, p.226) finds that in the UK context, negative features associated with Asian languages and their speakers in media and political discourse include civil disorder, racial violence, social segregation, threats to democracy and burdens on society. Such negative representations of speakers of minority languages are not restricted to the UK. For example, in the coverage of language testing in Sweden, Milani (2007, p. 125) highlights the ways in which media narratives are built on the assumption that migrants are unwilling or too 'lazy' to integrate into Swedish society. Similarly, in the German context, Stevenson (2006, p. 151) finds a politician represented in the press as associating the linguistic proficiency of non-native German schoolchildren to 'ghettoes in our cities, which lead to social conflicts'. What these studies have in common, despite analysing the press of different European nations, is the observation of the tendency of the press to promote ideals about monolingualism and a national language, while at the same time denigrating multilingualism and negatively representing individuals or groups whose main language is one other than the dominant variety.

The focus on problems related to speakers of non-dominant languages – languages other than English for our purposes – might be simply explained by the fact that problems make news; it is generally agreed that newsworthiness or 'news value' is found in negativity, rather than in positive stories (Galtung and Ruge 1965, p. 69). Therefore, rather than necessarily being driven by negative attitudes towards multilingualism, the unfavourable representations of non-English languages and their speakers may be an inevitable product of the news genre. However, as Bednarek (2016, p. 30) argues, negativity in the context of newsworthiness is subjective; while some audiences may consider a particular issue or happening as negative, others may not. Multilingualism, or people living in the UK and speaking languages other than English, may be a problem for some target audiences but not others. Therefore, in making choices regarding the topics and discourses that are drawn on alongside discussions of language, some parts of the media *construct* multilingualism as a problem. In other words, it is the media itself which renders issues to do with multilingualism as newsworthy by presenting those issues as negative and problematic. The media – and the press in particular – accomplishes this by associating languages other than English (and the speakers of those languages) with traits and social issues that are likely to be agreed upon as negative by most audiences, such as laziness, crime, violence, social segregation and conflict. This way, multilingualism and non-native English speakers become part of the problem, as the media exploit the newsworthiness of such social issues in making 'linguistic and cultural diversity visible' (Kelly-Holmes and Milani, 2011, p. 477). Once linguistic difference is made visible in this way, and is associated with negative social issues, speakers of minority languages become a target for distrust and discrimination by some members of the dominant English-speaking group. In this context, Blackledge (2002, p. 84) argues that close scrutiny of media discourse is required to 'understand how linguistic power relations operate to discriminate against linguistic minority individuals and groups'. The present study sets out to achieve just that, by examining how the right-leaning UK print media represent people living in the UK who cannot speak English, or at least who are perceived not to speak English well.

Methodology

In this study, we take a corpus-assisted approach to Critical Discourse Analysis (CDA). Corpus linguistics refers to a collection of methods that use specialist computer programs to study large collections of machine-readable text (a *corpus*, pl. *corpora*). CDA can be described as ‘a type of discourse analytical research that primarily studies the way social power abuse, dominance, and inequality are enacted, reproduced, and resisted by text and talk in the social and political context’ (van Dijk, 2015, p. 466; see also: Fairclough, 1995).

The combination of corpus linguistics and CDA can be mutually-enforcing, with each method able to overcome some of the obstacles associated with the other (Baker et al., 2008). For instance, because CDA research tends to be qualitative and interdisciplinary, analyses are often restricted in scope, focussing on only a few texts or examples. However, with corpus assistance, CDA can deal more effectively with larger and more representative datasets, often running into millions and sometimes even billions of words in size. In the present study, corpus assistance allowed us to examine 2,751 press articles (1,789,566 words) – a dataset likely to reveal a wider range of representations (and account for a longer time period) than would have been possible had we adopted a purely manual approach to CDA. Corpus assistance can also add a degree of objectivity to CDA approaches because it offers a set of predictable analytical techniques and advocates a spirit of methodological transparency underpinned by two guiding principles: (i.) no systematic bias in the selection of texts included in the corpus (i.e. do not exclude a text because it does not fit a pre-existing argument or theory) and (ii.) total accountability (all data gathered must be accounted for) (McEnery & Hardie, 2012). This notwithstanding, it should be noted that corpus approaches are only ever able to *reduce*, rather than remove completely, the influence of subjectivity on the analysis, as human users of corpus techniques are required to make a series of procedural decisions (discussed later) which shape the corpus output. That output must then be interpreted by the researcher (the computer does not *do* the analysis for us) – an analytical step which creates further space for human bias and subjectivity to shape the findings reported.

At the same time, our corpus-assisted approach also stands to benefit from its synthesis with CDA, for instance by being able to draw upon ideas and theoretical models advanced within existing CDA research. In this case, findings reported in existing CDA studies of media representations of minority groups and language debates (particularly studies by Reisigl & Wodak (2000, 2001), Blackledge (2005) and Hart (2010)) provide useful concepts and points of reference for interpreting the representations we identify (most notably here the CDA concept of ‘topoi’). Our corpus-assisted analysis is also enriched by CDA’s commitment to analysing social context (van Dijk, 2001), which includes providing arguably more robust theoretical accounts of the ways in which texts interact with the contexts in which they are produced and consumed. In the present study, we draw upon our understandings of the real-world social and political backdrop against which the articles in our corpus were published, paying particular attention to UK language policy and relevant events taking place within the time span of our data, in an attempt to re-situate the news stories and observed topoi within the

original contexts in which they were produced and consumed. In the subsections that follow, we introduce the corpus under study and the approach we take to analysing it.

Data

This study is based on a purpose-built corpus comprising right-leaning UK national newspaper articles concerned with the topic of speaking English, published between 2011 and 2016. The data start-point of 2011 was chosen because it was the year in which the language-related questions were introduced into the Census. The end-point of 2016 was chosen as this was the last full year that had passed when work on the project began. Since it was not possible, within the scope of this study, for us to examine *all* UK national newspapers, we decided to focus on right-leaning press articles only. This decision was based on two considerations: (i.) the right-leaning press accounts for the majority of newspaper discourse in the UK (of the ten national newspapers, six are right-leaning, three are left-leaning and one is centrist; Baker et al. 2013), and (ii.) the right-leaning press published over twice as many articles concerned with the topic of speaking English (as we defined this – discussed later in this section) compared with the left-leaning press (2,751 vs. 1,349 articles, respectively).

Articles were collected from the online searchable newspaper database *LexisNexis*. Our search criteria were that the words *speak** and *English* had to appear (contiguously or non-contiguously) in the headline and/or lead paragraph for an article to be included in the corpus.¹ The search included all right-leaning UK national newspapers – broadsheets and tabloids. Newswires were excluded, and similar results were grouped together to minimise the inclusion of duplicates (default search settings in *LexisNexis*).

The completed corpus is 1,789,566 words in size, comprising 2,751 articles sourced across all six UK national right-leaning newspapers. The data included more traditional-style articles (i.e. ‘hard news’), but also columns, editorials, reviews, letters to the editor and opinion pieces. We decided to include these publication types since they all occupy space on the pages and websites of the media outlets under study, are all consumed by readers, and all contribute to the language ideologies that we are interested in uncovering in this paper (see also: Gabrielatos, 2007, pp. 10–11). We did not filter the results for articles that we deemed not to be sufficiently ‘about’ the survey result or the topic of speaking English. The concept of ‘aboutness’ is complex in this context, as debates about language intersect, as we have already discussed, with numerous aspects of social life. Thus, deciding what does and does not count as an article relevant for our purposes presents difficulties (Scott, 2017) and would require us to make a series of subjective judgements which would impinge upon the replicability of our data collection procedure (Baker et al., 2013). Nevertheless, we performed a ‘sanity check’ on the corpus to ensure that the proportion of ‘false positives’ (i.e. articles that clearly had little to do with the English language) would not unduly influence our analysis. To do this, we extracted a random sample of 100 articles spanning all years and publication types. From this, we determined that just 3 articles had limited relevance to our purposes but the remaining 97 had something to do with the English language and those who can or can’t speak it (well). Table 1 provides a numerical breakdown of the corpus in terms of the number of words and articles per newspaper.

[Table 1 here]²

As this table shows, the corpus is imbalanced in terms of its representing tabloid and broadsheet newspapers, with broadsheet publications (*Telegraph*, *Daily Mail*, *Times* and *Express*) contributing 2,324 (84.4 per cent) of the articles and 1,678,664 (93.8 per cent) of the words in the data. This skew reflects the reality of right-leaning press publications in the UK, which tend to be broadsheet rather than tabloid. We decided against redressing the imbalance because we wanted the corpus to represent all available right-leaning press coverage relating to the issue of speaking English within our given date range, and this imbalance in our corpus arguably represents the real-life imbalance in terms of the types of media outlets that provided this coverage.³ It should also be noted that, procedurally this skew does not present much of an issue in terms of our analysis, since it is not our aim to compare broadsheet and tabloid representations. However, in terms of representativeness, it does mean that our analysis is likely to elucidate topoi and patterns in representation that are particularly prominent in broadsheet articles. Data downloaded from *LexisNexis* includes a large amount of metadata and indexing information, which meant that before they could be analysed, the texts in the corpus had to be pre-processed to remove such elements, leaving only the headlines and articles. The next section sets out the approach we used to analyse the data.

Approach: Corpus-assisted Critical Discourse Analysis

Our corpus-assisted approach to CDA can be resolved into two parts. The first part aims at identifying the social actor groups (van Leeuwen, 1996) that are frequently mentioned in the context of (not) speaking English in the newspaper articles in our data. To do this, we use the corpus linguistic technique of collocation – a word association measure which tells us how often two or more words occur alongside one another in the corpus (Baker, 2006, p. 95) – to identify words denoting social actor groups which frequently co-occurred alongside the expression ‘speak* English’ (contiguously). We decided to use the phrase ‘speak* English’ as an analytical entry point as it broadly corresponds to the wording of one of the Census items that motivated this study (i.e. How well do you *speak English*?) and was also the search term used to source articles for the corpus. Using *WordSmith Tools version 7* (Scott, 2016), we generated a list of collocates for the phrase ‘speak* English’, including words which occurred within the five words preceding and/or following ‘speak* English’ (L5>R5) at least five times throughout the corpus (default values in *WordSmith Tools*). We then extracted all collocates denoting social actors and manually grouped them into broader social actor categories. In determining social actor groups, we began by manually scanning the list of collocates to identify words that ostensibly denoted people or groups of people. To group those words into categories, and check that words were in fact used to denote social actor groups, we inspected a random sample of concordance lines in which the word in question was used alongside the expression ‘speak* English’. This also allowed us to resolve polysemous and ambiguous cases. To offer an example, the word *Britain* was a frequent collocate of ‘speak* English’ but tended to function within a prepositional phrase as an adverbial of place rather than a social actor, such as in this headline taken from the *Telegraph*: ‘If you don’t speak English you can’t belong in Britain’.

Having identified the most frequent social actor groups, we then analysed how each was represented across the articles using the concordance technique. Concordancing is a way of viewing the corpus data which displays all the occurrences of a particular word or phrase throughout the corpus within its wider contexts of use (Baker, 2006, p. 71). In the concordance view, each use of the word or phrase of interest is displayed running down the centre of the screen, with surrounding text displayed to the left and right. In the present study, this more contextualised view of the data provides us with a means through which to carry out more qualitative CDA and examine the topoi used to represent the social actors collocating with ‘speak* English’ throughout the corpus. For the identification of topoi, the analysis draws upon the taxonomy initially proposed by Reisigl and Wodak (2000, 2001) and elaborated by Blackledge (2005) and Hart (2010). While there is no exhaustive list of topoi, the topoi identified across these studies collectively provide a useful interpretive framework for our purposes. Although topoi are, as we have described, ‘parts of argumentation’, we do not analyse the texts in terms of their argumentative structure or rhetorical goals, but are instead interested in analysing topoi in terms of their representational power and critiquing them in relation to their underlying ideologies pertaining to linguistic competence, national identity and (im)migration.

To identify topoi, both authors independently examined texts in which one or more of the social actor collocates occurred alongside ‘speak* English’. In this qualitative step in our analysis, for each social actor collocate group (see Table 2) we examined a random sample of 30 articles, noting any observed topoi, and repeated the procedure until new topoi ceased to emerge (Sinclair, 2003).⁴ While the concordance display provided a useful starting point for this more qualitative step, to apprehend all of the topoi observable within each article (of which there tended to be multiple), it was necessary for us to go beyond the solitary concordance line and analyse each article in its entirety. Following the independent analysis, the authors then compared findings, which identified similar topoi, and agreed upon a set of interpretations for each article in the sample. Although existing CDA research provides a useful interpretive framework for this study, it should nevertheless be noted that topoi do not manifest in texts in any set or formulaic way and so their identification represents something of a subjective analytical process. Similarly, the process of labelling or ‘naming’ topoi can be equally subjective – a single topos might be interpreted differently by two researchers, while a third researcher might not even interpret it as a topos in the first place. For these reasons, corpus assistance provides an objective *starting point* for our analysis, but the identification of topoi requires that we make more subjective interpretations of the data. In the present study, we have endeavoured to describe all our methodological decisions and, in our analysis, have provided examples which we deem to be representative of the particular topoi we observed.

Analysis

As outlined in the previous section, our analysis uses the expression ‘speak* English’ as an entry point through which to examine how groups of social actors were represented in the context of the English language debates that were played out across the articles in our corpus. Before proceeding with our analysis, it is first worth noting that throughout the corpus the phrase ‘speak* English’ tends to be used as part of broader constructions that focus on people

not speaking English. To demonstrate this, we generated a list of left-sided collocates (L5>L1) of the phrase ‘speak* English’. From this list, we extracted the ten most frequent words which ostensibly served to pre-modify the search term in a way which expresses that a person or group either could not or should speak English (*not, don’t, can’t, cannot, must, should, doesn’t, unable, didn’t, couldn’t*). The combined co-occurrences of these ten items alone (n= 1,036) accounts for over half (51 per cent) of the 2,048 mentions ‘speak* English’ in the corpus. Furthermore, since we have only considered the top ten collocates of this type, this figure must be treated as a minimum. The focus on cases where people can’t (but should) speak English reflects the tendency of the press – and the mainstream media generally – to report stories about things and events that are problematic or in some way negative (Bourdieu, 1998). Yet, we might also interpret this focus as reflecting the tendency of the press to operate with a deficit model of linguistic competence, which foregrounds what a speaker *cannot* do, while backgrounding what they *can*.

The analysis in this article focusses upon the three most frequent social actor groups collocating with ‘speak* English’ in the corpus: EDUCATION, IMMIGRANTS and HEALTH (table 2).

[Table 2 here]

Although the three frequent social actor groups in table 2 are lexically distinct, they all tend to refer broadly to immigrants to the UK who are understood either to not speak English well or indeed at all. In the case of the EDUCATION and HEALTH groups, the immigrants are functionalised; that is, they are ‘referred to in terms of something they do, for instance, an occupation or role’ (van Leeuwen, 2008, p. 42). From the table above, we can see that, as well as being referred to as *immigrants, migrants* or *foreigners*, immigrants are functionalised in our data according to their roles as, amongst other things, schoolchildren, patients, doctors, nurses and the more generic ‘staff’ (note: teachers was used to refer to people born in the UK, rather than immigrants – as will be demonstrated later). This functionalising tendency marks an interesting trend in our data, whereby immigrants tend to be discussed, with regard to speaking English at least, in the contexts of public services, particularly to do with schools and healthcare. One consequence of such functionalising representations is that the immigrant referents are transported into readers’ worlds of immediate experience, so that they become ‘people “we” have to deal with in our everyday lives’ (van Leeuwen, 2008, p. 36). This helps such stories to satisfy the newsworthiness criterion of familiarity (Galtung & Ruge, 1965), as the social actors being discussed are situated within public service contexts with which most readers will have direct experience and indeed a ‘stake’ as UK taxpayers. This conceivably makes the representations and associated topoi more relatable to readers’ own lives and experiences – a discursive move that arguably serves to heighten the relevance and emotive force of the stories. Having identified the most frequent groups of social actors who cannot or do not speak English, at least according to the articles in our corpus, we then set about examining the topoi that were used to represent these groups in the context of language ideological debates. This analysis is reported in the next section.

To examine how the groups displayed in table 2 were represented in our corpus, we qualitatively analysed those articles in our corpus in which one or more of the social actor

words collocated with the expression ‘speak* English’. Our analysis identified seven recurring topoi: (i.) displacement, (ii.) finance, (iii.) burdening and weighting down, (iv.) danger or threat, (v.) culture, (vi.) advantage/usefulness and (vii.) responsibility. Throughout the course of the analysis we will explore each of these topoi in turn, providing textual evidence of their use in the form of representative examples taken from the corpus. For the facility of analysis, we will introduce each topos separately. However, it should be noted, as our analysis will demonstrate, that these topoi do in fact relate to and overlap with one another in the ways that they are used to represent the social actor groups featured in our data.

Topos of displacement

The first topos we identified in our corpus was the topos of displacement. The premise of this topos is that ‘the out-group will eventually outnumber and/or dominate the in-group’ (Hart, 2010, p. 67). Linguistically, we interpret this topos to be invoked through the use of aggregation (van Leeuwen, 2008) – that is, the representation of migrants as living in the UK and not speaking English in large (usually increasing) numbers. The examples below relate to immigrants and were taken from articles responding to the publication of the 2011 Census results in January 2013. Note that all corpus extracts provided in this article were selected because we deemed them to be representative of the linguistic manifestation of the particular topos being discussed.

800,000 migrants who cannot speak English

(Telegraph, 30th January 2013)

But if you only remember one statistic make it this: **more than four million migrants** do not speak English as their main language. The impact of this finding, which comes from newly released details in the 2011 Census, is profound.

(Express, 1st February 2013)

But **138,000 can speak no English and another 726,000 struggle with it**

(Sun, 5th March 2013)

Although these extracts report the same Census data, there is considerable discrepancy between the figures provided, attesting to propensity for the (at least right-leaning) media to circulate inaccurate and conflicting information about the number of immigrants who do not speak English living in the UK. The effect of this type of large-scale aggregation, we would argue, is that the statistics reported are rendered in as shocking and extreme terms as possible. For instance, the *Telegraph* headline reports that 800,000 migrants ‘cannot speak English’ (even though only 138,000 people self-reported as not being able to speak English – the others self-reported as speaking English ‘not well’). Though the *Sun* provides an arguably more nuanced picture of the statistics – distinguishing between those who struggle to speak English and those who cannot speak English at all – the linking expression ‘and another’ is likely, we would

argue, to encourage readers to interpret each statistic as contributing to the same over-arching trend: ‘But 138,000 can speak no English **and another** 726,000 struggle with it.’

The most extreme example of an article overlooking the nuances of the Census statistics can be found in the *Express*. The figure of four million+ immigrants does not represent either the number of people who self-reported as not speaking English or not speaking English well, but the number of immigrants living in the UK who do not speak English as their first language. This statistic reveals nothing about the linguistic competence of the four million immigrants to which it relates and overlooks the reality that millions of people the world over communicate effectively using a second or additional language in all aspects of their daily lives. This can be interpreted as reflecting a shift in the UK political agenda, observed by Sebba (2017), whereby it is no longer considered sufficient for people of migrant origin to have ‘satisfactory’ knowledge of English, but they are now expected to use English as their *main* household language in order to demonstrate their willingness to integrate into UK life.

As mentioned above, this topos extended to the representation social actors from the EDUCATION category, specifically immigrant schoolchildren who either don’t speak English or speak English as an additional language, who were consistently represented as growing in number and, in some cases, as actually outnumbering children whose native or main language is English, who are represented as being in the minority, as in this headline and lead paragraph taken from the *Daily Mail*.

Children who speak English as their main language at home **are now in the MINORITY** in 1,600 schools across Britain

(*Daily Mail*, 22nd March 2012)

The result of this trend is thus surely that native English speakers’ claim to dominant status is being threatened if not actually weakened by the increasing numbers of residents who speak other languages natively. In the extreme, by repeatedly emphasising the large and increasing number of people living in the UK who speak languages *other than* English, such articles could induce in their perceived native English-speaking readership the fear that (or at least suggest to them the possibility that) they and their children are increasingly becoming a linguistic minority in their own country. The logical extension of this argument is that the out-group will eventually replace the in-group as those with privileged and preferential access to social resources and will become ‘the majority whereupon they will gain control over the in-group and become the new cultural norm’ (Hart, 2010, p. 83).

Topos of finance

One consequence of the aggregating rhetoric explored in the previous section is that it allows people who can’t speak English to be rendered as what van Leeuwen (2008, p. 38) describes as ‘the object of “rational” economic calculation’ and, in turn, rendered as economically burdensome – in other words, presented through the topos of finance. Evidence of this can be found in articles relating to all three social actor groups. Articles drawing upon this topos tend to do so to allege that the UK Government and taxpayers have incurred substantial financial costs in order to support immigrants who have limited English skills. For example, the

following headline and leading paragraph, taken from the *Mail on Sunday*, purport to show the costs endured by police, councils and hospitals in translating documents for immigrants with limited English:

How you pay **£100m a year** to aid immigrants who can't speak English: Shock figures reveal huge sums spent on translators by police, councils and hospitals

- Police, hospitals and courts are all spending huge sums on translators
- Councils in England spent **£11 million** on language services in last year
- Polish is the most commonly translated language by courts and police
- Met spends nearly **£7million** a year on interpreters for crime suspects

(*Mail on Sunday*, 10th January 2015)

This example is representative of a broader discourse of financial burden that recurs across the articles in our corpus, and also illustrates the tendency for costs associated with crime and policing to be foregrounded in these contexts. Three of the four bullet points in the leading paragraph refer to costs relating to crime and policing, arguably drawing on a link, by now well-established in the right-leaning press, between immigration and crime (van Dijk, 1991). It is notable, also, that readers are addressed directly as the bearers of these purported costs – ‘How **you** pay £100m a year’ – enforcing further the sense of opposition between immigrants and readers, who are indexed here as non-immigrant, English-speaking tax-payers.

As well as constituting a financial burden on public sector services, immigrants are also represented as being economically burdensome as a result of the social welfare benefits they are purported to receive:

Benefits ‘to be stripped from immigrants who cannot speak English’; David Cameron intends to make the cuts in order to **slash the country's welfare bill**, according to reports

(*Telegraph*, 19th January 2014)

TORY plans to **strip benefits from migrants** who can't speak English have been blocked by, you guessed, Nick Clegg.

(*Express*, 21st January 2014)

The first extract, taken from the *Telegraph*, reports the intention of then-UK Prime Minister David Cameron to prevent immigrants who can't speak English from receiving social welfare benefit. However, the Government at the time was a coalition formed of Cameron's Conservative Party and Nick Clegg's Liberal Democrats, the latter of which opposed the introduction of such measures (much to the frustration of the right-leaning media; see extract from the *Express*).

As mentioned above, the representation of immigrants who can't speak English as constituting a financial burden was also evident in the discourse surrounding other groups. For

example, in the extracts below, schoolchildren who cannot speak English are presented as placing a financial strain on schools:

The set of perceived burdens brought by school children include economic ones, but are also wider-ranging and more general: Educating children who do not speak English as their mother tongue also puts a **huge financial strain** on schools.

(Daily Mail, 21st February 2014)

Like schoolchildren and immigrants generally, patients who can't speak English were also represented as constituting a financial burden to the taxpayer. For example, in the extract below, taxpayers (i.e. readers) are described as 'picking up' a translation bill for patients who can't speak English well:

TAXPAYERS have picked up a £13million translation bill for NHS patients who cannot speak English. The cost of translators has soared tenfold since 2006, while NHS boards have faced huge budget cuts in which thousands of staff will lose their jobs. And many NHS bosses admitted they do not know how many people are using the service, set up for immigrants and foreign tourists, despite its £13million drain on funds.

(Express, 21st March 2011)

Although the cost cited above is £13 million, other costs given vary between newspapers and over time, ranging from £4 million to £13 million, and are attributed not only to the need for translators, but also to the hiring of (particularly Polish) nurses to communicate with patients during their care, and developing picture charts for patients to point out the meals they would like.

This passage also demonstrates a curious pattern in our data. By situating translation bills and other costs associated with healthcare provision within a broader context of public services cuts (in the above example, NHS jobs), people with limited English skills are arguably presented as being in some way to blame for those cuts. It is not possible, in this instance, for the newspapers to argue that immigrants are 'taking jobs' from British workers, since this contradicts the argument that they cannot speak English and do not pay taxes. Nonetheless, immigrants are still construed as costing British people their jobs, with the reported costs associated with translation services causally linked, by the *Express*, to 'budget cuts in which thousands of staff will lose their jobs.' Rather than challenge or criticise the Government over its cuts to public spending and blame the enduring austerity economic model ushered in by the 2010 Conservative/Liberal Democrat Coalition Government, it is immigrants with limited English skills who are blamed for job losses.

The topoi that we have identified tended not to offer explicit solutions to the problems they propose in relation to people not speaking English – instead these tend either to be implied or left up to the reader to formulate for themselves. However, at some points, articles in our corpus offered quite explicit solutions as part of *this* topos – specifically that welfare benefits should not be given to people who aren't able to speak English. For example, the *Daily Mail* headline below was taken from an article reporting the outcome of a survey which claimed that

95 per cent of British people believed that immigrants ‘must speak English’, while 77 per cent believed that immigrants ‘should have access to benefits restricted’.

Immigrants must speak English and should have **access to benefits restricted**,
say Britons

(*Daily Mail*, 17th June 2014)

That these issues are bundled together for the purposes of this headline is evidence that, for the *Daily Mail* at least, immigrants’ rights to claim social welfare benefit should be contingent on their ability to speak English. Similarly, other articles proposed that immigrants should learn to speak English before they are granted access to healthcare services, as in the example below:

PATIENTS who cannot speak English cost the NHS £16million in interpreters last year - nearly double the 2008 figure. The 2012/3 bill could top £20million as 40 of 161 hospital trusts failed to respond to our Freedom of Information request for details. Polish speakers were in greatest demand across trusts which did answer, followed by Arabic and Slovak. The Pennine trust spent the most - £4.9million. Urdu was its top request. **Critics called for immigrants to learn English before being given access to services.**

(*Sun*, 29th July 2013)

The logical extension of this argument is that healthcare professionals should deny care to ill people if (and because) they have limited English skills. This is a quite cynical proposal, and such a policy would contradict directly one of the founding principles of the NHS – namely that care should be provided free for all at the point of use. We also find rather problematic the notion that ensuring accurate and effective communication between practitioners and patients is somehow a waste of taxpayers’ money, since the level of funding allocated to translation services is likely to impact directly on the health and wellbeing of the patients who rely upon those most essential services. This argument is also deeply contradictory, given that the argument that health professionals should speak English or be ‘struck off’ (explored later in our analysis) is itself underpinned by the understanding that patient safety is contingent on effective communication between patients and practitioners.

Topos of burdening and weighting down

The topos of burdening and weighting down is a causal topos which represents the minority out-group as a strain or drain on the members, resources and systems of the in-group (Reisigl and Wodak, 2001, p. 78). In some respects, this topos can be related to the topos of finance explored in the previous section. In that section, we saw how immigrants who can’t speak English are presented in the articles in our corpus as constituting a financial burden to English-speaking readers/taxpayers, who ‘foot the bill’ for the social welfare benefits claimed by this group, as well as more specific costs incurred in the contexts of healthcare provision. However, immigrants with limited English skills were also presented as burdensome in other ways – ways which ‘weighted down’ in-group native English speakers. This type of burden manifests most

explicitly in the consumption of public services and in the context of school education in particular. In the samples we analysed, this topos was particularly prominent in articles which mentioned social actors from the EDUCATION group, so we will focus on this context here. The examples below represent the different impositions that non-native English-speaking children are construed as placing on the British education system, which include increasing pressure on teachers and necessitating additional interpreting staff.

During a phone-in on LBC radio Sir Michael was told by one caller of a local school where **teachers were having to carry out “double the work”** due to a “huge influx” of children who could not speak English.

(Telegraph, 30th October 2014)

Astonishingly, one in nine English schools now has a majority of pupils who do not speak English as a first language. The official figures lay bare the **enormous strain** mass migration has placed on our schools

(Daily Mail, 22nd February 2014)

The burdens these children present are also linked to immigration generally in ‘influx’ and ‘mass migration’, implicitly placing blame on the children’s parents (who immigrated to the UK) as opposed to the children themselves. The same is the case in the 2014 example from the *Daily Mail*, in which ‘mass migration’ has placed some form of unspecified ‘strain’ on ‘our’ schools.

Non-native English-speaking pupils were presented not only as increasing the pressures on and workloads of teachers, but this was also construed as actively eroding the access of in-group English-speaking children to resources, thereby shifting from the topos of burden, to the topos of displacement (explored earlier), as in the examples below.

In a growing number of schools special classes are held for children who do not speak English **costing schools money, wasting teachers' time and harming the education of others.**

(Express, 24th November 2014)

Other children are getting less attention in class as teachers have to spend longer helping children who do not speak English well and preparing teaching materials for them, they [teachers] said.

(Times, 2nd April 2015)

Millions being spent in bid to provide support for eastern European pupils. Bilingual staff and textbooks in other languages among services provided. **British children turned away by schools** where more than 25 per cent of pupils do not speak English.

(Daily Mail, 23rd May 2016)

In the first two examples, the extra attention that non-native English-speaking children are afforded by teachers because of their limited English skills is represented as actively detracting from the attention that is given to ‘other’ (i.e. native English-speaking) children, essentially hampering their education. In other words, the resources at stake are teachers’ attention and, ultimately, a high-quality education, and the in-group’s access to these resources is being encroached upon by the out-group, to the extent that the in-group are now disadvantaged. This corresponds with Shuck’s (2006, p. 269) notion of the language of ‘special treatment’, whereby ‘members of the dominant group are excluded and members of the underrepresented groups are given undue advantage’. Shuck argues further that this kind of rhetoric is ‘projected onto a model of education as competition, which tends to construct the majority – the invisibly privileged – as an oppressed group’ (Shuck, 2006, p. 270). In the third example, this opposition is made more explicit through the direct contrasting of (mainly Eastern European) pupils who do not speak English against ‘British children’ in the competition for school places – the implication being that the allocation of school places to non-British pupils is in some way unjust, or that those children are undeserving of such places. Such anti-foreigner sentiment only makes sense in a discriminatory framework in which access to resources should be restricted on the basis of people’s nationality and, in this case, their native language.

Fears that children with English as an additional language will ascend to privileged status at the expense of native English-speaking children are likely to be further compounded by articles reporting that children with English as an additional language academically outperform native English-speaking children:

An educational landmark has just been passed that has rather striking implications for our school system. Over the weekend, reports confirmed that pupils who speak English as an additional language (EAL) - that is, not their first - **are now outperforming their native, English-speaking counterparts** for the first time.

(independent.co.uk, 6th March 2012)

Children who speak English as a second language are more likely to excel at primary school, official figures show, as they reveal that **white pupils are falling behind**. Figures released today by the Department for Education (DfE) reveal that non-native speaking children **perform better than their English speaking counterparts** at grammar, punctuation and spelling as well as mathematics.

(telegraph.co.uk, 15th December 2016)

Pupils who speak English as a second language are **outperforming native speakers in core GCSE subjects** as white British students fall behind 10 other ethnic groups, new research showed. **White British pupils are lagging behind** because other ethnic groups are receiving more help from their families.

(Telegraph, 4th April 2016)

As well as contradicting the notion that pupils are a burden on the school system, such reports are likely to be surprising for readers given the media and political backdrop in the UK, which typically propagates the normative view that languages other than English are associated with

academic *underachievement* (Blackledge, 2005, p. 226). This apparent over-performance may therefore be perceived as problematic or as threatening towards the in-group for a range of reasons, but primarily that better academic performance may lead to improved future prospects in terms of access to careers, resources and better living standards. In other words, if immigrant children achieve highly, this increases the likelihood of the out-group displacing the in-group over time. One might expect that having non-native English-speaking children working in good jobs is presumably more desirable for the in-group than having them rely on tax-payer funded state benefits (as per the topos of finance). However, this seems to be on the proviso that this achievement is not at the expense of native English-speaking, British children.

Another feature of both of the *Telegraph* examples above is an explicit conflation of native language and race, with children who speak English as a second or additional language being contrasted to ‘white’ children. Again, this discourse of competitiveness in school education is racialised, and discussions about language become discussions about race. Such representation serves to sow divisions between the in-group and the out-group, not just on the basis of native language, but also race and nationality. Children from each group are represented as different, separate and in competition with one another in terms of educational access and attainment.

The strong performance of non-native English-speaking children is framed as a loss of ‘assets’ rather than a gain for the country. Therefore, pupils with English as a second or additional language are discursively forced into a ‘no-win’ situation. Poorly-performing children are constructed as burdensome to the education system, to the detriment of native English-speaking (white) pupils. However, if they perform well, they are represented as threats, not only to the British (and white) children of the in-group, but also to the UK as a country, by securing the best jobs and living conditions and fundamentally challenging and potentially displacing the ideal and cultural norms of the linguistically (and racially) homogenous nation-state.

Topos of danger or threat

So far, our analysis has shown how the right-leaning press represent immigrants who can’t speak English as broadly constituting a threat towards native English-speakers and UK society as a whole through topoi of displacement, finance and burdening and weighting down. In this part of this analysis, we consider how immigrants with limited English skills could also be represented as constituting a quite literal threat, even danger, to the majority native English-speaking in-group. As Blackledge (2005, p. 68) puts it, the topos of danger or threat argues that if an action bears dangerous or threatening consequences, ‘it should not be performed; or, put another way, if there are specific dangers or threats, something should be done to prevent [them].’ We observed this topos in articles in which social actors from the HEALTH group collocated with ‘speak* English’, with doctors and nurses working in the UK reportedly unable to speak English (well) constructed as posing a threat to the health and safety of the public, for example:

ALARMINGLY, **thousands of doctors** cannot speak English properly
(*Express*, 22nd March 2011)

Doctors fears: PATIENTS could be at risk after health chiefs admitted **some doctors** can't speak English.

(*Daily Star*, 22nd March 2011)

Earlier, we observed the use of aggregation across articles in our data when discussing how many people living in the UK can't speak English. A similar kind of aggregation can be observed in the extracts above. However, these and other articles in our corpus also exhibit imprecision pertaining to exactly how many health professionals can't speak English (quantified above as *thousands* and *some*). This kind of vague quantification was only observable in relation to doctors and not the other groups (EDUCATION and IMMIGRANTS more generally). The term 'thousands' in the example from the *Express*, though vague, nevertheless quantifies the number of doctors who purportedly don't speak English in a sufficiently large amount so as to give the impression that the 'problem' they cause is widespread and should be a reason for concern (an interpretation further guided by the capitalised adverb 'ALARMINGLY' situated at the start of the clause). Vaguer still, the *Daily Star* quantifies 'some' doctors as not being able to speak English, which causes 'fears' and puts patients 'at risk'. We argue that this imprecise quantifying rhetoric is likely to induce in some readers a sense of panic, not just because there are purportedly 'thousands' of doctors who cannot speak English, but, perhaps more 'ALARMINGLY', because nobody quite knows how many doctors can't speak English (and by extension, which ones). In turn, this might give those readers the impression that the health authorities (to whom they entrust their health and lives) are not sufficiently monitoring or scrutinising the recruitment of healthcare professionals. However, in terms of specific cases, there is only one story published within the duration of our corpus which reports such an incident: the case of Dr. Daniel Urbani, a German doctor who caused the death of a patient by administering the wrong medication (and even in this case, it is not clear that Urbani's English skills were the cause of this mistake; see Baker & McEnery, 2014).

Topos of culture

Immigrants who can't speak English are also represented as having a disruptive impact on the communities and societies in which they live. Such representations are accomplished through what Reisigl and Wodak (2001, p. 80) describe as the topos of culture, wherein the rise in non-native speakers of English is framed as a threat to social cohesion in, and the cultural values of, the UK. We observed this topos in articles in which social actors from the IMMIGRANTS group collocated with 'speak* English', but not in relation to the other groups we consider in this study. This topos manifested textually in a number of different ways, with immigrants with limited English skills frequently described as separating themselves from the rest of society, for example as: 'living separately', 'living in segregated communities', 'failing to assimilate into British society', 'trapped in their homes' (particularly Muslim women), 'not full members of society', 'not contributing to society' and 'not respecting our way of life'. At other points, immigrants who can't speak English are represented as having an active but negative influence on society; described variously as: 'changing communities', 'driving a wedge through

communities’, causing communities to ‘break down’ and ‘ripp[ing] communities apart’. To offer an extended example, in the extract below taken from a 2013 article published by the *Express*, then-Secretary of State for Communities and Local Government, Eric Pickles, is quoted linking immigrants’ not speaking English to the breakdown of communities.

IMMIGRANTS must learn to speak English properly **to stop communities breaking down**, a Cabinet Minister said yesterday. Westminster Communities Secretary Eric Pickles warned that **newcomers with "broken English" could not play a full role in society**. And he renewed his attack on councils for wasting millions per year on translation services for migrants. **"A community of broken English is no community at all**. That's why we are taking action," Mr Pickles said in a newspaper article yesterday.

(*Express*, 15th November 2013)

In this example, Pickles goes so far as to forge a discursive link between fluent language skills and the very notion of community itself when he states that ‘A community of broken English is no community at all’. There is an implicitly responsabilising logic at work here. If limited English skills do indeed threaten social cohesion, as these articles suggest, then readers are left in little doubt that this is the fault of the immigrants themselves, who are consistently pitched in socially and grammatically agentive terms – i.e. they are the ones who ‘change’, ‘drive wedges between’ and ‘rip apart’ communities. There is no suggestion in these articles that any of the responsibility for the purported lack of social cohesion should be placed with the other members of those communities. We will return to this point in more detail when considering the topos of responsibility in the final section of the analysis.

Topos of advantage/usefulness

The topos of advantage/usefulness argues that if an action is considered useful, then it should be done. So far, our analysis has revealed some of the topoi on which the right-leaning press draws in order to present some immigrants’ purported inability to speak English (well) as threatening (in a number of respects) towards native-English speakers and UK society as a whole. However, in invoking the topos of advantage/usefulness, the merits of immigrants learning to speak English are framed as advantageous for those minority immigrant groups, rather than the native English-speaking majority. Consistent with the topos of culture examined in the previous section, the benefits of learning English tended to be framed in terms of helping immigrants to better integrate within UK society. To offer a representative example, in this extract from the *Daily Mail*, a Government-ordered review is reported as advising ministers to ‘force’ immigrants to speak English in order to ‘help them integrate’:

‘Immigrants should be forced to speak English to help them integrate,
government-ordered review set to advise ministers’

(*Daily Mail*, 2nd December 2016)

However, as Blackledge (2005, p. 68) demonstrates, although the topos of advantage/usefulness argues ostensibly for the benefit of the minority group, in the context of debates about multilingualism it can often be used to present a more illiberal agenda. It is difficult to make strong assertions about the attitudes and intentions of those who wrote and edited the articles in our corpus when we only have access to the articles themselves. However, the topoi identified elsewhere in our corpus, and which are collectively more pervasive than the topos of advantage/usefulness, suggest that it is the interests of the native English-speaking in-group (likely the perceived readership(s)) which underpin the agendas of the articles in our data. Indeed, there are some discursive clues as to this in the *Daily Mail* extract above. We might question the extent to which the Othering use of the pronoun *them* (i.e. us vs. them) reflects a truly integrative agenda. Moreover, the very notion of ‘forcing’ immigrants to speak English and not enjoining native speakers of English to learn other languages adopts something of a one-sided perspective on integration – that it is in the interests of immigrants that integration is successful and therefore that it is them (and not other members of the community) who should be responsible for ensuring social cohesion. This agenda thus seems to be one that is less concerned with the idea of integration but has more to do with maintaining the dominance of the English language and those who speak it fluently (natively) within UK society. Building on the idea of ‘forcing’ immigrants to learn to speak English, the next and final section of our analysis explores the topos of responsibility.

Topos of responsibility

The topos of responsibility argues that because a group or person is responsible for the way things are, that group or person should act to put things right (Blackledge, 2005, p. 72). According to this topos, is the immigrants themselves who are framed as being responsible for their (in)ability to speak English and thus, by extension, for the social and economic consequences that brings. We observed this topos in relation to the social actors in the IMMIGRANTS and EDUCATION groups. The example below is taken from the IMMIGRANTS group sample.

Last night Tory MP Nick de Bois said: “Councils have no legal obligation to provide translation services. They're currently having to make difficult funding decisions and funding excessive translation should not be a priority. The bill for interpreting and translating and telephone assistance was £43.5million. "Those needing translation, in many cases, have lived in this country for many years and therefore **have a responsibility to learn English.**”

(*Express*, 3rd May 2013)

A consequence of foregrounding immigrants’ responsibility for learning to speak English is that the roles and responsibilities of others in this process are backgrounded. In particular, there is no explicit acknowledgement in any of the articles across our corpus of the role either of other members of communities to learn other languages or of the government to provide support for immigrants to help them to learn to speak English and, by extension, to integrate more smoothly into society. In fact, in the above extract, Conservative MP Nick de Bois is

quoted as arguing quite the opposite – that against a backdrop of ‘difficult funding decisions’, funding ‘excessive translation should not be a priority’. Yet, these factors are important to bear in mind in the context of debates about multilingualism and immigration in the UK, not least because the UK has what Hunston (2017) rightly describes as ‘possibly the worst record in the world for learning languages’ (see also: Guardian, 2014), but also because the Government has imposed funding cuts which have resulted in reduced access to English language lessons for UK residents whose first language is not English (TES, 2017).

We also observed the topos of responsibility in relation to the EDUCATION category, with parents of immigrant children constructed as being responsible for teaching their children to speak English. Earlier in our analysis, we saw how the right-leaning press regards the commitment of resources to helping children of immigrants to learn English as constituting a burden. Consistent with this view, at other points in the corpus the responsibility for immigrant children learning to speak English is placed squarely at the feet of their parents:

The responsibility lies with the parents. If they don't work to make sure their children speak English at home, integration is impossible. It is not unreasonable to expect incomers to learn the language of their new country. Indeed, not to do so is unreasonable.

(*Sunday Express*, 24th July 2011)

Listen and learn: **parents need to be able to speak English to help their children to make progress at school.**

(*Telegraph*, 28th July 2011)

These examples were chosen because they broadly represent the arguments put forward in the articles in our corpus for why immigrants should be responsible for teaching their children to speak English: (i.) to help them to better integrate into UK society (see example from the *Sunday Express*) and (ii.) to help their children to make progress at school (see example from the *Telegraph*). We have already explored both arguments in previous sections of the analysis. However, it is worth pointing out that this type of responsabilising topos is likely to place immigrant parents into something of a double-bind. If they do not teach their children to speak English well, they can be blamed for their children’s failure to integrate into society, as well as for holding other, native English-speaking children back at school (since those children then require more attention from teachers). Yet, as we saw in the section on burdening and weighting down, if they *do* teach their children to speak English so that they perform well at school, this is likely to evoke the response – from sections of the right-leaning press at least – that those children were taking employment opportunities from other, native English-speaking children in the future.

Concluding discussion

In the years following the 2011 Census, right-leaning press debates about people who can’t speak English have focused on immigrants, with particular attention paid to public service contexts of education (as schoolchildren) and health (as patients and health professionals). Our

qualitative analysis uncovered the use of seven interconnecting topoi in the representation of these groups: (i.) displacement, (ii.) finance, (iii.) burdening and weighting down, (iv.) danger or threat, (v.) culture, (vi.) advantage/usefulness and (vii.) responsibility. Collectively, these topoi can be interpreted as weaving an over-arching media narrative which problematises and ‘Others’ non-native speakers of English living in the UK. We argue that this narrative is likely to invoke or intensify in some native English-speaking readers a kind of linguistic xenophobia – a sense of anxiousness about people living in the UK whose first language isn’t English, suspicion about their willingness to integrate into society, and fear that linguistic diversity will bring about negative consequences for them, their families and UK society as a whole. In light of this, we follow Blackledge and Pavlenko (2001, p. 243) in maintaining that such persistently negative representations of people on the grounds of their (perceived) English language competence is likely to be sufficient to legitimise the social exclusion and active discrimination of those who are perceived to be unable or unwilling to fit the linguistic ‘norm’.

We have argued that the focus on immigrants not speaking English in public services contexts is likely to create a sense in which readers feel personally impacted by immigrants’ purported linguistic deficiencies, both as users of those services themselves and as taxpayers who ultimately ‘foot the bill’ when it comes to increased staffing costs and translation bills. At the same time, this focus on public services contexts can also be interpreted as speaking directly to, and perhaps even contributing towards, public concerns surrounding immigration, education and health, all of which featured amongst the top voter concerns in the run up to both the 2015 UK General Election and the 2016 EU referendum (Jordan, 2015). It is difficult within the scope of a study such as this to hypothesise as to the influence that any section of the media (or its readership) has on events taking place in the world, including the outcomes of political elections and referenda. However, it seems likely that the discriminatory press discourses that we have identified will have shaped many voters’ opinions leading up to the point at which they entered the voting booth on these occasions.

Another consequence of the topoi identified across our corpus is that immigrants are responsabilised into ensuring their and their children’s English language competence and integration into UK society. A corollary of this responsabilisation, we observed from our data, is that immigrants are then accordingly blamed for having limited English skills, as well as for the economic, social, cultural and even health-related problems this is purported to bring about. As well as overlooking the difficulties that immigrants face in learning a new language and when attempting to integrate into unfamiliar communities, in order to situate any blame firmly at the feet of the immigrants, the topoi we have identified also serve to obscure the role (and any failings) of the Government when it comes to helping immigrants to learn English and assimilate into society, as well as the role of existing members of local communities to help newcomers to assimilate. Thus, although we would not argue that it is not beneficial for any person to attain a certain level of competence in the language that is spoken most widely in the place where they live, we do take issue with the discrimination against people on the grounds of their linguistic competence and their perceived unwillingness to learn a language, particularly when this discrimination is enabled by the active stripping of important contextual factors which serve to constrain immigrants’ abilities to learn English, including cuts in Government funding of English language classes.

As we discussed earlier, topoi are argumentation strategies or structures which postulate that if x is the case then y should/should not be done. In some cases, the solutions to the problems the topoi posed were stated explicitly within the bodies of the newspaper articles themselves, for instance calls for reductions in funding for translation services and for access to public services to be restricted for immigrants who can't speak English. However, solutions to the problems that topoi frame need not be explicitly stated as part of the discursive instantiations of the topoi themselves, but can instead be implied, or arise implicitly through some shared knowledge or common-sense agreement between members of a discourse community (i.e. newspapers and their target readerships). Although they did not feature as part of our analysis, many of the “below-the-line” comments posted online by readers in response to some of the articles in our data offered ‘solutions’ to the problems set up by the topoi in the bodies of the articles in the form of implorations – presumably to other readers – to vote at the 2015 General Election and the 2016 EU referendum in ways that would support an anti-immigration agenda (e.g. comments like ‘vote leave!’ and ‘vote UKIP!’). Future research should study such below-the-line comments as sites where readers can engage with, support and challenge the media representations they encounter and, in some cases, as we have observed, offer explicit “solutions” to the problems that articles frame. Another way in which future research could build on this study would be to compare right-leaning newspaper data against articles published by left-leaning media outlets. Although previous research of media portrayals of immigrants and minority groups has identified similar strategies of representation across the political spectrum (Baker et al., 2013), it is possible that left-leaning newspapers might offer counter arguments to the types of topoi we have identified here.

Notes

1. The asterisk following *speak* acts as a wildcard for any set of characters, so the search included words like *speaks*, *speaker*, *speakers* and *speaking*. We decided not to include *spoke* or *spoken* in our search term as only a minority of the results produced by these terms (less than 5 per cent) were about speaking English. Instead, these past tense forms tended to be to frame a quote or to indicate that someone had spoken to the newspaper or the media more widely about the topic of the story.
2. Sunday editions and sister publications are subsumed under the main newspaper name, for example *Times* also covers the *Sunday Times*. For most newspapers, the counts include articles published in print and online editions, for instance the category *Daily Mail* includes articles from the printed version of the newspaper as well as the website dailymail.co.uk, with duplicates removed where appropriate. The exceptions to this are the *Sun* and the *Times*, for which paywalls preclude the inclusion of online articles in LexisNexis results.
3. This skew also partly results from the fact that *LexisNexis* only stores articles which originally appeared in the print version of the *The Sun*, and not those which appeared on its website.
4. Although this method of analysing concordance lines is useful for carrying out qualitative analysis of large corpora, and is indeed well-established in corpus research,

it should be noted that this approach does not allow us to meaningfully quantify the prevalence of the topoi we observe. That said, we view the quantification of topoi in this instance as problematic. The conversion of the articles into a plain-text corpus forces us to treat (and count) all topoi as equal. However, we are of the view that any meaningful measure of the prominence of a topos should be sensitive to such factors as the location of the topos within the article (i.e. does it occur within the headline or towards the bottom of the page?), as well as where that article sits within the newspaper more broadly (i.e. does the article feature on the front page of the newspaper or is it buried on a page deep within it).

Acknowledgements

The authors would like to acknowledge the technical support provided by Ikechukwu Onyenwe during data collection and preparation. We would also like to thank Małgorzata Chałupnik for commenting on an earlier draft of this paper, as well as the three anonymous reviewers for their detailed and constructive feedback.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

David Wright is Lecturer in Linguistics in the School of Arts and Humanities at Nottingham Trent University, UK. His research interests include forensic linguistics, language in legal and evidential contexts, and discourses of abuse, discrimination and harassment. Mary Ann Evans Building, Clifton Campus, Nottingham Trent University, Nottingham NG11 8NS, UK.

Gavin Brookes is Senior Research Associate in the ESRC Centre for Corpus Approaches to Social Science (CASS) in the Department of Linguistics and English Language at Lancaster University, UK. His research interests include corpus linguistics, (critical) discourse studies, multimodality and health communication. FASS Building, Lancaster University, Lancaster LA1 4YD, UK.

References

- Anderson, B. (2006). *Imagined Communities: reflections on the origin and spread of nationalism*. London: Verso.
- Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.
- Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., & Wodak R. (2008). A useful methodological synergy? Combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, 19(3), 273–306.
- Baker, P. Gabrielatos, C., & McEnery, T. (2013). *Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press*. Cambridge: Cambridge University Press.
- Baker, P., & McEnery, T. (2014). Find the doctors of death: press representation of foreign doctors working in the NHS, a corpus based approach. In A. Jaworski & N. Coupland (Eds.), *The Discourse Reader* (3rd edition) (pp. 465-480). London: Routledge.
- Barbour, S. (2000). Nationalism, language, Europe. In S. Barbour & C. Carmichael (Eds.) *Language and Nationalism in Europe* (p. 1–17). Cambridge: Cambridge University Press.
- BBC (2013). 138,000 speak no English – Census. 30 January 2013. [online]. Available at: <http://www.bbc.co.uk/news/uk-21259401>. Accessed 15 August 2017.
- Bednarek, M. (2016). Voices and values in the news: News media talk, news values and attribution. *Discourse, Context & Media*, 11, 27–37.
- Bernard, R. (2009). The integration agenda in British migration law. In E. Guild, K. Groenendijk & S. Carrera (Eds.), *Illiberal Liberal States: Immigration, Citizenship and Integration in the EU* (pp. 277-298). London: Routledge.
- Billig, M. (1995). *Banal Nationalism*. London: Sage.
- Blackledge, A. (2002). The discursive construction of national identity in multilingual Britain. *Journal of Language, Identity and Education*, 1(1), 67–87.
- Blackledge, A. (2005). *Discourse and power in a multilingual world*. Amsterdam: John Benjamins.
- Blackledge, A. (2009). Inventing English as convenient fiction: Language testing regimes in the United Kingdom. In G. Extra, M. Spotti & Van Avermaet P (Eds.), *Language Testing, Migration and Citizenship: Cross-National Perspectives* (pp. 65–85). London: Continuum.
- Blackledge, A., & Pavlenko, A. (2001). Negotiation of identities in multilingual contexts. *International Journal of Bilingualism*, 5(3), 2432–2459.
- Blommaert J., & Verschueren J. (1998) The role of language in European nationalist ideologies. In: Schieffelin BB, Woolard K and Kroskrity P (eds) *Language ideologies: Practice and theory*. Oxford: Oxford University Press, pp. 189–210.
- Bourdieu, P. (1991). *Language and symbolic power*. Cambridge: Polity Press.
- Bourdieu, P. (1998). *On Television and Journalism*. London: Polity Press.
- Fairclough, N. (1995). *Media discourse*. London: Arnold.
- Fowler, R. (1991). *Language in the News. Discourse and Ideology in the Press*. London: Routledge.

- Galtung, J. & Ruge, M. (1965). The Structure of Foreign News. The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers. *Journal of Peace Research*, 2, 64–91.
- Gabrielatos, C. (2007). Selecting query terms to build a specialized corpus from a restricted-access database. *ICAME Journal*, 31, 5–43.
- Guardian (2014). UK students taking languages degrees at record low (Online article). Available here: <https://www.theguardian.com/education/2014/apr/10/students-taking-languages-record-low>.
- Hart, C. (2010). *Critical Discourse Analysis and cognitive science: new perspectives on immigration discourse*. Basingstoke: Palgrave Macmillan.
- Hunston, S. (2017). ‘Migrants should learn English’ – A response to the Casey Report (Online article). Available here: <https://blog.bham.ac.uk/susanhunston/2017/01/05/migrants-should-learn-english-who-could-object/>.
- Jordon, W. (2015). Health overtakes immigration as an issue for voters. *YouGov, UK*. [online] Available at <https://yougov.co.uk/news/2015/04/15/health-tops-immigration-second-most-important-issu/>.
- Kelly-Holmes, H. and Milani, T. M. (2011). Thematising multilingualism in the media. *Journal of Language and Politics*, 10(4), 467–489.
- Lippi-Green, R. (2012). *English with an accent: language, ideology and discrimination in the United States*. London: Routledge.
- Louw, B. (1993). Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies. In M. Baker, C. Francis & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 157–176). Amsterdam: John Benjamins.
- McEnery, T., & Hardie, A. (2012). *Corpus linguistics: method, theory and practice*. Cambridge: Cambridge University Press.
- Milani, T. M. (2007). A language ideology in print: the case of Sweden. In S. Johnson & A. Ensslin (Eds.), *Language in the media: representations, identities, ideologies* (pp. 111–129). London: Continuum.
- Milani, T. M. (2015). Language and citizenship: broadening the agenda. *Journal of Language and Politics*, 14(3), 319–334.
- ONS (2013). *2011 Census: Quick Statistics for England and Wales, March 2011* [online]. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/2011Censusquickstatisticsforenglandandwales/2013-01-30>. Accessed 15 August 2017.
- ONS (2016). *What is a Census?* [online]. Available at: <https://www.ons.gov.uk/Census/2011Census/howourCensusworks/aboutCensuses/whatisaCensus>. Accessed 15 August 2017.
- Paffey, D. J. (2012). *Language ideologies and the globalization of ‘standard’ Spanish*. London: Bloomsbury.
- Pavlenko, A. and Norton, B. (2007). Imagined communities, identity, and English language teaching. In J. Cummins and C. Davison (Eds.), *International handbook of English language teaching* (pp. 669–680). New York: Springer.

- Piller, I. (2001). Naturalization language testing and its basis in ideologies of national identity and citizenship. *International Journal of Bilingualism*, 5(3), 259–277.
- Reisigl, M. and Wodak, R. (2000). ‘Austria first’: A discourse-historical analysis of the Austrian ‘Anti-Foreigner Petition’ in 1993’. In M. Reisigl and R. Wodak (Eds.), *The Semiotics of Racism* (pp. 269–304). Vienna: Passagenverlag.
- Reisigl, M. and Wodak, R. (2001). *Discourse and discrimination. rhetorics of racism and antisemitism*. London: Routledge.
- Scott, M. (2016). *WordSmith Tools version 7*. Stroud: Lexical Analysis Software.
- Scott, M. (2017). News Downloads and Aboutness. Plenary lecture given at the International Corpus Linguistics Conference, University of Birmingham, Birmingham, UK.
- Sebba, M. (2017). ‘English a foreign tongue’: The 2011 Census in England and the misunderstanding of multilingualism. *Journal of Language and Politics*, 16(2), 264–284.
- Shuck, G. (2006). Racializing the nonnative English speaker. *Journal of Language, Identity & Education*, 5(4), 259–276.
- Sinclair, J. (2003). *Reading Concordances: An Introduction*. London: Longman.
- Stevenson, P. (2006). ‘National’ languages in transnational contexts: Language, migration and citizenship in Europe’. In :C. Mar-Molinero & Patrick Stevenson (Eds.), *Language Ideologies, Policies and Practices: Language and the Future of Europe* (pp. 147–161). Basingstoke, Palgrave Macmillan.
- TES (2017). Esol funding falls by £100m (Online article). Available here: <https://www.tes.com/news/further-education/breaking-news/esol-funding-falls-ps100m>.
- van Dijk, T. (1991). *Racism and the Press*. London: Routledge.
- van Dijk, T. (2001). ‘Multidisciplinary CDA: A plea for diversity’. In: R. Wodak & M. Meyer (Eds.), *Methods of Critical Discourse Analysis* (pp. 95–120). London: Sage.
- van Dijk, T. (2015). Critical Discourse Analysis. In D. Tannen, H. Hamilton & D. Schiffrin (Eds.), *The Handbook of Discourse Analysis* (2nd edition) (pp. 466–485). London: Wiley.
- van Leeuwen, T. (1996). ‘The representation of social actors’. In: C.R. Caldas-Coulthard & M. Coulthard (Eds.), *Text and Practices: Readings in Critical Discourse Analysis* (pp. 32–70). London: Routledge.
- van Leeuwen, T. J. (2008). *Discourse and Practice: New Tools for Critical Discourse Analysis*. Oxford: Oxford University Press.
- Vessey, R. (2017). Corpus approaches to language ideology. *Applied Linguistics*, 38(3), 277–296.

Table 1. Breakdown of newspapers in the corpus, ranked by number of articles and words.

Newspaper	Articles	Words
<i>Telegraph</i>	775	620,465
<i>Daily Mail</i>	682	549,315
<i>Times</i>	483	350,745
<i>Express</i>	384	158,139
<i>Sun</i>	308	86,342
<i>Daily Star</i>	119	24,560
Total	2,751	1,789,566

Table 2. Social actor groups and corresponding collocates, ranked by frequency.

Social actor groups	Corresponding collocates	Frequency	Texts
EDUCATION	<i>pupils</i>	116	115
	<i>children</i>	115	91
	<i>schoolchildren</i>	18	18
	<i>student</i>	16	13
	<i>child</i>	13	13
	<i>teachers</i>	12	12
	<i>pupil</i>	8	8
	<i>kids</i>	7	7
	<i>secondary</i>	5	5
	IMMIGRANTS	<i>immigrants</i>	60
<i>foreign</i>		56	47
<i>migrants</i>		45	36
<i>migrant</i>		13	12
<i>foreigners</i>		10	10
<i>immigrant</i>		6	6
HEALTH	<i>doctors</i>	44	32
	<i>nurses</i>	35	25
	<i>staff</i>	28	24
	<i>patients</i>	24	22