The Role of Language in the Transnational Formation of Romani Identity (Rom*Ident*)

Universities of
Manchester (United Kingdom), Graz (Austria), Aarhus (Denmark)
2010-2013

Final Report (publishable version), November 2013

Project objectives

With upwards of 3.5 million speakers living in dispersed communities, Romani is by far the largest language in the European Union that cannot be identified with a majority population in any particular region. Traditionally an oral language whose use was largely limited to informal interaction in the family domain, over the past few decades Romani has experienced a gradual expansion of its functions to include both formal and informal literacy and a variety of uses in public discourse. Having started in the 1960s, this development took on a new momentum since the 1990s following the democratisation process in central and eastern Europe, where the majority of the Romani-speaking population resides. The emergence of new Romani political, cultural and religious networks, support for those networks from trans-national organisations, and the rise of electronic communication have all contributed to a massive expansion of Romani into new domains. Romani is a fascinating test case for the role that language plays in the process of identity consolidation in a globalised world. Standardisation is no longer inherently connected to the 'territorialisation' of language. Instead, we witness a bottom-up process in which individuals take ownership of language and negotiate language practices. Status regulation and language planning can be instigated and even implemented by institutions other than national states. All this leads to pluralism of form rather than unification. Yet language is a key locus for political mobilisation. It allows players to claim authenticity, it offers opportunities for intervention by external sponsors, and it provides a discussion platform through which traditional images can be challenged and recognition can be awarded.

The project set out to investigate what roles language plays in forging a modernised Romani identity – as a cultural asset upon which identity is predicated, as an emblem of identity, and as a vehicle through which identity is presented and communicated. It asked which social conditions

and attitudes support pluralism in identity formation, and which forces promote uniformity. It explored the role that new technologies play in the innovative use and innovative shaping of language, and the role that European-led initiatives and processes have on raising the profile of Romani in individual countries. It tried to address the more specific question of how local initiatives -- NGOs, local authorities, individual activists -- react to the implementation of a Council of Europe initiative on Romani, and how members of Romani communities are inspired to engage in 'language activism' through trans-national encounters. Overall, the project set out to investigate how the Romani model in its pluralist and trans-national context can contribute to our theoretical understanding of trans-national and diasporic identity formation in the age of globalisation.

The work performed

In order to answer these questions, the project set out to compile and assess corpora comprising educational publications and translations of texts into Romani, resolutions and reports by international organisations pertaining to the status and protection of the Romani language, spoken language data on Romani from a variety of dialects, in order to obtain a clearer picture of structural variation within Romani and the distribution of structural variants, and electronic mediated communication in Romani, from websites and social networks. As a service to the community of users, the project also set a specific dissemination objective, namely the creation of a Romani Virtual Library containing a catalogue of educational material, text translations, reports and resolutions and project publications.

We investigated structural variation across dialects by carrying out an in-depth documentation of spoken varieties and comparing key variation features to those represented in writing in both coherent texts such as translations and educational materials, and in spontaneous online exchanges on social media. Special attention was given to the varieties of Burgenland and Mitrovica and to the Romani dialects of Romania as a sample of dialect variation. We investigated the way in which clusters of structural and orthographic variants can be mapped onto online user networks. We examined hybridity in Romani identity and language, including structural hybridity exemplifying contact at various levels, structure, discourse, and contact related considerations in language planning, such as choice of orthography and lexical innovations.
We examined strategies of lexical expansion in translations and websites.

In order to obtain a picture of the role that language plays in the process of identity consolidation through public debates as well as personal participation and networking, we carried out analyses of policy measures at the European level, covering the consideration of Romani in political resolutions, recommendations of the Charter for Regional or Minority Languages, and their implementation through state-based measures. We also examined the use of language in online and published narratives that deal with Romani identity and history, and the extent to which Romani identity is argumentatively predicated or supported by linguistic data and a discussion of language. Complementing this analysis,

we investigated how images accompany the promotion and delivery of Romani language teaching in textbooks and on websites.

To this end, the project developed a series of procedures for corpus compilation and corpus evaluation. Together, these various procedures lend the project its unique methodological character. We compiled a corpus of Romani language texts, many of them translations of political and information texts in other languages, and we developed an automated system for parsing and tagging these texts, which recognised established (previously recorded) vocabulary as well as grammatical formants, and on that basis is able to identify new lexical formations and thus provide a sub-corpus of what the project termed 'new vocabulary' -- one of the principal manifestations of domain expansion in a language the use of which is entering the public domain. We recorded samples of spoken data and developed a method of glossing and tagging items and retrieving structural variants into a table representation, thus automatically linking a descriptive structural overview of a dialect with audio and transcription examples from a corpus of connected speech. We adopted a computational method for assessing network connections among online users of social media, using the example of comments on YouTube videos, which we used to plot the distribution of dialect-structural and orthographic variants among clusters of users and so to link online social networking with bottom-up language codification. Finally, we used elements of critical theory and critical discourse analysis to interpret narratives on Romani identity and their reliance on information on language, and the representation of language in political resolutions, while also analysing shifts and the emergence of new ideas in concrete language policy with consideration to the contributions of agents, facilitators and audiences.

The main results

We have been able to characterise the structural changes that accompany the functional expansion of language use into new domains. They include the emergence of a more complex syntax, a higher number of nominalisations, new functionalities of cases as well as in the creation of words for new concepts, which are based on productive word-formation processes in the context of the integration morphology. Hybridity plays a role inasmuch as written Romani often replicates the dominant model code syntactically as well as lexically. We are thus able to identify a new dimension of contact-induced language change, one that is distinct from that found across spoken varieties of Romani. Our analysis of the written use of Romani shows that Romani on purely informative websites only has symbolic functions whereas the contact language, often English, covers the communicative ones. On the other hand, Romani covers communicative functions on interactive websites primarily used by Roma. In such cases written Romani is a transposition of oral usage from everyday life and private contexts into the informal style used among non-familiar interlocutors in public. Here, we identified the common liberty that users take to mix languages and follow what we have termed 'cosmopolitan practices'. In these online forums, we find two kinds of user networks, as far as the choice of dialect features and orthographic representation is concerned. The first clusters together users who belong to the same or similar Romani 'group' or 'na-

tion', who share a dialect, and so their online network tends to display consistent selection of dialect features. Orthography, however, varies, as it is a hybrid representation of contacts with majority languages, and these differ since the online communities often represent dispersed diaspora communities. We thus have consistency in dialect choice with pluralistic orthographic representations. In the other pattern, we find a common Romani theme (often musical or religious or both) but a pluralism of both dialects and orthographies. We can thus recognise a Romani 'virtual community' which tends to make use of the flexibility offered by the informal character of social media to transpose everyday multilingual practices such a code-switching from face-to-face, private and oral communication into remote, public communication in writing, and one that at the same time uses the opportunities offered by the electronically mediated communication to promote a codified form of a speech variety that has so far been limited to vernacular use. The analysis of Romani and its representation in identity narratives has shown us that Romani language is acknowledged as a key feature of the culture of the population referred to in the political discourse as 'Roma/Gypsies'. Language serves as the ideological basis from which to derive elements of nationhood such as territorial origin and similarity as well as a common cause (the promotion of language itself), which are crucial especially in the case of the Roma due to the absence of a contemporary home territory, collective historical awareness, leadership structures, formal institutions, and literacy, all of which as elements that would build an 'imagined community' in the sense of Anderson's (1991) substitute term for 'nation'. We find that the initiative to grant recognition is taken at the European level, and this is even followed by European level implementation initiatives. We also witness the emergence of a different kind of language policy discourse that departs from conventional language planning strategies and views pluralism of form as enabling domain expansion.

Wider implications: impact and use

Altogether, we see the emergence of a language policy that may be characterised as non-territorial in its outreach, transnational in its strategic approach, and pluralistic in its practical implementation. The project's resources as presented in its Romani Virtual Library will help understand and implement codification and teaching and learning strategies that follow this pluralistic process, i.e. cater to the language needs of a transnational audience and make use of the opportunities of digital technology for both language promotion and analysis.