

Chapter 2

The C.O.N.T.A.C.T. Methodological Approach

As already hinted at in the previous chapter, the C.O.N.T.A.C.T. project covered two main strands of research: the expression of hate speech and its perception. To these ends, a multi-method approach was adopted, encompassing different types of data. In this chapter, we will outline the shared procedures of data collection and analysis in relation to both the production data, i.e. online comments to news reports, and the perceptual data, i.e. interviews.

2.1 Harvesting and Analysing Online Comments to News Reports

Sharon Millar, Fabienne H. Baidier and Stavros Assimakopoulos

To investigate the expression of hate speech, the main source of data has been user-generated content as found in what is known in media circles as “*below the line*” comment fields on newspaper websites (Graham and Wright 2015: 139). The reason for this was that while there is a sizeable literature on how minorities are portrayed in the mainstream media,¹ a lot less attention has been given to the ways in which the general public reacts to this kind of discourse.² And indeed, as we will see in Chap. 3, the analysis of our collected data revealed a number of strategies that are used by commenters on news portals to communicate a negative stance toward the migrant and LGBTIQ minorities. However, given the multilingual character of our project, even the initial harvesting of such online comments in an

¹See, for example, van Dijk (1987, 1991), Reisigl and Wodak (2001), Baker et al. (2008), KhosraviNik (2010), KhosraviNik et al. (2012); as well as several papers in the *Journal of Language Aggression and Conflict* (cf. Kopytowska 2015; Musolff 2017).

²But see Erjavec and Kovačič (2012), Brindle (2016), among others.

attempt to compile manageable, and ultimately comparable databases for analysis in each national context was a challenge in itself.

As a first step, it was decided to use the publicly available Newsbrief web application,³ developed by the European Commission's Joint Research Centre. This web portal monitors online news reporting across the globe in 60 different languages and hence was especially appropriate for a multilingual project like ours. Its functionalities include the searching of archives using specific terms within designated time periods in countries and languages of one's choice. The tool then allows for a compilation of hits for keywords and from this, it is possible to search for news articles that are accompanied by comments.

A pool of keywords within the thematic areas of the project was established and, from this, partners selected those terms that were relevant for their national contexts. Many of these terms were shared across at least some partners, but to ensure the possibility of comparison, all partners were required to include the following keywords in their search: '*homosexual(s)*', '*immigrant(s)*', '*lesbian(s)*', '*LGBT*', '*Muslim(s)*' and '*refugee(s)*'. Some partners carried out automated searches in more than one language, either because the country has two official languages (e.g. Maltese and English in Malta) or because of a specific interest in certain minority languages within national borders (e.g. Russian in Lithuania).

The minimum number of keywords for the main themes of xenophobia and homophobia/transphobia was six per theme. Monitoring was carried out over two, non-consecutive time periods: 1.4.2015–30.6.2015 and 1.12.2015–29.02.2016. This was done to include a period where the contemporary refugee crisis might be less overwhelmingly predominant, at least in the media of some of the partner countries. The number of hits for each keyword per month was registered to give a quantitative mapping of what topics were most in focus in the media at the time.⁴

The Newsbrief tool provides no automated means to find articles with comments from the hits collected so this was approached manually. The baseline was that partners checked all hits per keyword per month, but to a maximum of 100–120. In cases where the number of hits exceeded this maximum, an appropriate ratio was applied; for instance, if a keyword produced 500 hits in a month, then 1 in 5 hits were checked. Following this method, most partners could find more than an adequate number of articles with comments, except for the Cypriot team, who turned to the Facebook sites of the newspapers that are part of the Newsbrief Cyprus database, and monitored reactions to the relevant articles there.

As the newspaper and comment data were to be analysed qualitatively, it was decided for reasons of feasibility to restrict the databases to approximately 5000–6000 words per keyword both for articles and for comments over the designated 6-month period. In cases where too many articles with too many comments had

³<http://emm.newsbrief.eu/overview.html>.

⁴Typically, across countries, issues to do with migration, immigrants and refugees dominated across both time periods with noticeably fewer articles and reports on matters of sexual orientation and gender identity.

been found, certain sampling criteria were applied. Firstly, for each month, only the week which attracted the most articles with comments for the keyword was selected. If that single week still provided too many words for comments, then the first 300–500 words were taken, but without cutting short an individual comment or comment thread. If the monthly word count for articles was exceeded, then those articles that had only very few comments were dropped, as were articles that had little actual relevance to the keyword.⁵ The objective was to obtain at least one article with comments per keyword per month where possible. Once compiled, the overall database for all keywords was checked for duplicate articles, which were removed and, if necessary, replaced.

It should be emphasised that we make no claim of representativeness for any of the partner databases. The Newsbrief web application, although comprehensive, has inevitably its own inherent biases and during data collection, it became obvious that certain newspapers, particularly the tabloids, tended to have articles with comments while others rarely were sources of reader comments.

As with data collection, the qualitative analysis of the comments in the databases was based around a common methodology, but partners were also free to subsequently develop their analyses further within their areas of expertise. We will describe here the shared approach, which aimed to identify the evaluative language used by authors towards the relevant target groups (immigrants, refugees, homosexuals etc.). In this context, evaluation is understood as “the expression of the speaker’s or writer’s attitude or stance towards, or view point on, or feelings about the entities and propositions that he or she is talking about” (Thompson and Hunston 2000: 5). While there are many aspects to evaluative language, the C.O.N.T.A.C.T. focus was on negative and positive evaluative polarity (Alba-Juez and Thompson 2014), which was then additionally related to speaker/writer strategies operating at phrasal, sentential and discourse levels, in terms of linguistic forms (e.g. lexical choice, metaphors, use of generics and argument strategies), as well as pragmatic functions (e.g. insults, threats, jokes, stereotypes and counter stances).

To some extent, inspiration was taken from the EU-funded Light On project,⁶ which collected (and continues to collect) racist expressions, providing their source and context as well as potential explanations as to why they are considered racist or discriminatory. In a similar fashion, the qualitative analyses conducted as part of C.O.N.T.A.C.T. also provided the discursive context, both in terms of the characteristics of the newspaper (e.g. tabloid or broadsheet, political orientation) and the interactional status of the comment (e.g. direct or tangential response to the article,

⁵For example, the Cypriot team had to disambiguate results for the keyword ‘*refugee(s)*’ as some referred to Cypriot refugees in 1974, a common issue in Cypriot newspapers, rather than the 2015 refugees, while a much commented upon article that was retrieved for the keyword ‘*black(s)*’ in Malta was an article about Darth Vader in Star Wars, which obviously has no connection to the issue of xenophobia.

⁶*Cross-community actions for combating the modern symbolism and languages of racism and discrimination*. For further information, visit <http://www.lighton-project.eu/site/main/page/project-en>.

response to another contributor). In addition, the reasons for the polarity categorisations of expressions as more or less negative or positive (or ambiguous) were stipulated by each group of analysts. In this way, what could be taken as subjective categorisations were given a degree of transparency.

The shared analytical approach resulted in lists of expressions with their categorisations that permit cross-country comparisons at a general level.⁷ In this setting, negatively-loaded expressions may be potential examples of hate speech as more broadly or narrowly defined whereas more positive-oriented language may exemplify counter speech. Obviously, more in-depth qualitative and quantitative analyses were then undertaken by individual partners to shed more light on the complexities of evaluative language (potential hate speech and counter speech) in relation to the target minority groups. As will be seen in Chap. 3, these included the use of corpus linguistic methods to investigate frequencies and collocational patterns, qualitative approaches dealing with specific forms and functions as well as interactional and co-constructural aspects of evaluative language use.

2.2 Approximating Perceptions of Hate

Sharon Millar, Fabienne H. Baider and Stavros Assimakopoulos

The second major strand of the C.O.N.T.A.C.T. research was a study of how the general public, and in particular young people belonging to the 18–35 age group, perceive hate speech in the local context of each partner country. This strand consisted of two phases. The first involved the online administration of a questionnaire across the consortium, and the second, which took place after the analysis of the questionnaire responses, comprised interviews intended to explore these responses in more depth. This combination of questionnaires and interviews is widely used in research wishing to capture broader perspectives and to pursue issues of interest with more targeted and in-depth questions (Adams and Cox 2008). Given this volume's aim of providing an overview of matters pertaining to the discourse analytic study of hate speech, the focus will be on the interview stage of this research strand.⁸ Nonetheless, it is still necessary to provide an overview of the

⁷Even though such comparisons are beyond the scope of the present work, just to mention one example, the use of the sickness and unnaturalness metaphors for homosexuality or non-conditional threats against refugees (e.g. *'torpedo the boats'*, *'electrify the fences'*, etc.) was present more or less across the C.O.N.T.A.C.T. datasets.

⁸Obviously, the interpretation of the questionnaire responses in each national context is also meaningful in itself and we intend to return to it on some other occasion, but given space limitations and, above all, the current volume's focus on mainly qualitative-based discourse analysis, we have decided to omit them from this section.

questionnaire design in order to contextualise the results of the interview analysis that follows in Chap. 4.

The questionnaire was intended to cover three major themes and comprised three sections. Firstly, respondents were requested to evaluate six authentic examples from each partner country's collected data in terms of acceptability, by marking their perceived acceptability on a 4-point Likert-type scale that included the options 'acceptable', 'somewhat acceptable', 'less acceptable' and 'not acceptable'. Each C.O.N.T.A.C.T. team selected three examples of negative polarity comments relating to migrants and another three relating to the LGBTIQ community from their national database. The examples for each category were chosen to represent different degrees of extremeness, ranging, for example, from obvious threats to insinuations. Finally, a further question asked whether the respondents would have assessed the six presented examples differently if they had been written in private, rather than public contexts online, for instance, in a private e-mail or during a casual chat with friends. It was hoped that this question would give an indication of how sensitive the general population is to the difference between public and private discourse when it comes to the expression of hate.

The second section of the questionnaire aimed to examine the respondents' attitudes towards—and experience of—reporting hate speech incidents. To contextualise the issue, we first asked participants to share their own experiences of hate speech as targets and as witnesses in their everyday life. Those participants who stated that they have some experience of the sort were prompted to indicate the place where the incident under question took place (i.e. at work, at school, in the street, etc.). The respondents were then asked whether they would report such incidents to the relevant authorities, and if they expressed unwillingness to do so, they were given a list of options to indicate why this might be the case (e.g. embarrassment, fear of reprisals, belief that police would not do anything, too much trouble to report etc.).

Finally, the third part of the questionnaire sought to investigate the respondents' perception of the concept 'hate speech' itself by asking them to indicate on a 6-point Likert-style scale the extent to which they agree with each of four definitions of hate speech that respectively equated the term with 'making negative prejudiced remarks', 'insulting', 'threatening' or 'encouraging other people to be violent or show hatred' towards people because of their race/nationality/ethnic origins/religion/gender and/or sexual orientation.

Against this backdrop, the aim of the interviews was to follow up on the questionnaire, by providing a better understanding of any interesting conclusions or particular issues that arose from the analysis of the questionnaire responses. Interviews were conducted either individually or in focus groups with young people aged between 18 and 35 who were normally residents of each partner country. Some of those interviewed had taken part in the questionnaire survey. At least 20 participants in total were interviewed per country (except for the UK, where only 12 took part in the interviews). Individual interviews lasted on average 15 min each, while focus group sessions had an average duration of 45–60 min each. Interviews were audio-recorded and transcribed orthographically.

The interviews followed a semi-structured format, using an interview guide which was structured around several themes to permit comparability across national contexts. Starting off with a brief presentation of the acceptability ratings given for the six examples of hate speech from the questionnaire, the interviewees were asked to give their opinion as to why each example received the overall ratings that it did. This was followed by a discussion of the concept of hate speech, which involved consideration of both the definitions given in the questionnaire and the need to legislate in relation to these definitions. A further theme covered was any experience interviewees may have had with hate speech or discriminatory discourse. Interviewers were also free to gear the discussion towards other issues identified from the analysis of the questionnaire responses as particularly important in the national context concerned. Each session was concluded by asking participants if they wished to add anything they deemed relevant to the discussion.

In terms of analysis, interviews were then subjected to “conventional content analysis” which is used to identify categories, patterns and themes that emerge from the data (Hsieh and Shannon 2005: 1279). Content analysis permits both qualitative and quantitative approaches (Bengtsson 2016) and, as we will see in Chap. 4, while the former was most generally adopted by partners, the latter was also used for pattern analysis by the team from Cyprus.

References

- Adams, Anne, and Anna L. Cox. 2008. Questionnaires, in-depth interviews and focus groups. In *Research methods for human computer interaction*, ed. Paul Cairns, and Anna L. Cox, 17–34. Cambridge: Cambridge University Press.
- Alba-Juez, Laura, and Geoff Thompson. 2014. The many faces and phases of evaluation. In *Evaluation in context*, ed. Geoff Thompson, and Laura Alba-Juez, 3–24. Amsterdam/Philadelphia: John Benjamins.
- Baker, Paul, Costas Gabrielatos, Majid KhosraviNik, Michał Krzyżanowski, Tony McEnery, and Ruth Wodak. 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 and Society* 19: 273–306. <https://doi.org/10.1177/09579265080888962>.
- Bengtsson, Mariette. 2016. How to plan and perform a qualitative study using content analysis. *NursingPlus Open* 2: 8–14. <https://doi.org/10.1016/j.npls.2016.01.001>.
- Brindle, Andrew. 2016. *The language of hate: A corpus linguistic analysis of white supremacist language*. London/New York: Routledge.
- Erjavec, Karmen, and Melita Poler Kovačič. 2012. ‘You don’t understand, this is a new war!’: Analysis of hate speech in news web sites’ comments. *Mass Communication and Society* 15: 899–920. <https://doi.org/10.1080/15205436.2011.619679>.
- Graham, Todd, and Scott Wright. 2015. A tale of two stories from ‘below the line’: Comment fields at the Guardian. *The International Journal of Press/Politics* 20: 317–338. <https://doi.org/10.1177/1940161215581926>
- Hsieh, Hsiu-Fang, and Sarah Shannon. 2005. Three approaches to qualitative content analysis. *Qualitative Health Research* 15: 1277–1288. <https://doi.org/10.1177/1049732305276687>.
- KhosraviNik, Majid. 2010. The representation of refugees, asylum seekers and immigrants in British newspapers: A critical discourse analysis. *Journal of Language and Politics* 9: 1–28. <https://doi.org/10.1075/jlp.9.1.01kho>.

- KhosraviNik, Majid, Michał Krzyzanowski, and Ruth Wodak. 2012. Dynamics of representation in discourse: Immigrants in the British press. In *Migrations: Interdisciplinary perspectives*, ed. Michi Messer, Renee Schroeder, and Ruth Wodak, 283–297. Berlin: Springer.
- Kopytowska, Monika (ed). 2015. Contemporary discourses of hate and radicalism across space and genres. Special issue, *Journal of Language Aggression and Conflict* 3:1.
- Musolff, Andreas (ed). 2017. Public debates on immigration. Special issue, *Journal of Language Aggression and Conflict* 5:2.
- Reisigl, Martin, and Ruth Wodak. 2001. *Discourse and discrimination: Rhetorics of racism and anti-semitism*. London/New York: Routledge.
- Thompson, Geoff, and Susan Hunston. 2000. Evaluation: An introduction. In *Evaluation in text: Authorial stance and the construction of discourse*, ed. Susan Hunston, and Geoff Thompson, 1–26. Oxford: Oxford University Press.
- van Dijk, Teun A. 1987. *Communicating racism: Ethnic prejudice in thought and talk*. London: Sage Publications.
- van Dijk, Teun A. 1991. *Racism and the press: Critical studies in racism and migration*. London/ New York: Routledge.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

