Lace avilen ko radio
Romani language and identity on the Internet

A thesis submitted to The University of Manchester for the degree of
Doctor of Philosophy in the Faculty of Humanities

2013

Daniele V Leggio

School of Arts, Languages and Cultures
## List of contents

- List of Abbreviations 4
- List of Figures 4
- List of Tables 4
- List of Graphs 5
- Abstract 6
- Declaration and Copyright Statement 7
- Acknowledgment 8

1 Introduction 9
  1.1 Diasporas on the Internet 10
  1.2 The Multilingual Internet 14
  1.3 Language choice in diasporic CMC 16
  1.4 The challenge for language planning theories 19

2 Practices and theories of language standardization 22
  2.1 The emergence of standard languages 22
    2.1.1 The Middle Ages linguistic situation 23
    2.1.2 The Modern Age and the emergence of the nation-state 23
    2.1.3 The Contemporary Age 25
  2.2 Language standardization in the post-colonial period 26
    2.2.1 Status and corpus planning 26
    2.2.2 The application of the model 28
    2.2.3 Reversing language shift 29
    2.2.4 Language rights 31
    2.2.5 Language planning in a global world 34
  2.3 The case of Romani 36
    2.3.1 Romani activists and intellectuals 36
    2.3.2 Rights recognition and state support 39
    2.3.3 RLS efforts 41
    2.3.4 The Internet 42
  2.4 Summary 44

3 The Mitrovica Roma 47
  3.1 The Mitrovica Roma community 47
    3.1.1 Life in Kosovo 47
    3.1.2 Migrations and dispersal 51
    3.1.3 Life in diaspora 55
  3.2 Radio Romani Mahala 57
  3.3 Summary 63

4 Languages and identities on RRM: ethnographic account 64
  4.1 Methodology and ethics 64
  4.2 RRM chat-room 66
    4.2.1 User-names, users' demographics and general behaviours 67
    4.2.2 CMC discourse features and net-iquette on RRM 72
    4.2.3 RRM chat-room as a diasporic space 73
  4.3 Language choice on RRM 76
    4.3.1 Greetings and goodbyes 77
    4.3.2 Community maintenance and strengthening 81
    4.3.3 Humour and strong personas 85
    4.3.4 Affection and intimate feelings 90
4.3.5 Self-policing and politeness

4.4 Summary

5 Languages and identities on RRM: quantitative analysis
5.1 Methodology
5.2 Language choice in user-names
5.3 Language choice in messages
   5.3.1 General distribution
   5.3.2 Distribution by communicative functions
5.4 Summary

6 The codification of Romani on RRM
6.1 Methodology
6.2 Choice of variety
   6.2.1 Fully shared features
   6.2.2 Contrasting features
   6.2.3 Partially shared features
   6.2.4 User consistency in choice of variants
6.3 Choice of spelling
   6.3.1 Postalveolar fricatives
   6.3.2 Postalveolar affricates
   6.3.3 Glottal and velar fricatives
   6.3.4 Semi-vowels
   6.3.5 Alveolar affricate
   6.3.6 Labial fricative
   6.3.7 Vowels
   6.3.8 Aspirated consonants
   6.3.9 Early Romani /ř/
   6.3.10 Spelling summary
6.4 Summary

7 Cosmopolitan practices, language development and the Internet
7.1 Language practice on RRM
7.2 Language development on the Internet
7.3 Representing and performing diasporas on-line
7.4 Cosmopolitan practices on RRM
7.5 Concluding remarks

References

Word count: 63281
LIST OF ABBREVIATIONS

CMC - Computer Mediated Communication
IM - Instant Messaging Software
RMS - Romani Morpho-Syntax Database
RRM - Radio Romani Mahala
1SG - First Singular
2SG - Second Singular
3SG - Third Singular
1PL - First Plural
2PL - Second Plural
3PL - Third Plural
ACC - Accusative
ABL - Ablative
COMPL - Complementizer
DAT - Dative
F - Feminine

FUT - Future
GEN - Genitive
IMP - Imperative
INTER - Interjection
M - Masculine
NEG - Negator
OBL - Oblique
PERF - Perfective
PL - Plural
PRES - Present
REF - Reflexive
REM - Remoteness
SG - Singular
SUBJ - Subjunctive
VOC - Vocative

LIST OF FIGURES

Figure 3.1: RRM homepage 1 58
Figure 3.2: RRM homepage 2 60
Figure 3.3: RRM homepage 3 61
Figure 4.1: Flatcast add-on, RRM not broadcasting 67
Figure 4.2: Flatcast add-on, RRM broadcasting 68
Figure 6.1: Location of Arli varieties 130

LIST OF TABLES

Table 2.1: Haugen’s (1972: 271) language planning matrix. 26
Table 2.2: Romani spellings 45
Table 5.1: Distribution of functions by languages 111
Table 5.2: Distribution of function by switches 111
Table 6.1: Arli varieties of former-Yugoslavia documented on RMS 130
Table 6.2: Features shared by Mitrovica Gurbet and all Arli varieties 131
Table 6.3: Features differentiating Mitrovica Gurbet from the Arli varieties 135
Table 6.4: Features partially shared between Mitrovica Gurbet and Arli varieties 142
Table 6.5: RRM alphabet 162
LIST OF GRAPHS

Graph 5.1: Language mixing in user-names
Graph 5.2: Language choice in user-names
Graph 5.3: Language choice in messages
Graph 5.1: Switching in messages
Graph 5.5: Position of switches
Graph 5.6: Switches for salutations
Graph 5.7: Distribution of languages for salutations
Graph 5.8: Distribution of switches for attention seeking
Graph 5.9: Distribution of languages for attention seeking
Graph 5.10: Distribution of switches for politeness
Graph 5.11: Distribution of languages for politeness
Graph 5.12: Distribution of switches for requests
Graph 5.13: Distribution of languages for requests
Graph 5.14: Distribution of switches for humour
Graph 5.15: Distribution of languages for humour
Graph 5.16: Distribution of switches for emphasis
Graph 5.17: Distribution of languages for emphasis
Graph 5.18: Distribution of switches for plain talk
Graph 5.19: Distribution of languages for plain talk
Graph 5.20: Distribution of switches for singing and shouting
Graph 5.21: Distribution of languages for singing and shouting
Graph 5.22: Distribution of switches for change of addressee
Graph 5.23: Distribution of languages for change of addressee
Graph 6.1: Forms of 1st and 2nd persons copula
Graph 6.2: Indicative and copula negators
Graph 6.3: Umlaut in 1SG copula and 1SG.PERF marker
Graph 6.4: Perfective markers
Graph 6.5: Prothesis of a-
Graph 6.6: Palatalization of alveolars
Graph 6.7: Palatalization of velar stops before front vowels
Graph 6.8: Short and long forms of 2SG.DAT
Graph 6.9: Genitive markers
Graph 6.10: Loan verb adaptation markers
Graph 6.11: Forms of demonstrative pronouns
Graph 6.12: Users consistency in choice of Romani variety
Graph 6.13: Unvoiced postalveolar fricative
Graph 6.14: Unvoiced postalveolar affricate
Graph 6.15: Voiceless postalveolar affricate
Graph 6.16: Voiceless glottal fricative
Graph 6.17: Voiceless uvular fricative
Graph 6.18: Semi-vowel /j/
Graph 6.19: Alveolar affricate
Graph 6.20: Aspirated consonants
The fall of the Eastern Block, the dissolution of former Yugoslavia and the subsequent enlargement of the European Union to include former socialist countries contributed to an increase in the movement of people from Eastern to Western Europe which began about a decade earlier. Among them, the Roma are probably the most clearly recognizable group and surely the ones that received, and keep receiving, more media attention. While their presence in the media as subjects of discussion is a topic worth analyzing, the present work is about their presence in a particular medium, the Internet, as actors and producers of content. As a population of Indian origin spread across Europe over the past five centuries, Roma have often been regarded as a diaspora. Ethnographic studies about diasporas and their usage of the Internet have often described diasporic websites as discoursive spaces in which new, hydrid identities are negotiated and stereotyping and marginalizing discourses about diasporic subjects are challenged. The role of languages in these websites, however, has often been neglected. On the other hand, sociolinguistic studies have highlighted how the Internet provides a space for vernacular language usage in which the relaxation of language norms and users’ creativity play a crucial role in overcoming the limitations in text transmission imposed by the medium. A partial bridge between these two trends of studies has been provided by the analysis of code-switching in diasporic websites, which has shown how meaningful language alternation is used to flag users’ hybrid identities. The study of the relationship between diasporic languages and identities on the Internet clearly appears to be in its infancy and only few case studies have looked at the interactions between each diaspora’s specific cultural and sociolinguistic settings and the usage of the Internet. Furthermore, many diasporas, including the Roma, speak unwritten languages which have not been or are just starting to be standardized. Processes of language standardization have always involved both identity and language policies and have often been pivotal in struggles for nationhood or minority rights recognition. While so far such processes tended to be mostly centralized and top-down, the Internet is offering a space for the spontaneous transition from orality to literacy. Thus, analyzing the interaction between diasporic, non-standardized languages and the identities of their speakers as manifested on the Internet can provide new insights into the relations between diasporic languages and identities and into language standardization processes. The present work investigates these issues by analyzing the on-line usage of Romani, the Indic language spoken by many Roma. The study draws on data collected through an online ethnography from Radio Romani Mahala, a website created and used by the recently dispersed community of the Mitrovica Roma. The data are analyzed both qualitatively, using discourse analytic methods, and quantitatively, using traditional sociolinguistic approaches. Combining such approaches allows drawing a nuanced picture of the phenomena under observation accounting both for micro level, individual patterns of usage and macro level trends shared by all users involved. Particular attention is also paid to the emerging Romani spelling and the role played by individual users in the establishment of shared writing norms. The interdisciplinarity of this approach will show how the interplay between diasporic identities and attitudes, non-standard language ideologies and the possibilities offered by the Internet is leading to effective language codification without the intervention of a central authority and outside the frame of any nation-state policy. Such findings call for a re-thinking of current notions on linguistic human rights. Based on the viability of the Romani model, I thus propose a theory of linguistic pluralism in transnational contexts centred around the notion of cosmopolitan sociabilities, non-utilitarian, everyday interactions creating open and inclusive relations across and even despite perceived cultural divides.
DECLARATION
No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

COPYRIGHT STATEMENT
i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s policy on Presentation of Theses.
ACKNOWLEDGEMENT

As any doctoral thesis, this work would have not been possible without the guidance of my supervisors, Yaron Matras, Adi Kuntsman, Iris Bachmann and Eva Schultze-Berndt. Thanks to all of you for the passion, efforts and care you have put into supervising me. I also have to thank all the colleagues at the Research Institute for Cosmopolitan Cultures and the participants in the language documentation and language contact seminars. Without the incredibly stimulating discussions we had I could not have put my ideas together in a clear way. A special mention is due to Dieter Halwachs and Peter Bakker for their nuggets of advice during the RomIdent project meetings. A special thank to Gertrud Reershemius and Yuni Kim for a really enjoyable examination and for their suggestions. And of course, I have to thank Emrah, Farija, Edi and everyone else at the Mahala for their friendship and for making me enjoy every single moment of my fieldwork. Faliminderes cavalen! Finally a special thank goes to la pandilla. Si no hubiera sido por San Güiven, San Jiulm, los asados y las otras fiestas este doctorado no lo iba a terminar ;)

1 Introduction

The fall of the Eastern Bloc, the dissolution of former Yugoslavia and the subsequent enlargement of the European Union to include former socialist countries contributed to an increase in the movement of people from Eastern to Western Europe which began about a decade earlier. Among them, the Roma are probably the most clearly recognizable group and surely the ones that received, and keep receiving, more media attention. While their presence in the media as subjects of discussion is a topic worth analyzing, the present work is about their presence in a particular medium, the Internet, as actors and producers of content.

As a population of Indian origin spread across Europe over the past five centuries, Roma have often been regarded as a diaspora. The classic definition of diaspora proposed by Safran (1991) rests on a number of core factors, namely migration from a historical homeland, preservation of the memory of this homeland and the perception by diasporic subjects that “they are not - and perhaps cannot be - fully accepted by their host countries” (83-84). As is the case for prototypical examples of diasporas such as the Jews and Armenians, the Roma are indeed conscious of their marginal status in the host societies, aware of their condition as an ethnic minority and have resisted complete assimilation. What sets them apart from classic diasporas is the lack, except in activist narratives, of any widespread consciousness of their common origin in India. However, as noted by Silverman (2012), even if India does not feature as a homeland for all of the Roma, individual groups regard the towns and countries they have left in the past forty years as their homelands. Furthermore, they maintain the memory of these recent homelands. Silverman therefore argues that the dispersal of such smaller groups constitutes a genuine process of diasporization. As a result, like many other diasporas1 Roma are caught in a process of continuous identity negotiations involving people, both Roma and non-Roma, in the various diasporic locations, including the homeland. Such interactions create dense networks of transnational communication, and Silverman vividly charts those created by Balkan Roma through travelling and face-to-face encounters. She also hints at the increasing importance of the Internet in creating and supporting such networks.

I will look precisely at this role of the Internet and, in particular, at the functions played by the languages employed in processes of identity formation and maintenance. My work is thus set against a growing body of interdisciplinary research focusing on the impact of the Internet on diasporas. This body of research comprises studies stemming from a socio-cultural tradition, dealing mostly with processes of identity formation, negotiation and

---

1 For an overview of different cases see the works in Braziel & Mannur (2003).
maintenance, and from a linguistic tradition, dealing with issues of language choice and expansion into media domains. The following review of this research will allow me to highlight some issues that have so far received less attention and that I intend to address in this thesis.

1.1 Diasporas on the Internet

In his introduction to The Media of Diaspora, Karim (2003a) noted how the dispersion of people and their developing of intercontinental networks of communications is not a recent phenomenon. However, since the late 19th century “there has been a technological explosion […] which makes the interactions of a print-dominated world seem as hard-won and as easily erased as the print revolution made earlier forms of cultural traffic appear” (Appadurai 1996: 29). The introduction of the telegraph, the radio, the telephone, satellite TV and the progressively increasing ease with which they allowed people to establish and maintain networks, each put diasporas apart from their historical predecessors. During the last three decades in particular, the Internet has proved to be a favourite medium among many diasporas and they “are changing in the light of new forms of electronic mediation” (Appadurai 1996: 196).

In the first place the Internet helps diasporas to “stay in touch with news and popular culture from the homeland” (Sinclair & Cunningham 2000: 15) by widening the reach of traditional mass media like TV programmes and newspapers. Furthermore, various other Internet tools and computer mediated communication (CMC) modes such as search engines, private e-mails, newsgroups, forums, chat-rooms and websites produced by diaspora members are also largely employed².

Hiller & Franz (2004), in their study of Newfoundland economic migrants, highlight the role played by Internet during the whole migratory process. Beside being instrumental in gaining practical information about the new settling location, the Internet is also used to establish ‘new ties’ both within the host community and the diaspora, to maintain ‘old ties’ in the homeland and to track ‘lost ties’ both within the diaspora and in the homeland. In this process diasporic Newfoundlanders come to constitute a ‘virtual community’ (Rheingold 1993). However, as opposed to virtual communities linked by a single issue of interest (for example the soap fans studied by Baym 2000) Hiller & Franz note how Newfoundlanders share varied interests because of the common origin and the shared migratory experience.

---

² As this work was planned and started before the explosion in popularity of social networks no literature about their usage by diasporas was still available. The papers in the special issue of Journal of Ethnic and Migration Studies ‘Migration and the Internet: Social Networking and Diasporas’ (38/9, November 2012) deal with the topic.
Similarly, in his study about the Indian immigrants newsgroup soc.culture.india, Mitra (1997) argues that the notion of virtual community is more pertinent when speaking of diasporic groups sharing not only a single issue of interest but a broad variety of cultural markers. In this kind of virtual communities, he notes, not only “familiar relationships that have been severed […] can […] be re-established” (Mitra 1997: 63) but they also constitute discursive spaces where the identities of participants are constantly debated and redefined (Mitra 2001). As opposed to traditional media that have spoken for and about migrants and diasporic subjects in stereotypical ways, Mitra notes how the Internet is redesigning the patterns of ownership of means of message production. This change offers diasporas the chance to voice their opinion, challenging dominant discourses about them.

The reestablishment of severed connections is echoed in Tsaliki (2003). She reports how Greek diasporics make use of Instant Relay Chats (IRCs) to reconstruct the fraternity and conviviality of the homeland. As a result, a computer-mediated ‘national’ community characterized by the emergence of normative features, constraints and distinctive languages is created. Grenglish (Greek typed in Latin characters and characterized by switches into English) is the distinctive language of this community, “a shared form of expression and communication that only Greeks, with an Anglophone background nevertheless, can take part in and appreciate and forms a flexible linguistics embodiment of the émigré Greek community” (Tsaliki 2003: 176).

The importance of the languages employed in diasporic virtual communities is also stressed by Kadende-Kaiser (2000). On Burundinet, a diasporic Burundian newsgroup, Kirundi (the national language) “is favoured in instances where information that is culturally specific to Burundi is needed” (Kadende-Kaiser 2000: 142). At the same time French (the colonial language) and English (the Internet lingua franca and the new contact language of many community members) are also used showing that users “take into consideration the demographics of the net as they attempt to reach as many net members as possible” (Kadende-Kaiser 2000: 142). Furthermore, Burundinet is a multiethnic forum in which solutions to the ethnic conflict between Hutu and Tutsi are debated. Importantly, and reminiscent of Mitra’s point about diasporic virtual communities challenging dominant discourses, the solutions debated are openly alternative to those proposed by the conflicting parties.

The fact that websites and CMC modes can by-pass “some of the hierarchical structures of traditional broadcast media” (Karim 2003a: 13) and are used by diasporas to promote alternative viewpoints and challenge dominant discourses, find confirmation in the works of Santianni (2003) and Yang (2003). In a study about the movement for a free Tibet, Santianni (2003) showed how the Internet has become an arena in which the West-Tibet
dialogue was reinscribed. In this arena, the Tibetan diaspora managed to “mitigate ‘cultural imperialist’ readings of Tibet” (Santianni 2003: 196) while at the same time negotiating Western imaginings of Tibet in order to raise political awareness of the Tibetan situation and gain support to the cause for its independence. Analogously, Yang (2003) in his on-line ethnography of ‘cultural China’ shows how in 1998 virtual communities of Chinese expatriates brought to the world attention a wave of ethnic violence targeting Chinese Indonesians in Jakarta by organizing a series of transnational protests against it. Yang also notes that this network of virtual communities, which he describes as a transnational public sphere, is contributing to the development of civil society in the homeland, a process held back by strict state control over traditional means of communication.

Another finding common in most research (Mitra 1997, 2001, Qiu 2003, Tsaliki 2003, Yang 2003) is that, beside offering discursive spaces such as forums and chat-rooms, diasporic websites also offer directories, event listings and news which contents resonate with their audience’s collective identity. Furthermore, as noted by Qiu (2003) in his study of on-line magazines produced and consumed by expatriate Chinese students, many diasporic websites make a point of stimulating community awareness. Diasporic websites, “by mobilising such shared values as culture, national identity and community awareness” (Qiu 2003: 155) then come to provide safe and comfortable spaces for diasporic people to ‘hang out’, share their experiences and thus create and maintain virtual communities. Considering all these elements, it appears that diasporic virtual communities and the spaces they inhabit are characterized by the active production of discourse and “the productive construction of new hybrid identities and cultures through the [...] process of maintenance and negotiation between [...] home and host culture” (Sinclair & Cunningham 2000: 15).

It must be said, however, that this continuous process “in which the identity, meaning and boundaries of a diasporic community are constructed, debated and reimagined” (Mandaville 2003: 135) does not necessarily lead to more tolerant, open attitudes. Mitra (1997), for example, describes nationalist conflicts characterized by verbal aggression on soc.cul.india. King (2003), analyzing the effort to maintain a Rhodesian identity on Inbada, a newsgroup for white Rhodesian expatriates from present-day Zimbabwe, notes how “Rhodesian on the Internet are stuck in an idealistic time-warp about the ‘old days’ in the mother country” (King 2003: 187). Such nostalgia for a romanticised entity does not result in overt racism, but at the same time white Rhodesians downplay the discrimination and inequalities of the past. Overtly nationalist discourse and narratives characterize the soc.cul.croatia newsgroup studied by Stubbs (1999). However, parallel to what Mitra (1997, 2001) noted about on-line diasporic discourses, Stubbs finds that these narratives and discourses are not a replication of mainstream nationalism. Rather than resorting to the
“cheap, unfashionable nationalist sentiments” (Stubbs 1999: 15.1) available through traditional media, diasporic Croatians create new nationalist narratives by creatively engaging with available meanings, images and themes.

Diasporic virtual communities are also not immune to internal conflicts which can often lead to a community split. Skinner (2005) vividly reports about one such clash on the Electronic Evergreen newsgroup. Created and used by diasporic subjects from Montserrat, one of the initial strength of this virtual community was that users had previous face-to-face knowledge of each other. On the long run, however, this very fact resulted in the polarization of the discourse between blacks and whites and ultimately in the fragmentation of the community. The case of the Electronic Evergreen also highlights how CMC in diasporic virtual communities does not necessarily stand in opposition to face-to-face interactions. Rather, as also noted by Tsaliki (2003) and Hiller & Franz (2004), CMC and face-to-face interactions (when possible) interact with, and often reinforce, each other.

A final point that has been made, for example by Mitra (1997, 2001), is that only the computer-literate segments of a diaspora have access to virtual communities. Furthermore, it has been noted that entire diasporas do not benefit from or even engage with the Internet. Ackah & Newman (2003), reporting about the case of the Ghanaian Seventh Day Adventists (GSDA) diaspora, caution against a ‘cyberoptimistic approach’ overstressing the centrality of CMC and forgetting about the users. Although a website has been produced by some members of the diaspora, most of them do not engage with CMC. Ackah & Newman attribute this fact to the GSDA’s preference for communication forms like dressing, singing and eating together which are better carried over in face-to-face encounters at the church. For this specific community, CMC is also unsuited to keep contacts with the homeland due to the difficulties in accessing modern technologies in Ghana. Furthermore, they argue, the GSDA’s desire to communicate in their home language is hindered both by the community low levels of literacy and English dominant position on the Internet.

This remark hints at the importance given by many diasporic subjects to the usage of their home languages and to the challenges posed by the Internet, at least in its early years, to the usage of minority languages. Arguably, as language is often regarded as a key identity element by diasporic subjects, these problems must have affected the way diasporas took to the Internet. Yet, with the exception of Tsaliki and Kadende-Kaiser (see above) the issue of language has often been neglected, leaving out from the picture the tool used to carry out the identity negotiation considered crucial in defining diasporic webspaces.
1.2 The multilingual Internet

The issue of language usage and choice on the Internet is, of course, not limited to diasporic subjects. Due to the medium’s origin in the United States, the Internet was initially an English-only domain. Text transmission, in fact, was originally based on the ASCII (American Standard Code for Information Interchange) character set. ASCII initially only supported the Latin characters used in English and thus discouraged usage of other languages, especially when their alphabets were drastically different from the English one. However, as the Internet spread across the globe many languages other than English started to be increasingly used on the Internet, leading to the development of UNICODE, a more refined protocol of text transmission allowing for the usage of more scripts, including many not based on the Latin alphabet. Although the technical issues were progressively addressed, the dominance of English remained a concern for many Internet users, yet more and more languages started to be used on it. Their increasing presence prompted the development of a trend of sociolinguistics studies on what has been termed ‘the multilingual Internet’ (Wright 2004c, Danet & Herring 2007) trying to understand how Internet users of non-English background were dealing with language choice and usage issues.

The papers in Wright (2004c) present the results of a research undertaken “within the framework of UNESCO’s multidisciplinary project Initiative B@bel […] concerned to promote cultural diversity in general and to seek ways to protect the lesser used languages of the world” (Wright 2004a: 5). The research investigated the position of local languages in respect to English. Quantitative data were drawn from university and high school students in Tanzania, Indonesia, the Unite Arab Emirates, Oman, France, Italy, Poland, Macedonia, Japan and Ukraine. Their analysis showed a good deal of convergence toward English, an increasing presence of prestigious standardized languages and difficulties in the usage of smaller, non-standardized ones. These results, although by the authors’ admission limited to the specific population studied, led them to conclude that the Internet is not “providing a space for diversity in language practice” (Wright 2004a: 12). At the same time, although noting the emergence of innovative and spontaneous spelling practices for Arabic and Japanese, they suggested that standardization would be needed for smaller, non-standardized languages to be effectively used on the Internet.

On the other hand, the studies in Danet & Herring (2007), although confirming the role of English as the Internet lingua franca, show that the Internet is offering spaces for different language practices. Furthermore, by deeply investigating innovative spellings they show how users creatively and playfully overcome the constraints imposed by the medium on smaller languages.
In a quantitative study of CMC features in forum postings and e-mails written in Greek and Greeklish (the same variety of written Greek referred as Grenglish by Tsali 2003, see above), Tseliga (2007) showed how users favour an innovative visual transcription pattern over established phonetic transcription norms to type Greek in Latin characters. In the visual transcription, Latin letters and even numbers are employed on the basis of their visual similarity with Greek characters (i.e.: Αθήνα /Athina/ Athens is transcribed A8hva, Tseliga 2007: 118). Tseliga integrated the quantitative study with semi-structured interviews with Internet users in order to highlight their attitude toward Greeklish. All of her informant expressed strong opinions against the standardization of Greeklish and rather showed a high degree of tolerance toward its variability and appreciation for its lack of strict norms. Furthermore, although aware that the Greek alphabet is now available thanks to the introduction of the UNICODE standard, Tseliga’s informants appeared unwilling to stop using Greeklish. Tseliga links this desire both to the now well established practice of writing Greeklish and to the emerging positive in-group values attached by users to this innovative spelling.

Similar conclusions emerge from the study of ASCII-ized Arabic (Palfreyman & Al Khalil 2007) and Romanized Cantonese (Lee 2007). In the case of ASCII-ized Arabic, Palfreyman & Al Khalil note, as in Greeklish, the usage of numbers resembling the shape of Arabic letters alongside the usage of established transliteration norms. In the case of Cantonese, lacking established transliteration norms, Lee reports that users favour a phonetic transcription. Romanized Cantonese mostly occurs in code-switches from English, particularly in sentence-final particles marking “speech-act types such as questions and requests, attitudinal factors, and emotional coloring” (Lee 2007: 198). In both cases, users expressed strong feelings toward these innovative spellings and stressed how they have come to index a sense of shared identity.

The authors also note how their informants’ creativity and playfulness matches the same attitudes shown in English CMC to compensate for the lack of extra linguistic clues (facial expressions, bodily attitudes, voice volume and so on). Research on English CMC has shown that emoticons (combinations of punctuation symbols to represent facial expressions and the writer’s mood; i.e.: :) for a smile ;) for a wink and :( for sadness), the usage of expressive capitalization (i.e.: to represent shouting) and of acronyms (i.e.: lol, laughing out loud) are employed to enrich textual messages (Herring 2002). This pool of resources is also used in non-English CMC, however, as noted by Katsuno & Yano (2007) on a study of Japanese housewives’ chat-rooms, they assume specific cultural meanings and usages based on users’ background.
Even the second conclusion put forward in Wright (2004a), concerning the need to standardize smaller languages to facilitate their usage on the Internet, has been challenged. Research on Jamaican Creole, Nigerian Pidgin (Deuber & Hinrichs 2007) and Mauritian Creole (Rajah-Carrim 2009) has shown how CMC is promoting the development of new forms of literacy for oral, non-standardized languages. In all these cases, the Internet has become the arena for spontaneous, creative and decentralised processes of language codification. In such processes “users can converge on novel codes […] without ever having to refer to a standard” (Rajah-Carrim 2009: 504). As a result shared writing norms are spontaneously emerging, often leading to spellings that increasingly differ from those proposed by language planning experts (Deuber & Hinrichs 2007).

Thus, in spite of the early dominance of English on the Internet, the medium has proven to be a space for vernacular language use. Furthermore, as it is the case for English, norms regulating language usage in more traditional domains are often flouted. This relaxation of norms, coupled with users’ creativity, allows people to experiment with the available resources in order to overcome the technical difficulties languages other than English might encounter on the Internet. Finally, the writing practices emerging on the Internet are often perceived by users as symbols of shared identities. Yet almost all the studies of the ‘multilingual Internet’ have gathered their data from students’ and young professionals’ websites and largely neglected diasporic websites. Thus, as in the case of the socio-cultural studies discussed above, sociolinguistic studies too do not help us understand how diasporic subjects make use of their linguistic repertoires on the Internet.

1.3 Language choice in diasporic CMC

As Androutsopoulos (2006b) notes, this gap in the understanding of diasporic language practices on the Internet has been partially bridged by the study of code-switching and language choice in diasporic CMC.

In her studies of e-mail written by first generation Greeks in England, Georgakopoulou (1997, 2004) shows how switching between Greek and English serves various discourse related functions such as reframing communication as informal, mitigating face-threatening acts and marking content as jocular (Georgakopoulou 1997). In the case of story telling, switching also helps separate narrations from their evaluation (Georgakopoulou 2004). Georgakopoulou notes how e-mail interactions among her informants are embedded and anchored in both on- and off-line relationships. This is made possible not only by her informants’ shared countries of origin and current settling but also by their condition as students turned professionals or academics. As such, Georgakopoulou notes further, the group is characterized by a common sense of solidarity. Thus, she
concludes that code-switches in e-mails, exactly as they do in oral communications among her informants, strengthen the solidarity among speakers and make shared cultural understanding and contact identities visible and relevant.

It is important to note, however, that “computer-mediated discourse is a site for interactionally meaningful use of language alternation, even in the absence of established offline relationships” (Androutsopoulos 2006b: 522). For example, Paolillo (1996, 1999, 2001) in his studies about Instant Relay Chats (IRC) and Usenet newsgroups catering for the Indian diasporas in UK, US and Canada, notes how, in spite of the prevalence of English, Indian languages are used. The relatively limited usage of ethnic languages and the dominance of English are attributed by Paolillo (1996, 2001) to four inter-relating factors: an on-going generational language shift, ambivalence toward home cultures, the high prestige of English in South Asia and Internet language norms, especially when a multi-ethnic audience is targeted. Ambivalence toward home cultures is particularly evident in a series of contributions to the soc.culture.punjab newsletter (Paolillo 1996) which clearly illustrate the kind of diasporic identity negotiations made possible in virtual spaces (see Mitra 1997, 2001). Concerning language alternation, Paolillo finds that switches to home languages are less frequent in pan-Indian environments than they are in ethnic-specific ones (i.e.: Punjabi, Bengali, Tamil, etc.), thus showing how English is considered an ethnically neutral language in multi-ethnic virtual spaces. Switches, both in multi- and mono-ethnic environments, occur however and are mostly used as displays of ethnic identity in fixed expressions (salutations and similar formulae), poetry and song lyrics. More creative usages of code-switching also occur, although less frequently, in jokes, attention-seekers, insults and appeals (Paolillo 1996).

A situation similar to the one described by Paolillo is described by McClure (2001) in the case of Assyrian. In her comparison of oral and written code-switching among diasporic Assyrians in the US, she notes how “[t]he Internet has become a very important means of uniting the Assyrian community in diaspora” (McClure 2001: 185). Although less frequent than in oral communication, switches on forums and chat-rooms occur and serve mostly to affirm ethnic identity, especially through the usage of fixed formulae in greetings and closings. Furthermore, she highlights how the Internet is also used to promote literacy in Assyrian. This once again challenges the conclusions proposed in Wright (2004a, see above) since standardization of Assyrian is still an on-going process in which three non-Latin alphabets are employed and that a standardized transliteration into Latin characters is currently not available.

In the case of Niuean, an endangered language spoken on the Pacific island of Niue and among its diaspora in Australia and New Zealand, Sperlich (2005) similarly notes that the
Internet is offering a new domain of usage for the language. Niuean is still limited to greetings and other formulae and to the performance of ethnic genres (songs, jokes, etc.), thus the Internet in itself does not appear able to reverse language death trends. However, Sperlich notes that the usage of Niuean alongside English signals the speakers’ desire to express a local identity in a global context.

With the exception of Georgakopoulou, through all the works discussed so far the home language and English are often put in a binary opposition. While the home language indexes indigenous culture, and thus implicitly authenticity, English marks the host cultures and values like assimilation and rootlessness. However, Androutsopoulos’ work (2006b, 2007) provides a more nuanced picture of diasporic multilingual practices on the Internet by pairing the more traditional sociolinguistics approach (quantitative with a focus on the medium and its modes of communication) with online ethnography (qualitative and with a focus on the users).

Looking at websites that cater for the Indian, Persian, Greek, Asian, Moroccan, Turkish and Russian diasporas in Germany, Androutsopoulos (2007) firstly operates a distinction between edited content, produced by the websites’ owners, and users’ content, produced by users in forums hosted on the same websites. This allows him to note how the language policies pursued by owners in the edited content directly influence users’ language choice. In the majority of the cases, edited content is dominantly in German due to the commercial nature of the websites and the owners’ orientation toward a multi-ethnic audience based in Germany. Such choice determines the prevalence of German in users’ content as well. As in the cases discussed above, however, home languages are still employed, showing again formulaic usage of language in salutations, songs and poems to mark identity and discourse oriented functions (addressing, relationship management, challenging other users language choice and so on) including the contextualization of discourse as serious (German) and playful (home languages).

Androutsopoulos (2006b), however, notes that German and home languages do not stand in a clean opposition, with home languages marking indigenous cultures and German marking host culture. The picture is complicated by the presence of English both in edited content (mostly websites’ names and mottos) and users’ content (mostly in nicknames and forum signatures). The presence of English, alongside home languages and German, in the edited content is chiefly attributed to a desire to align the websites with the so-called global culture for commercial purposes while at the same time retaining a sense of ethnic identity. In the case of users’ content, the usage of home languages does not simply index ethnicity and the usage of English index an uncritical adoption of global culture. Forum users, in fact, draw not from ‘global English’ but rather they “appropriate social styles of English
from the global flow of media discourse” (Androutsopoulos 2006b: 541). Switching between these styles of English, German and home languages thus come to index multiple cultural affiliations, each language used as an icon of individual identities, including but not limited to ethnicity.

Codeswitching thus seems to play a crucial role in the identity negotiations that take place in diasporic webspaces. Yet, except for a passing remark by Androutsopoulos (2006b) regarding the usage of the Latin alphabet for languages that traditionally do not use it, no mention is ever made regarding how diasporic subjects cope with the difficulties of using languages other than English outlined in the previous section.

1.4 The challenge for language planning theories

It clearly appears that the study of the relationship between diasporic languages and identities on the Internet is in its infancy and that only few case studies have looked at the interactions between each diaspora’s specific cultural and sociolinguistic settings and the usage of the Internet. Furthermore, many diasporas speak unwritten languages which have not been or are just starting to be standardized.

Processes of language standardization have always involved both identity and language policies and have often been pivotal in struggles for nationhood or minority rights recognition (Fishman 2006, Kaplan & Baldauf 1997, Haugen 1972, Skutnabb-Kangas & Phillipson 1995, Spolsky 2004, Wright 2004b). While so far such processes tended to be mostly centralized and top-down, CMC technologies, as I have noted above, are offering a space for the spontaneous transition from orality to literacy. Thus, analyzing the interaction between diasporic, non-standardized languages and the identities of their speakers as manifested on the Internet can provide new insights into the relations between diasporic languages and identities and into language standardization processes.

The present work investigates these issues by analyzing the on-line usage of Romani, the Indic language spoken by many Roma. Matras (2004) notes how Romani constitutes a perfect example of a primarily oral language which is however increasingly present, in writing, on the Internet. Furthermore, as I discussed at the beginning of this chapter, Roma can be regarded as a diaspora. However, as Silverman (2012) warns, in doing so there is a risk of erasing the differences that do exist between different groups of the whole Romani population. In order to avoid that, I will look specifically at a recently dispersed group originating from Mitrovica, Kosovo, while at the same time comparing their patterns of behaviour with those observed among other Roma.

Data on how the Mitrovica Roma make use of the Internet and how languages are employed to perform their identities in virtual spaces have been collected through an
online ethnography (Hine 2000, Miller & Slater 2000). These data are analyzed both qualitatively, using discourse analytic methods, and quantitatively, using traditional sociolinguistic approaches. As suggested by Androutsopoulos (2006a, see also above) combining such approaches allows to draw a nuanced picture of the phenomena under observation accounting both for micro level, individual patterns of usage and macro level trends shared by all users involved. Particular attention is also paid to the emerging Romani spelling and the role played by individual users in the establishment of shared writing norms. In order to avoid over-generalizing the findings that emerge from this approach, my data are contrasted with those collected by the RomIdent project “The role of language in the transnational formation of Romani identity” in other settings. The interdisciplinarity of this approach and the wide range of data analyzed will show how the interplay between diasporic identities and attitudes, non-standard language ideologies and the possibilities offered by CMC is leading to effective language codification without the intervention of a central authority and outside the frame of any nation-state policy.

Such findings call for a re-thinking of current notions on linguistic human rights. Their achievement has so far been strongly linked to the facilitating role of minority elites and of the state and to centralized efforts at language standardization (Skutnabb-Kangas & Phillipson 1995, Spolsky 2004, Fishman 2006). Based on the viability of the Romani model, I thus propose a theory of linguistic pluralism (Matras 2005) in trans-national contexts centred around the notion of cosmopolitan sociabilities, non-utilitarian, everyday interactions creating open and inclusive relations across and even despite perceived cultural divides (Glick Schiller, Darieva & Gruner-Domic 2011).

In order to do so, in chapter 2 I revise current theories on language standardization and linguistic human rights, looking at their historical development, their connections with nation-building processes and minority struggles. I also present the case of Romani, looking at how its codification is being pursued in different contexts.

In Chapter 3 I introduce the Mitrovica Roma, describing their life in Kosovo, their migratory trajectory leading to the current diasporic condition and showing how a cosmopolitan orientation already characterizes them. Radio Romani Mahala (RRM), the website investigated in the following chapters is also introduced here.

Chapter 4 is an ethnographic account of my fieldwork on RRM. The general characteristics of the chat-room hosted by the website are discussed and related to previous findings about simultaneous, multi-user CMC modes. Looking at a corpus of selected

---

3 The project, founded by the HERA initiative, involved research teams at the Linguistics Department at the University of Aarhus, the Romani Projekt at the University of Graz and the Romani Project at the University of Manchester. Intellectual and financial support from the above project and the Research Institute for Cosmopolitan Cultures at the University of Manchester were essential in carrying out my research.
examples, I qualitatively describe users’ discourse strategies, their patterns of language selection and their implications for the maintenance of group identity.

Following Androutsopoulos (2006a) suggestion that on-line ethnography and traditional sociolinguistic analysis should be combined to achieve “a more complex theorizing of the social and contextual diversity of language use on the Internet” (430), in chapter 5 I supplement the ethnographic account with a quantitative analysis. A larger corpus of examples is analyzed to show how the patterns observed in chapter 4 are shared by the majority of RRM users. Particular attention is paid to language choice in user-names, the virtual face of users, and to the relationship between language choice and communicative functions in the messages.

The same corpus is used in chapter 6 to look at the codification of Romani. I briefly describe the salient morpho-syntactic features differentiating the two varieties of Romani spoken by the community and provide quantitative data about their occurrence in the corpus. Concerning the spelling, I look at the solutions adopted for those sounds represented differently across languages using the Latin alphabet (fricatives, affricates, centralized vowels and semi-vowels) and for Romani specific sounds (aspirated consonants and reflexes of retroflexes) and at the level of consistency achieved by users.

In Chapter 7, I provide a summary of the findings and show how they can contribute to the definition of cosmopolitan sociabilities and in particular of their linguistic side. This allows me to then elaborate on the idea of linguistic pluralism thus showing how sociolinguistic theories can benefit from concepts such as cosmopolitanism and sociability practices.
2 Practices and theories of language standardization

In the previous chapter I have shown how the study of diasporas’ engagement with the Internet has often neglected the relationship between groups’ languages and identities. As a result, the efforts of diasporas, such as that of the Mitrovica Roma, to use unwritten, non-standardized languages in a dominantly written medium have largely gone unnoticed, and with them a series of practices that have the potential to provide a basis for new theorizing on language standardization.

In the present chapter I will sketch the historical processes that led to the emergence of standard languages and to the current theorization about them. In particular, I will discuss the links between language standardization and the emergence of the nation-state, focusing on the ideological unity between language, nation and state. The reactions to this ideology, leading to minority struggles for language maintenance and the recognition of linguistic rights, will also be discussed. Finally I will introduce the case of the on-going codification of Romani illustrating the ways in which it differs from previous attempts at language standardization.

2.1 The emergence of standard languages

It is generally agreed that writing emerged around the 4th millennia B.C. to help clerks in carrying out their duties when “the complexity of trade and administration [in the ancient kingdoms] reached a point at which it outstripped the power of memory” (Robinson 2011: 29). However, writing quickly became an instrument for the representation and legitimization of power as well. This connection between writing and power can already be seen in some of the most prominent attestations of early writing systems like legal and celebratory texts (Robinson 2011). Furthermore, writing also became fundamental in the fixation and transmission of religious texts. As Spolsky (2004) suggests, the need for standardization emerged in these contexts from the belief that “there is a correct and desirable form of the language, distinct from normal practice” (Spolsky 2004: 27). Furthermore, Spolsky links this belief to the idea that consistency and accuracy are required to avoid misunderstanding when communication is intended for a distant receiver. Indeed on the basis of these assumptions, standardization was encouraged by the ancient schools of scribes. However, scribes, clerics and other professionals remained the sole users of standardized languages and such languages were limited to written communication or very formal interactions, like religious practices, until a relatively recent time.
The contemporary concept of standard languages as the language of a nation-state began to emerge in Europe during the 16th and 17th centuries, when kingdoms like France, Spain and Britain began the shift from feudal to state organization.

2.1.1 The Middle Ages linguistic situation

During the Middle Ages, languages were not policed by rulers, resulting in a linguistic situation that was both more local and more ‘international’ than today. The majority of the European population was tied to the land they were born in and spoke the local dialect of one of the European language continua (Romance, Germanic, Celtic, Slavic, Baltic) or of one of the isolated European languages (Finnish, Estonian, Hungarian, Greek, Albanian, Basque). Such linguistic fragmentation did not constitute a problem since adjacent dialects in each continuum were mutually comprehensible and population movements were limited. On the cleavages between different continua, where problems might have arisen, multilingualism was common. Among the ruling and economic elites, multilingualism was common as well, mostly due to the merchants and nobility’s tendency to contract alliances on a continental stage through marriages. This favoured familial multilingualism, which was further strengthened by the higher mobility of these groups. Finally, the various churches acted across large spaces through the usage of their sacred languages: Latin for the Roman Church, Greek and Church Slavonic for the Orthodox Churches. In this fragmented scenario, rulers were not interested in policing linguistic behaviour since they always governed multidialectal, if not multilingual, entities and territories changed hand frequently (Wright 2004b).

2.1.2 The Modern Age and the emergence of the nation-state

However, since the beginning of the Modern Age monarchs started to fixate the borders of their kingdoms and also began to police language, imposing their own variety over other spoken languages and to the detriment of Latin in formal domains. Prescriptive grammars and vocabularies became crucial in the process as shown by the first modern grammar: appearing in Spain in 1492, it was dedicated to Queen Isabella and, according to the author, it would have helped spread Spanish along with the rule of the Catholic Kings (Haugen 1972).

The most emblematic example of the link between language standardization and the rise of the nation-state is probably that of France. Cooper (1989) shows how codification and standardization of the dialect of the capital begun with the 1539 Ordonnance de Villers-Cottêrets establishing French as the official language of the kingdom. Toward the end of
the century Henri IV made Paris the permanent home of the court and his administrative centre. As a consequence, aristocrats and the artists they patronized moved into the capital, forming a social and artistic elite. This elite, characterized by a high level of formal education and a keen interest in the work of classical authors, aimed at bringing the clarity, discipline and order of classical models into French literary practice. By discussing their writings and these issues in the various literary salons of the time, the elite managed to give French an “aura of high culture and to clothe its authority in its language” (Cooper 1989: 8). When in 1624 Cardinal Richelieu was appointed prime minister to Louis XIII, he saw the opportunity to turn one of these salons into an institution that could serve his own project of a modern, centralized state characterized by national unity. The Académie Française was established in 1634 and under Richelieu’s tight control its writers were mobilized to support the Cardinal’s policies in political tracts and laudatory poems. The Académie was also charged with the duty “to give rules to our language and to render it pure, eloquent, and capable of treating the arts and sciences” (Article 24 of the Académie’s statutes’, in Cooper 1989: 10), thus turning French into a cultivated language that could replace Latin in formal domains. However, until the French revolution the standardized language remained limited to the upper classes and only the introduction of mandatory education and conscription into the national army allowed it to spread to all social classes.

English, although imposed as the official language in 1536 with the Act of Union, achieved standardization not through state intervention but through a long process involving grammarians and printers. In the late 15th century two types of orthography could be found in England. Professional scriveners used the relatively consistent London-based standard employed by the scribes of the Royal Chancery in Westminster. Literate laypeople, on the other hand, were employing more idiosyncratic spellings as they tried to reflect regional characteristics in their private writings. Printers tended to follow the former, however since most of them came from Germany or the Netherlands, they occasionally deviated from it both because of the influence of their first languages and because of the typeset they imported. By 1660 however, the printers’ orthography had reached a certain consistency but it was still radically different from those employed in manuscripts. This prompted the publication of textbooks and wordlists as writers were complaining about the influence of non-standard speech and schoolmasters argued that such variations was preventing pupils to properly pronounce and spell English. This series of wordlist and textbooks culminated in 1755 with the publication of Dr. Johnson’s dictionary. As it provided the first authoritative model, it was consistently followed and progressively led to standardization (Salmon 1999).
The English case highlights the strong connection between language standardization and print capitalism. The desire, both among Protestant and Catholics, for direct access to the Bible in the late 16th century prompted its translation into the vernaculars. Printers profited from this demand and, since standardized languages offered a bigger market than the fragment linguistic landscape of the Middle Ages, they favoured them. This resulted in a circular process in which printers contributed to the standardization of the ‘national’ language on a ‘national’ territory by selling a Bible, and their other publications, in that language across that territory (Wright 2004b). As Anderson (1989) pointed out, print capitalism “created the possibility of a new form of imagined community, which [...] set the stage for the modern nation” (pg. 46). It did so by creating new, homogeneous fields of communication below Latin and above the spoken vernaculars, by fixing the language and thus helping to build the image of antiquity crucial to the idea of nation and, finally, by elevating the standardized variety above other related vernaculars and into positions of power.

2.1.3 The Contemporary Age

In France and Britain, the parallel processes of language standardization and the emergence of a national consciousness took place over nearly two centuries. In these cases monarchs were aiming at including and homogenizing disparate language groups living in the territories they had inherited at the time when kingdoms’ borders started to be fixated. The process of inclusion and homogenization accelerated with the shift in the status of people from subjects to citizens as absolutism gave way to democracy. The need to unify the citizens into a single community of communication that can be consulted on matters of state determined the emergence of a civic model of nationalism. In such model the state predates and is instrumental in creating the nation and equipping it with a language. The model, however, introduced the idea of the substantial unity of nation, state and language (Wright 2004b).

As Anderson (1989) notes, this idea could be consciously imitated and exploited by others willing to establish their own nation-state. The ‘pirating’ of this model partially explains the central role played by languages in the European nationalist movements of the 19th and early 20th centuries. The contemporary emergence of the ethnic model of nationalism as elaborated by Herder was also crucial in this stage as it postulated every nation as an independent entity and the right to govern itself. Standardized languages thus became the incontrovertible evidence of the unity of a people and, as anthems and flags, came to be regarded as necessary symbols and rallying points in the claim to nationhood.
Thus, as opposed to the lengthy process of standardization in nation-states characterized by a civic model of nationalism, in ethnic nationalistic movements languages were standardized by a small group of writers over few decades only (see for example Haugen 1972 on the cases of Serbo-Croatian’, Hungarian and Romanian). Furthermore, the spread of standardized languages was no longer a consequence of the emergence of a nation-state, but rather a preliminary condition for the formation of a nation and the establishment of its state.

In spite of these differences, the relative success of both kinds of European nationalist movements in standardizing and spreading a language to a whole polity constituted as a nation-state reinforced the ‘one nation, one language, one state’ ideology. During the second half of the 20th century this ideology was enthusiastically embraced by the new polities emerging from decolonization.

These polities, which reunited linguistically divergent groups under the same territorial unit, found themselves in the same situation of those early European nations that had developed a civic model of nationalism. At the same time they wanted to achieve the ‘one language, one nation, one state’ status as fast as those nations characterized by ethnic nationalism had done. They thus massively engaged in language standardization (Haugen 1972).

2.2 Language standardization in the post-colonial period

The efforts of emerging polities in Africa and Asia to equip themselves with a standardized language led to the development of the field of study know as language planning (LP). Various models intended to both describe and implement the processes needed to standardize and spread a language were developed. The most influential of them, and the one that was more or less explicitly followed by nearly all decolonizing states, was that developed by Einar Haugen (1972).

2.2.1 Status and corpus planning

Starting from the assumption that “every self-respecting nation has to have a language. Not just a medium of communication, a ‘vernacular’ or a ‘dialect’, but a fully developed language”, Haugen (1966: 927) examined the processes through which European nations had developed their own standardized languages. In their practices he identified two main processes which he termed status and corpus planning.

Status planning involves decisions about language selection and dissemination that impact on societal aspects (like perception, attitudes, etc.) of the language that is being
standardized, hence they cover aspects that are external to the language itself. On the other hand, during corpus planning the decisions taken are primarily linguistic and focus on the codification and elaboration of the selected norm, hence they impact on internal aspects of the standardizing language.

According to Haugen, the process can also be analyzed from the point of view of those decisions that impact on the form of a language and those that impact on its functions. Placing the four affected elements (society, language, form, function) in a grid, Haugen proposed the following matrix:

<table>
<thead>
<tr>
<th>Society (status planning)</th>
<th>Form</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection</td>
<td>Society (status planning)</td>
<td>Selection</td>
</tr>
<tr>
<td>Codification</td>
<td>Language (corpus planning)</td>
<td>Codification</td>
</tr>
</tbody>
</table>

Table 2.1: Haugen’s (1972: 271) LP matrix.

Haugen noted how the first step toward standardization has to be the selection of a norm. Generally, if a recognized elite is already established, its variety will prevail. If this is not the case, Haugen identified three main selection principles: the comparative, applicable to genetically related varieties for which a hypothetical mother tongue is reconstructed; the archaizing, in which the supposed mother tongue is identified in older, traditional writings, and finally the statistical in which the most widely used forms from different varieties are combined.

The following step requires the codification of the form: the fixation of an orthography, of a set of grammatical rules and of a basic lexicon able to deal with most of the everyday interactions in which the language would be used. The next phase identified is that of elaboration, during which the language is equipped with the tools required to work properly in formal domains such as administration, scientific and literary writing. Two main approaches can guide this process: importation, in which new materials are drawn from another established standard language or purism, the manipulation of native materials. Since codification aims at minimal variation in form while elaboration aims at maximal variation in function, the two processes often collide. To solve such conflicts, Haugen suggests that standards have to strike a balance between efficiency, the ease of usage and learning, and adequacy, the ability to answer all possible communicative needs of the complex national community. The final step in the standardization process is that of implementation, in which the standard language is spread through education and media into all possible new domains, and therefore to a whole community.
Another core assumption of Haugen’s model, although not explicitly formulated, is that language standardization is a centralized, top-down process in which a small group in power forces its decision onto a wider population. Indeed that had been the case in both civic and ethnic European nationalisms and decolonizing polities followed suite in that their governments, both democratic and dictatorial, took charge of standardization.

2.2.2 The application of the model

As mentioned before, one of the first difficulties decolonizing states had to face was the huge linguistic diversity of their population. Selection was made problematic by the fact that generally none of the native languages were spoken by a clear majority and hardly ever spoken by groups that had enjoyed any measure of power or prestige. Furthermore, native languages were often perceived as carrying over tribal affiliations and thus running contrary to the desired national unity ideal.

Even when it was possible to claim a native language was spoken by a clear majority or by a leading sector of the population, it often required enormous elaboration efforts as it lacked the lexicon and variety of styles required to cope with the need of a modern society. To avoid that, the alternative was to carry on using the colonial languages as they were already equipped for that. In certain contexts they were also perceived as ethnically neutral and could therefore break tribal affiliations (Haugen 1966).

Haugen himself, aware of such problems, suggested either to recognize multiple norms or to resort to the language of the colonizer. However, he considered both solutions as not ideal since in the first instance “a nation feels handicapped if it is required to make use of more than one language for official purposes” (1972: 245) and on the other hand, the languages of the colonizers could carry over an association with past oppression and convey an alien culture.

Whatever solution was chosen, implementation of the norm proved problematic as well. If a native language was chosen its introduction often encountered resistance by speakers of other languages. If colonial languages were chosen the effort required to teach them was considerable as only a very small minority of the population actually spoke them.

In spite of such difficulties, the governments of countries as different as Malaysia, the Philippines, Pakistan (Kaplan & Baldauf 1997), Tanzania and Indonesia (Wright 2004b) persisted in their attempts and a number of policies were drafted, leading to interventions in many sector of the affected polity. Where native languages were chosen elaboration was steadily pursued, although more often than not such languages were not equipped to be

---

1For some detailed accounts of different language planning efforts in postcolonial settings see Part 3 in Kaplan & Baldauf (2004b) and Chapter 4 in Wright (2004).
used as academic languages and higher education was instead carried out in the colonial language. Implementation as well was achieved with partial success: portions of the population did not acquire the national language, either because they actively resisted state measures, were not targeted by such measures or the measures themselves proved to be ill-conceived.

Kaplan & Baldauf (1997) argued convincingly that the partial success of such LP efforts can be largely attributed to one of their common traits: they generally targeted the single language chosen and only within the border of the state. Such policies and interventions hardly took into consideration the other languages spoken within the polity or the influence of developments the chosen language could undergo if it was the national language (when the colonial language was retained) or a minority language (when a native language was chosen) in another state. Furthermore, the authorities’ top-down approach often met with contrary bottom-up reactions. This combination of factors resulted in a series of unplanned development impinging on the desired outcome of achieving unity of nation, language and state.

One of these unplanned developments can be linked to the presence of groups speaking languages other than the standardized national one within the border of the state. Furthermore, the number of such groups in Western countries was increased by the rising presence of migrants. Both migrants and autochthonous minorities found themselves caught between the need to acquire the national standard not to lose social and economic opportunities and the desire to retain their own languages. On the other hand, both in postcolonial settings and in Europe, the nation-state pressured minorities to assimilate. As a result, the rate at which smaller languages were disappearing increased as their speakers progressively shifted to the majority language. Where smaller languages were not disappearing since their speakers tried to counteract their death, this created a friction between minorities and the nation-state and brought the issue of the right to language into prominence in the international debate.

These developments generated a vivid debate among scholars involved in LP, leading to the emergence of two strands within the field: one interested in reversing language shift and another going a step further and campaigning for the recognition of language rights.

2.2.3 Reversing language shift

The various reversing language shift (RLS) efforts undertaken by minorities provided a series of examples of bottom-up approaches to LP. Minorities trying to gain recognition for their communities from the state, in fact, initiated most of them. In some cases they
managed to gain some form of support from the state itself. In others the entire enterprise was carried out without state intervention, or even in open contrast to it.

Being involved or reviewing a wide range of such efforts involving, among others, native and immigrant languages in Australia, New Zealand and the United States, French in Quebec and minority languages in Europe, Fishman (2006) provided an influential model for the study and implementation of RLS programmes.

According to Fishman, RLS is a sub-category of status planning since its main objective is to change the attitudes of speakers toward the shifting language in order to develop language loyalty: a deep, sentimental attachment to the language that must be seen as the only appropriate, authentic and thus non-negotiable means to express the culture of the group. However, Fishman also acknowledged that corpus planning activities must be involved in RLS, but only in some of the 8 main stages into which he broke it down.

If the language is nearly dead its reconstruction is required, thus corpus planning must be undertaken, so that adults can acquire it (stage 8). Once the language has been reconstructed and acquired by adults it must be used in cultural interactions, like literary contests, poetry recitals and so on among the older generations in order to raise its profile and start changing the attitudes of speakers towards it (stage 7). Once this has been achieved the crucial step for a successful RLS programme is to guarantee the intergenerational transmission of the language through family, neighbourhood and community interactions carried out in it (stage 6). To further reinforce intergenerational transmission and guarantee the acquisition of additional varieties and styles, literacy schools, complementary but not alternative to mainstream education, must be established (stage 5). Although the main objective at this stage is to raise the profile of the minority language in respect to the majority one (status planning), some elaboration (corpus planning) is required to provide the language with a greater range of styles. Stages 8 to 5 aim at attaining a functional distribution of the two languages so that the one undergoing RLS is used for formal and informal purposes within the community (thus constituting a low variety in the society at large) and the national language (the high variety) is used in formal interactions outside the community. This situation, diglossia as Fishman defines it, is crucial in maintaining compartmentalization (Fishman 2006) of the two languages and for extension of the two cultures. Compartmentalization is considered the essential condition to maintain loyalty to the language and thus guarantee speakers will secure intergenerational transmission.

Once a stable diglossic situation has been attained, Fishman recognizes that it is possible to transcend it by progressively having the minority language accepted in higher
domains. The first would be mainstream education for the minority population, either as a subject, a means of instruction in part of the curriculum or as the sole means of instruction in dedicated schools (stage 4). Following mainstream education, the minority language should start to be used in the local, non-neighbourhood work sphere (stage 3), in the local, non-neighbourhood media and governmental services (stage 2) and finally in national education, work sphere, media and government activities alongside the national language (stage 1). From stage 3 onwards constant elaboration (corpus planning) would be required to keep the minority language at the same level of functionality of the national language while maintaining the two distinguished.

During the latter three stages, maintaining compartmentalization appears to be of even greater importance than in previous ones, since the rewards offered by access to the main culture through the national language can easily drive community members away from their language. Furthermore, all through the 8 RLS stages Fishman maintains that total devotion and dedication from all members of the community involved are of central importance for keeping the bottom-up spirit of the whole enterprise. In this regard, Fishman considers that “government too often needs to be held at arms length, to avoid the bureaucratization, interference, regulation and dependency that government tends to breed” (2006: 228).

In spite of Fishman’s aversion for the state, it is interesting to note that the ideological basis of RLS, language loyalty, and the accompanying compartmentalization of languages and cultures are predicated on the exact same unity of people and language at the core of the nation-state ideology. Furthermore, as RLS draws on Haugen’s model of LP, even the tools employed are the same used by nation-states.

Thus RLS replicates the top-down LP efforts enacted by nation-states, although it starts as a bottom-up, critical response to them. This continuity between the two can also be seen in the importance given by Fishman to the role of figures like teachers, administrators, group leaders, officials and committee chairs involved in RLS. RLS is therefore as centralized as any other LP activity enacted by the state. Furthermore, given this important role played by such an elite within the community involved in RLS, the process itself can be characterized as top-down in respect to the internal community dynamics, although it has to gain bottom-up support to be successful.

2.2.4 Language rights

The RLS efforts undertaken in Western Europe during the 1970’s by diverse minority groups (i.e.: Flemish in Belgium, Basques and Catalans in Spain), led to a series of
constitutional reforms in various countries that started to recognize the right of their minorities. Similar efforts took place in postcolonial settings and other Western countries with varied results. This incursion of minority language issues in the debates around the rights of both individuals and groups provided a fertile ground for the elaboration of a theory of language rights.

Ruiz Vieytez (2001) provided an historical reconstruction of the emergence of language rights in international treaties. He noted how, until the Final Act of the Congress of Vienna (1815), protection for minorities was limited to their religious rights (for example in the Treaties of Westphalia, 1648 and Oliva, 1660), thus explaining the role played by churches in the maintenance of minority languages during the formation of the first nation-states. Following the Congress of Vienna, and in concomitance with the rise of the ethnic nationalist movements of the 19th century, rights started to be recognized to those populations who were able to present themselves as nations. Given the importance of languages as symbols of national unity these rights always included the right to use the language itself. However, strictly linguistic minorities, those speaking a language different from that of the nation-state and who were unable or unwilling to attach themselves to a nationality, remained unprotected. These minorities were granted the right to use their languages only after World War I with the Paris Peace Treaties. The multilateral agreements and national constitutions that emanated from them stipulated that minority languages could not be banned in private and public, that the states would provide facilities to use such languages in courts and, where concentrations of speakers made it feasible, would promote at least primary education in them. Although both Western and Eastern European states signed the Treaties, the former did not implement them to the same degree of the latter. The crucial point of all the treaties up to this point was that they recognized the rights of groups rather than individuals. This protection of group rights was seen has having fostered the nationalist attitudes used to justify the atrocities committed during World War II. Therefore, within the framework of the United Nations, individual human rights were favoured over group rights to avoid the proliferation of further nationalisms.

However as language is as much a social as an individual phenomenon, the protection of individual rights alone often proved ineffective in helping minorities maintaining their languages (Ruiz Vieytez 2001). Thus, in the 1970’s alongside minority efforts at RLS, a new sensitivity to group rights developed and was enshrined in the constitutional reforms.

---

2 Since the focus of the present work is on a European minority I will only describe the development of language rights in Europe. For a complete overview of treaties not directly referring to Europe see Chapter 8 in Spolsky (2004). For an overview of their implementation in postcolonial settings see section III in Skutnabb-Kangas & Phillipson (1995).
of various European states. This renewed interest for minority rights turned into alarm following the ethnic tension in former Yugoslavia.

In an effort to strike a balance between the recognition and protection of minorities and the desire to avoid the violent fragmentation of nation-states, the Council of Europe promoted the European Charter for Regional or Minority Languages (1992) and the Framework Convention for the Protection of National Minorities (1995) while the Organization for Security and Cooperation in Europe (OSCE) issued the Hague recommendations on education rights of national minorities (1996) and the Oslo recommendations on the linguistic rights of national minorities (1998).

All these instruments recognize the right of minorities to use and develop their languages and to acquire the state language, to access public services in their languages and to use them in the media, private business enterprises and judicial processes. They also detail, either as suggestions or as actual provisions, the actions signatory states are required to undertake to guarantee such rights.

As multilateral agreements, all these instruments need to be first signed and then ratified before they can become effective. As a result, the application of the rights they enshrine has so far been relatively slow and uneven between different countries. The inter-wars split between Western and Eastern Europe in minority protection has been de facto replicated (Ruiz Vieytez 2001). Eastern countries have been in a sense obliged to comply with them since ratification and implementation were prerequisites for European Union accession. Western countries, on the other hand, tended to deal with their minorities through constitutional laws, often managing to avoid full compliance with the treaties or just not signing and ratifying them.

Finally, as they recognize the rights of groups and not of individuals, all these treaties contain a series of escape clauses and limitations such as ‘where feasible’, ‘where number of speakers allows for’ and ‘when there is request’. Thus, even when these treaties are ratified, there is no real obligation on the state to actually provide for the minorities. Therefore they suggested that linguistic rights should be considered a type of

---

3 For an overview on the implementation of such treaties see for example Holt & Packer (2001), Shuibhne (2001), Derhemi (2002) and the studies in Koenig & de Varennes (2001),
human rights. As for other types of human rights, the protection and promotion of linguistic rights must be a duty for the state. However, since linguistic rights are not enshrined in a declaration that will put them on par with other human rights, they campaigned for the adoption of a Universal Declaration of Linguistic Human Rights and possibly of a legally binding Charter or Covenant. According to them, this would guarantee that states not only respect the linguistic rights of minorities to intergenerationally transmit their languages but also make possible and guide the management and development of such languages as they do for the national one (Phillipson, Rannut & Skutnabb-Kangas 1995).

Thus, as opposed to Fishman’s RLS, Skutnabb-Kangas and Phillipson require the state not only to accept and recognize all the languages spoken in their territory but also to actively work so that they can be used in all possible domains. Here again, the assumption is that a central authority is always required to manage languages and guarantee their survival and spread. However, considering the tension between the state central authority, often unwilling to recognize the rights of minorities, and the minority elites campaigning for language rights, a supra-national decisional level is invoked as guardian of minority rights.

2.2.5 Language planning in a global word

The accumulation of LP efforts at national levels, of local efforts at RLS and other LP activities carried out at lower organizational levels than the state and the debate around linguistic rights led, towards the end of the 1990’s to the scholarly realization that the theory around LP required some revision.

Kaplan & Baldauf (1997), the first to voice this concern, noted that alterations in a language engineered through corpus planning have repercussions on the ways the language is used in various environments and also affect the societal arrangements around the language itself. Conversely, any intervention affecting society and engineered through status planning is reflected in changes in the language. Thus they argued that corpus and status planning although useful analytical concepts should not be seen as separate but as interdependent activities. Furthermore, they also noted how different polities and their languages had entered into Haugen’s planning paradigm not at the first stage (selection), suggesting that the model stages should not be seen as sequential.

They also distinguished between macro-planning, carried out at the nation-state level and led top-down by the state, and micro-planning, carried out by minorities, business

---

4 For some examples of such local efforts see chapter 6 and Part 3 in Kaplan & Baldauf (1997) and chapters 10 and 11 in Wright (2004b).
activities and other non-state actors with a bottom-up approach. They also noted the latest was, and predicted will be, practised more widely in the absence of macro-planning. Seeing that micro-planning was often more successful than macro-planning, they suggested that even at the macro level, planning and its implementation should follow a bottom-up approach through consultations with the communities of speakers of all the languages and varieties involved.

This last point brought them to argue that planning conceived in the terms of the ‘one language, one nation, one state’ myth was bound to fail or to create more problems than it solved. In this respect they suggested to take an ecological approach, keeping in mind that every language constitutes part of a larger ecosystem in which interventions that target one language will affect all the others. Furthermore as the language ecosystem is not coterminous with the nation-state and may be overlapping with other similar ecosystems irrespectively of political borders, changes in one nation-state will affect others who share some of the languages affected.

Kaplan & Baldauf therefore suggested that LP should not be conceived as a one-off activity undertaken by a single actor and imposed on everyone else. Rather they saw LP as a long-term, large-scale process involving a reticulated structure of many languages and many activities that require the attention of academic specialists, politicians, bureaucrats and the speaker communities involved.

Echoing Kaplan & Baldauf, Wright noted how both the practices and theories of LP at all level are still “framed by the nationalist paradigm even thought we may actually be moving beyond the national model” (Wright 2004b: 251). She sees this move away from the national model particularly in Europe where the state is sharing sovereignty with both local, sub-national and supra-, trans-national authorities in a layered power structure. In this layered power structure, thanks to the mobility of people and cultural flows afforded by globalization, the opportunities for social mobility are no longer limited to those that can be achieved within the nation-state. Therefore, English acting as a lingua franca is progressively replacing the national languages as the language of utility. On the other hand, the spreading of English and of the attached global culture has fostered a return to local languages as languages of identity.

In this situation, Wright sees an opportunity for minority language speakers. In the national model, monolingualism was seen as the norm, the national standard as the language of both utility and identity and minorities were then under pressure to shift to it as it offered more rewards. On the other hand even monolingual speakers of national languages are now increasingly becoming bilinguals as they need English as much as
minorities do. At the same time English as a *lingua franca* is hardly seen as a language of identity since supranational authorities and actors are less interested in group identities and loyalties than the national ones. Local languages, both national standards and minority ones, and English *lingua franca* bilingualism is becoming a viable alternative. Furthermore, unlike the standard in the national model, no language will be seen as having both utility and identity functions, thus this kind of bilingualism would be more stable as there would be no pressure to shift to English.

Wright admits that such a model is purely speculative, therefore she calls LP scholars to monitor what people are doing and want to do and how the local, national and global levels are developing and interacting linguistically. The accumulation of more evidence would then show if there is a move away from the national in language practice as in other fields of social life.

2.3 The case of Romani

The ongoing development of Romani offers precisely the chance to analyze this nexus between local, national and global, and, given the presence of Roma people in many national settings, requires an approach that looks beyond the national model of language.

Romani, the only Indic language spoken in Europe since the Middle Ages, remained an oral language until recently. The first systematic attempts at writing Romani date back to the 19th century, but were carried out by scholars attempting to document the language rather than its speakers. Over time an academic tradition characterized by the usage of the Roman alphabet has been established. For Romani, as it is the case for other languages, it is needed to represent sounds absent in Latin. Academics favoured the usage of the Slavic accented characters and introduced some grapheme combinations to represent such sounds.

Beside academic texts, the other early attestations of written Romani date to the early 20th century with the publication in Germany of a partial translation of the Gospels and on propaganda materials published in USSR in the 1930s. After War World II missionary translations in Romani appeared in Europe and North America. In all cases, however, Roma were not involved in the production of these texts (Matras 1999). It was only toward the end of the 1960’s and with the appearance of a Romani intellectual and activist scene that Roma started to write their language.

2.3.1 Romani activists and intellectuals

The two most emblematic early cases of activists and intellectuals involvement in the codification of Romani are those of the International Romani Union (IRU) and of the
Czechoslovak Union of Roma (Svaz Cikánù-Romù, SCR). In both cases, they tapped into the general change in awareness around minority rights characterizing the international debate around the 1970′s (see above).

IRU, founded in 1971 in London, united a circle of a few dozens Romani leaders comprising linguists, intellectuals and activists of both Romani and non-Romani origins. In their efforts to gain recognition for Roma rights, IRU embraced a typical nationalist ideology in that they tried to portray the Roma as a unified population characterized by a common language and all other symbols of nationhood. At its founding congress a flag and an anthem were adopted and in following congresses and conferences various drafts for a unified alphabets to be used in correspondence, literature, dictionaries, grammars and bible translations were discussed (Matras & Reershemius 1991).

This process culminated with the adoption in 1990 of Courthiade’s standardized Romani elaborated through a relatively strict adherence to Haugen’s model. According to its proponent, selection was based on a statistical approach, thus including the common forms more frequently found in various dialects. During codification a multi-dialectal alphabet had been developed in which the diacritics ě and ĺ are used alongside what Courthiade refers as ‘archegraphemic’ symbols. Ń would be read differently as /dž/, /ž/ or /ʑ/ depending on the reader’s dialect, while θ, q and ç respectively represent the sound variation /t~d/, /k~g/ and /s~c/ in the case endings of nouns and pronouns, thus for example /mange/ ‘for me’ would be spelled mange and /tuke/ ‘for you’ tuqe. In terms of elaboration, Courthiade discouraged the adoption of neologisms from other European languages and resorted to internal coinages or importation from Hindi, seen as a link with the Indic root of Romani (Courthiade 1989, 1992). For implementation, waving IRU’s endorsement of its standard Romani, Courthiade lobbied at both the European Commission and the EU. The first agreed to fund publications using his alphabet and the second funded a standardization commission (Matras & Reershemius 1991).

Although initially endorsed by IRU, members progressively grew unsatisfied with Courthiade’s standard Romani (for reactions to the proposal see for example Hancock 1993, Kochanowski 1995), mostly due to the complexity of the proposed spelling and to a series of selection choices that seemed unjustified. The top-down efforts of the standardization commissions also failed due to the widespread refusal of Romani writers to accept its authority. However, a number of texts, mostly IRU declarations and publications produced or supported by some IRU activists have been produced. Matras notes how all such texts have no real communicative function, but rather what he defines as a rallying
function: they flag the authors’ “ideological commitment and political allegiance and identification” (Matras 1999: 496) with IRU.

While Courthiade’s effort was informed by Haugen’s model and its implementation attempted top-down, the Czech/Slovak experience was much more bottom-up. Hübischmannová (1995) describes the whole process as a ‘trial and error’ one. Here again the process was initiated by a group of intellectuals and activists who, thanks to the opening of possibilities offered by the Prague Spring, founded the Union of Roma (Svaz Cikánů-Romů, SCR). The SCR published a bulletin, Romano l’il (Romani paper) in which texts in Romani started to appear. At this point, the linguistic commission of the SCR proposed an orthography. No debate occurred about the selection of the variety, simply Slovak Romani was chosen as the majority of contributors were Slovak Roma and Slovak Roma constituted the majority of the audience. Similarly, the orthography was based on the Czech alphabet as it was the writing system already known by both writers and audience, but combinations drawn from the academic conventions were introduced to represent Romani specific sounds.

Although the SCR orthography was never endorsed by the state, it became the blueprint for the majority of Romani publications. Elaboration of the language proceeded spontaneously as writers needed to introduce neologisms, generally borrowing from Czech/Slovak, and to broaden the stylistic repertoire of the language, generally by calquing the structures employed in the two majority languages.

The repression of the Prague Spring disrupted the process. However, after the fall of the Berlin Wall, the new political climate allowed for the establishment of various NGOs that included the promotion of Romani among their objective. Writing in Romani became increasingly popular and the SCR orthography continued to be used in the now independent Czech Republic and Slovakia. Considerable variation in spelling emerged as writers speaking other varieties than Slovak Romani tried to represent the specificity of their regional varieties. Elaboration as well continued to be carried out individually by writers. However, a purist tendency started to emerge and although calquing Czech or Slovak remained a popular choice, various authors started to introduce neologism by manipulating native materials or borrowing from other Romani dialects and Hindi. In spite of such increasing variation, Hübischmannová and Neustupný (1996) noted how comprehension was not affected.

Similar processes, always involving Romani NGOs as the main actors, took place in various countries including Hungary, Bulgaria, Poland and Serbia, leading to a proliferation of texts, mostly translations of world literature, ABC books, collections of
Romani fairytales and of original poetry. In all these cases, the writer’s immediate spoken variety was chosen as the basis for selection, although convergence on regional norms occasionally happened, and similar elaboration strategies were adopted in different contexts, namely the elaboration of native materials and borrowings from major languages including English-based internationalism. Similarly, different yet compatible spelling systems emerged, all characterized by the usage of the Latin characters for Slavic languages and the introduction of combinations drawn from the academic tradition in order to represent Romani specific sounds.

Even in these cases, however, the main function of the texts was not communicative, but rather emblematic: all texts served primarily to highlight the ability of Romani to function in formal domain, thus supporting the struggle for the socio-cultural equality of the Roma and symbolizing the need and demand for socio-political integration.

Yet the need to coordinate efforts among different NGOs in order to influence politics at a higher level than the national, namely the EU, led their members, whose command of English was relatively poor, to use Romani, the language they all shared, in correspondence among themselves. As a result the written language also started to acquire communicative functions (Matras & Reershemius 1991, Matras 1999).

2.3.2 Rights recognition and state support

The mobilization of Romani activists and their lobbying at both national and EU levels resulted in various states, particularly those that ratified the EU charters (see above), recognizing Romani as a minority language and providing some form of support for it.

Since 2000, Roma are recognized as a national minority in Sweden and Romani is granted the status of official minority language. The national radio broadcaster offers programmes in Romani and Roma pupils have the right to study in their own language. Following the spontaneous production of educational materials in the 1980’s, the Language Council at the Institute for Language and Folklore (http://www.sprakradet.se/) follows a multidialectal approach in carrying out LP activities. All their materials are published in the three main varieties of Romani spoken in the country (Kaderaš and Lovara by groups present since the 19th century and Arli by recent Balkan migrants) strictly following academic conventions in spelling.

Education in Romani also features prominently in Romania, where a national Romani language curriculum was adopted in 1999. Employing a slightly modified version of Courthiade’s standard (see above) the curriculum has since received widescale
implementation, through all levels, from pre-school to higher education (Halwachs 2011). However, Courthiade’s standard

is often criticised as artificial by local Roma activists and teachers; first of all, because neither pupils nor their parents are able to identify with this variety. In these accounts, the standard is described as distant to local varieties and – as it has almost no functions outside the classrooms – it is also valued as useless for the future life of the pupils. But there is no reliable evaluation of Romani teaching in Romania which proves these impressions as generally valid. (2011: 399)

A more successful and far reaching support for Romani can be seen in the case of the Republic of Macedonia. Following Macedonian independence in 1991 and under pressure by the EU and other transnational organization, Romani was declared one of the official minority languages. In 1992 a state sponsored conference gathered Romani activists, politicians and scholars to set the basis for Romani codification (Friedman 1995).

Writing Romani in Yugoslavia was a more extensive phenomenon than in other countries thanks to the usage of Romani by Romani intellectuals and singers and its presence in weekly radio and TV broadcast. A prescriptive bi-dialectal grammar covering Arli and Gurbet Romani, the more widely spoken dialects in Macedonia, had been published by Kepeski & Yusuf in 1980 and was chosen as the base for codification. The more marked Arli features were limited, resorting to elements selected from the other Romani dialects spoken in the region. The Latin alphabet for South Slavic languages (Serbo/Croatian) was preferred over Cyrillic and Courthiade’s alphabet in order to maintain a link with Latin-based spellings developed in other countries. The representation of Romani-specific sounds was again drawn upon the academic transcription (Friedman 1995).

In the early stages, both orthography and syntax showed a similar degree of variation to that encountered in the Czech/Slovak case. However, elaboration was carried out in a more centralized way, although resorting to the same range of sources for the introduction of neologisms and stylistic enrichment. Implementation, beside a continued presence of Romani in the media, also involved its usage in the 1994 Macedonian census. As a result, the forms used for the census and other materials published around that time showed increased consistency (Friedman 1996). This trend toward consistency was confirmed in the further step in the process, the usage of Romani in the periodical Romani Sumnal/Romski Svet (Romani World) (Friedman 1997).

As the process continued it was clear that norm selection was still ongoing and, apart from the 1992 conference, it was carried out informally. At the same time the production
of textbooks, as the language was to be used in schools, dictionaries, newspapers and in literature contributed, again informally, both to codification and elaboration. So, as in the Czech/Slovak case, consensus on the standard was emerging bottom-up through usage (Friedman 2005).

2.3.3 RLS efforts

The new sensitivity toward linguistic rights also meant that, in areas like Finland and in the Austrian province of Burgerland where Romani speakers were shifting toward the majority language, the authorities initiated or supported efforts at RLS.

Through legal changes introduced during the second half of the 1990’s, Finnish Romani, whose active knowledge was lost among the younger generations, was introduced as an optional subject in primary and secondary education. At the same time the Research Institute for the Languages of Finland received funding to carry out LP activities in cooperation with community representatives reunited in the Romani Language Board. Using the local variety of Romani and a spelling based on academic conventions but also including reduplicated letters to represent long articulation as done in Finnish, dictionaries and educational materials were produced and Romani also started to be used in public information materials, in radio broadcasts and in a couple of periodicals. Teaching of Romani has been occasional and has often reached only a minority of potential pupils, mostly due to the lack of qualified teachers. This prompted the institutions involved in LP to arrange annual courses for teachers of Romani. In spite of such efforts “the rights guaranteed by the law do not materialize or they are not fully utilized” (Granqvist 2006: 58) resulting in scarce usage of Romani in public domains and in continued attrition. Besides negative attitudes toward the Roma and their language still harboured by the majority population, the Roma’s own perception of their language as a secret one “has slowed down the written tradition, hindered the public use of the language and impeded the training of teachers and language instruction” (Granqvist 2006: 59).

Similar partial achievements occurred in the case of Roman, an endangered dialect spoken in Burgenland, Austria. After the Austrian government ratified the European Charter for Regional or Minority Languages in 1993, Roma were recognized as an ethnic minority and Roman was given the status of official national minority language. Resources were allocated for documentation and publications in the dialect and linguists based at Graz University cooperated with representatives of the community in order to reconstruct the language and to develop an orthography. Proposals to adopt a script similar to that developing at the same time in Macedonia or to use Hungarian conventions to represent
sounds borrowed from that language were rejected by community members to avoid any association with Croats and Hungarians, at the same time being recognized as national minorities in Austria. Therefore, a German based spelling was adopted but its inconsistencies were minimised (Halwachs 2012). A grammar, a dictionary and educational materials were produced and the language started to be used in periodicals, in extra-curricular lessons in primary schools and in radio broadcasts. Teaching of Roman, however, declined over time mostly due to the small number of potential pupils and since 2007 no courses are offered in schools, but a local NGO provides them using a mobile service centre. Thus, even if the language has seen an expansion into formal domains “neither revival nor reacquisition and maintenance of Burgenland Romani are guaranteed by the language-planning efforts” (Halwachs 2011: 387), although the whole enterprise positively affected self-consciousness and self-esteem among young Burgenland Roma and is now supporting their social integration.

2.3.4 The Internet

If in the various cases discussed above rallying and emblematic functions of written Romani largely outweighed the communicative ones, the embracement of the Internet in the late 1990’s fostered the communicative usage of the language.

E-mailing to and between NGOs representatives became a favourite means of communication. Furthermore, during the first half of the 2000s, the Next Page Foundation, sponsored by the Open Society Institute, financially supported the creation of Romani websites such as Romea.cz (www.romea.cz, where Romani can now only be found in the archive section at www.romea.cz/romanes/index.php) and Rrommedia (www.rromamedia.net). In most cases this resulted in NGOs periodicals moving from print to on-line format. As they continued publishing translations of original materials written in the majority language, the functions of such texts remained largely rallying and emblematic. However, as financial support was discontinued in 2007, most of them stopped publishing new materials. Newsletters hosted by Yahoo/Google, discussion forums and chat-rooms based on standard templates, on the other hand, largely improved communication among NGOs and individuals. Being free services, they are up to the present day largely used by activists to promote and coordinate NGOs activities, discussing the ongoing process of Romani codification and various other topics ranging from the serious to the more light-hearted (Leggio 2011b).

The increasing usage of Romani in CMC also contributed to the proliferation of new spellings as the accented characters commonly used were hard to transmit (see 1.2) and
English-like solutions (i.e.: ch, sh) started to appear. At the same time English-derived internationalisms started to become more common than borrowings from national languages and internal coinages (Matras 1999).

The preference for the Latin alphabet and English-based internationalisms are strikingly visible in the discussion pages of the Romani version of Wikipedia. Initiated by a single contributor, Desiphral, Romani Wikipedia was originally written using the Devanagari alphabet and heavily resorting to Hindi to introduce neologisms. As more people started to contribute, the number of articles using the Latin alphabet rapidly increased, although Desiphral repeatedly removed or edited many of them. Between 2007 and 2008, Desiphral’s actions generated a vivid discussion. He was accused of pursuing a personal LP project that was making Romani impossible to read. In his defence, Desiphral stated that the choice of Devanagari and Hindi roots was justified as they linked Romani to its Indic origins. The dispute escalated and was eventually settled by Wikipedia moderators who ruled Desiphral should not use Devanagari or Hindi roots without the pre-emptive consensus of other contributors. However, Desiphral insisted in his actions, leading in 2011 to his banning from Wikipedia and the progressive removal of all the Romani content using the Devanagari script (Leggio 2011b).

Finally, Romani is now also used on YouTube (Leggio 2012, Matras & Leggio 2012). Again, NGOs and Romani intellectuals publish videos to promote their activities and use Romani, often accompanied by translations and subtitles in English or other major languages, mostly for emblematic purposes. At the same time an increasing number of lay Roma have also started to use the platform. Posting videos depicting family celebrations, everyday activities, religious meetings of Pentecostal congregations or singers’ performances at various events, they create networks attracting other Roma sharing the same interests. In the comments, people exchange their views on the events depicted in the videos, joke with each other and if they migrated from the same areas keep in contact with distant relatives and friends. Romani is the main language of videos and texts, although other languages are regularly used.

On all the different communicative platforms provided by the Internet, a large variety of dialects is used, particularly on websites and social networks attracting Roma from

---
6 Romani is also increasingly used on Facebook. Due to the partially private nature of communication and to the relatively recent spread of Romani on it, such usage was not investigated at this time, but remains an issue of interest for future research.
different backgrounds. Both well established spellings and new ones developed spontaneously by individual users are used to represent such dialects. As it is the case for printed materials, such variation in dialects and spellings does not hinder users’ comprehension of the texts. The clear difference between on-line and printed materials is that communicative functions are overtaking emblematic and rallying functions.

2.4 Summary

In this chapter I have outlined how standard languages have become a fundamental staple of the nation-state ideology.

As a result of this historical process, the standardization of languages and the acquisition of literacy by an entire population have often come to be regarded as enterprises that can only be carried out successfully by a central authority, possibly the state itself or at least with some measure of support from it.

The case of Romani, however, shows how even in the absence of a central LP authority languages can successfully be used in writing. As Hűbschmannová & Neustupný (1996) and Matras (1999, 2004) noted the efforts to expand Romani into the written domain started as a decentralised process. In different locations and at different times, different actors started writing Romani, generally for emblematic and rallying purposes in the struggle for Roma rights. In all cases, Haugen’s standardization model was not followed strictly or completely ignored. Selection always fell on the immediate spoken variety of individual writers; codification targeted almost exclusively the orthography, generally modelled on the writing system of one of Europe’s national languages with some degrees of adjustment inspired by the academic writing tradition to represent Romani specific sounds resulting in a variety of different but compatible spellings (see Table 2.2); elaboration was carried out unsystematically as writers needed to introduce new terms in their work.

As nation-states get involved into the process of Romani expansion, its decentralized nature started to shift towards a pluricentric model (Friedman 2005, Halwachs 2011) as the intervention of states led to a more strict adherence to Haugen’s model, although the products of institutionalized planning programmes remained limited in their reach to single states or regions.

Mostly thanks to the usage of the Internet, however, new spellings continued to appear and continued to be used even where institutionalization occurred.
Through all these stages, written Romani changed from an emblem and a rallying point for activists campaigning for Roma right to an actual tool of communication used within the same circles. Written Romani thus acquired the status of language of both transcence, as it allowed Roma to communicate outside the limited circle of the local, face-to-face community, and identity, helping to maintain a link with their origins (Wright 2004b). This development was not limited to activists and intellectuals but, thanks to the embrace of the Internet, it also involved lay Roma. Romani, although granted official recognition only in a small number of countries, has thus acquired all the functions of a

<table>
<thead>
<tr>
<th>Sound</th>
<th>Aca</th>
<th>@</th>
<th>Bulg</th>
<th>Cze</th>
<th>Mac</th>
<th>Hun</th>
<th>Fin</th>
<th>Swe</th>
<th>Rom</th>
<th>Aus</th>
</tr>
</thead>
<tbody>
<tr>
<td>as in master</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>as in but</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>as in Betsy</td>
<td>c</td>
<td>ts</td>
<td>ts</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>c</td>
<td>c/q</td>
</tr>
<tr>
<td>as in chair</td>
<td>č</td>
<td>č</td>
<td>č</td>
<td>č</td>
<td>č</td>
<td>č</td>
<td>č</td>
<td>č</td>
<td>č</td>
<td>č</td>
</tr>
</tbody>
</table>

* as in church holiday  
  čh čh čh čh čh čh čh čh čh

*aspirated k as in cook house  
  kh kh kh kh kh kh kh kh kh kh

as in left  
  l l l l l l l l

as in man  
  m m m m m m m m m m

as in no  
  n n n n n n n n n n

as in onion  
  nj ny/nj ň ny nj nj n’ nj

as in pot  
  o o o o o o o o o o o

as in pin  
  p p p p p p p p p p

*aspirated p as in top hat  
  ph ph ph ph ph ph ph ph ph ph

as in rose  
  r r r r r r r r r r

roughly as in French rouge  
  ř rr rr

as in see  
  s s s s s s s s s s/s c s

as in ship  
  š sh šš š š š š š š š š sch

as in top  
  t t t t t t t t t t

*aspirated t as in pot hole  
  th th th th th th th th th th

as in tiger  
  tj ty/tj t’ ty tj t’ tj

as in pull  
  u u u u u u u u u u u

as in voice  
  v v v v v v v v v v v

as in comma  
  ö e w y ř ř

as in yes  
  j y/j y y j j j j j j j j j

as in zoo  
  z z z z z z z z z z

as in pleasure  
  ž zh zh ž ž ž ž ž ž ž ž ž ž ž

Table 2.2: Romani spellings: Aca-academic, @-Internet, Bulg-Bulgarian, Cze-Czech/Slovak, Mac-Macedonian, Hun-Hungarian, Fin-Finnish, Swe-Swedish, Rom-Romanian, Aus-Burgenland Romani. Based on RMS (http://romani.humanities.manchester.ac.uk/whatiss/status/codification.shtml)
standardized national language even if its expansion into the written domain was not carried out by a single authority.

To describe this model of language codification, Matras proposed the concept of linguistic pluralism, a bottom-up approach to language codification characterized by the willingness of writers to adjust their usage of the language to the requirements of the situation and that entails a free-enterprise approach to the use of language by individuals and groups, free from the control of power centres. Efficiency of communication [is] the only sanction or reward that is associated with the choice of variants in either phonological shape, lexicon, or spelling (2004: 13).

Linguistic pluralism is an approach that largely pre-dates the spreading of the Internet. Yet, as I have discussed in the previous chapter (see 1.2), the medium is removed from the control of power centres and efficiency of communication is of primary importance to its users. Thus, the Internet has provided a new, easily accessible domain well suited to the practices of linguistic pluralism. In order to provide a fine grained example of what the basic elements of these practices are, in the following chapters I will turn to the specific case of a single website used by the Mitrovica Roma.
3 The Mitrovica Roma

In the previous chapter, I have shown how written Romani changed from an emblem and a rallying point for activists campaigning for Roma rights to an actual tool of communication used within the same circles. In this process written Romani acquired the status of a language of both transcendence and identity: it allowed Roma to communicate outside the limited circle of the local community while at the same time maintaining a link with their origins. As I suggested, this development was not limited to activists and intellectuals but, thanks to the embracement of the Internet, it also involved lay Roma.

In order to study this process outside the channels of institutionalized activism, I have conducted a virtual ethnography of the web-spaces frequented by the Mitrovica Roma. The group was chosen as I had already been involved in traditional face-to-face ethnographic and linguistic investigations with the members of the community settled in Palermo, Italy (Leggio 2011a). Following their dispersal across Europe during the years of conflict in former-Yugoslavia, they have become extremely active on the Internet, thus making them a small-scale representation of the wider European Romani population. In this chapter I will introduce them and present some ethnographic information that will provide the background for understanding their on-line language choices (chapters 4 and 5) and their contribution to the process of Romani codification (chapter 6).

3.1 The Mitrovica Roma community

3.1.1 Life in Kosovo

As their name suggests, Mitrovica Roma originate from Mitrovica, Kosovo. Like the majority of the Roma in Kosovo they were fully urbanized and according to the 1981 Yugoslavian census (the last one to be answered by all the ethnic groups living in Kosovo prior to the ethnic conflict) constituted the second largest Romani community in the region (4299 people). The real figure, however, is believed to have been larger, due to assimilation pressures exerted on them by other groups, mainly Serbs and Albanians, and to the tendency not to disclose their ethnic background to the authorities (Pettan 2000, 2002).

---

1 Due to word count limitations all toponyms are given in the Serbo/Croatian spelling only rather than in the Albanian and Serbo/Croatian. This choice is also justified by the fact that my informants both pronounce and spell such terms following Serbo/Croatian conventions rather than Albanian ones.
The majority of the community lived in the Romani *mahala* (Romani district) along the river Ibar and mostly comprised Gurbet speakers and, in second instance, Arli speakers (Lapov 2004, Leggio 2011a). As Matras (2005) showed, this dialectal difference derives from different periods of settlement in Kosovo. Arli speakers, speaking a Balkan dialect of Romani, were present in the area since the 14th century and were most probably already urbanized in the 17th when Gurbet speakers, speaking a Vlax dialect, moved into the area from Walachia and Moldavia (Leggio 2011a). Beside these dialectal differences, Romani-Albanian-Serbo/Croatian 2 trilingualism was the norm within the community. Until the 1980’s it was also possible to encounter older individuals, particularly males in Arli families, still able to speak Turkish. As reported by Pettan (2002) speaking Turkish, the dominant language in urban settings during the Ottoman domination, was considered, both for Roma and non-Roma, a sign of pre-World War II urbanization, suggesting Gurbet speakers were not urbanized until relatively recent times. Earlier settlement and urbanization can explain why Boretzky (1995) reported Arli speakers in Kosovo enjoyed a higher social status, both within Romani communities and with other ethnic groups. Pettan (2002) as well confirms the higher status of Arli speakers and attributes it to their prominence as musicians. Similarly, Lapov (2004) reports that among the Mitrovica Roma settled in Florence, Italy, Gurbet speaking women tended to switch to Arli in the presence of Arli speaking males. This higher status of Arli speakers, and the attached greater social mobility, can in turn explain why in Mitrovica it was mostly Arli speakers who were shifting to Albanian.

In Mitrovica, the higher prestige of Arli speakers however did not discourage intermarriages between the two subgroups, and they regard themselves as a single community further unified by sharing Islam as their religion. Even one of their self-definition, *Xoraxane Roma* ‘Muslim Roma’, based on religious rather than linguistic affiliation, confirms the two subgroups regarded themselves as a single community.

Like all other Muslims in Kosovo, they belonged to the Sunna branch of Islam and some men were Sufis (dervishes) and belonged to a mystical brotherhood (*tariqa* in Qur’an Arabic) of the Ruфа’i order (Pettan 2002). Such brotherhoods were generally mono-ethnic and the Roma had their owns, led by a member of the community referred to as *šeo* (from Qur’an Arabic *sheik*).

The socio-economic profile of the community was also shared in spite of the linguistic differences. Concerning literacy, the majority was able to read and produce at least basic texts in Serbo/Croatian and Albanian as they generally attended schools until the

---

2 Since at the time Serbian, Croatian, Bosnian and Montenegrion were regarded as dialects of the same language I will refer to them as Serbo/Croatian, the official definition of the time.
mandatory age of 14. A discrete number pursued further education, particularly vocational training, and few individuals also attended university (Lapov 2004). As instruction in Turkish was reserved for Turks only (Pettan 2002), it is safe to assume that individuals speaking Turkish had no literacy, although there is no evidence for that. The professions practised were a mixture of traditional, itinerant Romani activities (musicians, craftsmen, peddlers, seasonal agricultural jobs) and of working class jobs (cleaners, low level employees in the public sector, builders, miners, skilled and unskilled workers) and some families even ran small businesses like restaurants, bars, clothes and textile shops generally catering for the community itself. It was not unusual for Roma men to be employed either in the public or private sectors and particularly the latter was highly regarded as it guaranteed a permanent income and a pension. However, many also engaged in independent work activities once their shifts were over (Pettan 2002). Roma women often worked as housemaids for families belonging to other ethnic groups (Lapov 2004). This range of activities was confirmed in 2006 during an interview I had with Seljo³, representative of the Mitrovica community with the authorities in Palermo from the mid 1990’s to the mid 2000’s:

**SELJO:** Everyone worked in the factories. […] There was a battery factory, a factory of…there were a lot of factories there! In Kragujevac [Serbia] there was a FIAT factory, do you know? In Mitrovica we produced all the parts for cars, trucks, this kind of stuff. […] The only big mine was the Treča … from the Treča they mined gold, metal, uranium, everything, everything! Some were sellers in the markets. My mother, my father were wholesalers, we sold stuff wholesale. Because we bought in Turkey and re-sell… […] And that’s how it worked, and people earned well, it was a good life. Some worked with plastic…they sold plates, stuff […] Do you know Rezak, right? [a neighbour in the camp for nomads of Palermo] They are masters, they do flooring, you know, tiles! He made the whole thing with his father, his brother […] They were called in every house. Someone build a new house and they were the ones who went and did the flooring, you got it? […] There were my uncles, they are in France now, they were all graduate, schooled, they all had jobs, welders, stuff […]

Although Roma were regarded as poorer and less educated, they shared the same socio-economic status as working class groups from other ethnicities. What set them apart from the rest of the population, in Mitrovica as in the rest of Kosovo, was their broader linguistic repertoire and their dominance in professional music-making (Lapov 2004, Pettan 2000, 2002). Pettan (2002) noted how both Roma and all other ethnic groups (the Orthodox Serbs and Montenegrins, the Catholic Croats and some Albanians, the Muslim

---

³ Since at the time my mastery of Romani was not good enough, the interview was conducted in Italian. All personal names have been changed to guarantee anonymity.
Albanians, Turks and Slavs - Muslimani) relied heavily on professional Roma musicians for the celebration of various festivities (Christmas, Easter, the closing of Ramadan, St. George’s Day and its Turkish equivalent Hıdırellez) as well as for the performance of various rituals (weddings, baptisms, circumcisions, soldiers send-off, funeral laments) and rural competitions like horse races and pelivani (bare-handed fight).

Many families relied almost exclusively on performing music at all such events as well as in non-ritual, urban contexts like the kafenave (coffee houses), as Seljo confirms:

SELJO: I was a musician! Listen, I earned more than a doctor! We played for everyone, Albanians, Serbs, we played all kinds of music, right? […] We needed a female singer able to sing in Albanian, Serbian, Romani, all the languages, right? And we went to pick her up in Skopje, in Macedonia…

VIKTOR: So you played for everyone? If an Albanian invited you to play, you went. If a Serb invited you, you went…

SELJO: Of course! As long as they paid us!

According to Pettan (2000, 2002), Roma musicians had managed to carve themselves a niche as custodians and performers of a single, wide musical repertoire deemed essential to the performance of all rituals and celebrations considered traditional by all Kosovo ethnic groups. In turn, Roma prided themselves of their ability to perform the traditional music of Kosovo better than anyone else.

However, as ethnic tensions started to mount during the 1980’s and audiences began to demand ethnically specific repertoires, Roma musicians had to adapt to the situation. Pettan (2000) reports that since they refused to play mono-ethnic repertoires, Roma musician dropped tunes that have come to be considered nationalistic by certain ethnic groups and replaced them by borrowing from sources as varied as Bollywood movies, Turkish pop music, Bosnian and Serbian newly composed folk music, Romani music from other Balkan regions and Western pop music. This broadening of the repertoire was also accompanied by a modernization of the instruments used. For example the čalgija, a Turkish-influenced ensemble comprised of clarinet, violin, accordion, banjo and darbuka (goblet-shaped drum) was turned into an electrified, western ensemble including guitar, synthesizer, bass guitar and drum set, and thus perceived by audiences as ethnically neutral. The musicians’ ability to constantly re-elaborate tunes and their virtuoso style made all audiences accept the innovations the Roma introduced. The same characteristics also allowed Roma musicians to constantly blur the dividing lines other groups tried to impose on the repertoire, to the point that they were able to sneak, for example, ‘Serbian’ tunes into a long stretch of ‘Albanian’ tunes without the audience noticing it (Pettan 2000).
According to Pettan’s informants such an approach was based on the principle of ‘universality’ which they defined as “central to their musicianship, with the clear objective that such a cosmopolitan orientation be superior to ethnically exclusive non-Gypsy repertoires” (Pettan 2000: 271). The peak of this process was reached with the introduction of improvised praise songs with lyrics addressing individuals which came to constitute the main body of a new genre known as tallava.

This open, inclusive and syncretic attitude was not limited to professional musicians only, but also extended to religious and cultural practices and celebrations. Muslim Roma, for example, openly attended the Catholic pilgrimage to Letnica on August 15 or the Orthodox pilgrimage to Gračanica on August 28, or both. The two pilgrimages celebrate the Ascension of the Virgin Mary, the difference in days due to the use of the Gregorian calendar by Orthodox and the Julian by Catholics (Pettan 2002). According to my informants in Palermo, Mitrovica Roma only attended the Letnica pilgrimage. There, as other Roma present (Pettan 2002), they considered participation in the official religious celebrations of marginal importance. What was more important to them was to pay homage to the Virgin Mary in order to secure her favour by touching and kissing her simulacrum, making halos of money around its head, presenting children to it and making sure they touched and kissed it as well. Interestingly, this manner of worshiping, shared with non-Roma Catholics in the Mediterranean, but also Roma in the same area and by those attending the celebration of Saint Sara in Saintes Maries de la Mer, France, was not practiced by non-Roma attending the Letnica pilgrimage.

Considering all these elements, Pettan argued that Roma in Kosovo were at the same time keepers of tradition and mediators, both between the different ethnic groups and of new cultural meanings and forms brought from outside the region. This double role placed them at the centre of Kosovo cultural dynamics, easing the discriminatory pressure they however faced and providing them with a source of income. At the same time, such positioning allowed them to maintain a distinct identity even if under assimilatory pressure. However, due to economic migration first and to the ethnic conflict later, this equilibrium was broken.

3.1.2 Migrations and dispersal

The change from planned to market economy started by the Yugoslav government at the end of the 1960’s resulted, at the end of the 1970’s, in a period of recession (Lapov 2004). On the origins of the tallava in entertainment, female-performed genres limited to female audiences in private settings and the role played by allegedly homosexual male performers in turning it into a dance music for mixed audiences in public spaces see Pettan 2003.
2004). This prompted a migratory wave firstly from south-east to north-west Yugoslavia and then from the whole of Yugoslavia to Western Europe. The Roma were among the first, but not the only ones, to participate in this migration. What singled them out, as it was the case for Roma migrating from other Eastern European countries, was the fact that they migrated as extended families rather than as individuals (Matras 2000).

Until the end of the 1980’s Mitrovica Roma used to migrate to Western Europe for periods of a few months up to a couple of years. They used to return to Mitrovica after having earned some money which they generally invested in the building of new houses for their families. During this period the preferred migratory route was via former-Yugoslavia, then into Northern Italy and from there to France, Belgium or Germany. The other route was via the Adriatic Sea and into Southern Italy (Lapov 2004). The words of Seljo, once again, confirm this:

SELJO: When my family and I came to Sicily it was 1977. We got here and we stayed, because before we were in Lecce, Francavilla Fontana, Brindisi [cities on the south-eastern Italian coast] We went around in trailers, you know? We really lived like Gypsies. We moved around, we camped in Catania, Gela [cities in eastern Sicily]. We were the whole family, my uncles with their family, with the children. It was about 12, 13 trailers, right? We moved from camp to camp, 15 days here, 15 there because the police didn’t want us to stop, right?

VIKTOR: And what you did during that time, as jobs?

SELJO: We begged!

VIKTOR: Just begging?

SELJO: Yeah!

It is interesting to note how Seljo, a teenager when migrating, distanced himself and the sedentary life he led in Kosovo from this experience by saying he and his family lived like Gypsies, while he normally referred to his group as Roma. He also linked this nomadic condition to the hostility of the local authorities. This was generally true in Italy where Yugoslav Roma were often denied living and work permits. As a result, they remained in Italy as illegal migrants moving from place to place and relying on begging and illegal activities as sources of income. The situation, however, was different in other countries were Yugoslav Roma generally managed to get work permits and remained in the same place for their entire stay in the country.

Beside the differences in working patterns depending on the country they went to, this first migratory experience impacted Mitrovica Roma in three ways. First, it opened them up to a series of cultural influences, particularly in music and religious practices which, as we have seen before, were then brought back to Kosovo. Second, as they came in contact with other Roma, both local to the countries of migrations and fellow migrants, they
became more aware of their origin in Kosovo and in Mitrovica in particular. The generation that grew up during the 1980’s in fact started to prefer Kosovače Roma ‘Kosovo Roma’ or Mitricače Roma ‘Mitrovica Roma’ (from the colloquial pronunciation of the town name, Mitrica) over Xoraxane Roma as their self-definition. Finally, the Mitrovica Roma’s linguistic repertoire and degree of education and literacy were affected. While those that had grown up in Kosovo had Serbo/Croatian and Albanian as their main contact languages and achieved education in these languages, the following generations hardly achieved any formal education. As they were constantly travelling around Europe they studied irregularly, often in different countries, and rarely completed mandatory education. They also added French, German and Italian to Serbo/Croatian and Albanian as their contact languages, though they often did not reach good proficiency in any such languages as they were generally exposed irregularly to them.

The migratory experience of the Mitrovica Roma drastically changed with the explosion of the conflict in Croatia and Bosnia-Herzegovina (1991-1992) and the simultaneous increase of the tension between Albanians and Serbs in Kosovo.

Due to two main factors the practice of periodically returning to Mitrovica ended abruptly. First, the preferred migratory route, via former Yugoslavia, was disrupted as the conflict isolated the south from the north and recently independent Slovenia tightly controlled its southern borderer. Second, both Serbs and Albanians each pressured the Roma to side with their respective side. The Albanians, who had structured a parallel society to circumvent the limitations imposed on them by the Serbs, forcefully registered Roma children in Albanian schools and pressured adults to declare themselves as Albanians in exchange of job opportunities. Serbs, in control of the police forces and of the official government structure were acting similarly. Those Roma openly opposing such forced assimilation attempts were regularly marginalized (Orhan 1999). As tension escalated into open conflict, both sides did not hesitate to use the force to recruit Roma, as Seljo referred:

**SELJO:** So, when the war started, the Albanians, you know what? They called you from one side and the Serbs called you from the other! “Because you’ve Yugoslavian passport, you’re a Yugoslavian citizen…” I was Rom! And there were the Albanians and they said “Look! You Roma have our customs, your surnames, names…” D’you get it? Because we have all names and surnames like the Albanians […] So, they sent you the latter “Come on our side!” And the others sent you the letter “Come on ours!” […] The Serbs, I know, I was told, they came armed and dragged you away! Instead the others, like that, they grabbed you and you’d never come back! It happened to a neighbour of mine, he was like my brother, they got him, father and son, and he never came back! He fought with the Albanians and in the end they shot him! They killed both father and son!
In Mitrovica, where Serbs had their stronghold, the Albanian forced the Roma out of the mahala accusing them of being Serbian collaborators and in some cases even destroyed their houses.

In this situation, those Roma who were outside Kosovo tried all they could to bring the rest of their families to Western Europe. The only way was by crossing the Adriatic Sea toward southern Italy on boats provided by human traffickers. Again, Seljo’s testimony is revealing:

SELJO: I brought her [the wife’s] family here. Sisters, nephews, dad, mom…Her brother had a carwash in Cosenza [southern Italy] but we lacked some money. So we collected some, because one person had to pay 2 millions Italian Lira to land in Italy, from Albania, right? So, they were a lot and the problem was you needed even a thousand Deutsche Marks for a child, even a baby. D’you get it? If he was more than 12-13, it was two millions, two thousand Marks! But we brought the whole family. Because mine was already here except Džeka [a brother] and I paid, I worked my finger to the bone to pay seven millions for him, to bring him here.

When in 1999 the Serbian positions in Mitrovica were bombed as part of the NATO intervention in Kosovo, the Romani mahala located close to them was completely destroyed. At this point, the Mitrovica Roma felt they could no longer return to their hometown. The trauma of the loss of the whole life Mitrovica Roma had led in Kosovo transpires clearly from Seljo’s words:

SELJO: A man fights all of his life, works, sacrifices himself, builds a house and then, you start from zero! How can you! You know, that’s the problem! I’ve always said, I swear to God, if I had the house, my own piece of land, I’d go there immediately! We too lost our houses, my father, my brothers, me…who’s going to have that again? […] Our mahala was very big, we were all there…it looked like a village, a town, you know? And we were together…but now one here, one in France, one there, one there…it’s sad, really! […] No, it was good there, it was good! They spoiled all, spoiled all! This war shouldn’t have happened, we’d have remained in peace! It was not needed, because look, half Yugoslavia is scattered across Europe, get it? What’s that worth for if you have no brother, no father, no friend?

However, although fantasizing to return, Seljo has no real intention to go back:

SELJO: It’s hard there! […] If I’d be there with my children, ten, I’d see them dying! They lack drugs, they lack everything, for God’s sake! You enter a pharmacy and what d’you buy? Nothing! There is nothing! […] I was there past August, I went to visit my father’s and my son’s graves…They cried while telling me! People search for food in the trash! Can you live with 60 Euros a month? No, Goddamn, no!
As reported by Lapov (2004), Mitrovica Roma were not aware of their Indian origins or, if they were, did not regard themselves as part of a large, pan-Romani diaspora. Yet Seljo’s words clearly show how the progressive dispersal of the community from the homeland is lived as a diasporic experience. Thus, exactly as suggested by Silverman (2012), through recent processes of dispersion the often forgotten diasporic condition all Roma share is revived. Those smaller groups who dispersed recently then become aware of their diasporic condition, although not necessarily link it to that of all Roma.

3.1.3 Life in diaspora

Currently the members of the once united Roma community of Mitrovica are dispersed in various cities mainly in Italy, France, Belgium and Germany.

In all countries except Italy, they generally managed to obtain refugee status or to maintain their status as labour migrants and as a result were allowed to settle in regular houses. In Italy, where they were often not granted refugee status, they were generally accommodated in camps for nomads, large caravan sites organized to accommodate the Roma’s trailers. Such sites were located on the outskirts of major urban centres, often on non-paved ground and not served with running water and electricity. In these areas, deliberately ignored by the authorities except for regular police control, the Roma often built their own housing with whatever materials they managed to collect. Also, due to their treatment during the 1980’s, begging and illegal activities remained the main sources of income for many. However, the situation progressively improved, particularly in Northern Italy, where many managed to regularize their positions and move out of the camps while at the same time the local authorities started to equip them with basic utilities and prefabricated housing units.

As a consequence of the forced settlement of the community in camps for nomads and the resulting marginalization, Arli speakers lost any prestige, both within the community and with outsiders, and the younger generations in their families have switched to Gurbet. In Florence, Lapov (2004) noted Arli was no longer spoken by individuals born and raised in Italy. Similarly in Palermo I noticed the same generation in the only Arli-speaking family, whose older members had partially shifted to Albanian, were unable to understand Albanian and only spoke the Gurbet Romani dialect used by the majority of the group.

The differing condition of the Mitrovica community in Southern Italy, Northern Italy and other countries resulted in the high mobility of those in the worst conditions. Up to now, many regularly leave Southern Italy for Northern Italy. Few also try to move to other
countries, but unless they have close relatives there who can apply for family reunion, they are often sent back to Italy.

On the other side, the presence of those Mitrovica Roma settled in France, Belgium and Germany, has since the mid 2000’s been challenged. All these countries signed repatriation treaties with Kosovo authorities (Knaus & Wildmann 2010, Visoka & Beha 2010) and started to send Roma back to Kosovo, either forcefully or for a monetary compensation. As we have seen from Seljo’s testimony however, Mitrovica Roma, like all other Roma targeted by such measures, have no intention of returning to Kosovo. Furthermore, the generations born since the 1990’s are growing up relatively well integrated in the Western European contexts and repatriation seriously affected those who were brought to Kosovo (Knaus 2011).

These generations, as families had become considerably less mobile then in the 1980’s, generally attended school regularly in the country of residence, which of course permits them to socialize with their peers and acquire the local culture. As a result of this relative stability, the linguistic profile of the Mitrovica Roma is changing again and is now differentiating as the younger generation is achieving a good level of literacy in a single national language rather than low levels in many languages. Furthermore, through the education system and a higher, regular exposure to Western media, they are also acquiring some form of English.

However, it is not only the younger generation that began acquiring Western European cultural features. Research conducted in the Italian camp for nomads showed that adults do so too (Lapov 2004, Staiti 2000). Just to mention an example I have personally witnessed, the men of the Palermo community have included in their diet a wide array of street food based on real entrails, knowledge and consumption of which is central to the working class Palermitano male identity. The community openness toward religious practice also remains unaffected. Mitrovica Roma remain Muslim but continue to attend Christian pilgrimages, for example to Lourdes. In Palermo this openness is particularly striking as not only Mitrovica Roma have embraced the devotion for Saint Rosalia, the city patron saint, and participate in the two yearly pilgrimages to her sanctuary, but they also visit it as part of their celebrations for Ramadan and Bajram (the celebration of Isaac’s sacrifice) and of Saint’s George Day/Hdırellez.

In a very similar manner to Pettan’s musicians (see above), Mitrovica Roma regard such openness as culturally superior to ethnic exclusivity and actually pride themselves in having mastered at least some aspects of the local cultures. But they also see such openness in utilitarian terms as it is often easing the relationships with their neighbours.
However and particularly when dealing with the authorities, this is still not eliminating discrimination completely. As a result, and as it was happening in Kosovo, the majority of the Mitrovica Roma keeps maintaining a distinct sense of their identity as Roma.

3.2 Radio Romani Mahala

Another change experienced by the community following its dispersal has been the enthusiastic and massive embracement of the Internet. Mitrovica Roma make sure to have at least an old laptop with a mobile internet connection even in the more deprived of Italian camps and those that can afford it do not hesitate to buy new machines and broadband connections. More often than not a single computer serves an extended family, but the wealthier can count on an average of 3-4 machines.

As I noticed during my fieldwork in Palermo, the Internet was used for utilitarian reasons like job searching or used goods trading, to keep in contact with relatives through instant messaging software, for free on-line gaming and video and music consumption. The Mitrovica Roma also established their own presence on the Internet, mostly on YouTube and on a few websites produced by community members. When the present research was planned (January 2009) Mitrovica Roma had no presence on Facebook and were hardly aware of its existence. As I conducted my fieldwork (October 2009-May 2011) they not only found out about it but started to use it as well, however the main hub of their Internet presence was a web-radio named Radio Romani Mahala (http://www.bEEPworld.de/members97/romani-mahala/, RRM from now on).

Young members of the Palermo community introduced me to RRM in summer 2007, but the site has been active since 2004. It had been created by Sultano and Elvis, two musicians living in France and owners of a small recording studio also offering photo and video footages at weddings and other celebrations.

The technology used to broadcast is provided by Flatcast (www.flatcast.com), a German-based portal for multimedia streaming offering free peer-to-peer facilities, including a live chat functionality. At the moment of writing (March 2012) Flatcast hosted nearly 3000 streams, the majority of them in Turkish (2210), German (280) and one of the languages of former-Yugoslavia plus Albanian (116). A quick look at the homepages of a few such streams suggests that the portal is popular among immigrants in German speaking countries, although about 200 streams are in various other languages (English, Dutch, French, Spanish, Slovak, Polish, Russian/Chechen and more). Of these, 18 including RRM are in Romani.
The peer-to-peer facilities offered by Flatcast have to be embedded into a pre-existing website, which in the case of RRM is based on the free, default platform offered by the German company Beepworld.de (www.beepworld.de). The RRM website is not professionally designed, rather RRM’s team has added element after element to the platform provided by Beepworld.de. This resulted in an unusually long homepage that stretches over about six average-size screens. As a consequence, accessing certain contents requires a lengthy scroll down. For practical purposes I have compressed it into three images, each thus comprising two average-size screen views.

At the time of my fieldwork, the homepage opened with the website name (Figure 3.1). In it, not only the *romani mahala* is mentioned, but the river Ibar as well. Furthermore
Kosovo (in this instance only spelled according to Albanian conventions) and Mitrovica are also mentioned. Immediately under the website name was a first address to the audience: *tumencar si i ekipa ko radio* ‘the radio team is with you’ followed by the DJs names. As I discovered during my regular visits to RRM, all of the DJs are members of Sultano’s or Elvis’ extended families and, like the two of them, they are musicians and are also involved in the studio activities. The welcome message, written in the community languages prior to its dispersal (in the following order: Serbo/Croatian, Albanian, Qur’an Arabic and Romani), further reinforced the website connection with Mitrovica and Kosovo. Mostly in Serbo/Croatian, but also using plenty of internationalisms was the phrase presenting the website: *original net kafe chat radio za sve nacije i genracije* ‘original net café chat radio for all nations and generations’. Finally, the message wishing good listening to the audience was entirely in Romani: *I ekipa taro o radio mandjol tumenje sukar asunipe ko nevo bersh -- 2011* ‘the radio team wishes you good listening for the new year -- 2011’.

The Flatcast’s add-on, the actual radio, was located immediately below the welcome message, but users generally had to scroll down the page to actually get to the add-on although they could hear the music even without seeing it. The add-on is active only when the radio is broadcasting and also contains the chat-room window that allows users to communicate while listening to the music. The flop-disk button located on the left of the Flatcast add-on allows users to save the chat-log and thanks to this functionality I was able to save all the sections I participated in.

On the left of the screen was the navigation menu, followed by a promotional banner by Beepworld.de. After the four default links to the homepage, to RRM’s e-mail and, with text in German, to a guestbook (with no messages left on it) and a feedback form, was a link to a now closed French-based NGO, URYD (*Union des Roms de l’ex-Yougoslavie en Diaspora* - Union of the Roma of former-Yugoslavia in Diaspora). Up until 2008, URYD coordinated the efforts of various smaller NGOs assisting Yugoslav Roma trying to settle in France. The following links were mostly to news sites, other on-line radios and on-line TVs.

All the radios are offshoots of RRM itself. I have never seen *Radio Kamlipe* ‘love’ active and it seems to be kept as a backup in case RRM should fail. The other two radios, *Romano Ilo* ‘Romani hearth’ and *Romano Tarikati* ‘Romani way’ (from Qur’an Arabic *tariqa* ‘way, road’) were active only during the Muslim festivities celebrated by the community. *Romano Tarikati*, in particular, only broadcasted *ilahije*, the prayers sung by Sufis during their rituals. A German-based Bosnian website offering a selection of MP3
and transcriptions (in Qur’an Arabic with translation in Serbo/Croatian both spelled using the Latin alphabet) of such prayers was linked along with the radios. Finally, except for a short lived channel produced by Čita a famous singer from the community, all the on-line TVs and news portals were produced by non-Roma (in Serbo/Croatian, Albanian and English) but originated from or were dedicated to Mitrovica and Kosovo. Along those external links there were a few internal ones leading to a couple of YouTube video galleries of singers from the community, to a photo gallery of Sultano, Elvis and the other DJs and to a selection of *sure* (Qur’an chapters) with Serbo/Croatian translation.

![Image](image-url)

**Figure 3.2**: RRM homepage 2 (Accessed 14-05-2011)

The Beepworld.de banner continued down the page on the side of more multimedia content (Figure 3.2). An animated picture of the European Union flag and the Romani phrase *ma bistren tumaro thand* ‘don’t forget your country’ preceded a video by an NGO documenting the situation of the *mahala* ten years after the conflict and a picture of a
NATO soldier during the military intervention. The animated picture of a credit card introduced the promotional section of the homepage (Figure 3.3). Here, along with more animated pictures there was a textbox containing the contact details of Sultano’s studio and a list of the services offered. Of particular interest was the presence, both inside the advertising box and immediately after it, of a stylised version of the Romani flag adopted by IRU (see 2.3.1) as a background for a darbuka, a frame drum and a violin, considered traditional instruments of the Kosovo Roma, and a white dove.

![Figure 3.3: RRM homepage 3 (Accessed 14-05-2011)](image)

During my fieldwork the homepage constantly changed and it still keeps changing while I am writing. The main changes were the addition or rearrangement of animated images or of pictures of Sultano and Elvis taken from the photo gallery. New links were
constantly added to the navigation menu: the ones leading to TVs and news portals for example, were not there in October 2009 and started to appear in late 2010. While new links were added, the old ones were not maintained and for example the ones to URYD and to the ilahiće are currently broken as the website of the first was removed and the one of the second rearranged.

I was struck by the fact that, in spite of its rather clumsy appearance and lack of maintenance, RRM clearly shows how Sultano and Elvis are following the principle of universality espoused by Kosovo Roma musicians during the 1980’s (see 3.1.1) while at the same time creating a memorial to their lost homeland.

While the use of Romani and the naming of the radio itself serve the latter function, the use of the main languages of Kosovo, the opening of the website to all nations and generations and the European Union flag hint at their cosmopolitanism. On the other hand, the usage of Qur’an Arabic and the links to radios and websites dealing with Sufi Islam together with the links to non-Roma websites dedicated to Mitrovica and Kosovo serve both as they simultaneously celebrate the Roma and their religion and place them into the wider context of Kosovo and of the world.

However, Sultano and Elvis are professionals and are thus also using the website to advertise the other services they offer. Yet, even when advertising, they flag both their Romani-ness and their cosmopolitan attitude by inserting the Romani flag and other symbols of Romani identity together with a symbol of peace.

This subtle balancing between cosmopolitanism and specificity is further reinforced by the choice of tunes broadcasted. RRM repertoire is, exactly as the website, based on the principle of universality and includes all the sources used by Roma musician in the 1980’s: South Asian, Turkish, Serbian and Bosnian pop and folk music, Romani music from all around Europe and Western genres such as hip-hop and R&B. The majority of tunes broadcasted, and by far the most popular with the audience, are however tallava songs in Romani and mostly performed by singers originating from the Mitrovica community. Interestingly, alongside tallava songs, now regarded as traditional, the younger musicians of the community have started recording hip-hop and R&B songs in Romani, showing the process of innovation started in Kosovo is continuing even in the diaspora.

Finally, even the spelling used for Romani shows that a culturally open attitude underpins the website design. In the first address to the audience (Figure 3.1), the sound /č/ in tumencar ‘with you’ is spelled with c as it would be in Serbo/Croatian and in the academic transcription. At the same time c represents the sound /čh/ in lace avilen ‘welcome’. Here two sources can be identified: the Italian usage of c before e and i to
represent /č/ and the common on-line Serbo/Croatian usage of plain letters when their accented counterparts are not available (Hentschel 1998). Whatever the source of such usage, it is interesting to note that aspiration is not marked. The sound /š/ on the other hand shows two different spellings. First it is spelled with plain s (again as in on-line Serbo/Croatian) in sukar alen ‘welcome’ and sukar asunipe ‘good listening’, but the same sound in bersh ‘year’ is represented following English usage. Finally the sound /dž/ is spelled dj following Serbo/Croatian on-line usage in mandjol ‘wishes’ and with an English-like j in tumenje ‘to you(PL)’. In the latter case a typo can also be supposed as an explanation for this difference, however, the fact that it has never been corrected suggests that, even if it is a typo, both Sultan and Elvis, as the site producers, and the audience are willing to accept it.

The design of RRM homepage clearly shows a first instance of how the cosmopolitan attitudes characterizing the Mitrovica Roma are continued on-line. It is also interesting to note how they interact with the relaxation of norms characteristic of language usage on the Internet (see 1.2), resulting in an example of a pluralistic approach to language codification (see 2.4).

### 3.3 Summary

In this chapter I have shown how the Mitrovica Roma, now dispersed across Europe, have not lost the cosmopolitan attitudes that characterized the whole of the Kosovo Roma population. Exactly as it was the case in Kosovo, such attitudes do not seem to lead them to embrace a cosmopolitan identity, rather their cosmopolitanism appears to be instrumental, a series of practices to successfully engage in the wider society while at the same time protecting and maintaining their distinct identity as Muslim Roma from Mitrovica. Even RRM homepage shows such attitudes and practices underpinning its design and they are also clearly visible in the differing spellings used for Romani. As the product of a small group of individuals, the pages of RRM tell nothing about the community at large and their engagement with the Internet. However, the chat-room available on it allowed me to observe exactly this, as the dispersed Mitrovica Roma can still be united through the interactions taking place in it. In the following chapters, I will look at these on-line interactions in order to show if and eventually how the users embrace the cosmopolitan attitudes and practices seen on RRM pages.
4 Languages and identities on RRM: ethnographic account

In chapter 1 I introduced some of the reasons why and ways in which diasporas engage with the Internet and the linguistic peculiarities of such usages. I noticed how, with few exceptions, the scholarly discussion about these two aspects has been conducted separately in different fields, namely social sciences, cultural studies and sociolinguistics. In order to bring these two aspects of the discussion together, this chapter will discuss how Mitrovica Roma interact on RRM chat-room.

4.1 Methodology and ethics

As I stated earlier, this thesis is based on a virtual ethnography on RRM. The ethnography took place between October 2009 and March 2011 and during this time I tried to go on-line for at least one hour once a day. However, virtual ethnography is necessarily “interstitial, in that it fits into the other activities of both ethnographer and subjects” (Hine 2000: 65). This meant that while some days I was unable to go on-line, at other times I was able to spend a few good hours interacting on the chat-room and/or browsing the website and following the links included on the homepage.

The first step in my engagement with RRM, from October to December 2009, resembled Herring’s (2004) ‘persistent observation’: a continuous monitoring of selected CMC environments in order to achieve insight onto local norms and discourse practices. During this time I collected a preliminary corpus of interactions while logged in as a guest user (see 4.2 for an explanation of the different log in possibilities) and without interacting on the chat-room. While this form of covert observation might appear ethically problematic, it was needed for two reasons. First, although I had been introduced to RRM and its chat-room, I had no idea of how to appropriately interact on it and did not want to rush into the field just to compromise it. Second, considering the variety of Romani spellings (see 2.3) and the possibility that writers adapt to each other’s choices, I felt I needed a dataset that I could later use to verify the impact of my presence and spelling choices on those of other users. Considering this corpus was collected without the people on RRM being aware of my research, I decided not to publish any data from this corpus and to just use it as a tool to better prepare my ethnographic investigation.

In late February 2010, once I had analyzed the preliminary corpus and felt confident I knew enough about the appropriate ways of participating on RRM chat-room, I contacted the website owners to inform them of my desire to conduct a research about their radio. I sent an e-mail in Romani to the address given in the homepage and added it to my list of
contacts on Windows Live Messenger, as I had noticed RRM users often exchanged this type of contact on the chat-room (see 4.2.3, 4.3.2 and examples 4.6 and 4.8). The e-mail, translated in English, read as follows:

Hi everyone on the radio!
I’m a non-Roma student at the University of Manchester, UK, studying Romani language and culture.
Right now I am conducting a research about how Mitrovica Roma write in Romani on Internet. I would also like to know why and how you started doing it and what is important for you in meeting at Radio Romani Mahala.
If you are fine with me doing this research with you I will regularly be on the chat talking with you about these issues. Parts of what you write will appear in my dissertation and in scientific publications. I will not use your nick-names so that no one can find out who you are and I will not give your contact details to other people.
I hope you will let me work with you and be your friend. If you want to know more about me and my research, write me an e-mail [my university e-mail address] or contact me on Messenger [my private e-mail address].
All the best!

The reply to this message came a few days later, through Messenger. It was Sultano, one of the two owners of the website (see 3.2) wanting to know how I knew Romani and RRM. I explained to him about my previous experiences with the Mitrovica Roma in Palermo and we found out we had friends in common. He allowed me to carry out my research and also to use his real name and the common DJs’ user-name (*ROMANI MAHALA* in bright blue) without anonymizing them. However, as the chat-room has many users, he also said he could not explain the situation to everyone. I therefore reassured him that I would have done that every time it was needed and negotiated permission to use individual user-names on a case-by-case basis.

The actual virtual ethnography, informed by the methods proposed by Miller & Slater (2000) and Hine (2000), took place between March and June 2010. During this period I participated as much as I could on RRM chat-room, logging in with my user-name and taking part into the chat-room discussions. When interacting on the chat-room, besides taking field notes on paper, I always saved the conversation log using the built-in function provided by the Flatcast’s add-on (the flop disk button on the top left corner, figure 4.2). Thanks to this function, the conversations can be saved as either .html or .txt files. While the black background and font colours are defining characteristics of RRM chat-room (see 4.2.1 and Examples 4.1, 4.14 and 4.15 for the role played by colours in the interactions), .html files, which would have allowed saving them, would have been too big for an easily manageable corpus and would have created problems when importing text into other
software. Therefore, I saved everything in .txt (see chapter 5 for information on corpus size and a quantitative analysis of it).

Such a prolonged immersion in the field allowed me to get a good understanding of RRM chat-room dynamics but obviously gave me no insight into the users’ ideas about their motivations, choices and behaviours. Therefore, in the period from August 2010 to March 2011, as suggested by Hine (2000) I attempted to complement immersion with online interviews, either on the chat-room itself or through instant messaging software (IM). However, I was unable to conduct such interviews. On the one hand, the festive and playful atmosphere characterizing the chat-room (see 4.2 and 4.3) prevented other users from taking my attempts at discussing their on-line behaviour seriously. On the other hand, when trying to conduct interviews through IM I was generally met with a polite refusal or my requests were simply ignored and the conversation veered toward more mundane subjects. I suppose both the open and covert refusal can be explained by the research participants understanding of IM as a channel to carry out private and intimate conversations (see 4.2). Therefore, since both on the chat-room and on IM the idea of an interview seemed inappropriate to the research participants, the final part of my fieldwork continued to be characterized by immersion only.

I ended the ethnographic participation once I noticed that nothing new was happening and all the interactions I was involved in or were happening around me started to have a *déjà vu* feeling to them. At this point I went through the various chat-logs, particularly those I had marked as interesting in my field notes.

I will now report about my experiences on RRM, providing an ethnographic overview of chat-room activities and general characteristics (4.2). This overview will provide the background needed to understand a selection of emblematic interactions that took place on the chat-room (4.3). The analysis of these interactions will not only enrich the overview provided earlier, but will also focus on user language choices and the resulting code-switches following the conversation-analytic framework model proposed by Auer (1998).

### 4.2 RRM chat-room

The following section provides an overview of the general characteristics of RRM chat-room. In 4.2.1 I will present chat-room users’ demographics and discuss certain behaviours they share with other CMC users, including the way they compose their user-names. In 4.2.2 I will discuss the usage of well attested CMC discourse features (capitalization, emoticons, etc.) on RRM chat-room. Finally, in 4.2.3 the more specific characteristics of RRM chat-room as a diasporic space will be discussed.
4.2.1 User-names, users’ demographics and general behaviours

As I anticipated when introducing RRM (see 3.2), the actual radio and chat-room facility is the Flatcast’s add-on embedded in the RRM homepage. If no DJ is broadcasting, the add-on is inactive and the space occupied by it shows an error message and the starting time of the last broadcast (figure 4.1). When on opening RRM homepage I found the add-on was inactive I browsed the site and followed the hyperlinks to external ones (see 3.2 for the description of the website).

When instead one of the DJs was broadcasting and the add-on was active (figure 4.2), I joined the chat-room. Since it is publicly accessible, any surfer visiting RRM page is automatically added to the list of participant and given a user-name. Software generated user-names take the form Guest or its German equivalent Gast followed by 8 digits (i.e.: Guest90762413, Gast90744197). The different languages indicate users are connected either through a German language server (as is Flatcast), in which case they receive Gast, or through a different language server, in which case they receive Guest. Although it is possible to both read and post as guest users, regulars on RRM prefer to create free accounts on Flatcast which allows them to participate with a user-name of their choice.

While collecting the preliminary corpus, I was aware that my automated user-name must have appeared to other users as belonging to a lurker and of the possible problems this might have meant when starting the ethnography. However, as I quickly noticed, users
are given a different automated user-name every time they enter RRM, and before I introduced myself no one was able to realize I had already been on RRM. Furthermore as I found out later on, the fact that RRM is simultaneously a radio and a chat-room allows for what in other settings might be regarded as lurkers to be accepted without problems. As I inferred from occasional comments, active users simply assume inactive guests are there just to listen to the music.

Before starting interacting on the chat-room, I had to choose a user-name to display during the ethnographic investigation. I was aiming at having a user-name that could convey both my position as an outsider and my knowledge of the Mitrovica Roma community and of its virtual space. While thinking of one, I remembered the moniker I was given during my fieldwork in Palermo. Besides doing my research, I was also cooperating with a group of voluntary youth workers engaging the community children in various activities. When playing with children I tended to take on the role of the foolish, funny-kind-of-guy and the children often referred to me as o dilo ‘the fool, the crazy’. The adults picked that up and, as I was the only Italian learning Romani and staying at the camp for entire days, started to jokingly compare me to Stephane, the main character of Toni Gatlif’s movie Gadjo Dilo (The Crazy Stranger, 1997), and I became affectionately known as ‘Daniele o dilo’. For my RRM user-name, therefore, I decided to pick that up
and to insert a reference to my origin, coming up with *DanieleItalianoDilo*. The user-name looked appropriate both to the characteristics of user-names (see below) I had observed during the collection of the preliminary corpus and to my desire to convey my position as an outsider. Furthermore, I thought it would have helped any Mitrovica Roma that knew me off-line to figure out who I was if we happened to meet on RRM. Although I did not meet anyone I had known before, the user-name was more than just appropriate: unless I mentioned who I was and what I was doing, other users assumed I was Roma.

Probably even because of my user-name appropriateness, most of those I interacted with more frequently allowed me to use theirs without anonymizing them. This was indicative of the level of trust I managed to achieve, particularly if we consider that user-names represent the virtual face of users.

Similarly to other CMC environments, usernames on RRM tend to include a personal name and show a playful usage of capitalization, symbols or both (Herring 2002). In usernames like _*_FARiJA*_, edy.boy, MiiLKaqoLLaDa* and __EmRaH StyLe__ the usage of capitals and symbols creates visually captivating designs that, as I experienced myself, help other users to immediately identify the messages posted by the given user. The recognisability of user-names is further reinforced by the usage of font colours. Users generally stick to a specific colour (but see Examples 4.14 and 4.15 about the negotiations on this issue) and tend to select it in order to complement the idea they wish to convey through the user-name, as Example 4.1 shows, where **ALPACINO-NR1** was using a warm, bright red (underlined text is in a language other than Romani).

<table>
<thead>
<tr>
<th>RRM Message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a Guest93643799: SVCIOLOPE MANDCE TI FARBA (19:51:51)</td>
<td>I like your colour</td>
</tr>
<tr>
<td>b <strong>ALPACINO-NR1</strong>: GANAV 799 MI FARBA SI SEXYY (19:52:08)</td>
<td>I know 799 my colour is sexyy</td>
</tr>
</tbody>
</table>

**Example 4.1:** Colours as identity

Another characteristic of user-names is that they often include a reference to the place where the user lives, either a town like in ::::Mohamed-Roanne::::, or a country like in ANNA-MISS-FR. This latter characteristic, alongside the reference to famous clothing brands (Dolce&Gabbana, Armani being the more frequent on RRM) or the insertion of English words (beside the ones already shown kiss, girl, mr, boss are frequent) is common in other diasporic CMC environments as well (Androutsopoulos 2007). Other elements commonly found in user-names are Romani words like princesa ‘princess’, čhaj (spelled
CAJ or CAI) ‘Romani girl’, kraljIJo ‘king’ or čhavo (spelled cavo) ‘Romani boy’. Both English and Romani words, brand names and references to popular culture (see **ALPACINO-NR1**, for example) are selected to give a hint about some of the user’s identity traits. I agree with Androutsopoulos’ conclusion about the same tendencies in Germany-based diasporic websites (2007), that this mixture of languages and elements from different cultural spheres signals the users’ desire to align themselves both to their heritage culture and to a wider youth culture, as it will more clearly appear from the analysis of specific interactions.

Besides their preference for youth culture, even the self declared age of users confirms that their bulk is composed of teen-agers and young adults (from 15 to 30 years old). Asking for someone else’s age or providing one’s own is quite common on RRM as it is part of the routine to find out if two users know each other or share relatives and acquaintances off-line (see 4.2.3 and Examples 4.6 and 4.8). An element I take as a confirmation of the users’ reliability is the fact that on regular working days the majority of self-declared teen-agers goes on-line only in the late afternoon and evening, therefore after school time. Similarly, in the morning when users in their 20s are dominant, women, normally staying home to take care of the house and of the younger children, are generally more numerous than men, normally going to work.

Since I mentioned the sex of users, numbers seem equally distributed between men and women. Again, as for age, there could be space for identity games when portraying one’s sex. However, since Romani inflects nouns and adjectives for grammatical gender, it is not only personal names or gendered elements in the user-names that give a hint about the users’ sex. Any self-referring noun or adjective used in the messages contributes to the representation of a user’s sex. Considering I have never seen a user inflecting self-referring nouns and adjectives in a grammatical gender different from that of the user-name, I have always assumed as all other users seem to do, that the grammatical gender of user-names mirrors the actual sex of users. The tendency, as generally documented in IRCs (Herring 2002), is for men to be more active and to normally start interactions, although only a minority of their attempts receives a response. Women, on the other hand, are less active and generally do not start interactions but when they do, they proportionally receive more responses. Reflecting the widespread hetero-normativity I have observed face-to-face among the Mitrovica Roma, RRM users, although engaging equally with same- and different-sex users in terms of frequency, keep sexually allusive messages between different-sex pairs only (see Example 4.10) while sexist remarks are limited to exchanges.
between men. This on-line continuation of an off-line behaviour confirms the impression that for sex, as for age, RRM users do not play identity games.

The explanation for the lack of a well established behaviour in early Instant Relay Chats (IRCs) (Bechar-Israeli 1997, Danet 1998) lies in the fact that playing identity games would interfere with one of the main usages of RRM chat-room. As I will show later (see 4.2.3 and 4.3.2), RRM users are constantly attempting to find out who the person behind a given user-name is. They do this in order to re-establish and maintain ties with distant relatives and acquaintances or to establish new friendships relationships across the diaspora. In both cases they often also aim at meeting face-to-face during some community celebrations in order to reinforce their relationships. Thus, playing identity games would run contrary to the possibility offered by RRM chat-room to reunite the community (see 4.4).

Apart from this difference, the general tone of the conversations on RRM chat-room is in all respects similar to that encountered generally on IRCs and is characterized by playfulness and disinhibition with much of the content being phatic and occasionally banal (Herring 2002).

In this playful atmosphere tensions and small conflicts can occasionally emerge, however they never escalate into full-out fights (flaming). In such cases, the DJs have moderator powers and can ‘kick’ (temporarily expel) or ‘ban’ (permanently expel) users. They, however, have to resort to such measures very rarely, as the users’ constantly attempt at self-policing the discussions. Furthermore, certain users take on the role of peace-keepers, thus avoiding the DJs’ intervention (see Examples 4.3, 4.7 and 4.11). Even when attempts at self-policing fail, the DJs at first delete all chat messages (see Example 4.11) in an implicit warning to people involved in the discussion. Only if after such a warning the discussion continues, the DJs intervene. Because of this co-operative policing of the discussions, during my fieldwork I have seen just a few users been kicked and none being banned.

Clearly in regard to demographics, visual properties of user-names, references to the wider youth culture and overall playful atmosphere, RRM users do not largely differ from users of similar IRCs and other CMC channels. However, RRM users do differ from them, as I mentioned in not engaging in identity play and flaming, but also in regard to the way in which they use certain features of CMC discourse and in their adherence to the so-called net-iquette.
4.2.2 CMC discourse features and net-iquette on RRM

A first difference with traditional IRCs is not due to the way users interact on RRM chat-room, but to how the Flatcast’s add-on works. New messages appear at the top of the screen rather than at the bottom, thus pushing older messages down rather than up (see the timing of messages in figure 4.2). Another difference is the lack of a function to engage in one-to-one conversations, thus users can only have public conversations on RRM.

Apart from these technical differences, the RRM chat-room can appear as chaotic as any other IRC with messages quickly scrolling out of the screen when the room is particularly crowded. As in other IRCs, the pattern of turn taking is disrupted as different conversations regularly overlap with each other and users can be engaged in multiple, unrelated conversations and thus might delay replying to one or more of them. To compensate for these problems, users keep their conversational turn short or break them into more messages to avoid the intended recipient to lose them amongst unrelated messages (see Herring 2002).

Common features of CMC discourse like abbreviations and non-standard spellings can be encountered on RRM. However, the concept of non-standard spelling can only be applied to languages other than Romani, since as we have seen in chapter 2 Romani does not have a standard spelling. However, it is sometimes problematic to clearly identify them. This is due to the fact that users on RRM are second-language learners and thus often do not master the orthographic rules of such languages. Similarly, abbreviations are far more commonly encountered when users employ these same languages (see chapter 6 for a more complete discussion of spelling variation and usage of abbreviations).

On the contrary, written-out laughs, screams and other sounds are used very frequently.

Emoticons, another well-known feature of CMC discourse, are almost non-existent on RRM. Although the Flatcast’s add-on offers a selection of pre-typed and graphically rendered ones (the button with three emoticons in figure 4.2) users never employ them. I have seen only a couple of users typing some of the more common emoticons like the smiley, the wink or the sad face, but even these users employed them very rarely.

Just as the above features of CMC discourse are only partially integrated into RRM usage, the same can be said of various norms of the so-called net-iquette. For example, the usage of capitalization, considered rude in other CMC environments as it represent abusive shouting and thus avoided, is quite frequent on RRM. At the beginning I supposed this norm was known to RRM users but that they, considering the high frequency of formalized shouting (see 4.2.3), tolerated it. However, I noticed that some users often switch, without any apparent reason, from small to capital characters or vice-versa while others, including
some of the DJs, only employ capitals. The same apparent lack of functional distinction between plain and capital letters also occurs on the comments on YouTube videos linked to on RRM. I therefore tried to ask about this, but the few times I did my messages were regularly ignored. Since ignoring a message is the most common strategy on RRM to signal to the poster that the message is irrelevant, I came to the conclusion that the norm about capitalization encountered elsewhere is unknown to RRM users.

Similarly, repeated typing is strongly employed and tolerated by RRM users. On other CMC environments such a practice is accepted as an expressive device up to a point, namely that users do not flood the chat window with long stretches of repeated characters. On RRM the functional usage of repeated typing is the same but it is also employed for singing along with the music played (by posting a long stretch of m, see Example 4.2m) or calling out someone, either as an attention-catcher or as a mocking (see examples 4.2f, 4.3n). Filling up many lines with long stretches of the same character is not just tolerated but actually encouraged by the DJs themselves who often do it, as this particular behaviour is crucial in the formalized shouting that, along with more practices and norms, helps marking RRM as a Romani diasporic space.

4.2.3 RRM chat-room as a diasporic space

That RRM chat-room is a diasporic space is not surprising, considering the frame provided by the website authors (see 3.2). What is interesting are the ways in which users make it so. Considering how overt identity negotiations featured prominently in the literature about diasporic webspaces (see 1.1 and 1.3) I was expecting to find people overtly talking about their lives and experiences as diasporic individuals, even if in the playful tone of chat conversations. Rather, I found out that no such discussion occurred and that users largely prefer to engage in mundane talks about the places they have been to and the persons they have met recently (often drifting into gossiping), their day-to-day struggles at or to find work and the ups and downs of family life. Yet, even these conversations constitute a relative minority of the messages posted and they are largely overwhelmed by greetings, goodbyes and formalized shouting (see Examples 4.2 and 4.3).

Addressing a salutation to the whole audience upon joining or on leaving the chat-room is a very strong norm on RRM, so strong that certain regular users even greet the audience before logging in, therefore as guest users. They then immediately log in and get their username. Even guest users who remain largely inactive generally say at least ‘hello’ when entering the chat-room.
Formalized shouts are words like *jasa, sa, hopa, hopsa, cek, joj* all translatable as ‘c’mon’ and generally show extensive repeated typing. On certain occasion I have even seen users posting a 3 lines long message only containing the word *jasa* with the final *a* taking up nearly all the space. In off-line communication, such words are shouted during the *abava*, any community celebration held for weddings, circumcisions or religious festivities. *Abava* always involve some live music, and these words, used by singers to punctuate and give rhythm to their singing, are also exchanged among audience members and between the audience and the musicians. They are intended to encourage people to dance and to express appreciation for the music played. On the chat-room they are used similarly to show appreciation for the music selected by the DJs and to encourage users to participate in the conversations. Sometimes, as a particularly popular song is played, one or even more users engaged in a conversation can suddenly burst into one such shout, and then quickly turn back to their interlocutors who just ignore the shouting.

Interestingly, in their role as moderators the DJs never sanction guest users who limit themselves to greetings and shouts. On the other hand they often, under the menace of kicking them out, encourage guests who try to interact beyond these limitations to get themselves a user-name.

Greetings and shouts thus constitute the minimum required of users to participate in the exchanges. Although apparently empty of any content, they however serve a very important function: they recreate the atmosphere that characterizes the community *abava*. In Mitrovica, *abava* took place in the street in front of the celebrating family’s house and any resident of the *mahala* was aware of it and could join in, even if just with a greeting and a shout. This is of course impossible today as families live in different places. However, when a family celebrates something, plans and arrangements are made plenty in advance so that at least part of those residing elsewhere can travel to the *abav*. RRM itself is used to let people know an *abav* is taking place. Generally, the *abava* organized by the DJs families are announced, but even other families, particularly if they have hired the services of Sultano’s studio, can do it. In such cases a few lines are added at the top of the homepage informing users about the town, date and invited musicians of the upcoming *abav* while the DJs keep reminding people on chat-room discussions.

*Abava* obviously constitute an occasions during which at least a portion of the community can be reunited. RRM, with its music and chat-room and its virtual presence in any family that owns a networked computer, offers this same opportunity on a daily basis.

As in any off-line festive gathering, people on RRM chat-room take the opportunity to catch up with friends and relatives, to have chilled down conversations and to build new
relationships. The fleeting nature of party talks is not affected by the breaking down of conversational turns typical of chat conversations and actually fit into it almost naturally. Similarly, the joyful atmosphere created by the music and shouts reinforces, and is mutually reinforced by, the light and playful tone of chat conversations. Given this festive atmosphere, maintenance of the community takes place through small talks interspersed with jokes and humour, in which, as I mentioned before, people however try to share bits and pieces of their lives.

RRM users tend to favour Romani when talking about their own community, while they favour contact languages when discussing topics relating to the wider society. This distinction is however not straightforward and as I will show in the next section many instances of code mixing and code switching complicate it.

It must be said that it is not just playful and light conversations RRM users want to engage into. However, the fact that Romani is the favoured language poses a problem if users intend to discuss private and intimate subjects. This is because everyone understands Romani and the chat-room lacks a function to carry on private conversations. In these occasions, users supplement RRM chat-room with IM, generally Windows Live Messenger to which they refer as MSN. It is not uncommon to see users knowing each other both on- and off-line inviting one another to go on IM because they have something they want to tell them in private. Similarly, friendships that start on RRM chat-room can be also sustained through IM. But what I consider one of the most important properties of IM is that it offers relatives and friends who have not seen each other for a long time and suddenly meet on RRM a chance to re-establish a connection.

Because people are on-line with their user-names, a lot of effort is put into discovering who the person behind a user-name is. When a user finds another one interesting based on what they do on the chat-room or when a user-name reminds of some acquaintance or relative, people immediately try to find out who, in off-line life, that user is. The accepted routine for doing that (see examples 4.6 and 4.8) involves asking where people live, who their relatives and friends are, how old are them and so on in the hope of finding out that they either know each other off-line or have some relatives in common. The asking and answering of questions is mutual and always carried out entirely in Romani, possibly punctuated with jokes and other small talks. When and if a point of contact is found, even if does not involve previous acquaintances or a kinship relation, and users desire to carry on their conversation in private, they exchange IM addresses.

This constant search for relatives and old and new friends, immersed in the on-line recreation of off-line community gatherings with their exchange of news and gossip is
what really makes RRM chat-room a diasporic space. In this space, differently from what has been discussed in the scholarship on online diasporas, the community and its essence are not debated and represented to insiders and outsiders. Rather, on RRM chat-room the Mitrovica Roma community is recreated and users actually perform their being members of such a community. This fact, together with the general light tone of the discussion, can also explain why serious discussions relating to the host countries and the wider society are avoided. RRM simply is not the place where to talk about being a Mitrovica Roma in any given country, but the place where to be a Mitrovica Roma.

Being a Mitrovica Roma, however, is not a monolithic, uncomplicated identity shared by all participants, but rather a dynamic set of overlapping practices enacted by each individual user in different ways and at different times. This is particularly evident in the behaviour of certain users, those ones I refer to as ‘strong personas’. They constitute a pool of regular users who over time have established themselves as dominant characters on the chat-room. Each developed their unique personal on-line identity by blending different behavioural facets of what can be regarded as stereotypical characters, like the prankster, the ladies’ man, the independent woman and so on, into the norms and practices that makes RRM chat-room a diasporic space.

Given the textual nature of interactions on RRM, the use of various languages plays a crucial role, both in interactions overall and in the efforts of strong personas in sustaining their on-line identity. In turn, the mixing of and switching between various languages help to further characterize RRM chat-room as a diasporic space in which multiple identities are enacted.

4.3 Language choice on RRM

In order to understand the role of language choice in RRM interactions, I will now analyze a sample of selected conversations. The examples have been chosen as they strongly show the strategies and norms that surround language choice and also highlight the issues discussed in the previous sections. Greetings, goodbyes and formalized shouting are discussed (4.3.1) as are processes of community maintenance (4.3.2). Specific sections are dedicated to the role of languages in humour and in sustaining strong on-line personas (4.3.3), in creating a private space within the public chat-room (4.3.4) and in self-policing (4.3.5).

The analysis of language choice is based on Auer’s conversation-analytic framework (Auer 1998) which distinguishes between code switching as a locally meaningful process in which a given language is favoured at any given time, and language mixing as a 76
meaningful pattern of frequent alternations which however lack clear discourse functions. Furthermore in terms of sequential structure, alternation between stretches entirely in a language is distinguished from insertion of single items from one language into a sentence in another one. In terms of function a distinction is introduced between discourse- and participant-related switches. By sequentially analyzing the interactions I will show how switches align with or challenge previous code choices of other speakers and how they index users’ identities, background knowledge and language norms.

In all the examples I have rearranged the messages for ease of reading, so that the older ones are at the top rather than at the bottom. The black background and font colours have been lost when saving the discussion log. Those user-names for which I had no permission to publish are replaced by U (for user) followed by a number. Underlined text is in a language other than Romani. When interactions took place over about 25 messages, I did not remove the ones irrelevant to the issue been discussed in order to give a feeling of the overlapping nature of conversations. When interactions stretched over more than 25 messages, I removed the irrelevant ones to fit the interaction into a single page. A question mark following the line number signal doubts about the translation due to unusual spellings or usage of slang terms I am not familiar with.

4.3.1 Greetings and goodbyes

As I mentioned (4.2.3) greetings constitute some of the more frequent messages. A typical sequence of greetings and responses is illustrated in Example 4.2.

<table>
<thead>
<tr>
<th>RRM message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a</strong> Admir: selam aleykum sarinenge ko radio (18:22:25)</td>
<td>selam aleykum to everyone on the radio</td>
</tr>
<tr>
<td><strong>b</strong> U1: JASAAAAAAZAAAAA (18:22:28)</td>
<td>c’moooooooorrrrrrrrrrrrrrr</td>
</tr>
<tr>
<td><strong>c</strong> Admiral: zdr zdr dj (18:22:29)</td>
<td>hello hello dj</td>
</tr>
<tr>
<td><strong>d</strong> <em>ROMANI MAHALA</em>: ALEKIM SELMA ADMIR (18:22:33)</td>
<td>aleykum selam Admir</td>
</tr>
<tr>
<td><strong>e</strong> Admir: zdr zdr Guest94423807 (18:22:36)</td>
<td>hello hello Guest94423807</td>
</tr>
<tr>
<td><strong>f</strong> U2: hav ci korii jak emrakhhhhhhhhhhh (18:22:36)</td>
<td>I eat your blind eye emrakhhhhhhhhhhh</td>
</tr>
<tr>
<td><strong>g</strong> MiILka_qoKoLLaDa*: POZ ADMIR (18:22:40)</td>
<td>hello admin</td>
</tr>
<tr>
<td><strong>h</strong> U2: zdr admirr (18:22:43)</td>
<td>hello admin</td>
</tr>
<tr>
<td><strong>i</strong> U3: cekkkkkkkkkkkkk emrah (18:22:45)</td>
<td>c’mooooooorrrrrrrrrrrrrrr</td>
</tr>
<tr>
<td><strong>j</strong> <em>EmRaH StyLe_</em>: NA NA NAMANGAV TU MER PARO Guest94423807 HEHE (18:22:48)</td>
<td>no no I don’t want you to die heavy Guest94423807 hehe</td>
</tr>
<tr>
<td><strong>k</strong> U3: zdr admir (18:22:52)</td>
<td>hello admir</td>
</tr>
<tr>
<td><strong>l</strong> MiILKa_qoKoLLaDa*: MMMMMMMM (18:22:53)</td>
<td>Mmmmmmmmmm [singing noise]</td>
</tr>
<tr>
<td><strong>m</strong> <em>ROMANI MAHALA</em>: TEHRA SI ABAV AMENDE O ERMINI IBRAHIMIIIJOJJJJJ (18:22:53)</td>
<td>there’s a party tomorrow at our place ermin ibrahim c’mooon</td>
</tr>
</tbody>
</table>

Example 4.2: Greetings
Admir, a male born and living in Macedonia, had just joined the chat-room. In quick succession he first addressed the Islamic greeting inserted in a Romani sentence (4.2a), to the whole chat and then a Romani greeting (zdr a shortening for zdravo ‘hello’) specifically to the DJ (4.2c). The first of Admir’s choices can be regarded as the marked one as it flags his identity as a Muslim Roma. The DJ quickly replied (4.2d) and although having been addressed specifically with a Romani, un-marked greeting, chooses the Islamic, marked one. Doing so, he acknowledged both the presence of Admir and his apparent desire to forefront his identity as a Muslim Roma. However, Admir then turned to a guest user and greeted him using zdr again. At this point MiiLKa_qoKoLLaDa* (4.2g), U2 (4.2h) and U3 (2.1), the latter two at the time teasing _–_EmRaH StyLe_–_ (4.2f, 4.2j)1 greeted Admir. Contrary to the DJ, they used Romani greetings (zdr and poz, another shortening for pozdravo ‘hello’). They, therefore, preferred Admir’s un-marked choice and aligned themselves to it by greeting him simply as a Roma.

This different behaviour between DJ and other users is quite systematic. When a user addresses a marked greeting, the DJ, no matter who he actually is, aligns himself to that choice. Other users instead favour un-marked choices, as it can be seen in Example 4.3.

_* FARiJA *_, a female living in Italy, joined the chat-room at a very busy time. As it is common on RRM, her first address (4.3a) was to the whole audience, but also included a specific call to me, to which I did not respond since I missed it because I was starting a conversation with U4 (4.3e). While _* FARiJA *_ first greeting is entirely in Romani, her following one (4.3d) addressed to the DJ, which she knows personally, uses English hello. Here again, the switch to another language is the marked choice as it flags a specific sense of identity, namely that of a strong, modern woman, a persona _* FARiJA *_ is constantly trying to represent on RRM chat-room (see discussion of Example 4.8). The DJ’s response (4.3p) came a mere 7 seconds later, however the number of messages appeared in this short time seemed to push him to open with a surprised oooooo. As we have seen in Example 4.1, the DJ aligned himself with _* FARiJA *_ marked choice, also reinforcing it by spelling hello in an unconventional way that reminds of French pronunciation and thus signalling the fact that he lives in France. U1, engaged in a conversation with U2 (4.3f, 4.3.i and 4.3k), replied a little later (4.3q) but instead favoured a plain, un-marked Romani greeting, thus again ignoring the marked choice.

---

1 Analyzing every instance of teasing would be beyond the scope of the present work. It is important to notice that teasing is largely used in child-directed speech in Romani and features strongly even in adult speech (Réger 1999). Obviously, the playful tone of chat conversations is conducive of such a practice.
It is worth noting that the Romani greetings are, in the dialect of the Mitrovica Roma, well established borrowings from Serbo/Croatian. On RRM chat-room, they are also the only Romani words that people abbreviate. Since the abbreviation of *zdravo* to *zdr* and *pozdravo* to *poz* is common on Serbian-only CMC (Hentschel 1998), it looks like RRM users have not only borrowed the words themselves, but also the norms for writing them on CMC. I regard this as another instance of the Mitrovica Roma ability, illustrated in chapter 3, to acquire and skilfully use forms and norms from other cultures.

Examples 4.2 and 4.3 also clearly show how heavily repeated typing features on RRM. While in 4.2f repeated typing is used for singing, in 4.2f and 4.3m it serves to call out someone. Formalized shouting also occurs in the above examples (4.2a, 4.2f, 4.2i, 4.2m, 4.3l, 4.3n) showing how this practice does not interfere with simultaneous conversations. Finally, in 4.2n it is possible to see one of the announcements about *abava* I mentioned in 4.2.3.

As entering the chat-room offers a chance to provide a first hint about oneself, similarly leaving can offer a chance to strengthen one’s on-line identity. Example 4.4 shows an

---

2 The analysis of sexist comments and other gendered practices will be beyond the scope of the present work. Suffice to say that they generally replicate off-line community norms (see also 4.2.1).

3 Note that all languages other than Romani used on RRM clearly show second language learner characteristics. Analyzing them will again be beyond the scope of the present work but it is interesting to notice that no user sanctions others for such usages.
interminable goodbye performed by _-EmRaH°StyLe-_—, one of the strong personas portraying himself as a prankster without equals and a self-styled ladies’ man.

As _-EmRaH°StyLe-_ announced he was leaving, but will come back later at night (4.4α), most of the women present at that moment (U3, U13, and MiiLKa_qoKoLLaDa*) wished him goodbye but at the same time tried to make him stay (4.4b, 4.4e, 4.4f, 4.4g).

<table>
<thead>
<tr>
<th>RRM Message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a <em>-EmRaH°StyLe-</em>—: &lt;&lt;&lt;&lt;&lt; CIAOO KO RADIO MANDAR KA SUNDIJA PALO 12H00 (18:34:31)</td>
<td>&lt;&lt;&lt;&lt;&lt;&lt; goodbyes to the radio from me catch you at around 12H00</td>
</tr>
<tr>
<td>b U3: lele emrah (18:34:45)</td>
<td>[mocking sound] emrah</td>
</tr>
<tr>
<td>c <em>-EmRaH°StyLe-</em>—: CIAOO SVIMA DanieleItalianoDilo BYE BYE PI PAJ HEHE (18:34:45)</td>
<td>goodbye everyone DanieleItalianoDilo goodbye drink some water hehe</td>
</tr>
<tr>
<td>d <em>* FARiJA *</em>: Ciao Emrah Ga Muzi e gurunmnjen (18:34:55)</td>
<td>Goodbye Emrah Go Torment the cows</td>
</tr>
<tr>
<td>e MiiLKa_qoKoLLaDa*: NAA EMRAHJ KAJ DJAAAAAAA (18:34:55)</td>
<td>noo emrah where are you gooooooooooing</td>
</tr>
<tr>
<td>f U3: emrah (18:34:56)</td>
<td>emrah</td>
</tr>
<tr>
<td>g U13: CIAO EMRAH (18:35:04)</td>
<td>goodbye emrah</td>
</tr>
<tr>
<td>h <em>-EmRaH°StyLe-</em>—: FARiJA AVAV TE MUZIV TU HEHE (18:35:06)</td>
<td>farija I come to torment you hehe</td>
</tr>
<tr>
<td>i <em>* FARiJA *</em>: HOOOOOOOOOOOOOOOOOOOOOOO 18:35:12</td>
<td>hoooooooooo</td>
</tr>
<tr>
<td>j <em>* FARiJA *</em>: DIK CE KO SUNO (18:35:14)</td>
<td>look joh in your dreams</td>
</tr>
<tr>
<td>k U3: ahhahahahahaahahahahhah (18:35:14)</td>
<td>ahhahahahahaha</td>
</tr>
<tr>
<td>l U14: bes emrah kaj ca tu (18:35:15)</td>
<td>stay emrah where do you go</td>
</tr>
<tr>
<td>m <em>-EmRaH°StyLe-</em>—: CIAOO KO RADIOO KA SUNDIJA POSLEM DanieleItalianoDilo BISOU POUR LES FILLES POUR LES MEC BOX (18:41:24)</td>
<td>goodbye to the radio catch you later DanieleItalianoDilo a kiss to the girls to the guys a punch</td>
</tr>
<tr>
<td>n U7: HAJT AKANA MO EMRAH CA LELE SODE DROMA KA PENE CAO (18:41:42)</td>
<td>c'mon now bro emrah go [mocking sound] how many time will you say goodbye</td>
</tr>
</tbody>
</table>

Example 4: Goodbyes

He ignored all of them to wish another general goodbye and, with a switch into English he teased me about the fact I had just said I was having a beer (4.4c). Then _* FARiJA *_, with whom he is good friend, intervened and true to her persona (see Example 4.8) pushed him away, suggesting he could take his pranks somewhere else (4.4d). He instead took her message as a hook for staying and threatened to make her the victim of his jokes (4.4h). _* FARiJA *_ answer came shortly after and reminded him of who he was dealing with (4.4j). The hook _-EmRaH°StyLe-_— was waiting for came from U14, another woman, who asked him where he was going (4.4l). He stated he was going to have some kebab and
for the following five minutes he kept inviting people to join him just to drop them with a joke. Then (4.4m) he wished another general goodbye and a menacing ‘see you later’ to his latest victim, me. Finally, to show how good he is with women, he switched into French to send kisses to the girls and punches to us boys. However, he was still not done and as a girl asked him to stay a bit more he hanged around, silently, until U7, which often acts as a peace keeper, jokingly invited him to actually leave (4.4n). As in other cases (see 4.3.5) this act of policing was carried out in Romani.

The few switches in _-
EmRaH°StyLe_- goodbye are again strategic as they either try to create or reinforce a bond through some shared knowledge or experience (his using English to wish me goodbye, 4.4c) or serve to depict the user’s persona (his using French, the stereotypical language of love to appear as a lady’s man, 4.4m).

4.3.2 Community maintenance and strengthening

As I mentioned in 4.2.3, maintenance of the community takes place through small talks, in which people also share bits and pieces of their lives, as it can be seen in Example 4.5.

<table>
<thead>
<tr>
<th>RRM Message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a</strong> Sadi Styler: oooo gott diesen senger hat hals schmerzen (18:37:21)</td>
<td>oooo god this singer has a sore throat</td>
</tr>
<tr>
<td><strong>c</strong> Sadi Styler: hahaha (18:37:24)</td>
<td>hahaha</td>
</tr>
<tr>
<td><strong>d</strong> MiIlKa_qoKoLLaDa*: ABER DER ANDRE VOLL DIE STIMME (18:37:25)</td>
<td>but the other one has a powerful voice</td>
</tr>
<tr>
<td><strong>e</strong> MiIlKa_qoKoLLaDa*: MAN (18:37:26)</td>
<td></td>
</tr>
<tr>
<td><strong>f</strong> <em>ROMANI MAHALA</em>: ASUN CERKUSI SAR CILABOL O ADMIR HAHAAHA (18:37:41)</td>
<td>you bitter people listen at how Mirsad sings hahaha</td>
</tr>
<tr>
<td><strong>g</strong> MiIlKa_qoKoLLaDa*: JASHAAAAA (18:37:42)</td>
<td>jashaaaaa</td>
</tr>
<tr>
<td><strong>h</strong> MiIlKa_qoKoLLaDa*: JOI MANCEN TE ASUNEN MO EX SAR GILABOL (18:38:13)</td>
<td>hey do you want to listen at how my ex sings oooo god</td>
</tr>
<tr>
<td><strong>i</strong> Sadi Styler: oooo gotttt (18:38:13)</td>
<td></td>
</tr>
<tr>
<td><strong>j</strong> MiIlKa_qoKoLLaDa*: PISIN PO YOUTUBE [ex-boyfriend’s name] (18:38:43)</td>
<td>write [ex-boyfriend’s name] on youtube</td>
</tr>
<tr>
<td><strong>k</strong> <em>ROMANI MAHALA</em>: MILKA ASUN SO CEROL KATE O DEMAJLI ALLAHILE (18:39:12)</td>
<td>milka listen to what demajli does here oh my god</td>
</tr>
<tr>
<td><strong>l</strong> MiIlKa_qoKoLLaDa*: SI 2 JEK LOLI MAICA JEK KALI MAICA MUNRO EX SI KOJA MIT KAI MAICA (18:39:18)</td>
<td>there are 2 one red shirt one black shirt my ex is the one with the black shirt</td>
</tr>
</tbody>
</table>

*Example 4.5: Sharing life*
As the DJ played the song named by U1 (4.5b), Sadi Styler and MiiLKa_qoKoLLaDa*, respectively a young man and woman at the time engaged in a conversation predominantly in German (see Examples 4.9 to 4.11) commented on the quality of the two singers involved (4.5a, 4.5c to 4.5e)

The DJ, noticing their comment, played another song performed by one of them and, in Romani, challenged Sadi Styler and MiiLKa_qoKoLLaDa* statements about the singer’s poor vocal qualities (4.5f). The choice of Romani, although the other two were using German, serves two functions. First, it reframes a discussion relating to the community in the community language. Second, it marks the DJ’s challenge as a non-threatening one, even if the address ‘you bitter people’ might have been interpreted that way. The light tone of the challenge is further reinforced by the written-out laugh closing it. A shout by MiiLKa_qoKoLLaDa* (4.5g) welcomed the song and showed she accepted the DJ’s point. Indeed, as the singer is actually good and the particular song very popular, everyone on the chat-room joined in the shouting (all 16 messages following 4.5g, they have been cut for pagination purposes). Furthermore, MiiLKa_qoKoLLaDa*, temporarily abandoning the conversation with Sadi Styler, switched to Romani (4.5h) and offered the whole audience to listen to her ex-boyfriend. Sadi Styler, probably disappointed for being ignored, dropped in an aside comment (4.5i) marked as such by his sticking to German even if the language of the conversation had been switched to Romani. Amidst the continuing shouts (all cut messages) MiiLKa_qoKoLLaDa* gave instructions on how to find and recognize her ex-boyfriend on a YouTube video (4.5j and 4.5l). The DJ however, was more interested in getting MiiLKa_qoKoLLaDa* attention on the next song he played (4.5k). He continued to use Romani, but to give emphasis to his message closed it with a switch into Qur’an Arabic.

Since in the following minutes no one commented on MiiLKa_qoKoLLaDa* ex-boyfriend, it is clear that no one checked out the video. What is interesting is how consistent she was, the moment she decided to address the whole audience to share her life, in using Romani. Excluding her usage of *ex, which I regard as an established borrowing since Romani lacks a term to express this relationship, the only instance of another language influence is her usage of the German preposition *mit ‘with’ in 4.5l. To express the same concept, MiiLKa_qoKoLLaDa* could have inflected the word *maika ‘shirt’ in the instrumental case. Because of this I do not consider this switch a conscious choice for communicative purposes. While *mit is attested in other dialects of Romani historically spoken in Germany (Matras 2002) this is the only case attested in the corpus, therefore I am not regarding it as a contact induced language change. Rather, as German was, at that
given moment, the pragmatically dominant language for MiLKa_qoKoLLaDa*, I will suggest this is a local instance of what Matras termed fusion, “the mental nonseparation of linguistic systems while executing a linguistic operation” (2000: 520).

<table>
<thead>
<tr>
<th>RRM Messages</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a U6: VOUS AVEZ A MANGE OU LA FRANCE DOIT VOUS AIDEZ DANS LE FINANCE (12:30:18)</td>
<td>do you[PL] have an income or France must help you with money</td>
</tr>
<tr>
<td>b U6: ??????????????? (12:30:58)</td>
<td>??????????????</td>
</tr>
<tr>
<td>c <em>ROMANI MAHALA</em>: oui se la franc (12:31:09)</td>
<td>yes that’s France</td>
</tr>
<tr>
<td>d U6: DJ SAJ DEMA CO MSN (12:36:26)</td>
<td>dj can you give me your MSN</td>
</tr>
<tr>
<td>e U6: BATE (12:36:32)</td>
<td>bro</td>
</tr>
<tr>
<td>f U6: TE PUCAVTU KANCIK (12:36:40)</td>
<td>to ask you something</td>
</tr>
<tr>
<td>g <em>ROMANI MAHALA</em>: ko san tu (12:37:01)</td>
<td>who are you</td>
</tr>
<tr>
<td>h <em>ROMANI MAHALA</em>: U6 vacar so mange te puceman U6 (12:41:20)</td>
<td>U6 talk what do you want to ask me U6</td>
</tr>
<tr>
<td>i U6: TU KATAR SAN (12:41:34)</td>
<td>where are you from</td>
</tr>
<tr>
<td>j <em>ROMANI MAHALA</em>: fr (12:41:47)</td>
<td>fr</td>
</tr>
<tr>
<td>k U6: KAJ AN FR (12:41:50)</td>
<td>where in fr</td>
</tr>
<tr>
<td>l <em>ROMANI MAHALA</em>: troyes (12:41:53)</td>
<td>troyes</td>
</tr>
<tr>
<td>m <em>ROMANI MAHALA</em>: a tu (12:41:58)</td>
<td>and you</td>
</tr>
<tr>
<td>n U6: ME ANO NANCY (12:42:05)</td>
<td>me in nancy</td>
</tr>
<tr>
<td>o <em>ROMANI MAHALA</em>: ok (12:42:14)</td>
<td>ok</td>
</tr>
<tr>
<td>p U6: MENSURI ME KAKO MPINCARE (12:42:30)</td>
<td>do you know my uncle Mensuri</td>
</tr>
<tr>
<td>q <em>ROMANI MAHALA</em>: luanoso kako va (12:42:39)</td>
<td>yes Luano’s uncle</td>
</tr>
<tr>
<td>r U6: SO PEROL CE O LUOANO TUCE DJ (12:43:07)</td>
<td>what is Luano to you dj</td>
</tr>
<tr>
<td>s <em>ROMANI MAHALA</em>: amal (12:43:09)</td>
<td>a friend</td>
</tr>
<tr>
<td>t U6: SODE BESH SITU DJ (12:43:29)</td>
<td>how old are you dj</td>
</tr>
<tr>
<td>u <em>ROMANI MAHALA</em>: U6 so mange (12:43:33)</td>
<td>U6 what do you want</td>
</tr>
<tr>
<td>v U6: PA VACAR MUABETI (12:43:49)</td>
<td>just have a chat</td>
</tr>
</tbody>
</table>

Example 4.6: Exchanging IM addresses and looking for acquaintances
Example 4.5 showed how, when talking about their life in relation to the whole community, RRM users favour Romani. Similarly, they prefer Romani to mitigate messages that can be interpreted as threatening. Finally, they might insert elements from other languages to give emphasis to their messages.

As I mentioned in 4.2.3, favouring Romani poses a problem if users intend to discuss private and intimate subjects. On these occasions, users supplement RRM chat-room with IM. Example 4.6 will illustrate the sort of exchanges occurring when people try to get someone else’s IM contacts.

This exchange took place as the radio started broadcasting and the only users connected were the DJ, U6, a relatively new user, and me. As usual we had exchanged greetings and then, as it tend to happen when there are few users logged in, fell silent. After a while U6, rather abruptly and only based on the widespread knowledge that the DJs live in France, asked about our economic situation and if we received benefits from the French welfare (4.6a). As it is often the case even in off-line communication, he favoured French as the language of interaction with the authorities and anything related to them. Both the DJ and I ignored him. I felt the question was asked rudely and furthermore, it did not relate to me. Most likely for the DJ it was only the rudeness of the question that stopped him from replying. U6, however, restated his question with a stretch of question marks (4.6b), to which the DJ, sticking to French, replied in a rather evasive manner and without giving out any personal information (4.6c), in fact ending this part of the conversation.

We fell silent again while the chat-room was slowly getting populated. During this time U6 must have realized his approach was not appropriate. Therefore, I suppose in order to save face, U6 switched into Romani and asked the DJ for his IM address, explaining he wanted to ask him something. The DJ, apparently willing to give him a second chance but not knowing U6, asked who he was (4.6d to 4.6g). However, as a good sized crowd had logged in, the usual shouting and small talks had started and U6 had joined in, thus probably not noticing the DJ’s question. The DJ then prompted him to go ahead (4.6h) and after a while U6 started the accepted routine for finding out who the person behind a user-name is (4.6j to 4.6t, see also 4.2.3). In this example, however, probably because of the impoliteness of U6 first question, the DJ remained evasive and finally asked quite bluntly what exactly U6 wanted (4.6u). Realizing the conversation was going nowhere, U6 saved face by dismissing his attempt at starting a serious discussion as just a casual conversation (4.6v).

It might seem from Examples 4.5 and 4.6 that Romani and other languages stand in a polar opposition. Romani is the language of community identity and affection, used to talk
about life in the community, to mitigate perceived threats or to save face. Other languages instead seem to act as languages of utility and relate to the discussion of life in the wider society. However, this distinction is not so straightforward. It is the lack, in the above examples, of any humour and the presence of very few strong personas that lands this impression.

4.3.3 Humour and strong personas

To better show how humour and the attempt at sustaining the on-line identity of a strong persona affect language choice I have selected two examples that closely resemble the above ones. Example 4.7, focusing on humour, started as a discussion about work opportunities, while example 4.8, introducing one of the chat-room strong personas, is an attempt at getting another user IM address.

### RRM messages and translations

<table>
<thead>
<tr>
<th>RRM messages</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a ANNA-MISS-FR: rien sar coripe kate naj so cerav (18:49:06)</td>
<td>nothing is like poverty here there's nothing for me to do</td>
</tr>
<tr>
<td>b :::::Mohamed-Roanne::::: JA DE LA FRANCE DE MERDE WALLAH (18:49:23)</td>
<td>yes France is shit for god’s sake</td>
</tr>
<tr>
<td>c ANNA-MISS-FR: bravo edy ka cav tu ani buci samo za kilometra te racuni (18:49:35)</td>
<td>bravo edy I’ll drop you to work only to count the kilometres</td>
</tr>
<tr>
<td>d ANNA-MISS-FR: kikikikikikikikiki (18:49:40)</td>
<td>hahah</td>
</tr>
<tr>
<td>e :::::Mohamed-Roanne::::: HAHAH (18:49:47)</td>
<td>ahahahah</td>
</tr>
<tr>
<td>f edy.boy: AHAHAHAHA (18:49:50)</td>
<td>yes mohamed</td>
</tr>
<tr>
<td>g ANNA-MISS-FR: oui mohamed (18:49:52)</td>
<td></td>
</tr>
<tr>
<td>h :::::Mohamed-Roanne::::: SE VOV CERDA MATEMATIKA BUT BESA ANNA O EDY (18:50:08)</td>
<td></td>
</tr>
<tr>
<td>i :::::Mohamed-Roanne::::: HAHAHAHA (18:50:09)</td>
<td>hahahaha</td>
</tr>
<tr>
<td>j edy.boy: MA BRE C PAS DELA MERDE LA FANCE PASKE ELLE VU DON A MANGE CAFF ASEDIK (18:50:21)</td>
<td>don't mates France isn’t sh*t because it feeds you rancid coffee</td>
</tr>
<tr>
<td>k ANNA-MISS-FR: hahahahahahaah (18:50:29)</td>
<td>hahahahahah</td>
</tr>
<tr>
<td>l :::::Mohamed-Roanne::::: HAHAHAH (18:50:31)</td>
<td>hahahah</td>
</tr>
<tr>
<td>m :::::Mohamed-Roanne::::: LE RESTO (18:50:36)</td>
<td>the other half</td>
</tr>
<tr>
<td>n :::::Mohamed-Roanne::::: DU COEUR (18:50:39)</td>
<td>of the heart</td>
</tr>
<tr>
<td>o edy.boy: C VRE NON (18:50:39)</td>
<td>it's true no</td>
</tr>
<tr>
<td>p edy.boy: AHAHA (18:50:41)</td>
<td>ahaha</td>
</tr>
<tr>
<td>q ANNA-MISS-FR: hahahahahahaah gja si edy (18:50:46)</td>
<td>hahahahahah ah so it is edy</td>
</tr>
<tr>
<td>r edy.boy: EHEEH (18:50:53)</td>
<td>eheeh</td>
</tr>
<tr>
<td>s ANNA-MISS-FR: vec ceramince muhabeti (18:50:55)</td>
<td>we’re only having a cocky chat</td>
</tr>
<tr>
<td>t edy.boy: LE FRANCCCE MANGE LE FORMAGE KAI TE AVOL LEN ZURALE KOKALA (18:54:36)</td>
<td>the French eat the cheese so that they’ll have strong bones</td>
</tr>
</tbody>
</table>

Example 4.7: Humour
The above conversation occurred between three users: two men, ::::Mohamed-Roanne:::.. and edy.boy, and a woman, ANNA-MISS-FR. By their user-names, it was clear that Mohamed lived in France, more precisely in Roanne, as did Anna, although she only referenced the country (FR) and not the exact town. edy.boy did not give any hint about where he is living, but he seemed to know both others very well, possibly also off-line, suggesting that he is, or at least was, living in France as well.

I was logging in when the first message by ANNA-MISS-FR was displayed. I was not aware of how the conversation started, but the fact that the content seemed surprisingly serious immediately caught my attention and stopped me from actually logging in.

ANNA-MISS-FR (4.7a), using mostly Romani, was complaining about poverty and the lack of opportunities in France. ::::Mohamed-Roanne:::.. (4.7b), switched to French, thus suggesting the discussion was about life in relation to the wider community. However, he addressed a strong insult to France, showing a rather negative feeling towards the country. The deepness of the feeling was further reinforced by the final emphatic switch into Qur’an Arabic.

ANNA-MISS-FR did not answered immediately to ::::Mohamed-Roanne:::.. as she was addressing a work-related joke to edy.boy. The joke was in Romani and to mark it as non-offensive it was shortly followed by a laugh (4.7c, 4.7d). The other two welcomed ANNA-MISS-FR joke with a burst of laughs (4.7e, 4.7f). Only at this point ANNA-MISS-FR noticed ::::Mohamed-Roanne:::.. message and openly expressed her agreement (4.7g), switching from Romani to French, thus accepting ::::Mohamed-Roanne:::.. choice of language and displaying her sharing of a negative feeling about France.

The conversation was interrupted for a moment by greetings and shouts addressed to a regular coming back to the chat-room, with all three of them cheering him. After that pause, ::::Mohamed-Roanne:::.. resumed ANNA-MISS-FR joke about edy.boy adding his own personal punch to it (4.7h, 4.7i). He switched to Romani marking a different tone from his last message and, as ANNA-MISS-FR, added a laugh to make clear the joke was not meant as offensive.

edy.boy, however, was more interested in taking the side of France. He opened his message (4.7j) with an aggressive attention seeker in Romani, then switched to French to defend France, arguing it is giving his friends a chance to live decently (because France feeds you). However, with a final switch back to a Romani expression borrowed from Turkish he added a vein of humour to the message. This final switch achieved two things:
it mitigated the aggressive opening and suggested a partial agreement on the negative feeling about France. The laughs by the other two (4.7k, 4.7l) confirmed the effectiveness of edy.boy move. Even more, ::::..Mohamed-Roanne:::: felt obliged to apologise (4.7m, 4.7n) and he used French to do so. However, his messages can have a double, humorous reading, created by the unconventional spelling in *le resto* (4.7m). As it is common in French CMC and SMS (Anis 2007), ::::..Mohamed-Roanne:::: omitted the final Ss from both words, creating confusion between *les restos* ‘the restaurants’ and *le reste* ‘the rest, what is left’. *Les Restos du Cœur* ‘the Restaurants of the Heart’ is a French charity active in France big cities and distributing meals to the needy. Thus ::::..Mohamed-Roanne::::, although accepting edy.boy’s point, jokingly kept his distance from France. In the following message still in French, edy.boy accepted both ::::..Mohamed-Roanne::::’s apologies (4.7o) and as his laugh suggest (4.7p) the mocking of France. Joining in the communal laugh, ANNA-MISS-FR (4.7q, 4.7s) used Romani to relieve the tension by bracketing their discussion as ‘cocky’.

Unfortunately at this point my Internet connection faltered for a few minutes, but when it resumed I realized that it was now edy.boy (4.7t) who was joking about the classical stereotype of French people as cheese eaters. Interestingly, he started the message in French but switched into Romani when introducing the complement clause expressing the actual punch. At the same time ANNA-MISS-FR and ::::..Mohamed-Roanne:::: had separately engaged other two different users, thus edy.boy message went unnoticed and the conversation between the three ended.

Through the whole interaction a sense of shared understanding and complicity was created not by the exclusive usage of Romani, supposedly the language of identity and affection, but rather through the switching between the various languages in the speakers repertoire. As suggested by Gumperz (1977) the juxtaposition of languages conveys meaning, in this specific case it reinforces the humorous content of messages by defying the expectation that French, the language of formal interaction, should be used when talking about France. In 4.7j and 4.7t in particular, the switching to Romani, the language of informal interaction, signals a change of tone which, coupled with the meaning of the lexical content of the switched segments themselves, results in the creation of a playful atmosphere.

Once this atmosphere is created, even face-saving, which looking at Example 4.6 seemed related to the usage of Romani, is carried out in the contact language, although a vein of humour is maintained by playing with words (4.7m, 4.7n). Most importantly, the switching between languages and the resulting humour allowed the three to express rather
ambivalent feelings towards the host country. That would have not been possible if they had stuck to a single language, since as I mentioned when analyzing Example 4.6, serious discussions about the host countries are avoided.

Another instance in which the whole linguistic repertoire of users is mobilized is the portrayal of their on-line personas. Much like in greetings, the sustained insertion of single words from languages other than Romani is used to convey a certain sense of personal identity by suggesting an association between the user and the cultural values attributed to the languages used.

Example 4.8: Representing on-line personas

<table>
<thead>
<tr>
<th>RRM messages</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>U1: zdr farija (18:31:47)</td>
</tr>
<tr>
<td>b</td>
<td>U1: sar san tu (18:31:48)</td>
</tr>
<tr>
<td>c</td>
<td>_ * FARIJA *_: ZDR U1 (18:31:56)</td>
</tr>
<tr>
<td>d</td>
<td>_ * FARIJA *_: SUPER SIJUM (18:31:58)</td>
</tr>
<tr>
<td>e</td>
<td>_ * FARIJA *_: thx(18:31:59)</td>
</tr>
<tr>
<td>f</td>
<td>_ * FARIJA *_: a tu (18:32:00)</td>
</tr>
<tr>
<td>g</td>
<td>U1: super (18:32:06)</td>
</tr>
<tr>
<td>h</td>
<td>U1: katar san tu farija (18:32:13)</td>
</tr>
<tr>
<td>i</td>
<td>_ * FARIJA *_: U1 Me Italia sem (18:33:40)</td>
</tr>
<tr>
<td>j</td>
<td>_ * FARIJA *_: a tu (18:33:42)</td>
</tr>
<tr>
<td>m</td>
<td>U1: mesem tari fr farija (18:34:04)</td>
</tr>
<tr>
<td>n</td>
<td>U1: sode bos si tut farija (18:34:14)</td>
</tr>
<tr>
<td>o</td>
<td>_ * FARIJA *_: U1 BUT PURI SEM TUCE (18:35:38)</td>
</tr>
<tr>
<td>p</td>
<td>_ * FARIJA *_: PARDOSNIK (18:35:45)</td>
</tr>
<tr>
<td>q</td>
<td>U1: farija si tut msn ce korije (18:41:58)</td>
</tr>
<tr>
<td>r</td>
<td>_ * FARIJA *_: VA U1 SIMAN SE ACIVE N I AČILO KONI BI Msn GN E VI M MAMI SILA MSN (18:42:22)</td>
</tr>
<tr>
<td>s</td>
<td>U1: ka dema co msn farija ako madje (18:42:48)</td>
</tr>
<tr>
<td>t</td>
<td>_ * FARIJA *_: MO U1 SAFAR MSN MANCE BREE (18:43:10)</td>
</tr>
<tr>
<td>u</td>
<td>~ danijela ~: pozzi svima (18:43:13)</td>
</tr>
<tr>
<td>v</td>
<td>_ * FARIJA *_: NAJ MAN TUCE MO MSN (18:43:16)</td>
</tr>
<tr>
<td>w</td>
<td>_ * FARIJA *_: SORRY (18:43:17)</td>
</tr>
</tbody>
</table>


A few users are particularly skilled at that, such for example \_* FARiJA *\_, which I already introduced in Examples 4.3 and 4.4. \_* FARiJA *\_ persona is that of a strong, modern woman, but still tied to Romani values and culture, as clearly shown by her preference to write in Romani. She always shows herself as self-confident and independent. Because of that, she tends to receive much attention from male users. At the same time she is quite selective about whom she is interacting with and often rebukes unwanted approaches with sharp comments. During my fieldwork, it took me quite a long time to go beyond the exchange of greetings with her. Example 4.8, picking up exactly from where example 4.3 ended, shows just one of her rebukes, this time to U1 who was trying to get her IM address.

For the previous hour or so, U1 had been trying, with some success, to get women’s IM addresses. However, most of the women he had interacted with had left the chat-room and most likely also their IM. He then turned his attention to \_* FARiJA *\_ and asked how she was, a tried and tested icebreaker on RRM (4.8a, 4.8b). \_* FARiJA *\_ reciprocated the greeting in Romani (4.8c) then, using super, an internationalism common in young people speech in various countries, followed by sijum, the Romani copula, stated she was feeling great (4.8d). She also thanked U1 using a shortening for thanks (4.8e). Then, in Romani, she gave him a chance to start a conversation (4.8f), to which U1 replied using the same internationalism employed by \_* FARiJA *\_ (4.8g). This was a smart move by U1 as he managed to retain \_* FARiJA *\_ attention and find out where she is from (4.8h to 4.8m).

However, \_* FARiJA *\_ attitude changed as soon as U1 asked for her age (4.8n). Asking for the age of someone from the opposite sex is generally regarded on RRM as an attempt at starting a flirtatious conversation, provided the two people involved are more or less of the same age. Therefore \_* FARiJA *\_, who was clearly not interested, first let U1 wait for more than a minute then, without saying her age nor asking for his, just stated she was too old for U1 (4.8o). She anyway sweetened the refusal (4.8p) by switching to a slang expression common among Italian teen-agers and young adults. This expression, French pardon with a Slavic-sounding end (-ski), is however meant to be slightly sarcastic and anyway conclusive of any previous discussion.

Apparently U1 accepted that and for the next six minutes he did not try to interact with \_* FARiJA *\_. However, he later tried another approach and asked if she had an IM account (4.8q) closing his request with a very common Romani moniker, generally reserved for unruly girls. \_* FARiJA *\_, after making him wait again, tried to shove him off and at the same time displayed her persona of a modern woman and ridiculed him (4.8r). Yet, U1 insisted in his request, although now he looked less self-confident (4.8s).
this point _* FARiJA *_ reacted aggressively and, using only Romani, violently dismissed U1’s request. However, she then switched to English to close the interaction and mitigate her previous message (4.8t to 4.8w).

Clearly, while in her first messages _* FARiJA *_ was using Romani, English and internationalisms to engage in a friendly conversation with U1, in the ones from 4.8o onwards, Romani became the language of distance and even aggression and English and Italian slang were used to mitigate these feelings and to save face. Even in spite of this change, the portrayal of _* FARiJA *_ persona remained constant, thus it is not a single language that conveys it, but rather the whole of her linguistic repertoire.

Comparing these last two examples with the previous ones (4.5, 4.6) it should be clear how humour and the desire to represent oneself as a strong persona drastically complicate the association of Romani, as language of identity, with the community and that of other languages, as languages of utility, with outsiders. Rather, all the resources in the user linguistic repertoire are mobilized depending on the specific and situated communicative goals each user is pursuing at a given moment.

4.3.4 Affection and intimate feelings

To further reinforce the previous observation, it is interesting to look at how affection and intimate feelings can be expressed publicly on RRM without offending other users or exposing oneself too much. To do this, I will go back to the dominantly German conversation between Sadi Styler and MiiLKa_qoKoLLaDa* that opened Example 4.5. Sadi Styler, although quite often on RRM, is a rather plain user, not trying to present a particularly strong personality. MiiLKa_qoKoLLaDa*, on the other hand, regularly presents herself as a sweet, flirtatious girl and frequently teases male users. During my fieldwork the two developed a rather strong relationship, and the discussion I am going to analyze marked its beginning.

Sadi Styler entered the chat-room and chose the Islamic greeting to introduce himself (4.9a). However, as the chat-room was quite busy and many people at the same time were requesting songs to the DJ, nobody noticed him. He then, switching to German as the language of their previous interaction, addressed MiiLKa_qoKoLLaDa* (4.9b). As part of her strategy to present herself as flirtatious, MiiLKa_qoKoLLaDa* always adapts to the language chosen by her interlocutor. Since she lives in one of the German-speaking regions of France, she is most of the time able to do so. However, as much as _* FARiJA *_ she is not keen to interact with anyone, therefore not recognizing Sadi Styler, she asked who he was (4.9c). Once Sadi Styler explained who he was (4.9d, 4.9e)
and MiiLKa_qoKoLLaDa* realized they already knew each other, she called him with an affectionate moniker she then kept using all the time (4.9g).

<table>
<thead>
<tr>
<th></th>
<th>RRM messages</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Sadi Styler: selami gredenke (18:27:08)</td>
<td>selam everyone</td>
</tr>
<tr>
<td>b</td>
<td>Sadi Styler: milka schokolade kenn ich dich nicht heheheh (18:27:30)</td>
<td>milka schokolade don’t I know you heheheh</td>
</tr>
<tr>
<td>c</td>
<td>Miil.Ka_qoKoLLaDa*: SADII WER BIST DU (18:27:41)</td>
<td>sadii who are you</td>
</tr>
<tr>
<td>d</td>
<td>Sadi Styler: ich bin sadi wo wir letzens geschrieben haben (18:28:05)</td>
<td>I’m Sadi we wrote to each other recently</td>
</tr>
<tr>
<td>e</td>
<td>Sadi Styler: wo wir die ganze zeit deutsch gereddet haben (18:28:13)</td>
<td>when we were talking German all the time</td>
</tr>
<tr>
<td>f</td>
<td>Miil.Ka_qoKoLLaDa*: SADII DU PENNER (18:28:45)</td>
<td>Sadii you bum</td>
</tr>
</tbody>
</table>

Example 4.9: Affection and intimate feelings

From that moment on, the two interacted primarily with each other. After exchanging some news about their lives, they started talking about the music played by the DJ, as I have shown in Example 4.5. In that occasion and in a few others, either of them abandoned their mutual exchanges to interact with other people. Normally it was MiiLKa_qoKoLLaDa* to stray away from their interactions. In such cases, Sadi Styler turned to someone else, occasionally dropping an aside comment (see 4.5i), until MiiLKa_qoKoLLaDa* came back to him. However, if he turned to some other woman, MiiLKa_qoKoLLaDa* was quick in getting his attention back with a teasing or supposedly menacing message.

This went on for about 20 minutes during which they mostly used German. During that time no one else interfered with their conversation unless, as in the case of the comments about the singers in Example 4.5, it was touching upon non-private subjects.

During the whole conversation, MiiLKa_qoKoLLaDa* had been teasing Sadi Styler on different occasions and at a certain point he had started to tease her back. Example 4.10 shows the peak of this mutual teasing.

To understand Sadi Styler’s first message, it is important to notice that both fair complexions and extremely dark ones can occasionally be found among the Mitrovica Roma. The community shares the widespread norm that fair skinned women are more desirable, while dark skinned individuals are often considered naughty and hot blooded. Sadi Styler, although he did not know MiiLKa_qoKoLLaDa* off-line, was mocking her
for her dark skin colour but at the same time complimenting her for her character.

Interestingly, while the sentence is dominantly in German, the key attributes of MiiLKa_qoKoLLaDa*, as they draw on a culture specific set of values, are expressed through Romani adjectives.

MiiLKa_qoKoLLaDa* reaction was dismissive of the mockery (4.10e, in German) yet appropriated it, and thus admitted being of dark complexion (4.10f, in Romani). Coupled with the association of dark skin with a hot blooded personality and the switch from German to Romani, this admission assumes a sort of provocative tone.

<table>
<thead>
<tr>
<th>RRM messages</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>a</strong> Sadi Styler: JA MAN HAB DICCH DIGGA BISSCHEN KALI ABER EGAL KUDLI (18:46:11)</td>
<td>yo man I think you mate are a bit too black but sweet all the same</td>
</tr>
<tr>
<td><strong>b</strong> Sadi Styler: HEHEHE (18:46:12)</td>
<td>hehehe</td>
</tr>
<tr>
<td><strong>c</strong> MiiLKa_qoKoLLaDa*: PENNER (18:46:21)</td>
<td>bum</td>
</tr>
<tr>
<td><strong>d</strong> MiiLKa_qoKoLLaDa*: HAHAHHA (18:46:23)</td>
<td>hahahaha</td>
</tr>
<tr>
<td><strong>e</strong> MiiLKa_qoKoLLaDa*: JA WAS SOLL ICH MACHEN (18:46:30)</td>
<td>well what should I do</td>
</tr>
<tr>
<td><strong>f</strong> MiiLKa_qoKoLLaDa*: KAJ SEM KALII (18:46:33)</td>
<td>since I'm a black woman</td>
</tr>
<tr>
<td><strong>g</strong> Sadi Styler: SI POMATE BRE ANI ITALIA AJISCHA UND SO (18:46:49)</td>
<td>mate they've creams in Italy ajscha and the like</td>
</tr>
<tr>
<td><strong>h</strong> Sadi Styler: SPASS BEBI SEIN NICHT SAUER VALA ICH MACH NUR SPASS (18:47:01)</td>
<td>it's a joke baby not an insult for god's sake I'm only kidding</td>
</tr>
<tr>
<td><strong>i</strong> MiiLKa_qoKoLLaDa*: DIK DIK (18:47:03)</td>
<td>look look</td>
</tr>
<tr>
<td><strong>j</strong> MiiLKa_qoKoLLaDa*: DU PENNER (18:47:12)</td>
<td>you bum</td>
</tr>
<tr>
<td><strong>k</strong> MiiLKa_qoKoLLaDa*: AVAV CE KOTE (18:47:15)</td>
<td>I come there for you</td>
</tr>
<tr>
<td><strong>l</strong> Sadi Styler: JASHAAAAAA (18:47:16)</td>
<td>jashaaaaa</td>
</tr>
<tr>
<td><strong>m</strong> Sadi Styler: MIILKA (18:47:18)</td>
<td>milka</td>
</tr>
<tr>
<td><strong>n</strong> MiiLKa_qoKoLLaDa*: IKLAV UPRAL TUTE (18:47:20)</td>
<td>I crawl over you</td>
</tr>
<tr>
<td><strong>o</strong> MiiLKa_qoKoLLaDa*: MARAV TU (18:47:22)</td>
<td>I kill you</td>
</tr>
<tr>
<td><strong>p</strong> Sadi Styler: KNUSCTHAAAAAAA (18:47:24)</td>
<td>kitryyyyyyy</td>
</tr>
<tr>
<td><strong>q</strong> MiiLKa_qoKoLLaDa*: ICH LIEBE DICCH AUCHHHHHHHH (18:47:36)</td>
<td>I love you toooooooooo</td>
</tr>
<tr>
<td><strong>r</strong> Sadi Styler: HOOOOOOOOOOOOOP (18:47:37)</td>
<td>hooooooooooooop</td>
</tr>
<tr>
<td><strong>s</strong> Sadi Styler: WEN HAST DU DAS GESAGT (18:47:44)</td>
<td>who are you talking to</td>
</tr>
<tr>
<td><strong>t</strong> MiiLKa_qoKoLLaDa*: DIR (18:47:50)</td>
<td>you</td>
</tr>
</tbody>
</table>

The tease, however, was replicated by a joke, carried out in Romani, about the availability of skin bleaching creams (4.10g). At the end of this message Sadi Styler used an utterance modifier (and the like) which, as is often the case when languages are in contact (Matras 2000), is a borrowing, in this case from German. Either because of this or
in an attempt to clearly mark it as different from the previous, the following message (4.10h) mitigating the potential offensive content of the joke was in German. As in other examples (4.5k, 4.7b) emphasis was given to the whole message by the insertion of a Qur’an Arabic word.

MiiLKa_qoKoLLaDa* reply (4.10i to 4.10k) was in Romani, but for the use of the German moniker she addressed to Sadi Styler when she recognized him (see 4.9g), and kept playing with the association of dark skin with hot temper as these sort of phrases are generally uttered when menacing someone at the start of a fight. In an apparently paradoxical reaction to a menace, Sadi Styler replied with a shout (4.10l) clearly addressed at MiiLKa_qoKoLLaDa* (4.10m), thus revealing he had understood the sexual allusion. Sticking to Romani, MiiLKa_qoKoLLaDa* added two further messages (4.10n, 4.10o) which kept playing with the allusive connection between violence and sex. An allusion Sadi Styler picked up, although switching to German, the moment he called MiiLKa_qoKoLLaDa* ‘kitty’ (4.10p). Then the two, maintaining usage of German, played the part of the romantic lovers (4.10q to 4.10t).

For the following minute, and just using German, they kept exchanging messages very similar to these last ones. MiiLKa_qoKoLLaDa* even started to play the part of the possessive, jealous woman giving herself to one and one only man. During these exchanges, still no one else interfered. This suggests that the dominant and sustained usage of a language other than Romani between limited numbers of users marks their interactions as private. The shared understanding that such conversations are private gives users a chance to express those feelings and emotion that would otherwise be expressed on an IM chat. Such feelings and emotion can even include, like it was the case between Sadi Styler and MiiLKa_qoKoLLaDa*, sexually allusive jokes.

The playful nature of the whole exchange and the role of German in keeping it private were revealed a little later. As a song by Ćita, the most popular singer from the community, was playing, Sadi Styler was shouting madly. At the same time, MiiLKa_qoKoLLaDa* came out with an irreverent sentence in Romani about Emran, another regular user who was not on at that time (4.11a).

Most likely, because of the shouting and the comments about the song posted in Romani by nearly everyone on the chat-room at that moment, MiiLKa_qoKoLLaDa* was carried away and posted her message in Romani. Lacking any recipient name, it appeared directed to no one in particular. The message actually targeted Sadi Styler only, but it should have been in German to flag that. This interpretation is confirmed by the fact that in
her next message (4.11d) MiiLK_qoKoLLaDa* switches back to German, starting to explain why it was good that Emran was not on the chat.

However, her first message had been read as a public one and immediately followed by an enraged shout by U7 (4.11b) and a surprised one by Sadi Styler (4.11c). U7 reaction became clear a few seconds later (4.11e, 4.11h) when he took the defence of his friend. U7 used Romani, with a final emphatic switch into Qur'an Arabic, so that it was clear he intended everyone to know his strong stance. Sadi Styler must have noticed that as he broke on a series of embarrassed laughs (4.11f, 4.11i). MiiLK_qoKoLLaDa*, on the other hand, did not notice U7 intervention and, sticking to German, started confessing her secret to Sadi Styler (4.11g).

<table>
<thead>
<tr>
<th>RRM messages</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a MiILKa_qoKoLLaDa*: KUKUUU SO LACE KAJ NAJ O EMRAN KATE (18:48:58)</td>
<td>peekabooooo how good [is it] that emran is not here</td>
</tr>
<tr>
<td>b U7: ELAAAAAAAYAAAAA (18:49:01)</td>
<td>elaaaaaaaaaaaaa</td>
</tr>
<tr>
<td>c Sadi Styler: HIOOOOOOOOOOOOO (18:49:01)</td>
<td>hioooooooiooooooo</td>
</tr>
<tr>
<td>d MiILKa_qoKoLLaDa*: DER WÜRDE MICH UMBERINGEN (18:49:02)</td>
<td>he would kill me</td>
</tr>
<tr>
<td>e U7: MA SADIN E EMRAN BRE (18:49:07)</td>
<td>don't laugh at emran mates</td>
</tr>
<tr>
<td>f Sadi Styler: AHA (18:49:09)</td>
<td>aha</td>
</tr>
<tr>
<td>g MiILKa_qoKoLLaDa*: TUT MIR LEID MUSS DIR WAS GEDEHEN (18:49:11)</td>
<td>I'm sorry I must confess something</td>
</tr>
<tr>
<td>h U7: MO AMAL SI VALLAH (18:49:11)</td>
<td>he's my friend for god's sake</td>
</tr>
<tr>
<td>i Sadi Styler: OKE AHA (18:49:12)</td>
<td>ok aha</td>
</tr>
<tr>
<td>j <em>ROMANI MAHALA</em>: *** All chat deleted *** (18:49:14)</td>
<td>*** All chat deleted ***</td>
</tr>
<tr>
<td>k MiILKa_qoKoLLaDa*: DU BIST NICHT DER EINZIGE (18:49:17)</td>
<td>you are not the only one</td>
</tr>
<tr>
<td>l Sadi Styler: OKE AHA (18:49:24)</td>
<td>ok aha</td>
</tr>
<tr>
<td>m MiILKa_qoKoLLaDa*: EMRAN IS AUCH MEIN SCHATZ (18:49:28)</td>
<td>emran is my sweetheart too</td>
</tr>
<tr>
<td>n Sadi Styler: AHA OKE (18:49:30)</td>
<td>aha ok</td>
</tr>
<tr>
<td>o U7: HEY MA SADIN E EMRAH VOV SI MO AMAL VALLAH (18:49:31)</td>
<td>hey don't laugh at emrah he's my friend for god's sake</td>
</tr>
<tr>
<td>p MiILKa_qoKoLLaDa*: U7 O EMRAN SI MO SCHATZ (18:49:42)</td>
<td>U7 emran is my sweetheart</td>
</tr>
<tr>
<td>q Sadi Styler: O O (18:49:44)</td>
<td>O O</td>
</tr>
<tr>
<td>r U7: GENAU HAHA (18:49:45)</td>
<td>that's it haha</td>
</tr>
<tr>
<td>s Sadi Styler: ERWICHTE (18:49:46)</td>
<td>caught</td>
</tr>
<tr>
<td>t MiILKa_qoKoLLaDa*: HAHAHA (18:50:07)</td>
<td>hahaha</td>
</tr>
<tr>
<td>u Sadi Styler: SCHREIB DOCH MAL (18:50:11)</td>
<td>drop me a line</td>
</tr>
<tr>
<td>v Sadi Styler: HAHAHA (18:50:15)</td>
<td>hahaha</td>
</tr>
<tr>
<td>w Sadi Styler: ALLE AM PENNEN (18:50:18)</td>
<td>[when] all are sleeping</td>
</tr>
</tbody>
</table>

Example 4.11: Affection and private feelings 3

At this point the DJ, noticing mounting tension, exercised his powers as moderator (see 4.2.1) and deleted all chat messages (4.11j).
As it was possible to post again, MiiLK_a_qoKoLLaDa* resumed her explanation to Sadi Styler and in German explained she had another man, namely Emran (4.11k, 4.11m). Sadi Styler, aware of being watched, replied with more embarrassed laughs (4.11l, 4.11n). As Emran was mentioned again, U7 jumped in, literally repeating his previous messages in defence of Emran, including the emphatic switch to Qur’an Arabic (4.11o). Finally, MiiLK_a_qoKoLLaDa* realized what was going on and, by way of apologizing, translated her previous message into Romani, basically informing everyone that Emran is, at least online, her boyfriend (4.11p). Sadi Styler probably showed shock at the whole situation by using what looks like an emoticon (4.11q), but given the limited usage of emoticons (see 4.2.2) I can not be sure about this interpretation.

However, MiiLK_a_qoKoLLaDa* translation was enough for U7 who accepted the excuse, but used German rather than the expected Romani (4.11r). By doing so, he showed Sadi Styler and MiiLK_a_qoKoLLaDa* some sympathy and complicity, as he made clear he was able to understand the long conversation that had been going on between the other two and yet did not interfere. MiiLK_a_qoKoLLaDa* laugh (4.11t) and the suggestion by Sadi Styler (4.11s to 4.11w) about continuing the conversation when other users are not connected (sleeping) cemented this feeling of complicity, again flagged through the sustained usage of German.

Affectionate and intimate feelings, even if playful and possibly offensive of other user sensibilities, can thus be expressed on the common space provided by RRM chat-room and are not only limited to IM. Their expression, however, is made possible by the implicit norm that sustained usage of a language other than Romani marks a conversation as private. However, in these situations users can involuntary switch back to Romani, like MiiLK_a_qoKoLLaDa* did in 4.11a, and open-up their conversations to the whole audience.

4.3.5 Self-policing and politeness

The intervention of U7 in Example 4.4n and Examples 4.7 and 4.11, can also be regarded as instances of self-policing. However, as they occurred in the midst of other discussions they did not show the formalized politeness that often accompanies them.

Politeness appears to be a highly regarded norm on RRM and can be clearly seen in the request for songs addressed at the DJs, as in Example 4.12.
The first message by U8 is the most frequent way of asking for a song using Romani šaj ‘it’s possible’. Users may add a further ‘please’ or use it instead of šaj. The further or replacing ‘please’ is always in a different language, the most common being German bitte, French stp (s’il te plait), Serbo/Croatian prosim and English pls, and generally occurs at the end of the message. Users also ask for specific songs or singers, replacing the generic jek gili ‘one song’ with a title or kanč andar ‘something by’. Whatever forms the requests take, if either a šaj or ‘please’ in another language are not present, the DJs just ignore them. The eventual negotiation about which song exactly is to be played and when, if the DJ is receiving a lot of the request at the same time, are always carried out in Romani, with the occasional insertion of ‘please’ and ‘thanks’ from other languages.

Politeness and self-policing come together in the negotiations about who is to use a certain font colour for their messages. As I mentioned (4.2.1) colours, along with user-names, represent the virtual face of users. However, they are not assigned permanently and uniquely to one user, so it can happen that a newcomer chooses a colour already taken by another user. In such cases, the first holder of the colour immediately asks for the other to change it (Example 4.13).
Normally newcomers immediately change colour, but sometimes they refuse resulting in longer negotiations like the one in example 4.14, where U6 arrived later, yet initially asked U8 to change colour.

As in the requests for songs, certain users add a non-Romani ‘please’ at the end of their requests (4.13Bc, 4.14k). Furthermore it is interesting to notice the variation in the word for ‘colour’, a borrowing in all Romani dialects. On RRM three words are used: boja/bojava from Serbo/Croatian (4.14h, 4.14m), farba from German (4.13Bc) and color/colori/kolor/coleur from French or Italian (4.13Af). In spite of such variation, users have no problems in understanding what the others mean, even if they do not use the given word.

It is clear from the last examples that self-policing on RRM tends to be associated with the usage of Romani. Yet in this kind of interactions, but also more generally, users switch into other languages when aiming at being polite, even if only inserting single, highly formalized words to mark their intention.

<table>
<thead>
<tr>
<th>RRM Message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a U8: ZDR ROMANI (12:20:42)</td>
<td>hello romani</td>
</tr>
<tr>
<td>b 2 messages cut</td>
<td></td>
</tr>
<tr>
<td>c <em>ROMANI MAHALA</em>: zdr U8 (12:21:03)</td>
<td>hello U8</td>
</tr>
<tr>
<td>d U8: OPSAAAAAAAAAAAAAA (12:22:06)</td>
<td>c'moooooooooooooooooo</td>
</tr>
<tr>
<td>e U6: JASAAAAAAAA (12:22:12)</td>
<td>c'moooooooooo</td>
</tr>
<tr>
<td>f U8: SSAAAAAASSSSSSSSAAA (12:22:24)</td>
<td>c'moooooooooooooo</td>
</tr>
<tr>
<td>g U6: JASAAAAAASSSSSSAA (12:22:34)</td>
<td>c'mooooooooooooooooon</td>
</tr>
<tr>
<td>h U6: LE AVER BOJAVA (12:22:42)</td>
<td>get another colour</td>
</tr>
<tr>
<td>i U8: U6 AVER (12:23:30)</td>
<td>U6 another</td>
</tr>
<tr>
<td>j U8: BOJA (12:23:33)</td>
<td>colour</td>
</tr>
<tr>
<td>k U8: PLIZZ (12:23:35)</td>
<td>plizz</td>
</tr>
<tr>
<td>l <em>ROMANI MAHALA</em>: zdr U6 (12:23:37)</td>
<td>hello U6</td>
</tr>
<tr>
<td>m U6: TU LE AVER BOJAVA (12:23:38)</td>
<td>you get another colour</td>
</tr>
<tr>
<td>n U8: NA TU SE ME (12:23:42)</td>
<td>no you but if I</td>
</tr>
<tr>
<td>o U6: SAR SIJAN MANCE DI (12:23:45)</td>
<td>how are you my dj</td>
</tr>
<tr>
<td>p U8: MAJANKLAL LE MLA (12:23:46)</td>
<td>got it before</td>
</tr>
</tbody>
</table>

**Example 4.14**: Colour negotiations 2

4.4 Summary

It should be clear by the examples discussed how even the users of RRM chat-room are embracing, almost certainly unconsciously, the principle of universalism espoused by Kosovo Roma musicians during the 1980’s (3.1.1) and underpinning the choices of RRM owners and DJs in designing the web-site and selecting the music to be broadcasted (3.2).

As I noted, such choices frame RRM as both a memorial to the lost homeland and a cosmopolitan place. The users of RRM chat-room take this a step further and turn RRM
into what might be considered as a new homeland in cyberspace. Into this new, virtual homeland they recreate the now dispersed community. As in the case of other diasporas (Hiller & Franz 2004, Mitra 1997, Tsaliki 2003) this virtual community is characterized by the recreation of the fraternity and conviviality of the homeland, the re-establishment of lost ties and the creation of new ones. On RRM, this aspect is clearly shown in the recreation of the abava atmosphere and the constant attempts at obtaining other users’ IM contacts. The inclusion of adverts about community gatherings, both in the homepage and within chat-room discussions, replaces the function of directories and event lists found in more professional diasporic web-sites (Qiu 2003, Mitra 2003) in providing contents that resonate with their audience collective identity. Similarly, the links to news sites about Kosovo provide a way to stay in contact with news and popular culture from the homeland (Sinclair & Cunningham 2000).

At the same time RRM differs from other diasporic virtual communities in that it is not a space where stereotyping and marginalizing discourses about the community are challenged through open discursive negotiations and representations (Mitra 2001, Santiasianni 2003, Yang 2003). Rather, users of RRM chat-room perform multiple identities, as they do in off-line life, in a series of interactions that helps constituting the users as a community. This peculiarity of RRM can be attributed, first and foremost, to the lack in the frame provided by the web-site designers of any overt invitation to talk about the community identity, something that is constantly found in previously studied diasporic websites. Furthermore, the medium provided for interaction, being synchronous and multi-user proves to be, as it did in various other instances (Herring 2002) unconducive to this kind of serious, sustained discussions. Finally, the level of education of RRM users is generally lower than that of the diasporic CMC users documented in the literature (which has largely focused on university students and young professional) thus, I would argue, making them less inclined toward introspection.

The frame provided by the web-site designers, with its usage of multiple languages also act as the licensing factor for users to freely employ their whole linguistic repertoire. I have shown through the examples how Romani is the dominant language and the preferred choice to talk about life within the community and to police the on-line discussions. Other languages, on the other hand, tend to be used when talking about life in the wider society. Code switching, however, is largely employed to mitigate face-threatening acts by marking content as jocular or by inserting politeness markers, to display particular identities through fixed expressions (salutations and emphasis markers) and in jokes, teases, attention-seekers and appeals.
These patterns of code-switching are common in many diasporic on-line communities (Georgakopoulou 1997, 2004, Sperlich 2005, Androutsopoulos 2006, 2007, McClure 2001) although in such cases English or the languages of the host countries were dominant. The dominance of English or national languages has always been attributed to on-going processes of language shift and to the role of English as the Internet lingua franca. The dominance of Romani on RRM points at the health of the language even in the diaspora. This resilience can be attributed to two factors that favour inter-generational transmission of the language: the Mitrovica Roma tendency at migrating as familial units rather than individuals and their condition as a marginalized and often segregated minority (see 3.1).

However, as in other diasporic web-spaces, code-switching serves RRM users to display their alignment with multiple, hybrid identities. Such identities, however, have at their core a sense of Romani-ness clearly flagged by the dominant usage of Romani.

This usage of Romani is characterized by playfulness (as I have shown in this chapter) and creativity (as I will show in chapter 6), as much as the registers of standardized languages employed in CMC. Speakers of standardized languages, coming from literate cultures, display these characteristics in the way they re-appropriate established writing norms and introduce new ones. Georgakopoulou (1997, 2004) and Palfreyman & Al Khalil (2007) suggest that this re-appropriation and creation of writing norms flags the CMC users’ shared understanding of themselves as a community. For Romani speakers, coming from an oral culture with no writing norms to refer to, it is the act itself of writing the language that displays such characteristics. This further difference between Roma and non-Roma CMC users can also explain why RRM users have embraced only a small portion of established CMC norms.

The different set of norms RRM users abide to, coupled with the on-line continuation of off-line practices, their spontaneous policing and the usage of their wide linguistic repertoire constitute them as a virtual community and simultaneously keep RRM a decidedly exclusive space for Mitrovica Roma. Other Roma who do not share the same broad linguistic repertoire or are unfamiliar with certain off-line norms of the Mitrovica community generally spend little time on RRM or interact very little on the chat-room.

In spite of the exclusiveness of the space they contribute creating, RRM users display a surprising openness when it comes to the markers employed in this process. They do not abide to a preconceived and essentialist notion of who and what they are that would limit them to the use of a limited array of resources. Rather, they employ and accept everything that their condition as a diaspora engaging with CMC technologies offers them in order to recreate the conviviality of the homeland.
In the previous chapter I have analyzed the language choice patterns on RRM chat-room. I deliberately took a micro-level, discourse analytic and ethnographically informed approach in order to give the reader a qualitative appreciation of the exchanges occurring on the chat-room. From this analysis it emerged that Romani is the dominant language and that code switching is employed to mitigate face-threatening acts by marking content as jocular or by inserting politeness markers, to display particular identities through fixed expressions (salutations and emphasis markers) and in jokes, teases, attention-seekers and appeals. It also emerged that code mixing dominates the choice of user-names where, I suggested, it is used to provide a first indicator of users’ alignment with multiple, hybrid identities.

Following Androutsopoulos (2006a) suggestion that on-line ethnography and traditional sociolinguistic analysis should be combined to achieve “a more complex theorizing of the social and contextual diversity of language use on the Internet” I will now turn to a quantitative analysis of the corpus. Combining qualitative and quantitative approaches will allow me to account both for micro level, individual patterns of usage and macro level trends shared by all users involved.

This will show that the patterns identified in the previous chapter are not limited to the selected examples and will give an impression of just how representative they are of the exchanges on RRM chat-room.

5.1 Methodology

The log save function built in the Flatcast add-on (see 4.1) allows saving up to one hour of conversation at a time. Over an estimated participation time of 100 hours, this resulted in a total of 32.40 hours of conversation, equivalent to 32,597 messages, being logged. For the quantitative analysis I selected a sample of 1,000 messages, roughly equivalent to one hour of conversation. As during the fieldwork I noticed differences in the numbers, age and sex of logged-on users at different times of the day (see 4.2.1), I made sure that the selected messages were equally distributed between those posted in the morning, afternoon, evening and night. Apart from this criterion, the messages were selected randomly, at this stage without looking at the content of the exchanges or at the languages used.

Firstly, I identified all the user-names and the languages in which they were composed. Excluding anonymous users (see 4.2.1) I counted about 100 users during my fieldwork. Of these, 10 were regulars, present on the chat-room nearly every day; another 15 were
frequent users, participating at least 3-4 days a week; the remaining were occasional users, present once a week or less. The 54 user-names identified thus constitute a fairly representative sample of the whole population.

Looking at the findings in the literature about code switching in diasporic CMC (see 1.3) it seems to me that, besides displaying users’ embracement of multiple identities, switching is triggered\(^1\) by the users’ desire to achieve certain communicative functions. What I am suggesting, based on Gumperz's (1977) assertion that the juxtaposition of languages can be meaningful, is that CMC users, lacking in extra-linguistic cues like intonation and body language, are prompted to code-switch in order to enrich their messages. This can range from the addition or reinforcement of a specific meaning or ‘tone’ not necessarily clear from the words users are employing, to the flagging of their intention to change other users’ behaviour, whatever by simply catching their attention or by having them do something.

On RRM chat-room exchanges, the communicative functions that seemed most prominently linked to switching, and to which I will be looking, were the expression of politeness and humour (including teasing), salutation, attention seeking, request, singing and formalized shouting. A further communicative function I noticed, and to which I will refer as emphasis, was the addition of a sense of urgency or the attempt at highlighting the importance of a message, both showing an emotional involvement of the user.

It can be noted that these functions regularly feature in the literature on diasporic CMC as strongly connected to code switching together with insulting, change of addressee face-saving and policing (see 1.3). While I have included the first two in my analysis in order to better compare RRM with other diasporic CMC environments, I have not included face-saving and policing. On RRM chat-room, face-saving is always achieved through humour or the insertion of politeness markers. Similarly, policing always involves a request and possibly politeness and emphasis markers. Since humour, request, politeness and emphasis were already covered, specific categories for face-saving and policing appeared redundant. Other functions such as narration, description, interrogation and so on contribute, within RRM interactions, to the structuring of what (Ehlich & Rehbein 1980) define homileic discourse. According to them, homileic discourse is used to socialise and provide entertainment. Often characterized by ludic aspects, it normally occurs in spontaneous and casual conversations (Ehlich & Rehbein 1980: 343). Since on RRM users are constantly socializing and providing entertainment to each other through spontaneous conversations.

(see chapter 4), it seemed appropriate to merged all communicative functions contributing to homileic discourse into a single category I refer to as plain talk. Based on my fieldwork observations, plain talk did not seem to be accompanied by code switching.

In order to verify if the communicative functions listed above do or do not trigger switching, I coded each message for the languages used in it, the presence or absence of switches and the communicative functions served by the message as a whole.

Taking messages rather than conversational turns as units of analysis is justified by the nature of the medium itself: users tend to spread their conversational turns over more than one message and are often engaging in parallel conversations whose turns, spread over many messages, regularly overlap (see 4.1.2 and all the examples discussed). Finally the fact that messages are generally short, sometimes even comprising just one word, further complicates establishing if a switch is insertion or sequential, as defined by Auer (1998). Therefore, in order to verify how code switching is realized in synchronous CMC environments, I coded switches on the basis of their position in the flow of messages: within a single message or from a previous message by the same user. Finally, I coded switches for their direction: from Romani into a contact language, from a contact language into Romani, from a contact language into another contact language.

Concerning the languages, I found myself struggling to classify messages which only contained written-out sounds (i.e.: hahaha for a laughter). Similarly problematic were certain abbreviations and expressions used in many languages (i.e.: ok, np, Nr I) as well as many internationalism, mostly related to popular culture (i.e.: brand names, star, boss). Such items proved particularly problematic when the same user was posting them in isolation between two messages in different languages. I therefore classified all these elements as a further language, which I will refer to as Net-talk. It can be argued that it constitutes a CMC register rather than a distinct language, and that would indeed be the case if RRM was a monolingual environment. However, following Blommaert’s description of multilingual language practice as drawing on “truncated repertoires” (2010: 106) in which limited resources from different languages are used side by side, Net-talk can be regarded as a CMC-specific pool of resources. As such, it constitutes a further code to and from which users can switch or from which they can draw to insert single items in sentences in one of the other languages they use.

I will now introduce the results of the quantitative analysis, starting with the distribution of languages in user-names (5.2). The general distribution of languages and switches in the messages will then be discussed (5.3.1) followed by the discussion of their distribution by communicative functions (5.3.2)
5.2 Language choice in user-names

In 4.2.1 I noted how, similarly to other multilingual CMC environments, users on RRM chat-room tend to mix languages in their user-names. Example 5.1 shows some of the user-names on which I based my observation.

<table>
<thead>
<tr>
<th>User-name</th>
<th>Languages used</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ROMANI MAHALA</em></td>
<td>Romani</td>
<td></td>
</tr>
<tr>
<td><em>FARiJA</em></td>
<td>Romani</td>
<td></td>
</tr>
<tr>
<td><strong>ALPACINO-NR1</strong></td>
<td>Net-talk</td>
<td></td>
</tr>
<tr>
<td>::::Mohamed-Roanne::::::</td>
<td>Romani+French</td>
<td>Roanne: city in central France</td>
</tr>
<tr>
<td><em>-EmRaH'StyLe</em>-_</td>
<td>Romani+English</td>
<td></td>
</tr>
<tr>
<td>edy.boy</td>
<td>Romani+English</td>
<td></td>
</tr>
<tr>
<td>CIbAlO_bOY</td>
<td>Romani+Net-talk</td>
<td></td>
</tr>
<tr>
<td>ANNA-MISS-FR</td>
<td>Romani+Net-talk</td>
<td></td>
</tr>
<tr>
<td><em><strong>ArMaNy-CaJ</strong></em></td>
<td>Net-talk+Romani</td>
<td>caj ‘Romani girl’</td>
</tr>
<tr>
<td>MiiLK$a_qoKoLLaDa*$</td>
<td>Net-talk+Albanian</td>
<td>Milka: chocolate brand</td>
</tr>
<tr>
<td></td>
<td></td>
<td>qokollada: Albanian spelling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>q = voiceless palatal stop</td>
</tr>
</tbody>
</table>

Example 5.1: User-names. Personal names have been considered Romani

As shown in graph 5.1, the quantitative analysis of user-names confirms language mixing in user-names is indeed common. In the composition of 30 out of 54 user-names (56.6%) two or more languages were used, while in 24 (44.4%) a single language was employed.

Graph 5.1: Language mixing in user-names

I also suggested that, just like in the cases documented by Androutsopoulos (2006b, 2007), mixing languages in user-names constitutes a strategy to show the users’ embracement of multiple, hybrid identities. To understand of what elements these
identities are composed, it is interesting to look at the choices made by the users, as exemplified in graph 5.2.

Of the user-names in a single language, 19 (35.2%) are in Romani, 3 (5.6%) in Net-talk and 2 (3.7%) in English. Choosing Romani only is a clear identity statement, forwarding users’ identity as Roma. Similarly, the choice of English or Net-talk, as the languages of modernity and of popular culture, signals a strong alignment with them.

As noted by Halwachs (2011) when it comes to TV and other traditional media, Roma at large tend to be consumers of mainstream media. The Mitrovica Roma are no exception to this and looking at brand names (Armani, Dolce&Gabbana, Hugo Boss) or at the name of actors (Al Pacino) chosen for their user-names and at the content of most YouTube videos linked to on RRM pages indeed confirms that the popular culture they want to align with is that propagated by mainstream media.

![User-names](image)

**Graph 5.2:** Language choice in user-names

Going back to graph 5.2, the other large slices confirm that being Roma, modern and participation into popular culture seem to be the main components of the identities RRM users are embracing. 13 users (24.1%) in fact combine Romani and English in their user-
names and a further 5 (9.3%) combine Romani and Net-talk. 4 more (7.4%) combine all
three languages.

A further confirmation of the centrality of these three components to the identities of
RRM users comes from the fact that the remaining 8 users (14.8%) employ one of the
Mitrovica Roma community contact languages together with either Romani, English, Net-
talk or a combination of the latter three.

Another interesting point is that although Qur'an Arabic is occasionally used in the
exchanges (see 5.3 and examples 4.2, 4.5, 4.7, 4.11) it does not appear in user-names. This
suggests that the religion of the wider community is not considered an identity aspect RRM
users want to forefront through their user-names.

Finally, it is important to notice that only in 8 user-names Romani is not used at all.
Combined with the fact that user-names exclusively in Romani constitute the majority of
the sample, this suggest that Romani-ness is the most important facet of identity RRM
users want to portray through their user-names.

The analysis of language choice in the user-names thus supports the conclusion put
forward in the previous chapter that RRM users embrace multiple, hybrid identities that
retain Romani-ness at their core. It has also shown that the desire to appear modern and to
show participation into popular culture is not limited to some users like _ * FARiJA * _ or _ _ EmRaH°StyLe_ _, whose interactions I analyzed.

5.3 Language choice in messages

Language choice in user-names tells us about the identities facets users want to
forefront, but not how they actually perform such identities. The analysis of language
choice in the messages and of the role played by switching between them will cover
precisely this aspect of RRM interactions. At first, I will introduce the general distribution
of languages and switches, in order to identify the main language of interaction on RRM.
Subsequently, the relationship between languages, switches and communicative functions
will be discussed.

5.3.1 General distribution

As it is the case with user-names, even in messages Romani is the most frequently used
language, as shown in graph 5.3.

Two thirds of the messages (671 – 67.1%) contain Romani. Of these, 580 (58%) are
entirely in Romani and in the remaining 91 (9.1%) Romani is used alongside at least one
more language. Net-talk is the language most frequently used with Romani, occurring in
37 messages (3.7%). Serbo/Croatian, with 13 messages (1.3%) is the first among the contact languages to be used with Romani. Interestingly, English with 11 messages (1.1%) out-ranks the remaining contact languages, all occurring alongside Romani in less than 10 messages (below 1%). More complex messages containing Romani, a contact language and Net-talk are even less frequent (below 0.5%).

Looking at the usage of other languages in isolation, Net-talk is again the favoured one occurring in 216 messages (21.6%). German and Italian are the only others totalling more than 1% (German 4.8% with 48 messages; Italian 1.1% with 11 messages), while the remaining contact languages, English included, score low (below 1%).

Combinations of two or more contact languages are extremely rare, all occurring in less than 5 messages (below 0.5%).

Graph 5.3: Language choice in messages
Finally, 22 messages (2.2%) were either empty, containing meaningless strings of characters or used such unusual spellings to make establishing the language used impossible and were therefore excluded from analysis.

It is interesting to note that, of the pre-dispersal contact languages, Serbo/Croatian is used in the messages, while Albanian is not. However, Albanian was used in one of the user-names (see Example 5.1) and in RRM home page welcome message (see 3.2). This seems to suggest that while some users retain some, probably passive, knowledge of Albanian, the majority of users has lost any active knowledge of it.

Looking at the other languages, German seems to be the post-dispersal contact language users are more comfortable with, as suggested by its relatively high usage in isolation. English, although used relatively frequently in conjunction with Romani is seldom used in isolation. Coupled with the usage of well know words alongside Romani seen in the previous chapter, this suggests that RRM users’ competence in English is not very high, but the usage of single words allows them to show their participation into popular culture.

The latter points will be investigated more in depth in the next section where I analyze the relationship between languages, switching and communicative functions. Before moving to the next session, however, I will discuss the frequency of switches and their position in messages.

Graph 5.4 clearly shows how switching is employed in slightly more than a third of the messages (370 – 37.8%), confirming that it is a widely employed strategy although it does not dominate RRM exchanges.
Regarding the position of switches in the flow of messages, users can switch within a single message, as shown in 5.1a, in which a French word opens a Romani sentence, and in 5.1b, where a Qur’an Arabic word closes a French sentence.

<table>
<thead>
<tr>
<th>RRM messages</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a ANNA-MISS-FR: #rien# sar coripe kate naj so cerav (18:49:06)</td>
<td>#nothing# is like poverty here there's nothing for me to do</td>
</tr>
<tr>
<td>b:::.Mohamed-Roanne:::. JA DE LA FRANCE DE MERDE #WALLAH# (18:49:23)</td>
<td>yes France is shit #for god’s sake#</td>
</tr>
<tr>
<td>c ANNA-MISS-FR: bravo edy ka cav tu ani buci samo za kilometra te racuni (18:49:35)</td>
<td>bravo edy I'll drop you to work only to count the kilometres</td>
</tr>
<tr>
<td>d ANNA-MISS-FR: <em>kikikikikikikikiki</em> (18:49:40)</td>
<td><em>kikikikikikikiki</em></td>
</tr>
<tr>
<td>e:::.Mohamed-Roanne:::. <em>HAHAH</em> (18:49:47)</td>
<td><em>hahah</em></td>
</tr>
<tr>
<td>f edy.boy: AHAHAHAH (18:49:50)</td>
<td>ahahah</td>
</tr>
<tr>
<td>g ANNA-MISS-FR: <em>oui</em> mohamed (18:49:52)</td>
<td><em>yes</em> mohamed</td>
</tr>
</tbody>
</table>

Example 5.2: Position of switches in the flow of messages. Dashes mark beginning and end of internal switches. Stars mark beginning and end of switches from a previous message.

Another possibility is for users to switch in their following message as shown in 5.1d, where the written-out laugh, a Net-talk element, is emphasising the humorous tone of the Romani sentence in 5.1c.

Graph 5.5 shows how this type of switch, from a previous message by the same user, occurring 73.8% of the times (273 tokens) clearly outranks switches within a message (internal switches, 76 tokens - 20.5%).

This notable difference is surely influenced by the fact that users are often engaged in multiple conversations and a switch from their previous message can be motivated by
accommodation to a different user’s choice of language. This is clearly shown in 5.1g where ANNA-MISS-FR switches to French to accommodate to ::::Mohamed-Roanne:::. usage of French in 5.1b, while her previous message in Net-talk (5.1d) was addressed to edy.boy.

<table>
<thead>
<tr>
<th>RRM messages</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>Sadi Styler: JA MAN HAB DICH DIGGA BISSCHEN #KALI# ABER EGAL #KUDLI# (18:46:11)</td>
</tr>
<tr>
<td>b</td>
<td>Sadi Styler: <em>HEHEHE</em> (18:46:12)</td>
</tr>
<tr>
<td>c</td>
<td>MiiLKa_qoKoLLaDa*: PENNER (18:46:21)</td>
</tr>
<tr>
<td>d</td>
<td>MiiLKa_qoKoLLaDa*: <em>HAHAHHA</em> (18:46:23)</td>
</tr>
<tr>
<td>e</td>
<td>MiiLKa_qoKoLLaDa*: <em>JA WAS SOLL ICH</em> MACHEN (18:46:30)</td>
</tr>
<tr>
<td>f</td>
<td>MiiLKa_qoKoLLaDa*: <em>KAJ SEM KALII</em> (18:46:33)</td>
</tr>
</tbody>
</table>

Example 5.3: Splitting of conversational turn at point of switch

However, this difference in frequency of the two types of switches also confirms that the medium itself is influencing the code-switching patterns of users. Considering that users of synchronous CMC modes split their conversational turns over more than one message not to loose the attention of their interlocutor (Herring 2002), I suggest that switching between messages serves a similar function. Switching between messages, in fact, creates a sharp contrast between subsequent messages by juxtaposing two languages. Such a contrast puts the switching itself in a prominent position thus enhancing its communicative functions, for example signalling a change of addressee as in 5.1g or emphasising the humorous content of a previous message as 5.2b. Furthermore isolating switches in a different message can reinforce the splitting of conversational turns over more messages as a device to capture the interlocutor’s attention, as shown in 5.2e and 5.2f where the Romani subordinate clause is separated from the German main clause.

5.3.2 Distribution by communicative functions

As observed in chapter 4, certain languages seemed to be preferred for certain communicative functions and I have suggested that such communicative functions can trigger switches when users want to achieve them. Furthermore, in the previous section I have hinted at the fact that RRM users might not be equally competent in all the languages they employ in the chat-room exchanges. Analyzing the distribution of languages and switches in relation to communicative functions will allow verifying if these two observations are in fact founded.
When I counted the instances of language usage and switches in relation to communicative functions it was often the case that a single message or switch served more than one function at a time. This is shown in 5.4a, where U13 addressed one of the formalized shouts (see 4.2.3) to Guest60347471 to get his attention. In 5.4b, another anonymous user, Guest60446136, tried to get U13 attention by describing his position. Finally, in 5.4c the switch into Net-talk served to add humour and give emphasis to the previous message by Guest60446136.

<table>
<thead>
<tr>
<th>RRM messages</th>
<th>Translation</th>
<th>Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>a U13: JASHAAAAA... Guest60347471 (13:26:08)</td>
<td>c'mooooon</td>
<td>Formalized shouting/Attention seeking</td>
</tr>
<tr>
<td>b Guest60446136: U13 pale tu kate (13:26:17)</td>
<td>U13 [I'm] here close to you</td>
<td>Attention seeking/Plain talk</td>
</tr>
<tr>
<td>c Guest60446136: <em>hehe</em> (13:26:18)</td>
<td><em>hehe</em></td>
<td>Humour/Emphasis</td>
</tr>
</tbody>
</table>

Example 5.4: Mapping of communicative functions onto messages

As a result of this, the total count of instances of the various functions is higher than the actual number of messages and switches.

The data in Table 5.1 clearly show that Romani, besides being the overall more frequently used language, is also the only one used in all the communicative functions identified.

Net-talk, although being second in overall frequency, has a much more restricted usage and is almost exclusively limited to emphasis and singing. German, although with a much lower count, is however used in nearly all functions, with the exclusion of emphasis, singing and insult. Excluding the high number of instances of singing in Serbo/Croatian, this language is not used significantly in any other functions. English usage, on the other hand, is evenly distributed in salutations, attention seeking, humour and plain talk. Qur’an Arabic like Serbo/Croatian shows a high specialization, but for emphatic purposes. It is also occasionally used in salutations. Similarly, Italian is used mostly for salutations and attention seeking and is nearly absent in other functions. French, although less frequent than English shows usage in exactly the same functions, namely salutations, attention seeking, humour and plain talk. Finally, Turkish also appeared in the sample but in just two messages. One, introduced by DJ šaj (Romani ‘DJ can you [play]’), contained the title of a Turkish song the user wished to listen to, while the other one contained the refrain of the same song.

The distribution of switches (Table 5.2) confirms some of the observations from the previous chapter. Switching for emphatic purposes, the addition of a sense of urgency or the attempt at highlighting the importance of a message, is clearly dominant while it ranks equally with its avoidance for politeness and change of addressee purposes.
Table 5.2: Distribution of function by selects. % of the total occurrences of function X endorsed by each of the 5070 XOT - Other Language

<table>
<thead>
<tr>
<th>Percent</th>
<th>No OL</th>
<th>11 OL or More</th>
<th>Relevant</th>
<th>Emphasis</th>
<th>Humour</th>
<th>Request</th>
<th>Politeness</th>
<th>Attention</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 5.1: Distribution of function by selects. % of the total occurrences of function Y endorsed in language X.

<table>
<thead>
<tr>
<th>Percent</th>
<th>No OL</th>
<th>11 OL or More</th>
<th>Relevant</th>
<th>Emphasis</th>
<th>Humour</th>
<th>Request</th>
<th>Politeness</th>
<th>Attention</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
On the other hand, my observation from the previous chapter that humour, salutations, attention seeking and singing seemed to trigger switching does not seem to apply to the larger sample.

Given the overall dominance of Romani it is not surprising that the most frequent direction of the switches is from Romani into another language, shortly followed by switches from another language to Romani. Switches between two languages other than Romani, although used, are very rare, suggesting that users tend to return to Romani from a contact language before switching into another one.

The dominance of Romani on RRM chat-room confirms that, as I observed through the examples in chapter 4, users want to forefront Romani-ness as a core element of their identity. Furthermore, it suggests that Romani-ness is the only identity facet shared by all users. Similarly, it is safe to assume that Romani is the only language shared, at least to some degree, by all users while other languages might be completely absent from the repertoire of individual users as they have never spent enough time in countries where a given language is spoken. Thus, using Romani as the default language of communication in this setting flags the shared understanding of users as members of a single community. The diasporic nature of this community and the individual differences among users, on the other hand, are made manifest by switching into the other languages they are experiencing in their daily life.

It is also worth noting that some of my conclusions regarding the functions of switching are apparently challenged by the general overview of communicative functions distribution by switches. However, not all the differences in frequency of a given language or switching in certain functions are relevant, as chi-square tests for each function showed.

**Salutations**

In the qualitative analysis switching for salutations, as shown in example 5.5, appeared as a well established pattern.

However, on the larger sample the preference for not switching for salutations (graph 5.6) is relevant ($p < 0.02$).

As I have observed, however, switching for salutations constituted a marked choice as it helped to signal specific aspects of users’ identities (4.3.1). It is then not surprising that on a larger sample switching is not dominant. Considering the presence in the examples of many users who were trying to present a strong on-line persona, it can be concluded that switching for salutations is a strategy limited to this minority of users.
<table>
<thead>
<tr>
<th>RRM message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Admir: #selam aleykum# sarinenge ko radio (18:22:25)</td>
<td>#selam aleykum# to everyone on the radio</td>
</tr>
<tr>
<td>B Admir: zdr zdr dj (18:22:29)</td>
<td>hello hello dj</td>
</tr>
<tr>
<td>C <em>ROMANI MAHALA</em>: <em>ALEKIM SELMA</em> ADMIR (18:22:33)</td>
<td><em>aleykum selam</em> Admir</td>
</tr>
<tr>
<td>D <em>* FARiJA*</em> : ZDRAVOOO ZDRAVOOO!! SVIMAA DanieleItalianoDilo (18:30:56)</td>
<td>hellooo helooo!! to everyonee DanieleItalianoDilo</td>
</tr>
<tr>
<td>E <em>* FARiJA*</em> : <em>HELLO</em> DJ (18:30:58)</td>
<td><em>hello</em> dj</td>
</tr>
<tr>
<td>F <em>ROMANI MAHALA</em>: <em>OOOOOO</em> #HALLO# FARIJA (18:31:05)</td>
<td><em>oooooo</em> #hello# farija</td>
</tr>
</tbody>
</table>

**Example 5.5:** Switching for salutations.

![Graph](image)

**Graph 5.6:** Switches for salutations. OL = Other Language

Looking at the languages used for salutations (graph 5.7) can give an idea about what identity aspects these particular users want to forefront.

The higher preference for Romani with its 63 over 89 ($p < 0.01$) instances of usage is immediately clear. Among the remaining languages used, the post-dispersal contact languages together with English outrank Qur’an Arabic, while Serbo/Croatian and Turkish are never used.

This preference for the more recent contact languages of the Mitrovica community over the older ones can be taken as an indicator of the progressive loss of active knowledge of the older contact languages among RRM users.
Furthermore, it complements the preference for Net-talk and English for user-name. Switching for salutations, although a marked behaviour, reserved for the purpose of foregrounding particular identity facets by a minority of users, is realized in such a way to strengthen the common tendency of RRM users to present themselves as modern and participating into popular culture.

**Attention seeking**

Based on instances such as in example 5.6, where U7 intervened in a German conversation between other users starting his message with a Net-talk element, I suggested that attention seeking as well triggered switching.

<table>
<thead>
<tr>
<th>RRM Message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a MiILKa_qoKoLLaDa*: EMRAN IS AUCH MEIN SCHATZ (18:49:28)</td>
<td><em>emran is my sweetheart too</em></td>
</tr>
<tr>
<td>b Sadi Styler: <em>AHA OKE</em> (18:49:30)</td>
<td><em>aha ok</em></td>
</tr>
<tr>
<td>c U7: <em>HEY</em> #MA SADIN E EMRAH VOY SI MO AMAL# #VALLAH# (18:49:31)</td>
<td><em>hey</em> #don't laugh at emrah he's my friend# #for god's sake#</td>
</tr>
</tbody>
</table>

Example 5.6: Switching for attention seeking

However, the same pattern observed for salutations can also be seen in attention seekers.

Avoidance of switches (graph 5.8) is again dominant with 69 attention seekers realized without switching against 15 realized by switching ($p < 0.001$).
Romani is again the favoured language (graph 5.9) with 65 over 84 instances of usage ($p < 0.01$) and the more recent contact languages, together with Net-talk and English, outrank Serbo/Croatian and Qur’an Arabic.

Thus, the conclusions put forward for salutations can perfectly explain the distribution of switches and languages in attention seekers.

*Politeness*

In discussing the usage of politeness markers (see Example 5.7) I noted how both Romani and other languages seemed to be used regularly without any particular preference for one or another and thus switching seemed common.
<table>
<thead>
<tr>
<th>RRM Message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>a U12: U2 AVER FARBAA #STPPP# (19:04:01)</td>
<td>U2 another colour #plss#</td>
</tr>
<tr>
<td>b U8: DJ SAJJ JEK GILI (12:08:31)</td>
<td>dj please one song</td>
</tr>
<tr>
<td>c U14: dj saj jek gili #bitte# (20:10:12)</td>
<td>dj please one song #please#</td>
</tr>
</tbody>
</table>

**Example 5.7:** Switching for politeness

The analysis of the larger sample confirms these observations.

Politeness was achieved with or without switching exactly the same number of times, namely 30, showing there is not a favourite pattern among RRM users (graph 5.10).

**Graph 5.10:** Distribution of switches for politeness

**Graph 5.11:** Distribution of languages for politeness
Similarly, although Romani is more frequently used (graph 5.11), with 32 over 60 instances of usage, this difference does not appear relevant ($p > 0.05$). Again as we have seen for salutations and attention seekers the more recent contact languages together with English and Net-talk are more frequently used than Serbo/Croatian.

**Requests**

In the case of requests for personal details (4.3.2, 4.3.3), for change of behaviour and for songs (4.3.4, 4.3.5) I observed how Romani appeared to be the favourite language and how switching was generally avoided.

![Graph 5.12: Distribution of switches for requests](image)

![Graph 5.13: Distribution of languages for requests](image)
The same emerged from the analysis of the larger sample, where 36 over 45 \( (p < 0.02) \) requests did not include a switch (graph 5.12) and Romani was used in 40 of them \( (p < 0.001, \text{graph } 5.13) \).

**Humour**

Humour, as discussed in 4.3.3 (see also Example 5.8), appeared to be an area were switching was dominant, but the analysis of the larger corpus shows that in 193 cases against 124 humour is achieved without switching. However, this difference proved to be irrelevant \( (p > 0.05) \).

![Graph 5.14: Distribution of switches for humour](image)

<table>
<thead>
<tr>
<th>RRM Message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>edy.boy: LE FRANCCE MANGE LE FORMAGE #KAI TE AVOL LEN ZURALE KOKALA (18:54:36)</td>
<td>*the French eat the cheese <em>so that they'll have strong bones</em></td>
</tr>
</tbody>
</table>

**Example 5.8:** Switching for humour

In the case of languages as well, the apparent prominence of Romani with 197 instances of usage against the 120 of all other languages (graph 5.15) proved to be irrelevant \( (p > 0.05) \) confirming that no language is favoured over the others for humorous purposes.

It is important to notice that the high occurrence of Net-talk in humorous functions is entirely due to the usage of written out laughs to mark otherwise ambiguous messages (see 5.4c). Conversely, all other languages are used as Romani to actually perform jokes and other humorous acts. Furthermore, the high occurrence of German in this function suggests
that, among the new contact languages, it might be the one users are more comfortable with.

**Languages**

Graph 5.15: Distribution of languages for humour

*Emphasis*

In all examples discussed emphasis, the addition of a sense of urgency or the attempt at highlighting the importance of a message, both showing an emotional involvement of the user was always achieved by switching, generally into Qur’an Arabic.

<table>
<thead>
<tr>
<th>RRM Message</th>
<th>Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ROMANI MAHALA</em>: MILKA ASUN SO CEROL KATE O DEMAJLI #ALLAHILE# (18:39:12)</td>
<td>milka listen to what demajli does here #oh my god#</td>
</tr>
</tbody>
</table>

Example 5.9: Switching for emphasis

The prominence of switching for emphatic purposes clearly emerged from the quantitative analysis: in 73 cases against 8 ($p < 0.001$), emphasis was achieved by switching (graph 16).

Regarding languages, however, Qur’an Arabic, although scoring its highest usage in this function, is largely outranked by Net-talk which, with its 61 over 81 occurrences most of them written-out sounds, is the favourite language for emphatic purposes ($p > 0.001$).
These results, besides confirming the observations regarding switching for emphasis, strengthen the impression that RRM users, although still retaining an attachment to Islam are more inclined towards those identity markers that link them to modernity and popular culture.

**Plain talk**

In the case of plain talk, the umbrella term I am using for functions that contribute to the structuring of homileic discourse (Ehlich & Rehbein 1980) Romani is clearly the favoured language (graph 5.19) and switching is avoided (graph 5.18). In both cases the differences appear highly significant ($p < 0.001$).
Again, as it was the case with humour, German not only scores high but its usage is also qualitative similar to that of Romani, reinforcing the impression that it is the preferred among the new contact languages.

*Singing and formalized shouting*

In 4.3.1 I have explained how singing and formalized shouting contribute to the recreation of the *abava* atmosphere on RRM chat-room. It will therefore be expected that Romani is the favourite language in this function and that switching is avoided.
Although graph 5.20 shows that avoidance of switching is more frequent in singing and shouting, the difference is not relevant ($p > 0.05$).

The explanation for the non-relevance of switching avoidance in singing and shouting comes from the distribution of languages in such function (graph 5.21).

Romani and Net-talk score almost equal, with the latter being slightly more frequent (81 instances for Romani, 84 for Net-talk) a difference that, however, appears relevant ($p < 0.001$). This is due to the high number of written out sounds not assimilable to any of the formal shouts (jaša, opsa, ček, joj) found in the sample.
Furthermore, as occasionally songs in Serbo/Croatian and, although very rarely, in Turkish are played by the DJs, users also post lines from the lyrics of such songs, thus resulting in an overall count of other languages used for singing higher than that of Romani, in spite of the high number of formalized shouts posted by users.

**Change of addressee**

Although changes of addressees did not feature prominently in any of the examples discussed in chapter 4 (but see example 5.2g), it seemed necessary to analyze the distribution of languages and switches in this function as it has been noted switching tend to accompany it in other diasporic CMC environments (Androutsopoulos 2007, McClure 2001, Sperlich 2005).

![Graph 5.22: Distribution of switches for change of addressee](image)

As it was the case with politeness, the instances in which a change of addressee is accompanied by a switch are exactly the same as those in which the opposite happens (45 in each case).

As for languages, although new interlocutors are more often addressed in Romani (62 instances over 90) this difference does not appear significant ($p > 0.05$). Finally, the relatively high number of instances of German (19) reinforces the impression that users are quite comfortable in using this language.
Insult

The extremely low occurrence of insults in the sample, just 6, all in Romani and not accompanied by switches, is statistically irrelevant. However, the datum is telling about the standards of politeness on RRM chat-room and confirms that users try to avoid conflict on it through self policing as I discussed in 4.3.5.

5.4 Summary

It should be clear by now that the trends illustrated through the ethnographic account (chapter 4) are overall shared by the whole of RRM users and not limited to those featuring in the examples discussed.

The pattern of code mixing in user-names confirmed that they desire to forefront Romani-ness, modernity and participation into popular culture in their self-presentations. These identity traits feature strongly in their interactions as well, as the usage of the various languages shows.

Romani is the main language of interaction and its usage in all communicative functions suggests all users are competent in it. Furthermore, its dominance can be taken as an indicator of the positive feelings of users toward it.

Turkish, active knowledge of which was already lost to the younger generations prior to the community dispersal (see 3.1.1) is employed on RRM only when songs in this language are requested or sung.

Qur’an Arabic is used on RRM in fixed expressions for salutations and emphatic purposes, confirming a lack of active knowledge among RRM users. As the language of
religious practice, memorized for the purpose of praying, such lack of active knowledge is not surprising.

Albanian is never employed in the messages and only rarely in user-names, pointing in this case to an even greater lack of knowledge among RRM users. Based on my fieldwork experience with the Mitrovica Roma settled in Palermo, where only the elders and few young adults who had left Kosovo in the 1990s knew Albanian, I would suggest that this is the case not just for RRM users but for the whole community. Considering Mitrovica was, and still is, an area of Kosovo with a Serbian majority and that Serbo/Croatian was the dominant contact language, it is not surprising that Albanian has been the first among them to fade away from speakers’ knowledge.

Serbo/Croatian, however, is employed prominently for singing along when songs in this language are broadcasted and only marginally for humour. Thus it would seems that active knowledge of it, although not as rapidly as that of Albanian, is also declining.

Knowledge of Italian among RRM users also appears limited and the language is almost exclusively used in fixed expressions in salutations and attention seekers. However, Italian is the primary contact language of those Mitrovica Roma settled in Italy, and as I have observed in face-to-face interactions in Palermo, they have a fairly good knowledge of it. It would then seem that the majority of RRM users are not settled in Italy. Considering Mitrovica Roma in Italy have limited access to the Internet as they are mostly settled in camps for nomads (see 3.1.3), this is not surprising. Given the harsher treatment received by Mitrovica Roma in Italy when compared to other countries, it can also be the case that RRM users settled in Italy are not particularly well inclined to the language and avoid using it extensively on RRM chat-room.

A similar case can be made for French. Although slightly less frequent than Italian, it is used in more functions, also including humour and plain talk. This can be taken as an indicator that similar numbers of users are settled in Italy and France, but that knowledge of the national language is better among those settled in France than among those settled in Italy. However, it would seem that, as for Italian, a somewhat controversial alignment with France (see also example 4.3) prevents the extensive usage of French.

Following this line of reasoning, the fact that German is the most frequently used among the post-dispersal contact languages and the one showing the broadest functional usage (all functions but emphasis, singing and insult) indicates both that most RRM users are settled in German-speaking areas and that they have a more positive orientation toward the German language.
It may be the case that these preferences are specifically relevant to RRM chat-room but that off-line language choices are quite different. However, verifying if this is the case would require extensive off-line fieldwork and it is beyond the scope of the present work. What is relevant is the fact that, no matter how many users are proficient in each language and what their feeling toward them are, once a licence for multilingualism is given (see the opening of RRM homepage in 3.2 and the conclusions in chapter 4) all languages available to the community are accepted and used. Furthermore English, not part of the linguistic repertoire of the Mitrovica Roma community at large, and Net-talk, a medium specific language, are also used frequently. All these languages, however, are not accessed randomly, and users show rather clear patterns of usage.

While Romani is clearly the dominant language overall, in expressing humour or emphasis and singing and shouting it is not dominant. Furthermore, these functions also show a positive correlation with code-switching. It is important to notice that this kind of functions is accompanied in face-to-face communication by cues such as intonation and body language. It would then seem that attempts at achieving communicative functions requiring extra-linguistic clues trigger switching. The juxtaposition of languages thus compensates for the disembodied nature of CMC communication. This is confirmed by the fact that RRM users stick to Romani and avoid code-switching in salutations, attention seekers, requests and plain talk, where the availability of transparent lexical items, formalized expression and punctuation compensate for the lack of extra-linguistic cues. Finally, expressing politeness and changing addressee sit somewhere in between these two possibilities. For both this functions transparent lexical markers are available in all the languages in the RRM users’ repertoires. However, in face-to-face communication such functions can also be achieved purely through intonation. Mirroring these possibilities, on RRM chat-room no language is preferred over the others for politeness and change of addressee, but switching and maintenance of the choice of any given language occur with equal frequency.

Moreover, the synchronous nature of chat-room communication offers a further opportunity for reinforcing this suppletive role of switching. As the prominence of switches from a previous message by the same user shows, users tend to raise the prominence of their switches by placing single switched words in a different message or by splitting sentences over more turns just before switching into another language (as shown in examples 5.2 and 5.3).
If a communicative trigger is behind switching on RRM, the choice of languages into which users switch seems to instead be governed by their desire to flag particular aspects of their identities.

As Sebba & Wootton suggest (1998) “[s]ocial identities are made manifest through talk, not just through the actual language or ‘code’ used but also through content and context” (284, emphasis in the original). Speakers draw on a “varied range of linguistic attributes - and through these, the extralinguistic structures they symbolise - in the creation of those identities” (277). On RRM chat-room ‘talk’ is obviously computer-mediated however, users do construct their identities through interactions exactly as the British Caribbean studied by Sebba & Wootton did.

The context in which RRM interactions take place is, as I have shown in chapter 3, a memorial to the users’ lost homeland where members of the younger generations congregate, as I have shown in chapter 4, to recreate the conviviality of the homeland itself. Here, through sustained usage of Romani, users clearly flag their attachment to Romani-ness. At the same time, by using Fench, German, Italian and Serbo/Croatian, even when their knowledge of them seems to be passive, they also flag their diasporic condition. English and Net-talk, on the other hand, are mobilized to enact the same identity aspects portrayed through user-names, namely modernity and participation into popular culture. Finally, Arabic, although employed occasionally by a minority of users, shows their religious affiliation.

Furthermore, Sebba & Wotton also point to the fact that identities are “flexible constructs, created, negotiated and constantly changed in the course of interaction” (1998: 277). This dynamism of identity performances is clearly visible in the behaviour of those users that deviate from the common patterns I have shown in this chapter. This group of user comprises the ones I have introduced in chapter 4 and a few others like them constantly trying to display strong on-line personas. To do so, they draw on well established behaviours and switch between languages in those situations in which the majority of users avoids doing so, namely in salutations and attention seekers. Rather than sticking to Romani in these functions, they switch to other languages and, again, they prefer those languages that will further highlight their alignment with modernity and popular culture.
6 The codification of Romani on RRM

In the previous two chapters I have shown how Romani is the main language of interaction on RRM chat-room. However, as I mentioned in 3.1.1, two dialects are spoken by the Mitrovica Roma: a Gurbet variety and an Arli variety. Furthermore, when describing RRM homepage (see 3.2) I noted how the spelling employed for Romani in the welcome message showed variation in the representation of certain sounds due to the influence of different writing systems.

Considering the history of Romani codification (see 2.3), the question arises concerning the choices made by RRM users in employing Romani in written communication. What variety are they using? Does their spelling show the same variation visible in the homepage welcome message? Are the same attitudes underpinning their choice of languages (as shown in chapters 4 and 5) also influencing the transition of Romani from an oral to a written language? To what extent are their choices compatible with other attempts at writing Romani?

In the following sections I will address these questions, drawing attention to RRM users’ choice of variety (6.2) and spelling (6.3) and relating them both to other instances of Romani codification and to the wider theoretical debate around language planning (6.4).

6.1 Methodology

I carried out the analysis of the chosen Romani variety and of the spelling used for it on the same 1000 messages sample employed in the previous chapter.

Concerning the spelling used, I focused my attention on those sounds that are not transcribed in the same way across the various contact languages experienced by the Mitrovica Roma and on Romani specific sounds: the postalveolar affricates /č/ and /dž/, the postalveolar fricatives /š/ and /ž/, the uvular and glottal fricatives /x/ and /h/, the semi-vowels /j/ and /w/, the alveolar affricate /c/, the labial fricative /v/, the centralized vowels /ɨ/ and /ǝ/, the aspirates /ph/, /th/, /čh/ and /kh/ and the reflexes of Early Romani /ř/. This selection of sounds will allow ascertaining to what degree RRM users are following the established norms of the major languages they know and how aware they are of other attempts at writing Romani.

In order to ascertain what variety of Romani is employed on RRM I counted the occurrences of those morpho-syntactic and phonological features relevant in classifying Romani dialects in a geographic-historical perspective (Matras 2005, 2002: ch. 9). According to this perspective, relations between dialects are not genetic, and thus absolute,
but relative and based on the adoption by speakers of a set of similar features (Matras 2005: 21). Such features emerged through processes of change and innovation that took place between the 15th and 17th century, the period of settlement following the Roma dispersal from the Balkans (Matras 2005: 7). These innovations originated in different geographical centres and their spread from one community to the other through reciprocal contacts was limited by geographical and historical borders, resulting in the emergence of coherent diffusion spaces (Matras 2005: 8). Dialects can therefore be said to belong to a given diffusion space based on the number of shared features typical of that diffusion space. Furthermore, dialects located at the boundaries between diffusion spaces can share features typical of two or more diffusion spaces. Similarly, features from a diffusion space other than the one where a dialect originated from could have been acquired through contacts when communities moved from one space to another.

Based on these criteria, I have shown how one of the Romani varieties spoken by the Mitrovica Roma originated in the Vlax diffusion space, centred around Walachia and Moldavia and then moved into the Balkan diffusion space, acquiring some Balkan features, around the 17th century (Leggio 2011a). This historical trajectory is shared by various other Romani groups in the Southern Balkans generally referred to as Gurbet, a Balkan Turkish word meaning ‘foreigner’. These groups are often contrasted with groups referred to as Arli, another Balkan Turkish word meaning ‘settled’ (see Boretzky 1986, 2003, Matras 2002). Most of the dialects spoken by Arli Roma do not share the Vlax innovations found in the dialects of the Gurbet Roma. Some Vlax features might have been acquired in some Arli varieties through contact, but generally they show a number of archaisms and a very limited number of innovation originating from the Balkan diffusion space, pointing at their continuous presence in the region since the 13th century (Matras 2005, 2002: ch. 9).

Although my informants in Palermo rejected the term Gurbet as a self-denomination (Leggio 2011a: 59), I will use it to refer to their Romani variety to differentiate it from the other variety spoken by Mitrovica Roma which, following their own usage, I will refer to as Arli.

In order to establish what variety is used on RRM, I looked for occurrences in my sample of the Vlax features of Mitrovica Gurbet. The occurrences of such features needed to be contrasted with those of the Arli variety. Unfortunately, a description of the Arli variety spoken by Mitrovica Roma is at present not available. However, it is safe to assume that it would share many features with Arli varieties spoken in former-Yugoslavia. Six varieties from the region labelled Arli by the speakers are documented in the Romani Morpho-Syntax Database (www.romani.humanities.manchester.ac.uk/rms, RMS from now
A further two, labelled Kovački by speakers, share a large number of typical Balkan features and have therefore been included. The varieties in question are listed in Table 6.1, also showing the RMS codes I will use to refer to them in this work. Figure 6.1 shows where the dialects are spoken and their position relative to Mitrovica.

<table>
<thead>
<tr>
<th>RMS Code</th>
<th>Self-attributed name</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>MK-002</td>
<td>Arli</td>
<td>Skopje, Macedonia (blue on map)</td>
</tr>
<tr>
<td>MK-003</td>
<td>Arli</td>
<td>Skopje, Macedonia (blue on map)</td>
</tr>
<tr>
<td>MK-004</td>
<td>Kovački</td>
<td>Kumanovo, Macedonia (red on map)</td>
</tr>
<tr>
<td>MK-005</td>
<td>Arli (Gautnikane)</td>
<td>Kumanovo, Macedonia (red on map)</td>
</tr>
<tr>
<td>MK-012</td>
<td>Kovački</td>
<td>Skopje, Macedonia (blue on map)</td>
</tr>
<tr>
<td>YU-011</td>
<td>Arli</td>
<td>Beočin, Serbia (pink on map)</td>
</tr>
<tr>
<td>YU-014</td>
<td>Arli</td>
<td>Zrenjanin, Serbia (black on map)</td>
</tr>
<tr>
<td>YU-016</td>
<td>Kosovan Arli</td>
<td>Gnjilane, Kosovo (yellow on map)</td>
</tr>
</tbody>
</table>

**Table 6.1**: Arli varieties of former-Yugoslavia documented on RMS

![Figure 6.1: Location of Arli varieties. Map from RMS](image)

I did not expect to encounter dialect features originating from diffusion spaces other than the Vlax and Balkan ones as speakers of dialects from Central, North-western and North-eastern Europe did not seem to participate on RRM conversations. Therefore I first looked at features originating in the Vlax and Balkan diffusion spaces and encountered both in Mitrovica Gurbet and in the Arli varieties (fully shared features, 6.2.1). Then, in order to establish what variety is employed on RRM, I looked at the contrasting features (6.2.2) distinguishing the Vlax diffusion space (and thus Mitrovica Gurbet) from the Balkan diffusion space (and thus the Arli varieties). Finally, a number of Mitrovica Gurbet
features were shared by some but not all the Arli varieties. Such features are either Vlax features adopted into some Balkan varieties or contact developments common to the two groups as they are both influenced by Slavic languages. I looked at those partially shared features (6.2.3) in order to provide a more detailed picture of the choices made by RRM users.

6.2 Choice of variety

The nature of communication on RRM (see chapter 4), being extremely contingent and generally occurring between two individuals, combined with the general light tone of conversations, resulted in a style closely reminiscent of oral communication (a characteristic common to synchronous CMC modes, see Herring 2002). Because of this reason, tokens for some of the diagnostic features investigated occurred rarely in the sample. In particular, perfective verbs or verbs inflected for plural persons (both present and perfective) were relatively rare. Similarly, very little comparative or superlative adjectives occurred in the sample. This resulted in extremely low token counts for certain important features like the usage of Romanian comparative/superlative marker maj (Mitrovica Gurbet)\(^1\) versus Slavic po or inherited -eder (Arli varieties), the takeover of -en as a 2PL perfective marker (shared between Mitrovica Gurbet and Arli varieties) or the usage of relative pronouns and complementizers. However, the number of features encountered and their token count allow for clearly identifying which variety is used on RRM chat-room.

6.2.1 Fully shared features

Table 6.2 lists the features fully shared by Mitrovica Gurbet and the Arli varieties.

<table>
<thead>
<tr>
<th>Fully shared features</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2SG perfective marker -an</td>
</tr>
<tr>
<td>2</td>
<td>analytic future tense marking using ka</td>
</tr>
<tr>
<td>3</td>
<td>adjectival inflection of 3SG perfective (gelo ‘he went’, geli ‘she went’)</td>
</tr>
<tr>
<td>4</td>
<td>loss of word final s</td>
</tr>
<tr>
<td>5</td>
<td>short 1/2SG possessives (mo, čo/to) alternating with full forms</td>
</tr>
<tr>
<td>6</td>
<td>retention of distinction -un(d)˘r- vs -ir- in 1/2SG possessive markers</td>
</tr>
<tr>
<td>7</td>
<td>2PL perfective marker -en</td>
</tr>
</tbody>
</table>

Table 6.2: Features shared by Mitrovica Gurbet and all Arli varieties. Note that 1 to 3 are encountered both in the Vlax and Balkan diffusion spaces. 4 to 6 are development common to various dialects of former-Yugoslavia irrespectively of their group belonging. 7 is the only innovation exclusive to the Balkan diffusion space.

\(^1\) The usage of maj in Mitrovica Gurbet is not a case of code-switching (Leggio 2011a: 72). As in all dialects originating in the Vlax diffusion space the particle has been permanently borrowed from Romanian (Matras 2002: 78).
The shared feature that occurred most frequently was the 2SG concord marker in the present tense copula and in perfective lexical verbs. All of the tokens found in the sample (44 2SG present copula, 19 2SG perfective lexical verbs) employed -an (Example 6.1. In all examples the spelling employed by RRM users has been maintained while phrases split between more than one messages have been reconstructed). This form is found in dialects spoken South East of the so-called Great Divide (Matras 2005) separating south-eastern from north-western Europe.

<table>
<thead>
<tr>
<th>a)</th>
<th>sar</th>
<th>sijan</th>
<th>mange</th>
<th>tu</th>
</tr>
</thead>
<tbody>
<tr>
<td>how</td>
<td>be.2SG</td>
<td>1SG.DAT</td>
<td>2SG</td>
<td></td>
</tr>
<tr>
<td>b)</td>
<td>tu</td>
<td>pendan</td>
<td>kaj</td>
<td>san</td>
</tr>
<tr>
<td>2SG said.2SG</td>
<td>COMPL</td>
<td>be.2SG</td>
<td>the.M</td>
<td>enzo</td>
</tr>
<tr>
<td>c)</td>
<td>kaj</td>
<td>san</td>
<td>korije</td>
<td>araklan</td>
</tr>
<tr>
<td>where</td>
<td>be.2SG</td>
<td>blind.VOC.F</td>
<td>found.2SG</td>
<td>your.OBL</td>
</tr>
</tbody>
</table>

How are you? You said that you are Enzo, right Emra?

Example 6.1: 2SG copula and perfective marker

Another frequent feature, common to all the dialects of south-eastern Europe, was the analytic formation of the future tense by preposing a particle derived from the verb ‘to want’ or ‘to like’ to the verb in the present tense. Both Mitrovica Gurbet and the Arli varieties from RMS use the particle ka from kam- ‘to want’. All the 18 instances of verbs in the future tense found in the sample also used ka (Example 6.2).

<table>
<thead>
<tr>
<th>a)</th>
<th>vi</th>
<th>me</th>
<th>ka</th>
<th>hav</th>
<th>mangro</th>
</tr>
</thead>
<tbody>
<tr>
<td>too</td>
<td>1SG.FUT</td>
<td>eat.1SG</td>
<td>bread</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b)</td>
<td>ka</td>
<td>cerv</td>
<td>mance</td>
<td>aver</td>
<td>msn</td>
</tr>
<tr>
<td>FUT</td>
<td>do.1SG</td>
<td>1SG.DAT</td>
<td>other</td>
<td>msn</td>
<td></td>
</tr>
</tbody>
</table>

I’m going to have food as well I’ll get myself another IM account

Example 6.2: analytic future tense marking

Finally, a further feature common across all dialects located south-east of the Great Divide, the retention of adjectival inflection in the perfective third persons of intransitive verbs (gelo went.3SG.M, geli went.3SG.F), was found in the sample. However, it is one of the features that occurred extremely rarely (see 6.1) and a single token (example 6.3) was found. By contrast, no form showing assimilation into the inflection of transitive verbs (geljas ‘he/she went’) was found.

---

2 Note in 6.1.b an attempt at identifying the person behind a user-name (see 4.3.2) and in 6.1.c a case of teasing (see footnote 1 in chapter 4).

3 Note 6.2.b is part of an exchange aiming at obtaining the user’s IM contacts (see 4.3.2).
The occurrence of the above features, particularly the first two given their high token count, points to the absence of Romani speakers from north-western Europe among RRM users, but still does not rule out the presence of speakers from the entire south-eastern European region. However, four more shared features, whose co-occurrence is limited to the area of former Yugoslavia, were encountered in the sample. They are the dropping of word final -s (2.SG.PRES marker -as > -a; ACC.M markers -es/-os > -e/-o), the alternation between long and short forms of the 1SG and 2SG possessive pronouns (respectively 1SG mungro~mo, 2SG čiro~čo in Mitrovica Gurbet and 1SG munro~mo, 2SG tiro/čiro~to/čo in Arli varieties), the retention of Early Romani distinction between 1SG (-un(d)r-) and 2SG (-ir-) possessive markers and the take over of the 2PL.PRES marker -en as 2PL marker in perfective verbs and in the copula.

The loss of word final -s, being a phonological feature, can only be inferred by the spelling employed by users, therefore not spelling -s when expected (mostly in 2.SG.PRES verbs, ACC.M nouns and 3SG pronouns) was taken as an indicator of the presence of this feature. In all the 17 instances where -s was expected none of the users spelled it (example 6.4, see also 6.5.a).

With the exception of a single token of a 1SG long possessive pronoun (6.5.a), all 1SG and 2SG possessives occurred in their short forms (39 in total, example 6.5).

As I have observed among the Mitrovica Roma in Palermo (Leggio 2011a), the short forms of the possessives are the preferred choice. Long forms are generally used for emphatic purposes or when the possessed noun is omitted (exactly as in 6.5.a). This general preference for the short forms is very likely reinforced by the need for fast typing.
typical of synchronous CMC modes of communication (Herring 2002), thus explaining the single occurrence of a long possessive pronoun. 6.5.a, however, shows the user was replicating one of the possible pronunciations attested in Mitrovica Gurbet of the Early Romani cluster -n(d)ř- (see 6.3.9 for a discussion of spelling and the problem posed by this particular cluster). This attempt clearly shows the continuation of the distinction between Early Romani 1SG possessive marker -un(d)ř- and the 2SG possessive marker -ir-.

6.5.a, however, shows the user was replicating one of the possible pronunciations attested in Mitrovica Gurbet of the Early Romani cluster -n(d)ř- (see 6.3.9 for a discussion of spelling and the problem posed by this particular cluster). This attempt clearly shows the continuation of the distinction between Early Romani 1SG possessive marker -un(d)ř- and the 2SG possessive marker -ir-.

<table>
<thead>
<tr>
<th>a)</th>
<th>b)</th>
<th>c)</th>
<th>d)</th>
<th>e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTER want.2SG my.M</td>
<td>NEG give.1SG my.M</td>
<td>Listen to my song</td>
<td>my.PL eyes</td>
<td>gone.out.3SG your.M</td>
</tr>
<tr>
<td>Do you want mine [IM address]?</td>
<td>I don’t give my IM address to anyone</td>
<td>Listen to my song</td>
<td>blinded.PL</td>
<td>head</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>My eyes have gone blind</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>You’re out of your mind</td>
</tr>
</tbody>
</table>

**Example 6.5:** 1/2SG possessive pronouns

Finally, as a further confirmation that all users on RRM come from the Southern Balkans, two tokens of the only innovation spreading from this area, namely the take-over of the 2PL present tense marker -en as the 2PL perfective and copula marker, were found (example 6.6). As it was the case for the adjectival inflection of 3SG.PERF intransitive verbs, no tokens of forms from other areas were found in the sample.

<table>
<thead>
<tr>
<th>a)</th>
<th>b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tumen o mursa sien hovavne</td>
<td>lace avilen</td>
</tr>
<tr>
<td>2PL the people.PL be.2PL liar.PL</td>
<td>well came.2PL</td>
</tr>
<tr>
<td>You people are liars</td>
<td>Welcome (you all)</td>
</tr>
</tbody>
</table>

**Example 6.6:** 2PL copula and perfective marker

This overview of the features fully shared between Mitrovica Gurbet (Vlax variety) and all the Arli (Balkan) varieties from RMS confirms the impression I had during fieldwork that no Romani speakers from outside former Yugoslavia participate on RRM interactions.

---

5 Note again the instances of teasing (6.5.e, 6.5.d, 6.5.g), parts of conversations about IM contacts (6.5a, 6.5b, 6.5h), comments on user’s font colour (6.5.i) and also talks about off-line life (6.5.f).

6 In translating 6.6.a, based on the interaction from which it was taken, hovavne has been considered a typo for hohavne ‘liars’. 6.6.b was posted several times by the DJ who copied and pasted RRM welcome message (see figure 3.1) as users logged-in.

---

134
6.2.2 Contrasting features

It is now possible, by analyzing the features differentiating Mitrovica Gurbet from the Arli varieties (Table 6.3), to identify which of the two varieties spoken by Mitrovica Roma is used on RRM chat-room. In some cases only tokens of Gurbet forms were found. I will first discuss these and then move onto the features for which both Mitrovica Gurbet and Arli tokens were found.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Gurbet forms</th>
<th>Arli forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>future and subjunctive copula</td>
<td>av- ‘to come’</td>
<td>ALL: ov- ‘to become’</td>
</tr>
<tr>
<td>indefinites</td>
<td>person khon(i), thing khanc(i)</td>
<td>ALL: person nik(0), thing niš(t)0</td>
</tr>
<tr>
<td>prothesis in 3 person pronouns</td>
<td>v-</td>
<td>ALL: no</td>
</tr>
<tr>
<td>Early Romani -ndi-</td>
<td>continued as -ngr-, -nrr- or -nr-</td>
<td>ALL: continued as -r- or -nr-</td>
</tr>
<tr>
<td>copula</td>
<td>1/2/3 persons s-</td>
<td>MK-002/MK-003: 1/2 persons Øinj-, 3 persons (i)tan-, MK-004/MK-012: 1/2 persons sij-, 3 persons s- MK-005: 1/2 persons Øij-, 3 persons Ø- YU-011/YU-016: 1/2 persons hij-, 3 persons tan- YU-014: 1/2 persons sij-, 3SG s-, 3PL sin-</td>
</tr>
<tr>
<td>negators</td>
<td>indicative ni, copula naj</td>
<td>ALL: indicative na, copula na(ne)</td>
</tr>
<tr>
<td>umlaut in 1SG copula and 1SG.PERF markers</td>
<td>yes (*jom &gt; -em)</td>
<td>ALL: no (-um)</td>
</tr>
</tbody>
</table>

Table 6.3: Features differentiating Mitrovica Gurbet from the Arli varieties

For the quantifier interrogative only the Mitrovica Gurbet form sode was found in the sample (9 tokens, example 6.7).

<table>
<thead>
<tr>
<th>ta</th>
<th>cane.Ø</th>
<th>sode</th>
<th>si</th>
<th>mandar</th>
<th>o Basel and know.2SG how.much is 1SG.ABL the Basel</th>
</tr>
</thead>
<tbody>
<tr>
<td>and do you know how far Basel is from me?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Example 6.7: quantifier interrogative

<table>
<thead>
<tr>
<th>Ex</th>
<th>sode</th>
<th>bes</th>
<th>si</th>
<th>tut</th>
</tr>
</thead>
<tbody>
<tr>
<td>how.much year is 2SG.OBL</td>
<td>How old are you?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

135
Also in the case of the future and subjunctive copula, only the Mitrovica Gurbet form *av- ‘to come’ was found (7 tokens, example 6.8) while no Arli *ov- ‘to become’ were encountered.

\[
\begin{array}{llllllllll}
\text{a)} & \text{valahi} & \text{te} & \text{avol} & \text{lestar} & \text{konzerti} & \text{ani} & \text{pusti} & \text{Franzuska} \\
\text{INTER} & \text{COMP} & \text{be.SUBJ.3SG} & \text{3SG.ABL} & \text{concert} & \text{in.F} & \text{bleak.F} & \text{France} \\
\text{Good God, if there was a concert by him in bleak France} \\
\text{pa te avol i katra 400 euro} \\
\text{but COMP be.SUBJ.3SG the.F ticket.F 400 euros} \\
\text{but the ticket will be 400 euros} \\
\text{b)} & \text{av} & \text{sasto} & \text{comp} & \text{te} & \text{avol} & \text{len} & \text{zurale} & \text{kokala} \\
\text{be.FUT} & \text{health.M} & \text{COMP} & \text{COMP} & \text{be.SUBJ.3SG} & \text{3PL.OBL} & \text{strong.PL} & \text{bones} \\
\text{Be well!} & \text{So that they will have strong bones} \\
\end{array}
\]

**Example 6.8:** Future and subjunctive copula

For the indefinite pronouns again no Arli forms *niko(j)~neko(j)/džiko, person indefinite and *ništo–nešto/diso, thing indefinite were found. All the six indefinite pronouns encountered matched the forms used in Mitrovica Gurbet, *khoni(k), OBL khanikas– person indefinite and *khanč(i) thing indefinite (example 6.9 and 6.5.b for a person indefinite in the dative case).

\[
\begin{array}{llllllllll}
\text{a)} & \text{ni} & \text{asunav} & \text{kanc} & \text{NEG} & \text{hear.1SG} & \text{nothing} & \\
\text{I hear nothing} \\
\text{b)} & \text{saj} & \text{pucav} & \text{tut} & \text{kaci} & \text{can} & \text{ask.1SG} & \text{2SG.OBL} & \text{something} & \text{Can I ask you something?} \\
\text{c)} & \text{ka} & \text{astren} & \text{la} & \text{koni} & \text{FUT} & \text{find.3PL} & \text{3SG.F.OBL} & \text{someone} & \text{Someone will find her} \\
\end{array}
\]

**Example 6.9:** Indefinite pronouns

Another feature for which Gurbet forms only were found was the insertion of prothetic sounds before third person nominative pronouns. In Mitrovica Gurbet *v- is the prothetic sound found, while the Arli varieties do not show prothesis. As Romani is a pro-drop language only three pronouns occurred in the sample (example 6.10).

\[
\begin{array}{llllllllll}
\text{a)} & \text{nonaci} & \text{si vaj} & \text{nona.GEN.F} & \text{is she} & \text{She is Nona’s [daughter]} \\
\text{b)} & \text{miss de Kosovo 2000 si vaj} & \text{miss of Kosovo 2000 is she} & \text{She is Miss Kosovo 2000} \\
\text{c)} & \text{se vov cerda matematika but besa} & \text{but if he did.3SG maths many years} & \text{But if he did the maths for so many years} \\
\end{array}
\]

**Example 6.10:** Prothesis in 3rd person pronouns
Finally, again with three tokens only, the continuation of the Early Romani cluster -ndř- only appeared as -ngr- (see *mangro* ‘bread, food’ in 6.2.a and *mungro* ‘my.M’ in 6.5.a) which is not found in any of the Arli varieties.

The first feature showing both Gurbet and Arli tokens is the form of the 1st and 2nd persons present tense copula. Romani dialects show variation in the selection of the copula root (*s*~*h*~) and in the insertion of a perfective stem (*-inj*~*no insertion*) between the root and the person concord markers (Matras 2002: 230). In Mitrovica Gurbet the copula is *s*- for all persons, showing selection of *s*- and no insertion. The Arli varieties show variation in the selection of the root: MK-002, MK-003, MK-005 have Ø-; YU-011 and YU-016 have *h*-; MK-004, MK-012 and YU-014 have *s*-. Regarding the perfective stem MK-002 and MK-003 have *-inj*-, while all the others have *-ij*-.

Furthermore, in all the Arli varieties 1st and 2nd persons show insertion, but 3rd persons do not or are replaced by presentative forms (in MK-012 3PL shows insertion)\(^7\).

On the RRM sample all tokens had *s*- as root. Furthermore all the 37 tokens of 3SG copula (no 3PL were found) encountered in the sample were spelled as *si*, suggesting that the Arli variety used on RRM behaves like MK-004, MK-012 and YU-014 in its treatment of the singular copula.

---

\(^7\) Note that all the Arli varieties follow the markedness hierarchy proposed by Elšik & Matras (2006), in that 3rd persons forms are the first to undergo changes.
the copula are more frequent (see graph 6.1, example 6.11, for Gurbet and Arli 2SG copula see example 6.1, for Arli 2PL see 6.6.a).

<table>
<thead>
<tr>
<th>a)</th>
<th>b)</th>
<th>c)</th>
</tr>
</thead>
<tbody>
<tr>
<td>but</td>
<td>super</td>
<td>sar</td>
</tr>
<tr>
<td>purani</td>
<td>semi</td>
<td>sen</td>
</tr>
<tr>
<td>tuche</td>
<td>tuche</td>
<td>tuche</td>
</tr>
<tr>
<td>too old.F</td>
<td>super be.1SG</td>
<td>how be.2PL</td>
</tr>
<tr>
<td>I’m too old for you</td>
<td>I’m super</td>
<td>How are you (all)?</td>
</tr>
</tbody>
</table>

**Example 6.11:** Gurbet and Arli forms of the copula

A further feature distinguishing the Mitrovica Gurbet from the Arli varieties from RMS is the form of the negators: indicative *ni* and copula negator *naj* in Mitrovica Gurbet; indicative *na* and copula negator *na(ne)* in the Balkan ones. As it was the case with the previous feature, the Gurbet forms were more frequent (graph 6.2).

**Graph 6.2:** Indicative (*ni* Mitrovica Gurbet, *na* Arli Varieties) and copula (*naj* Mitrovica Gurbet, *na(ne)* Arli varieties) negators.

All the 8 instances of the Balkan copula negator occurred in the short form *na* (see examples 6.12.e and 6.5.h), probably, as it was the case for possessive pronouns, due to users’ need to speed up typing.

Note that both in Mitrovica Gurbet (6.12.a, 6.12.b) and in the Arli tokens from RRM (6.12.e), the negated 3SG copula is expressed by the negator alone. All Arli varieties from RMS also show this characteristic.
Another feature distinguishing Mitrovica Gurbet from the Arli varieties is the shape of the 1SG concord marker in perfective verbs and in the copula. In Mitrovica Gurbet historical jotation at morpheme boundary leads to umlaut (*-jom > -em) while in the Arli varieties, there is no umlaut (-um). Again Mitrovica Gurbet forms constitute the majority of tokens found in the sample (graph 6.3 and example 6.13 for lexical verbs, 6.11.a and 6.11.b for the copula).

Note that in example 6.13.c, as g is used on RRM to represent/g/ and /dž/ (see 6.3.2 for notes on spelling), the perfective markers used with verbal roots in dental sonorants can match either the form used in MK-005, MK-012 and YU-011 (-dž-) or in MK-002 (-gj-). Variation in the shape of this perfective marker is due to jotation of the historical marker -d-. In the Arli varieties the jotated perfective marker -dž- generally undergoes affrication.
to *-dž/-dz* eventually leading to de-affrication (as in MK-004 where *-dž/-dz* > *-z*), substitution through a jotated velar *-dj* > *-gj* or palatalization to *-d’* (as in YU-014/YU-016). In Mitrovica Gurbet, the yod is dropped giving *-d*-, (note that in this respect MK-003 behaves like Mitrovica Gurbet).

Due to the overall low occurrence of perfective verbs in the sample, only 9 tokens were found. Of these, *civgum* ‘put.PERF.1SG’ above was the only Arli form. The remaining tokens all showed *-d*- (graph 6.4, example 6.14, see also 6.1.b and 6.10.c).

**Graph 6.4: Perfective markers**

<table>
<thead>
<tr>
<th>a)</th>
<th>se</th>
<th>me</th>
<th>bicaldem</th>
<th>tuce</th>
<th>invitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>but/if</td>
<td>1SG</td>
<td>sent.1SG</td>
<td>2SG.DAT</td>
<td>invitation</td>
<td></td>
</tr>
<tr>
<td>But if I sent you an invitation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b)</th>
<th>so</th>
<th>cerdam</th>
</tr>
</thead>
<tbody>
<tr>
<td>what</td>
<td>did.1PL</td>
<td></td>
</tr>
<tr>
<td>What have we done?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Example 6.14:** Perfective marker *-d*-

The last feature clearly showing a higher frequency of Gurbet forms is the prothesis of *a*- (graph 6.5). This phenomenon is common to all Romani dialects, but in Mitrovica Gurbet (as in all Vlax varieties) it is extensive and targets words like *šun* > *ašun* ‘to listen/to hear’ and *lav* > *alav* ‘word/name’, which are not affected in other varieties (example 6.15, also 6.5.c and 6.9.a).
Contrary to what had seemed from the analysis of quantifier interrogatives, the future and subjunctive copula, indefinite pronouns, prothesis in 3rd person pronouns and the continuation of Early Romani -ngř-, the last few features (1st and 2nd persons of the present tense copula, negators, umlaut in 1SG copula and 1SG.PERF markers, prothesis of a-) clearly show that both Mitrovica Gurbet and an Arli variety are employed on RRM chat-room. Mitrovica Gurbet is overall used thrice as frequently as the Arli variety, however, some users, as I will more clearly show in 6.2.4, employ both Gurbet and Arli forms. This suggests that there is some degree of convergence among Gurbet and Arli users of RRM. This convergence might be due to off-line influences between speakers, as the partial overlap between Mitrovica Gurbet and some of the Arli varieties from RMS (MK-004/MK-012/YU-014 for 3SG.PRES copula, MK-003 for perfective marker -d-) seems to suggest. This impression is further reinforced by the set of features partially shared between Mitrovica Gurbet and the Arli varieties I am going to discuss in the next session.

6.2.3 Partially shared features

As I mentioned in 6.1, speakers from different diffusion spaces that have come in contact because of migration might acquire features from other dialects. Indeed, the takeover of the 2PL present tense marker -en as the 2PL perfective and copula marker is a feature Mitrovica Gurbet acquired as it moved into the Balkan diffusion space (see 6.2.1
and Leggio 2011a:111). Similarly, a number of Balkan dialects have been shown to be prone to Vlax influence following contact with Vlax migrants (Matras 2002: 236). More recently, process of convergence in face-to-face communication among Arli and Gurbet speakers have been documented in Kosovo (Boretzky 1986, 1995).

<table>
<thead>
<tr>
<th>Feature</th>
<th>Mitrovica Gurbet</th>
<th>Arli varieties</th>
</tr>
</thead>
</table>
| palatalization of alveolars in selected lexemes | extensive \((\text{buti} > \text{bući} \text{‘work’}, \text{tiro} > \text{čiro} \text{‘yours.M’})\) | MK-002/MK-003/MK-005: no  
MK-004/MK-012: \(\text{bući} > \text{bući} \text{‘work’}\)  
YU-011/YU-014/YU-016: \(\text{buti} \sim \text{bući} \text{‘work’}\) |
| palatalization of velars before front vowel | extensive \(\text{mandže} < \text{mange ‘me.DAT’, } \text{džili} < \text{gili ‘song’}, \text{čer-} < \text{ker-} \text{‘to do}, \text{čin-} < \text{kin-} \text{‘to cut’, } \text{čher} < \text{kher ‘house’}\) | MK-002/MK-003: no  
MK-004/MK-012: partial \(\text{mange ‘me.DAT’, } \text{džili} < \text{gili ‘song’}, \text{čer-} < \text{ker-} \text{‘to do}, \text{čin-} < \text{kin-} \text{‘to cut’, } \text{čher} < \text{kher ‘house’}\)  
YU-011/YU-014/YU-016: extensive \(\text{mandže} < \text{mange ‘me.DAT’, } \text{džili} < \text{gili ‘song’}, \text{čer-} < \text{ker-} \text{‘to do}, \text{čin-} < \text{kin-} \text{‘to cut’, } \text{čher} < \text{kher ‘house’}\) |
| forms of 2SG.DAT                               | \(če \sim \text{tuče}\) | MK-002/MK-003/MK-004/MK-012/YU-014: \(\text{take only}\)  
MK-005/YU-011/YU-016: \(če \sim \text{tuče}\) |
| masculine singular dative marker              | \(-e\) only                      | MK-002/YU-014: \(-e\)  
MK-003/MK-004/MK-012: \(-e\)  
MK-005/YU-011: \(-e\)  
YU-016: \(-e\) only |
| singular genitive marker                     | masculine \(-\text{Ø-}\)  
feminine \(-\text{e-}\) | MK-002: feminine \(-\text{kVr-}\)  
masculine \(-\text{Vr-}\)  
MK-003: \(-\text{kVr-} \sim \text{Vr-}\)  
MK-004/MK-012: \(-\text{kVr-}\)  
MK-005: masculine \(-\text{kVr-} \sim \text{Vr-}\)  
feminine \(-\text{kVr-} \sim \text{Vr-}\)  
YU-011: masculine \(-\text{k} \sim \text{r} \sim \text{Ø-}\)  
feminine \(-\text{kVr-} \sim \text{k} \sim \text{e-}\)  
YU-014: masculine \(-\text{Ø-}\)  
feminine \(-\text{kVr} \sim \text{Ø-}\)  
YU-016: masculine \(-\text{Ø-}\)  
feminine \(-\text{k} \sim \text{e-}\) |
| loan verb adaptation markers                 | \(-\text{isar-} \text{and -isav-}, \text{optionally reduced to -i-}\) | MK-002/YU-016: \(-\text{in-}, \text{optionally reduced to -i-}\)  
MK-003/MK-005/YU-011/YU-014: \(-\text{in-}\)  
MK-004/MK-012: \(-\text{iz-}\) |
| 3SG present tense marker                     | \(\text{-ol}\)                    | MK-002/MK-004/MK-005: \(-\text{el only}\)  
MK-003/MK-012/YU-011/YU-014/YU-016: \(-\text{el} \sim \text{-ol}\) |
| demonstratives                               | \((\text{V})\text{k-}\)           | MK-002: \(-\text{k-}/\text{d-}\)  
MK-003/MK-005: \((\text{V})\text{k-}/(\text{V})\text{d-}\)  
MK-004: \(\text{kVd}/\text{kV}\)  
MK-012: \(\text{(V)k-}/\text{kVd}/\text{kV}\)  
YU-011: \((\text{V})\text{k-}/\text{kVd}/(\text{V})\text{gVd-}\)  
YU-014: \((\text{V})\text{k-}/\text{d-}\)  
YU-016: \((\text{V})\text{k-}/\text{kVd}/\text{kV}\) |

Table 6.4: Features partially shared between Mitrovica Gurbet and Arli varieties
The partially shared features I am going to discuss now can be regarded as cases of such convergence or of common developments Mitrovica Gurbet and the Arli varieties from RMS are undergoing together.

The first feature for which some Arli varieties behave like Mitrovica Gurbet is the palatalization of alveolars in selected lexemes (i.e.: buti > buči ‘work/thing’, tiro > čiro ‘yours’, dive > džive ‘day’), a Vlax innovation adopted, to some extent, in some Balkan dialects. Mitrovica Gurbet, as a Vlax variety, shows palatalization in all possible lexemes. In YU-011, YU-014 and YU-016 the sound change is apparently ongoing as they show alternation between buti and buči ‘work/thing’ but not in other lexemes. Similarly in MK-004 and MK-005 only buti is affected and even shows affrication to bući. MK-002, MK-003 and MK-005 on the other hand do not participate at all in this sound change.

In the sample no instances of dive/džive ‘day’ or other words containing /d/ were encountered. Similarly, no cases of Balkan tikno ‘small’ versus Vlax cikno (in which palatalization also leads to affrication) were found. However a total of 24 items for buti–buči ‘work’ and to–čo ‘2SG.POSS’ were found. In such cases, spelling c, ci or ch for /č/ rather than t for /t/ was taken as an indicator of the presence of this feature (see 6.3.2 for notes on spelling).

Only in 3 cases, all of them 2SG.POSS pronouns (see 6.5.e to 6.5.i) t was used, while all tokens for buti–buči showed palatalization (graph 6.6, example 6.16).

Graph 6.6: Palatalization of alveolars in buti–buči ‘work’ and to–čo ‘2SG.POSS’
Example 6.16: Palatalization in buti

A similar feature, due however to contact with South Slavic languages (Matras 2002: 49), is the palatalization of velar stops before front vowels (g+e/i > dž+e/i, k+e/i > č+e/i, kh+e/i > čh+e/i).

In Mitrovica Gurbet, MK-005 and YU-016 the phenomenon is extensive and words like mandže < mange ‘me.DAT/want.2SG’, džili < gili ‘song’, ěr- < ker- ‘to do’, čin- < kin- ‘to cut’, čher < kher ‘house’ are systematically affected. Among the other Arli varieties MK-002 and MK-003 do not participate in this sound change. In YU-011 and YU-014 /k/ and /kh/ are less prone to undergo palatalization and they show mandže < mange ‘me.DAT/want.2SG’ and džili < gili ‘song’ but čer- ~ ker- ‘to do’, čin- ~ kin- ‘to cut’, čher ~ kher ‘house’. MK-004 and MK-012 instead do not show palatalization with /e/ (mange ‘me.DAT’, ker- ‘to do’, kher ‘house’) but with /i/ palatalization occurs and, as it was the case for alveolars, also leads to affrication (zi < gili ‘song’, cin- < kin- ‘to cut’).

As in the case of palatalization of alveolars, spelling was taken as an indicator of the presence of this feature. Spellings like gjeli in example 6.3 clearly show an attempt by the user to represent a palatalized sound. Similarly, using c, ci or ch to represent /č/ instead of k for /k/ serves the same function. Although some potential tokens involving g > dž had to be excluded due to unclear spelling (see 6.3.2 regarding the problems posed by the representation of these two sounds), the results clearly indicates that the majority of velar stops are palatalized (graph 6.7).

Example 6.17: Un-palatalized velars

Words not showing palatalization are instances of khel- ‘to dance/play’, tuke 2SG.DAT pronoun and adžiker- ‘to wait (example 6.17; for tokens of palatalized /k/ see 6.2.b, 6.10.a,
6.10.c, 6.14 and 6.16.b; for tokens of palatalized /g/ see 6.2.b, 6.3, 6.5.a and 6.5.i). The Arli variety used on RRM thus seem to pattern with YU-011 and YU-014 in that /k/ and /kh/ appear to resist the sound change.

Graph 6.7: Palatalization of velar stops before front vowels mandže < mange ‘me.DAT’, džili < gili ‘song’, čer- < ker- ‘to do’, čin- < kin- ‘to cut’, čher < kher ‘house’

MK-005, YU-011 and YU-016 also share the alternation between short and long forms of the palatalized 2SG.DAT pronoun (tuče~če). This feature is also present in Mitrovica Gurbet. In all of them the short form occurs post verbally as a semi-clitic, exactly as it was found in the sample (graph 6.8 and example 6.18, for long forms see also 6.14.a) where the short forms were often typed in a single word with the verb.

Graph 6.8: Short and long forms of 2SG.DAT
18.c also shows a further feature shared by Mitrovica Gurbet and some of the Arli ones (MK-003, MK-005, YU-011, YU-014 and YU-016): the dropping of k in the masculine dative marker (-ke > -e). The only token of a lexical item showing this feature is shown in example 6.19. The remaining 10 tokens found in the sample were instances of the reason/goal interrogative pronoun sose < soske ‘what for/why’ from so ‘what’ (as in 6.18.c, no tokens of soske where found).

A further instance of case markers reduction involves the genitive. In Mitrovica Gurbet only the short marker -k- is used. Furthermore, it is reduced to -Ø- in masculine nouns and it is palatalized before front vowels in feminine nouns. In the Arli varieties the long marker -kvr- alternates with short -r- or -k-. YU-014 and YU-016, however, do not have long -kvr- and -k- undergoes the same processes of reduction and palatalization as in Mitrovica Gurbet (graph 6.9; example 6.20 and 6.10.a)
In 6.20.b it is also possible to see the treatment of borrowed verbs (kostol ‘it costs’ from German kosten). Loan verbs are adapted in Romani by the usage of markers derived from the Greek present tense inflectional endings (-in-, -is-, -iz-) or inherited transitivity and intransitivity markers (-ker-, -ar-, -al-) or combinations of them (Matras 2002: 128-129). In Mitrovica Gurbet the markers, common among Vlax varieties, are -isar- for transitive and -isav- for intransitive. The Arli varieties instead have -iz- or -in- (both for transitive and intransitive depending on the variety), and -isal-/isaj- (for intransitives in MK-004, MK-005, MK-012 and YU-011). MK-002 and YU-016, like Mitrovica Gurbet, also have a contracted form -i- occurring with transitive verbs. Both -in- and -i- were found in the sample (example 6.21), while there were no instances of -isar- or -isal- (graph 6.10).

Interestingly, a -Ø- marker also emerged with recently borrowed intransitive and medio-passive verbs (examples 6.21.c from French se bloquer ‘to crash’ and 6.20.b from German kosten ‘to cost’). It would seem that, as observed by Matras (2002: chapter 6), in the varieties used on RRM -ol (derived from the contraction of the inherited intransitive marker and the 3SG.PRES marker -jov-el > -(j)ol) has become an independent inflection class to which inchoatives, intransitives and recently borrowed verbs are assigned.
Furthermore -ol, in this case unrelated to the intransitive marker -(j)ov-, is the 3SG present concord marker for consonantal verbs in Mitrovica Gurbet (Leggio 2011a: 86). In MK-003, MK-012, YU-011, YU-014 and YU-016 the 3SG present concord marker shows alternation between -el and -ol in consonantal verbs. This latter feature occurred prominently in my sample, with all 3SG present consonantal verbs taking -ol (example 6.22; see also 6.8.a, 6.8.c, 6.12.b and 6.16.b).

The form of demonstratives is the last feature partially shared between Mitrovica Gurbet and the Arli varieties encountered in the sample. In Mitrovica Gurbet all demonstratives are based on the root k-. The distinction between specific and non-specific demonstratives typical of all Romani dialects (Matras 2000:103) is achieved in Mitrovica Gurbet by the reduplication of the carrier vowels, -a- for situational referents and -o- for discourse referents. The Mitrovica Gurbet four-way system of demonstratives is therefore kava ‘this.M’, kova ‘that.M’, akava ‘this.one.M’, okova ‘that.one.M’. However, I noted how among the younger generations the specific forms with reduplicated vowel seems to be disappearing and the four-way system is undergoing a simplification into a two-way

---

**Example 6.22:** 3SG present tense concord marker

Furthermore -ol, in this case unrelated to the intransitive marker -(j)ov-, is the 3SG present concord marker for consonantal verbs in Mitrovica Gurbet (Leggio 2011a: 86). In MK-003, MK-012, YU-011, YU-014 and YU-016 the 3SG present concord marker shows alternation between -el and -ol in consonantal verbs. This latter feature occurred prominently in my sample, with all 3SG present consonantal verbs taking -ol (example 6.22; see also 6.8.a, 6.8.c, 6.12.b and 6.16.b).

The form of demonstratives is the last feature partially shared between Mitrovica Gurbet and the Arli varieties encountered in the sample. In Mitrovica Gurbet all demonstratives are based on the root k-. The distinction between specific and non-specific demonstratives typical of all Romani dialects (Matras 2000:103) is achieved in Mitrovica Gurbet by the reduplication of the carrier vowels, -a- for situational referents and -o- for discourse referents. The Mitrovica Gurbet four-way system of demonstratives is therefore kava ‘this.M’, kova ‘that.M’, akava ‘this.one.M’, okova ‘that.one.M’. However, I noted how among the younger generations the specific forms with reduplicated vowel seems to be disappearing and the four-way system is undergoing a simplification into a two-way

---

**Note in 6.22.a a call to the DJ to sanction an abusive user and in 6.22.c a case of self policing (see 4.2.1 and 4.3.5).**
system (Leggio 2011a: 77). The Arli varieties also have forms with \textit{k}- root and reduplicated carrier vowels which are used for non-specific referents. For specific referents they have forms in \textit{Vd}-, \textit{kVd}-, \textit{kVk}- or \textit{gVd}- depending on the dialect. However, MK-003, MK-012, YU-011, YU-014 and YU-016 also have \textit{k}- demonstratives without reduplicated carrier vowels alternating with those with reduplicated carrier vowel. As shown in graph 6.11, \textit{k}- demonstratives without reduplicated vowel constituted the majority of tokens found in the sample.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.11.png}
\caption{Graph 6.11: Forms of demonstrative pronouns}
\end{figure}

None of the other roots attested in the Arli varieties from RMS were found on the sample, but 4 tokens of root \textit{g}- were found (example 6.23.d and 6.23.e). Root \textit{g}-, although unattested in the Arli varieties from RMS, is found in MK-001, a Gurbet variety from Skopje, Macedonia. Unfortunately, it was impossible to establish where the users who employed such forms were from.

<table>
<thead>
<tr>
<th>a)</th>
<th>\textit{mo}</th>
<th>\textit{pral}</th>
<th>\textit{si}</th>
<th>\textit{kova}</th>
</tr>
</thead>
<tbody>
<tr>
<td>my,M</td>
<td>brother is that,M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>That is my brother</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>b)</th>
<th>\textit{akava}</th>
<th>\textit{radio}</th>
<th>\textit{isto}</th>
<th>\textit{sur}</th>
<th>\textit{o}</th>
<th>\textit{radio roma}</th>
</tr>
</thead>
<tbody>
<tr>
<td>this.one,M</td>
<td>radio</td>
<td>same</td>
<td>like</td>
<td>the,M</td>
<td>radio roma</td>
<td></td>
</tr>
<tr>
<td>This radio here is the same as Radio Roma</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>c)</th>
<th>\textit{emrah}</th>
<th>\textit{ka}</th>
<th>\textit{pahravala}</th>
<th>\textit{po}</th>
<th>\textit{upas}</th>
<th>\textit{kola}</th>
</tr>
</thead>
<tbody>
<tr>
<td>emrah</td>
<td>FUT</td>
<td>break.1SG.3SG.F.OBL</td>
<td>in,M</td>
<td>half</td>
<td>that.F</td>
<td></td>
</tr>
<tr>
<td>Emrah will break her in two, that woman</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>d)</th>
<th>\textit{pendasa}</th>
<th>\textit{agja}</th>
<th>\textit{vi}</th>
<th>\textit{mance}</th>
</tr>
</thead>
<tbody>
<tr>
<td>told.3SG.REM</td>
<td>this.one.F</td>
<td>too</td>
<td>1SG.DAT</td>
<td></td>
</tr>
<tr>
<td>She was telling me exactly this</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>e)</th>
<th>\textit{aj}</th>
<th>\textit{sikova}</th>
<th>\textit{gia}</th>
</tr>
</thead>
<tbody>
<tr>
<td>and</td>
<td>learn.1SG.REM</td>
<td>this.F</td>
<td></td>
</tr>
<tr>
<td>And I’m learning this</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textbf{Example 6.23:} Demonstrative pronouns. See also 6.3, 6.4.a and 6.22.b
Therefore I am not in a position to claim either that the g- root is used in the Mitrovica Arli variety or that the users were in fact from Macedonia, although, considering the roots employed in the other Arli varieties the latter seems more likely.

It is important to note that I documented Mitrovica Gurbet in Palermo in 2009 while the samples for the Arli varieties on RMS were collected in former-Yugoslavia between 2003 and 2006. It should, then, be clear by the above partially shared features that processes of convergence between the dialects started before the dispersal of the Mitrovica Roma. However, concerning convergence as seen on the RRM sample, the question remains whether users have pluralistic repertoires in which both Gurbet and Arli forms are present or whether speakers with distinct dialects interact on the chat-room.

6.2.4 User consistency in choice of variants

To address the above question I have mapped the choice of features attributable to either Mitrovica Gurbet or the Arli variety used on RRM to individual users. I had to exclude from my analysis users that only posted formalized shouts, written-out sounds or whose Romani did not show any feature relevant for distinguishing Mitrovica Gurbet from the Arli variety. This left me with a sample of 46 users, nearly half of the total users I counted during my fieldwork (see 5.1).

Some users, like U15 for Arli features and U16 for Gurbet features, showed usage of a single variety, while others, like _­_EmRaH°StyLe_­_, employed both Gurbet and Arli forms (example 6.24).

<table>
<thead>
<tr>
<th>User</th>
<th>RRM message</th>
<th>Translation</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>U15</td>
<td>na sian ko to msn</td>
<td>you’re not on msn</td>
<td>Arli copula sij-</td>
</tr>
<tr>
<td></td>
<td>tumen o mursa sien hovavne KO SIAN</td>
<td>you people are liars</td>
<td>Arli perfective marker -g/–dž- and 1SG.PERF concord marker</td>
</tr>
<tr>
<td></td>
<td>civyum tu ko mo msn</td>
<td>I’ve added you to my msn</td>
<td></td>
</tr>
<tr>
<td>U16</td>
<td>iklol co sero</td>
<td>You’re out of your mind</td>
<td>Gurbet palatalized /t/ in short 2.SG.POSS</td>
</tr>
<tr>
<td></td>
<td>savi si ci kamli</td>
<td>Which is your girlfriend?</td>
<td>Gurbet copula s- and 1SG.PERF marker</td>
</tr>
<tr>
<td></td>
<td>me sem koreja</td>
<td>I am blind</td>
<td></td>
</tr>
<tr>
<td></td>
<td>me sem koja</td>
<td>I am that one</td>
<td></td>
</tr>
<tr>
<td><em>­_EmRaH°StyLe_­</em></td>
<td>NAJ MUZIK KAJ SAN TU RAKLIJE</td>
<td>there’s no music where are you girl?</td>
<td>Gurbet copula negator</td>
</tr>
<tr>
<td></td>
<td>SUNAV BRE</td>
<td>I’m listening bro</td>
<td>Gurbet palatalized /t/ in short 2.SG.POSS</td>
</tr>
<tr>
<td></td>
<td>HAV CE VUSTA ME</td>
<td>Me, I’ll eat your lips</td>
<td></td>
</tr>
</tbody>
</table>

Example6.24: Users’ choice of Romani variety
As it can be seen in graph 6.12, users that employed Gurbet forms only are the majority (39%, 18 users), shortly followed by users who have a pluralistic repertoire but employ Gurbet forms 60% of the time or more (26%, 12 users). Users employing only Arli forms constituted the third largest group (15%, 7 users). Users with an Arli dominant repertoire (Arli forms used 60% of the time or more) and with a balanced repertoire (using features of each variety 50% of the time) accounted for another 4.5% (2 users) each.

![Choice of variety](image)

**Graph 6.12:** Users consistency in choice of Romani variety

Finally, for 5 users (11%) only tokens of partially shared features were found, preventing me from establishing if they had a Gurbet, Arli or mixed repertoire.

Generally, users with a mixed repertoire were systematic in that certain features were Arli while others were Gurbet. For example _EmRaH°StyLe_, a Gurbet dominant user, systematically showed lack of a- prothesis in bijav ‘wedding’, šun- ‘to listen’ and lav ‘word/name’ (see example 6.24). Occasionally, however, some users had Arli and Gurbet tokens for the same feature. For example _FARiJA_ *, a Gurbet dominant user, alternated between Arli and Gurbet forms of the copula (see examples 4.8.d for Arli sijum ‘I.am’ and 4.8.o for Gurbet sem ‘I.am’).

It is clear from the analysis of the dialectal features that, just like in other attempts at writing Romani (see 2.3), RRM users are employing their immediate spoken varieties. Mirroring the prevalence of Gurbet speakers in the Mitrovica Roma community at large
(see 3.1), users employing only Gurbet forms largely outnumber those using only Arli forms. The number of users with mixed repertoires (35%, 16 users combining Gurbet-dominant, Arli-dominant and balanced ones) nearly equals that of Gurbet-only users indicating that processes of convergence among the two varieties started off-line, as observed in 6.2.2, continue even on RRM.

Contrary to many assumptions concerning written communication (see 2.1 and 2.2) the usage of two varieties, as the examples in chapter 4 have shown, does not hinder communication. More importantly, I was unable to find a single message containing comments about other users’ choices in my entire corpus. However, on a couple of occasions I witnessed users asking for clarification when some minor misunderstandings occurred, but even in those cases no purist attitudes or concepts about the ‘appropriate’ way of writing were put forward.

6.3 Choice of spelling

As mentioned in 2.3, various different yet compatible spellings for Romani have emerged over time, all of them based on the alphabet used for the national language spoken by the writers and showing various degrees of adjustment in order to represent Romani specific sounds.

In the case of the Mitrovica Roma, as seen through chapters 3 to 5, the major languages they are exposed to were, prior to the community dispersal, Serbo/Croatian, Albanian and marginally Qur’an Arabic plus, following the community dispersal, German, French, Italian and marginally English. Furthermore, the educational trajectories of individuals vary across generations and depend on the countries they have been and are currently living in, leading to different degrees of literacy in different languages.

The first thing to note about the spelling on RRM is that the Latin script is the only one employed. This is not surprising considering that those members of the community who achieved literacy in Serbo/Croatian and know the Cyrillic alphabet are from the older generations (40 or more years old) and do not generally interact on RRM chat-room. Furthermore, even if they do and although the chat-room supports different scripts, they only have access to western keyboards displaying the Latin script. The same applies to the Arabic script, which is however likely to be known just by a tiny educated minority of users.

The exclusive usage of the Latin script can therefore be attributed to generational and educational factors and to technical constraints. However, as the Latin alphabet is the
favoured one in all attempts at writing Romani, this choice also makes RRM spelling compatible with other Romani spellings.

The choice of the Latin script, however, poses some problems regarding the representation of those sounds that were absent in Latin and which show variation in spelling across all languages that employ the Latin script. Considering that the Mitrovica Roma are exposed to many languages that employ the Latin script, I will start my analysis of RRM spelling from these sounds. Furthermore, I will also look at those sounds that are specific to Romani and which have often been under the attention of those language activists which embarked upon more systematic attempts at writing Romani.

6.3.1 Postalveolar fricatives

The postalveolar fricatives are unvoiced /š/ (as in ship) and voiced /ž/ (as in pleasure).

No tokens of /ž/ were found in the sample, but this is unsurprising as the sound is generally rare in Romani and usually found in loan-words.

/š/, on the other hand, is quite frequent and in the sample was found in words like (a)šun- ‘to listen’, akuš- ‘to insult’, barš ‘year’, murš ‘person’. Exactly as in the welcome message (see 3.2) no instances of Italian {sci/sce}, French {ch} or German {sch} were found, reducing the options for the representation of this sound to {s} and {sh} (graph 6.13).

![Graph 6.13: Unvoiced postalveolar fricative](image)

While usage of {sh} can be attributed both to Albanian and English, I am inclined to consider it an English influence given the total absence of Albanian from the languages
used in the chat-room (chapter 5) and the lack of any Albanian-only spelling solutions for other sounds (see below).

The usage of {s} for /š/ is common on CMC in Serbo/Croatian (Hentschel 1998) when Latin {ś} is not available. RRM users’ preference for this solution shows that they are aware of both the norms for writing Serbo/Croatian off- and on-line and of Romani spellings in which š is the favoured choice following its adoption in academic writing.

6.3.2 Postalveolar affricates

The postalveolar affricates are voiced /dʒ/ (as in joy) and unvoiced /č/ (as in chair).

As was the case for /š/, even for the unvoiced postalveolar affricates /č/ (found in words like čer- ‘to do’, bičal- ‘to send’, čor- ‘to steal’, tuče-če ‘for you’) the most commonly used spelling, {c}, is derived from Serbo/Croatian CMC conventions in which it is used to represent Latin {č} (graph 6.14), reinforcing the point made previously about awareness of Serbo/Croatian norms and of other Romani spellings.

This choice is further reinforced by Italian influence, where {c} represents /č/ before /e/ and /i/. Awareness of the Italian writing norms is furthermore evident in the usage of {ci} followed by {a} in ciao ċao/ ‘goodbye’ (also occurring as cao) and o in a single token of svčolpe /svčolpe/ ‘to like’.

The English-like solution, {ch}, is again encountered, however it is used very rarely, while no instances of German {tch} or Albanian {ç} were found (/č/ is absent from the French sound inventory).

<table>
<thead>
<tr>
<th>/č/</th>
<th>92%</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td></td>
</tr>
<tr>
<td>3%</td>
<td></td>
</tr>
</tbody>
</table>

**Graph 6.14:** Unvoiced postalveolar affricate

98 {c}
5 {ci} (4 ciao)
3 {ch} (1 chat)
Given the preference seen so far for solutions derived from Serbo/Croatian CMC norms, it was expected that the voiced postalveolar affricate /dž/ (found in words like dža- ‘to go’, džan- ‘to know’) would be represented as {dz}, used in place of Latin {dž}. However, this solution was not employed, and /dž/ shows high variability in its representation (graph 6.15).

![Graph 6.15: Voiced postalveolar affricate](image)

Probably {dz} is not used because one of its constituents, {z} for {ž}, is seldom used as the voiced postalveolar fricative /ž/ rarely occurs in both Mitrovica Romani varieties (see above).

Serbo/Croatian influence, however, is visible in the tokens of {d}{ in Demajli /Džemajli/ personal name) and {dj} (in djav /džav/ ‘I go’) as they are used in CMC to replace Latin {d}, representing the alveolo-palatal affricate, a sound very similar to /dž/.

A high number of tokens of {c}, as I have shown representing /č/, were found where some way of representing /dž/ was expected. Considering that in German /dž/ (a rare sound found only in loan words) shows a tendency to merge with /č/, I would suggest that the RRM users employing {c} to represent /dž/ are settled in German speaking areas and are sharing this sound change. However, rather than using German spellings – {dsch} for /dž/ or (Hentschel) for /č/ - these users stick to the norm for /č/ emerging on RRM and discussed above.

The usage of {g} in items like ganav /džanav/ ‘I know’ or ga /dža/ ‘go!’ on the other hand derives from Italian influence, where {g} represents /dž/ before /e/ and /i/. However, in Italian (as in all other languages experienced by the Mitrovica Roma) {g} also
represents /g/. This same variation is found on RRM spelling in words like mangav /mangav/ ‘I want’, tigaj /tigaj/ ‘frying pan’ and ciganski /ciganski/ ‘Gypsies’.

This double usage of {g} makes impossible to establish what sound users wanted to represent in words that might undergo palatalization of velar stops before front vowels (see 6.2.3). Because of this reason, I excluded a number of words from my analysis, namely 9 tokens containing the dative marker -ge > -dže, 7 tokens of gili > džili ‘song’ and gilav- > džilav- ‘to sing’, 2 of mange > mandže ‘you want’ and 2 of gel- > džel- ‘went’.

By the point of view of users, employing the same letter for the two sounds can be helpful as it makes representing variation easier: users palatalizing /g/ in the target words will read {g} as /dž/, those that do not palatalize will read {g} as /g/.

For some users, however, this overlap seems to be confusing as they clearly look for alternative ways to represent /dž/, as spellings {gj} and {dc} clearly show.

{gj} can be regarded as a compromise between Italian and Serbo/Croatian norms with Italian {g} modified, as in Serbo/Croatian {dj}, by {j} to represent /dž/. It must be said that {gj} is used in Albanian and Serbo/Croatian to represent /dž/ (as in argue). However, a number of tokens of gja- /dža-/ ‘to go’ and gjan- /džan-/ ‘to know’ clearly indicate that on RRM spelling {gj} represents /dž/.

The single token of {dc} can be similarly regarded as a compromise, although idiosyncratic, between either Serbo/Croatian {dj} or German {dsch}, but employing the character used on RRM for /č/. In either case it would seem to be an attempt at representing a transitional sound between /dž/ and /č/.

Other solutions available to users to avoid confusion between the two values of {g} would only be English {j} or Albanian {xh} as French does not have /dž/ in its sound system and borrows the spelling for it from other languages. However, as I’ll show in 6.3.4 plain {j} is exclusively used to represent the semi-vowel /j/. The total absence of Albanian {xh}, on the other hand, confirms the lack of active knowledge of this language I noticed in chapter 5 when analyzing language choice.

6.3.3 Glottal and velar fricatives

The voiceless glottal fricative /h/ (like in ham) is found in words like hem or haj ‘and’ and her ‘thigh’. The voiceless uvular fricative /x/ (like in loch) is found in words like xa- ‘to eat’ and xoxav- ‘to lye’. Both in Mitrovica Gurbet and in all the Arli varieties from RMS /x/ is progressively merging with /h/, resulting in pronunciations like /hohav/ for xoxav ‘I lye’. 

156
Graphs 6.16 and 6.17 would at a first glance suggest that this merger has reached nearly full completion, as only in a minority of cases the two sounds are spelled in different ways. However, it is important to note that in the Latin alphabet for Serbo/Croatian /x/ is represented by {h} and that the language does not have /h/ in its sound inventory and does not have a way to spell it. At the same time {h} is used in French, German and English to represent /h/. It is therefore safer to assume that users are drawing from the Serbo/Croatian norm in order to represent /x/ (although a single instance of German {ch} was encountered in the sample) and from French, German and English norms to represent /h/. Such choices, in turn, make the issue of representing the /x/ > /h/ merger irrelevant.
The complete avoidance of \{x\} for /x/ is furthermore telling about RRM users’ lack of exposure to academic and activist literature, in which \(x\) is instead used systematically. Finally, regarding the \{Ø\} tokens for /h/, all 6 were instances of *hem* or *haj* ‘and’. Since I have noted a tendency among young speakers in Palermo to drop /h/ word initially unless it was the result of /x/ > /h/, I am inclined to think that these spellings come from users settled or that have spent a relatively long time in Italy.

6.3.4 Semi-vowels

Across all Romani dialects, the semi-vowel /j/ (as in *yes*) is found mainly in positions following vowels (čhaj ‘Roman girl’, kaja ‘this.F’, šaj ‘to can’) and in morphophonological jotation phenomena, like the insertion of -ij- in the copula discussed in 6.2.2 (Matras 2002: 52). In my sample it also occurred frequently in word initial position, but only in the formalized shouts *jaša* and *joj*.

RRM users, as it can be seen in graph 6.18, are quite systematic in following the Serbo/Croatian norm by employing \{j\} in word final and initial positions and \{ij\} (occasionally typed as \{ji\}) word internally.

The usage of \{i\} in word internal positions is also frequent. As the high token count of borrowed words (international *radio* and Italian *zio* ‘uncle, mate’) suggests, this is a clear influence of French, German, English and Italian norms.
Interestingly {y}, although used frequently in English elements in the nick-names (i.e.: boy) and in Net-talk items (yeah) is hardly ever used in inherited lexical items.

The semi-vowel /w/ (as in whisky) is absent from the inherited Romani sound inventory and can only be found in borrowed items or emerging through contact induced sound changes (Matras 2002: 52). In my sample the only word that might contain /w/ was wallah ‘for God’s sake’ from Qur’an Arabic. However, among the Mitrovica Roma in Palermo, I heard this word variably pronounced as /vallah/ or /wallah/ and the same variation occurred on the RRM sample, with 4 tokens spelled as vallah and 1 spelled as wallah, suggesting this single user is aware of the international conventions for spelling Arabic in the Latin alphabet.

6.3.5 Alveolar affricate

The alveolar affricate /c/ (as in Betsy) occurred rarely in the sample (mostly in the instrumental marker -nca) and, with the exception of the 3 tokens of borrowed zio ‘uncle, mate’ from Italian, was always spelled as {c} as in Serbo/Croatian and Albanian (graph 6.19).

![Graph 6.19: Alveolar affricate](image)

6.3.6 Labial fricative

Spellings for the labial fricative /v/ (as in voice) were checked to ascertain if any of the users employed German {w} to represent this sounds rather than the solution common to all other languages, {v}. All of the 147 tokens founds (words like voj ‘she’, vačar- ‘to
speak’, 1SG.PRES concord marker -av) showed usage of {v}, showing no influence of German spelling.

6.3.7 Vowels

The inherited Romani vowel system only comprises 5 sounds, therefore their representation is straightforward: {ə} for /a/ (as in master), {e} for /e/ (as in pet), {i} for /i/ (as in easy), {o} for /o/ (as in pot) and {u} for /u/ (as in pull).

However, in Mitrovica Gurbet, YU-014 and YU-016 two centralized vowels /ɔ/ (as in comma) and /ɨ/ (as in roses) are attested. Such sounds occur in Romani dialects of the Balkan region as a result of contact influences and alternate with their uncentralized counterparts /e/ and /i/ (Matras 2002: 59).

In my sample only 7 potential tokens of /ɔ~/-/e/ (all of them in berš ‘years’) occurred and they were always spelled as {e}. If the users intended to represent /ɔ/, they were following either French or German norms (both using {e}), rather than the Serbo/Croatian norm of not spelling it. However, as both in French and German /e/ and /ɔ/ are spelled {e} it is impossible to say what sound the users intended to represent. Just like in the case of /h/ and /x/ (see 6.3.3) the choice of {e} makes the problem of representing variation redundant.

6.3.8 Aspirated consonants

The presence in the Romani sound inventory of a set of aspirated consonants /ph, th, kh, čh/, pronounced as their plain counterpart and followed by aspiration, is the main phonological feature marking Romani out as an Indic language (Matras 2002: 54). Because of this reason, activist have always took great care in representing these sounds as a way to flag the Indian origin of the Roma people. As aspirated consonants are not distinctive in the main European languages and are represented as their plain counterparts, activists have always relied on the academic tradition of combining the plain consonant with {h} to represent aspirates.

On RRM spelling, however, aspirated consonants are systematically represented as their plain counterparts, without marking aspiration (graph 6.20). The only instance of an attempt at marking aspiration, shown in 6.23.c from pharav- ‘to break’, shows a typo suggesting the user was not sure about how to exactly mark aspiration.
Therefore, just as in the case of not using \{x\} for /x/, not marking aspiration indicates RRM users did not have any sustained exposure to academic or activist literature.

6.3.9 Early Romani /ř/

Early Romani /ř/, the actual sound quality of which remains however unclear, was a continuation of retroflexes /ḍ/ and /ṭṭ, ḍḍh/ (roughly as English /d/ and /t/ pronounced by South Asian speakers). In many dialects /ř/ merged with /r/ (like Scottish r) but in a few it became /ʁ/ (like French r) or a prolonged trill /rr/ (Matras 2002: 50).

Just as in the case of the aspirates, these sounds represent a clear connection with India, therefore some activists systematically represent the continuations of /ř/ as \{rr\} to stress this link. However, even if the long trill /rr/ is attested in Mitrovica Gurbet in words like rom /rrom/ ‘Rom’, pharav- /pharray- ‘to break’ no users employed \{rr\} in their spelling, again indicating a lack of awareness of activist literature.

/ř/ was also part of the consonant cluster -ndř-, which, as seen in 6.2.1 (examples 6.2.a and 6.5.a) only occurred in the Gurbet realization, showing spelling \{ngr\} closely replicating the actual pronunciation /ngr/.

6.3.10 Spelling summary

The spelling developed on RRM is largely phonological, in that generally each letter represents a single sound and that grapheme combinations representing a single sound are avoided. However, as illustrated in table 6.5 (spellings occurring 5 times or less are not indicated) and as discussed through this section, there are some exceptions. Considering
that none of the standardized scripts based on the Latin alphabet which served as models for RRM users is entirely phonological, the fact that they have achieved a relatively high degree of phoneticity for their spelling is striking.

<table>
<thead>
<tr>
<th>RRM spelling</th>
<th>Sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>/a/</td>
</tr>
<tr>
<td>b</td>
<td>/b/</td>
</tr>
<tr>
<td>c</td>
<td>/c, ċ, čh/</td>
</tr>
<tr>
<td>d</td>
<td>/d/</td>
</tr>
<tr>
<td>f</td>
<td>/f/</td>
</tr>
<tr>
<td>g</td>
<td>/g, dž/</td>
</tr>
<tr>
<td>gšt</td>
<td>/dž/</td>
</tr>
<tr>
<td>e</td>
<td>/e, ǝ, ǝ/</td>
</tr>
<tr>
<td>h</td>
<td>/h, x/</td>
</tr>
<tr>
<td>i</td>
<td>/i, j/</td>
</tr>
<tr>
<td>j, ij</td>
<td>/j/</td>
</tr>
<tr>
<td>k</td>
<td>/k, kh/</td>
</tr>
<tr>
<td>l</td>
<td>/l/</td>
</tr>
<tr>
<td>m</td>
<td>/m/</td>
</tr>
<tr>
<td>n</td>
<td>/n/</td>
</tr>
<tr>
<td>o</td>
<td>/o/</td>
</tr>
<tr>
<td>p</td>
<td>/p, ph/</td>
</tr>
<tr>
<td>r</td>
<td>/r, rr/</td>
</tr>
<tr>
<td>s, šh</td>
<td>/s, š/</td>
</tr>
<tr>
<td>t</td>
<td>/t, th/</td>
</tr>
<tr>
<td>u</td>
<td>/u/</td>
</tr>
<tr>
<td>v</td>
<td>/v/</td>
</tr>
<tr>
<td>z</td>
<td>/z/</td>
</tr>
</tbody>
</table>

Table 6.5: RRM alphabet

Regarding the influences at the base of RRM spelling, users do not seem to be aware of the academic and activist traditions of writing Romani, as I have shown in the cases of aspirated consonants, /x/ and the reflexes of /ř/. Nevertheless, their spelling is largely compatible with other spellings developed for Romani as it is prominently based on standard Serbo/Croatian norms which largely overlap with academic- and activist-developed spellings.

Where RRM spelling diverges from other Romani spellings, this is due either to the limitations imposed by the medium (see the embracement of CMC Serbo/Croation norms) or to the influence of the new contact languages experienced by the community.

Italian influence is clearly visible in the double value of {g} (used for /g/ and /dž/). At the same time, it is also reinforcing the CMC Serbo/Croatian usage of representing /č/ with {č}. French and German influence, on the other hand, are visible in the solutions adopted for the representation of /h/ and of /a/. Finally, English influence is prominent in the usage of {sh} for /šh/.
Example 6.25: Self-repair of typos

In spite of all these different influences, and of the differing degree of formal education achieved by the users, the other striking feature of RRM spelling is the relatively high consistency of its usage. With the exception of /dž/, all sounds which have more than one possible spelling are represented in the same way 78% of the time (on average). Furthermore, each user generally employs a single solution among those available for sounds with multiple representations. Some of the regular users are even going a step further in that they self-repair their own typos wherever or not they involve these problematic sounds (example 6.25).

Attention to one’s own spelling, however, does not turn into attention to other users’ spelling. As it is the case for the choice of the Romani variety used (see 6.2.4), I could not find any case of users reproaching others for spelling choices in my corpus. However, on a couple of occasions during fieldwork I noticed users suggesting a different spelling for /dž/ if the one employed might create confusion. Exactly as for misunderstandings around dialectal features, the suggestions were never framed in terms of the ‘appropriate’ way of writing.

6.4 Summary

If we were to describe the transition of Romani from oral to written form on RRM resorting to Haugen’s standardization model (see 2.2.1), we would immediately realize that the first step in his matrix, the selection of a norm, has been by-passed by RRM users. As the analysis of dialectal features has shown, they use the two varieties spoken by the community without trying to reduce the variation between them in any way.

Because no variety has been selected, it is impossible for users to establish which paradigmatic patterns are the ‘correct’ ones, thus creating a prescriptive grammar, and which lexical items are ‘appropriate’, thus creating a prescriptive lexicon. Therefore the following step, codification of the form, can be carried out only in regard to the language orthography. As I have shown, RRM spelling draws from the orthographies of the various major languages known by the Mitrovica Roma. However no user is taking the lead in this
process establishing which spellings are correct and which are not. Rather a consensus pattern is emerging through repeated action.

Concerning the further steps envisaged by Haugen, elaboration and implementation, only the first is relevant to RRM users. As far as I know, no steps have been taken by any of the community members to have Romani recognized outside the private sphere of communication. Similarly on RRM no discussion ever emerged about extending language use outside the on-line community. As no need to implement language usage in other contexts is felt, implementation is not undertaken.

Some form of elaboration, on the other hand, is needed, particularly in relation to the terminology referring to the technologies employed for communication. In these cases, what users are doing is switching to a contact language. For example, in 6.14.a it is possible to see the usage of French invitation ‘invitation’, the official translation for friend request used by Microsoft on the French interface of Messenger, while in 6.22.c English ban is used. Such switches are also leading to permanent borrowings as the Romani inflection on French se bloquer ‘to crash’ (6.21.c) clearly shows.

Furthermore, as Kaplan & Baldauf (1997) noted for other cases, codification and elaboration are happening simultaneously and not in an actually planned way. Rather, the whole transition form oral to written form is happening spontaneously on RRM as users look for practical solutions that will allow them to effectively communicate across the diaspora.

This practical reason seems to be the only motivation behind the attempt at writing Romani. The lack of any purist and prescriptive attitudes shows that RRM users do not share the type of ideological background linking an idealized standard language to a monolithic representation of the group culture and identity commonly seen in centralized processes of language planning (see 2.2). Rather, as Matras (2004) has suggested, the only criterion they follow is efficiency of communication. As the lively discussions described in chapter 4 have shown, they are achieving this with ease, making the case of Romani codification on RRM a perfect example of linguistic pluralism.

Even more importantly, the free approach to language usage characteristic of linguistic pluralism is not limited to Romani, but is extended to all the languages employed on the chat-room. As described above, both the spelling and certain grammatical features of the major languages on RRM chat-room show typical language learners’ traits, which in a prescriptive perspective should be regarded as mistakes and sanctioned. In some cases, language learners’ traits overlap and mix with unconventional spellings typical of natives’ usage in CMC environments. Just to mention a few cases, in example 4.7.j French pasque
‘because’ is spelled paske, vous ‘you.PL’ is spelled vu, donne ‘she gives’ is spelled don and manger ‘to eat’ is spelled mange, all of them being simplified spellings found in French text messages (Anis 2007). Just as no reproaching comments are posted about Romani usage, the same applies to such cases.

All these elements point at a continuity with the attitudes and practices common in the whole Mitrovica Roma community. Both before and after their dispersal, they have constantly been exposed to different yet, depending on the context, equally appropriate sets of socio-cultural norms. Furthermore, their condition as a marginalized, diasporic group often engaging in itinerant service economy forced them into a position in which they had to manoeuvre with and between such sets of norms. This flexible attitude is clearly visible, for example, in their manoeuvring between the ritual practices of different religions or in their appreciation for and, in the case of musicians, the ability to conjugate genres coming from different music traditions I described in chapter 3. On RRM chat-room, this flexibility is applied to language usage: the norms governing each of them are regarded as an accessible repertoire of guidelines to be, depending on the situation and the goals to achieve, followed strictly or adapted. As no set of norms is fully embraced and anyone is constantly moving between them, purist and prescriptive attitudes do not emerge. As a result, communication is effectively achieved without the burdens of an orthography, a grammar and a lexicon defining what the correct language must be and the attendant requirement that everyone uncritically embraces and takes pride in it.
7 Cosmopolitan practices, language development and the Internet

I have argued (see chapter 3) that, through the sophisticated usage of languages and images displayed on the homepage and the choice of music broadcasted, RRM DJs have created a memorial to the lost homeland. Furthermore, I have shown how the audience of RRM is recreating the conviviality of the homeland by recontextualizing well-established community norms and practices, such as formalized shouting, high standards of politeness, avoidance of insult, playful teasing (see chapter 4), within the new domain offered by the Internet. At the same time, new practices such as searching for old acquaintances, exchanging IM contact details to carry out private conversations, using a language other than Romani to mark a conversation as private (see chapter 4) and, most strikingly, the regular usage of written Romani emerged spontaneously within this new domain. Many of the norms governing these new practices seem to have emerged as a response to the technical limitations imposed by the Internet. Some, like the mapping of letters to phonetic values (see 6.3), are clearly drawn from or based on the cultural repertoires of the majority cultures encountered by the Mitrovica Roma. Others, like code switching in communicative functions usually connected with strong extra-linguistic cues (see chapters 5) rely on the ability to move between forms drawn from various cultural repertoires.

I have also pointed out that, although recreating the conviviality of the homeland, users do not display purist and exclusionary attitudes by subscribing to a monolithic idea of what being Mitrovica Roma should be. Rather, as the interactions analyzed in chapter 4 and the choice of languages (chapter 5) and of Romani varieties (chapter 6) have shown, each RRM user performs their own personal, hybrid identity. Such personal identities retain a shared sense of Romani-ness at their core, flagged through the dominant usage of Romani, but combine it with elements pointing to individuals’ participation in wider circles (for example Islam and the global popular culture) flagged through the usage of languages other than Romani.

Finally, in chapter 6, I have shown how this absence of a purist attitude extends to the choice of the Romani varieties employed and the spellings used for them without hindering the transition from orality to literacy.

Following Pettan (2002’, see chapter 3) I defined RRM as a space inspired by a cosmopolitan perspective, it is now my intention to show how the chat-room users’ practices can also be characterized as cosmopolitan. I will therefore summarise and further analyse these practices (7.1 to 7.3) and then relate them to various definitions of cosmopolitanism (7.4). This will in turn allow me to link cosmopolitanism and the
possibilities offered by the Internet to language development occurring outside any formalized and institutionalized language planning.

### 7.1 Language practice on RRM

As I have shown, RRM users access and deploy forms from all the different linguistic repertoires they are exposed to. Their preference for certain forms over others, as shown in chapters 4 and 5, is guided by communicative targets (i.e.: achieving humour, asking politely, etc.) and at the same time by the desire to portray oneself simultaneously as a Roma from Mitrovica living in the diaspora but also as participating in wider circles (i.e.: the larger Muslim community, global popular culture). Similarly, when it comes to the orthography employed, the official spellings of standardized languages alternate with non-standard spellings common in CMC and elements from all such spellings are used to represent the sounds of Romani (6.3).

Although the various elements employed can be attributed to for example Romani or German, the boundaries between given languages as defined by prescriptive perspectives are blurred. Furthermore, as I have shown in chapter 5, users show differing levels of competence in the different languages they employ, so that for example Qur’an Arabic only appears in formalized salutations and emphatic particles while Romani is employed in a much wider range of functions. RRM linguistic practice is then characterized by the usage of what Blommaert & Backus (2013) define as superdiverse repertoires. As they suggest, superdiverse repertoires are by no means limited to diasporic, mobile populations but rather constitute the basis of most linguistic practices in the contemporary, globalized world. Superdiverse repertoires are acquired through individual life trajectories of language learning and are therefore limited and unique to singular individuals. However, they share the common characteristic of being made up of ‘truncated’ resources “often derived from a variety of languages, and with considerable differences in the level of development of particular resources” (Blommaert 2010: 106). However, the usage of any linguistic resource is valued on the basis of its indexicality, its capacity to point “towards socially and culturally ordered norms, genres, traditions, expectations” (Blommaert 2010: 33). As a result, depending on the given socio-cultural context some elements “are systematically perceived as valuable, others as less valuable and some are not taken into account at all” (Blommaert 2010: 38).

The various elements in the repertoires employed by RRM users have clear indexical values and are displayed “in order to make sense to others, that is, in order to operate within the norms and expectations” (Blommaert & Backus 2013: 29) characteristic of RRM chat-room. Romani is indexical of the ethnic ties shared by users, Serbo/Croatian,
French, German and Italian index their condition as a diaspora, Qur’an Arabic salutations and emphatic particles point to belonging to Islam. English words, particularly those written following CMC conventions, and Net-talk items display a positive orientation towards global popular culture. On the other hand, the absence of any prescriptive attitude suggests that no element is valued more than the others. As RRM users are aiming at recreating the conviviality of the homeland in their mundane practices, the cultural flexibility characterizing the Mitrovica community at large (see 3.1) manifests itself in the usage of all the available linguistic resources. As long as successful communication, and thus conviviality, is achieved the usage of any resource in their repertoires, even if not fully developed, is accepted without questions.

This does not mean that RRM users are unaware of the values each resource in their repertoires can assume in different settings. In fact, given RRM users’ more or less overt refusal to engage in any critical reflection about their practices (see 4.1 and 4.2), I am not in a position to describe what their broader attitudes towards the elements in their repertoires might be. Nor can I claim that their attitude erases the stratification of languages based on the socio-cultural values attributed to them. Rather, the point is that this stratification is temporarily ignored.

As a result the linguistic interactions emerging through users’ practices on RRM are extremely complex, comprising many overlapping and not clearly bounded linguistic forms.

This peculiarity reminds us of the fractal nature of cultural forms suggested by Appadurai (1996). According to him, all cultural forms are in a constant flux across various dimensions he refers to using the suffix –scape. Thus, the ethnoscape is “the landscape of persons who constitute the shifting world in which we live” (33), the technoscape comprises the various technologies currently available, the mediascape is the combination of physical media and of the images they circulate and the ideoscape is a concatenation of images and narratives “often directly political and frequently have to do with the ideologies of states” (35). All these scapes are perspectival constructs “inflected by the […] situatedness of different sorts of actors” (33) and the cultural forms comprised in each of them stand in a mutually influencing relationship with the forms comprised in the other scapes. As such, just like the languages used on RRM, cultural forms do not posses “Euclidean boundaries, structures, or regularities” (46). It might then be useful, following Coupland’s (2003) suggestion, to introduce a further dimension of cultural flows
which we can call *linguascape*¹. Composed of truncated repertoires, the linguascape comprises the linguistic resources through which people interact and express themselves.

In the case of RRM, when users employ forms attributable to Romani, Qur’an Arabic or one of the languages of the countries they live in, the linguascape reflects the ethnoscape of the community at large as a group of recently dispersed Muslim Roma. Images and personas drawn from the mediascape of mainstream music and cinema are represented and enacted through the sophisticated usage of linguistic forms associated with them, mostly English and Net-talk. All of this is happening within the framework provided by the relatively easy to access and open technoscape of the Internet. The ideoscape of Internet usage, marked by its flouting of established rules, transpires in the usage of abbreviations and non-standard spellings. Similarly, the lack of elements characteristic of the activist literature in RRM linguascape points to the absence in the users’ ideoscape of elements drawn from the activist narratives about who and what the Roma are. This in turns relates to a different way of imagining the Roma ethnoscape, but I will return on this later on in the chapter. What is important to notice now is that this relationship between the users’ ideoscape and RRM linguascape can add a further explanation to, and a condition for, the fact that no attempt is made to regiment language usage on the chat-room. As users are not making any political statement, at least not through RRM, they do not need to produce an essentialized image of language that will fit into the nation-state ideoscape and could be used as a tool in support of claims for recognition.

The absence of a political dimension in the users’ ideoscape is thus what makes it possible for RRM linguascape to remain fluid and irregular, with no degrees of value attributed to the different elements of the repertoires used in it. Finally, as outsiders are only able to participate if they can reproduce this fluidity, a convivial and friendly environment is created. In such environment there is, as far as I have seen, no space for the emergence of essentializing discourses that would aim at reducing linguistic variation in the name of a supposedly unique and true identity.

### 7.2 Language development on the Internet

The interaction between the technology and media of the Internet and the ideology and ethnicity of users has, however, led to a change in the linguascape experienced by RRM users, namely the emergence of a fully functional writing system for Romani (see chapter 6). Since its emergence was not planned but rather happened spontaneously as users looked for practical solutions to communicate across the diaspora, I prefer to describe this

---

¹ The term has also been used by Jaworski, Thurlow, Lawson & Ylänne-McEwen (2003), Jaworski, Ylänne-McEwen, Thurlow & Lawson. (2003) and following them Reershemius (2011a, 2011b). Even if their usage of the term is reminiscent of Appadurai’s concept of *-scapes*, none of them links it explicitly to his work.
change in the linguascape as an instance of language development rather than of language planning.

What appears to be crucial in this development is the role played by the Internet. As a relatively easy to use and, in the urbanized context of Western Europe, readily accessible technology it offered the Mitrovica Roma a quick and cheap way to remain in contact. However, being a text-based technology it forced them to write Romani if they wanted to maintain a link between their ethnic identity and their language. At the same time it offered access to varied media in which various languages are used. Most importantly, certain sections of the Internet are removed from the traditional channels of language planning and policing. In such environments, even in the case of languages with highly formalized and well established norms, the rules regarding language usage are often flouted (see 1.2 and 1.3). The fact that these on-line practices mirrored certain practices and attitudes of the Mitrovica Roma could be assumed to have acted as a catalyst for them to feel free in experimenting with Romani. Because of their apparent lack of interest for any form of political recognition, and thus the drive to turn language in a political-ideological tool as done by many other ethnic minorities, the development from orality to literacy in Romani took place through everyday interactions, in the performances enacted on RRM chat-room.

The fact that similar language developments are taking place among other communities of Roma who have started to use the Internet (as discussed in 2.3.4), but also for other previously unwritten languages (for example Jamaican Creole and Nigerian Pidgin, Deuber & Hinrichs 2007 and Mauritian Creole, Rajah-Carrim 2009) suggests that what is happening on RRM is not an isolated case, but rather an emerging pattern of changes in various languages. What puts such changes apart from the planned attempts at initiating language change enacted in the past century is, I will suggest, the different direction of flows between the various scapes involved in the changes.

In traditional attempts at language planning, the elites of de-colonizing polities aimed at gaining better access to modern technologies and international markets. Seeing how the unity of nation, state and standard language seemed the only way to gain access to greater resources, they planned to change the ethnic and linguistic profile of the emerging polities to match the dominant image of the nation-state. In this sense, language planning can be regarded as a top-down act of representation portraying the identity of emerging polities as that of legitimate actors on the global stage and thus entitled to access its resources.

On the other hand, in language developments occurring on the Internet we have individuals, often coming from diasporic backgrounds and not members of the political-economical elites, aiming at maintaining forms of conviviality typical of their community. In order to do so, they take to the Internet, one of the most easily accessible
communication technologies. Through their everyday, mundane interactions they create virtual spaces in which their community is replicated without consciously attempting to change the community itself. However, in order to cope in the environment provided by the Internet, a change in the language is required. The technology itself offers a veritable host of resources that can be drawn upon to solve the problem. Very often these resources also include written examples of different varieties of the same language, either produced by other individuals engaging in similar practices but also by activists using the Internet to support their campaigning. If, like in the case of RRM the ideology of the individuals involved are characterized by a disregard, however temporary, for value hierarchies and no desire to gain recognition, the change is achieved smoothly and spontaneously through the very same practices that created the conditions for its occurrence. Marked as they are by cultural flexibility and acceptance of variation the practices of any individual or group can easily sit alongside those of others without conflict arising around the correctness of anyone’s practice. This type of language developments then takes place bottom-up through identity and togetherness performances aiming at maintaining a particular group identity.

The fact that the emergence of a written language could be achieved through everyday performances assisted by modern communication technologies without necessarily resorting to the essentialized representations of structured language planning efforts confirms the viability of a model of linguistic pluralism (Matras 2004, see also 2.3.5). Yet, I would suggest that if, as it is the case for Romani, some users of a developing language want to see its written form employed beyond the sphere of private, convivial communication some form of traditional language planning would be required. However, the cultural flexibility and acceptance of variation seen in the performances should inform any representation produced to achieve the desired goal. Failing to do that could lead to language users being faced with an essentialized representation of what the ‘correct language’ must be. If users are then asked to change their varied practices to match such a prescriptive representation, they might completely reject attempts at gaining more recognition for the language.

7.3 Representing and performing diasporas on-line

The contrast between structured language planning efforts and the linguistic performances leading to language change discussed above has an interesting counterpart in RRM users’ refusal to discuss and reflect upon their practices and performances with me.

In their pursuit of the conviviality of the homeland, RRM users engage in a series of practices, through which they display their individual and group identities (see chapter 4) apparently without thinking about what these practices and performances might come to
represent for outsiders. Considering how much theorizing has been produced around the issues of representations and self-representations of diasporas, particularly regarding their engagement with the Internet (see 1.1 and 1.3), it is worth reflecting upon the absence of this characteristic from RRM interactions.

My early remark that there might be a different way of imagining the Roma ethnoscape between Romani activists and RRM users (see above) will be a starting point for this reflection. As a quick overview will show, the majority of Romani activist website or newsgroup closely match the common academic description of diasporic webspaces as discursive spaces in which marginalizing and stereotyping discourses are challenged and new, hybrid identities are negotiated (Leggio, 2011b, 2012). The narratives representing the Romani ethnoscape produced in such spaces, might vary in details but often converge in presenting the whole Romani population as a single diaspora originating from India, united by its language and striving to overcome a century long condition of marginalization and discrimination. In these narratives, distinctions based on religious affiliation, country of residence and class are very often put aside. On RRM, on the other hand, the Romani ethnoscape is the very specific community of Romani-speaking Muslims that left Mitrovica and now live all around Europe. Furthermore, this identity is not talked about but comes to life through engagement in everyday practices.

In her work about diasporic Balkan Roma in the US, Silverman (2012) similarly notes how “Roma constitute a multiplicity of cultures that neither intermarry nor identify as one group; this variation is erased by conceiving of the Romani diaspora as a unified cultural unity” (40). Furthermore she stresses how “enactment of identity via performative genres, especially music, is a visible, audible symbol in the Balkan Romani diaspora. Through performances, identity is conceptualized” (41). Yet she also notes that musicians, as much as activists, sometime capitalize on unified notions of the Romani diaspora when, in the pursuit of their goals, they have to represent it to non-Roma (see also Georgieva 2006).

We can then assume that in every diaspora, on a regular basis, people perform their diasporic identities in everyday practices. Only in some occasions some of them, often in a position of material or cultural privilege, represent these identities to outsiders, often essentializing them in a single master narrative. It would thus seem that scholarly attention to the on-line presence of diasporas has so far only focused on the latter few.

Indeed in the early years of the Internet it is likely that, as Gajjala (2002) noted, few diasporic subjects using it were “likely to be materially or culturally ‘underprivileged’” (188). This being the case, the subjects of study had “access to the same (or similar) power structures as the researcher. They, as much as [the researcher, were] anthropologists, reporting not just about [their] own diasporic communities […] but also […] about [their]
host society/culture” (189). The discourses produced in such diasporic webspaces were thus similar and resonating with the academic discourses produced about diasporic subjects and, I would say unsurprisingly, caught the attention of researchers. Yet, over the years the Internet has become more accessible to underprivileged diasporic subjects.

As we have seen on RRM, rather than engaging in the same processes of representation documented in earlier research, underprivileged diasporic subjects replicate their everyday performances on the Internet. This performative engagement with the medium allows them to maintain a sense of unity and shared belonging among community members. At the same time, it allows individuals to display varying orientations towards cultural elements drawn from the repertoires of the majority populations they are in contact with.

7.4 Cosmopolitan practices on RRM

In order to show how all the above practices can be regarded as an instance of cosmopolitanism it is the case to revise some of the ways in which the term has been defined and used so far.

One way in which the term cosmopolitanism has been used can be summarized as political. In this usage cosmopolitanism is the opposite of nationalism. Cosmopolitans are therefore individuals claiming to be citizens of the world and advocating equality and human rights while cosmopolitanism is regarded as part of a philosophical, ethical and political project alternative to neo-liberal ideologies (see for example Appiah 1996, Archibugi 2004, Beck 1998, Lu 2000, Nussbaum 1996, Waldron 2000).

The term cosmopolitanism has also been used in a cultural sense, as a particular stance towards the diversity of human cultures. In this usage, definitions of cosmopolitanism are generally inspired by Hannerz’s conceptualization of cosmopolitanism as “a willingness to engage with the Other […] an intellectual and aesthetic stance of openness toward divergent cultural experiences” (1990: 239). Furthermore, Hannerz regarded cosmopolitanism as an attribute of individuals displaying an ability to appropriate elements drawn from other cultures and what he termed cultural competence “a built-up skill in manoeuvring more or less expertly with a particular system of meanings and meaningful forms” (ibid.). Finally, Hannerz contrasts cosmopolitans, as relinquishing their belonging to any given culture, with locals, which instead cherish their being part of a particular culture. Because of the importance he gives to the individual’s desire to experiment with different cultures and the absence of a sense of belonging, Hannerz identifies as good examples of cosmopolitans the professional expatriates who choose to work abroad and take pleasure in their engagement with cultural diversity. By the same token he rejects migrants and diasporic subjects since their engagement with different cultures is forced
upon them and they tend to create surrogate homes “with the help of compatriots, in whose circle one becomes encapsulated” (1990: 243).

As noted by Werbner (1999) the identification of cosmopolitan individuals with professional expatriates originating from the wealthier classes in the Western world is implicitly elitist and Eurocentric\(^2\). Along the same lines Calhoun notes that linking cosmopolitanism to intellectual and aesthetic values can lead to underestimate “the extent to which the cosmopolitan appreciation of cultural diversity is based on the privileges of wealth and perhaps, especially, citizenship in certain states” (Calhoun 2002: 108).

Following these critiques scholars have started looking for instances of cosmopolitanism precisely among the kind of individuals that Hannerz rejected as examples of cosmopolitan individuals. Variously termed as working class (Werbner 1999), vernacular (Diouf 2000) or migrant (Kothari 2008), the cosmopolitanism of migrants and diasporic subjects, they suggested, is characterized by cultural competence, but is not based on an intellectual and aesthetic stance of openness. Rather, it stems out of the necessity of individuals and groups to ease their insertion in the migratory context. Furthermore in these cases cosmopolitanism is not an identity trait, but rather a series of practices, something that people do. As Kothari notes, however, “over time and in quotidian spaces, through mundane experiences and everyday encounters cosmopolitan attitudes and behaviours can and do become embedded” (2008: 514). These practices lead individuals to embrace and make use of a multiplicity of cultural features (languages, religious beliefs, dietary customs, dress codes, etc.) that “create new forms of cosmopolitanism whose manifestations no longer refer necessarily and obligatory to the acquisition of an identity” (Diouf 2000: 702) as cosmopolitans. Yet the sense of identity of subjects engaging in cosmopolitan practices is changed in that “belonging to more than one ethnic and cultural localism simultaneously” (Werbner 1999: 34) is not perceived as a contradiction. All these authors also noted how, in these cases, ethical and philosophical orientations leading to the embracement of cosmopolitanism as a political project are generally absent.

Glick Schiller et al. (2011) concur with Werbner, Diouf and Kothari in maintaining that cosmopolitan practices entail cultural competence and “do not necessary lead to the embracing of a universalistic self-identification or single global political project” (404). Furthermore, Glick Schiller et al. contribute with a clear definition of cosmopolitan practices to the debate on cosmopolitanism. They see cosmopolitan practices as instances of sociability: everyday interactions in which “people gather in the same place or in cyberspace around some point of shared interest that is not primarily utilitarian” (402).

\(^2\) Hannerz later revised his position, see Hannerz (2004).
such interactions people are “not passive consumers but active participants in the creation of common places” (403) where “common and sometimes fluid identities within larger collectives and transnational social fields” (ibid.) are made manifest.

Now, going back to RRM users and their behaviours on the chat-room, it is possible to see how they neatly fit into Glick Schiller et al. definition of cosmopolitan practices as instances of sociability. RRM users gather around the web-site as they share an interest for the music broadcasted by the DJs, but the availability of the chat-room allows them to actively participate and not just passively listen.

The other crucial element characterizing cosmopolitan practices, cultural competence, the ability to appropriately use different cultural features, is clearly visible in RRM users’ skilful usage of the various languages available to them and in the patterns of code-switching allowing them to supplement the lack of extra-linguistic clues typical of written communication and clearly achieving their communicative targets (as seen in chapter 5) or to mark out conversations as private (as seen in examples 4.9 to 4.11). Cultural competence is also visible in the way users drew on the mapping of sounds to letter of the major languages they are familiar with to develop a spelling for Romani (see chapter 6).

Concerning the possibility that cosmopolitan practices might lead to the embracement of a global political project centred around human rights and equality, Sultano and Elvis, the owners of the web-site, clearly stated in the homepage that RRM is a space ‘for all nations and generations’ and even included the European Union flag and various symbols of peace (see 3.2). The fact that they used various internationalisms and Serbo/Croatian words, but not Romani, in writing the website description is an indication that they were not addressing this message just to the Mitrovica Roma, but rather to a global audience in an apparent call to get united in spite of differences. However, the absence of any discussion on RRM chat-room about racism, human rights, equality and so on shows that, just like in other instances of non-elite cosmopolitanisms, RRM users are not embracing a global political project.

In order to explain this contradiction it is worth remembering that Sultano and Elvis, as well-established migrants in France and owners of a small recording studio (see 3.2), are in a position of relative wealth among the Mitrovica Roma. The majority of RRM users, on the other hand, are often either living in precarious economic conditions or under threat of forced repatriation, sometimes even both (see 3.1). It would then seem that, as suggested by Calhoun (2002, see above), the privileges of wealth and possibly citizenship are indeed a pre-condition for people engaging in cosmopolitan practices to also embrace a global political project.
A further reason why racism and discrimination are never discussed is that while RRM users are often confronted with such issues off-line as they engage with non-Roma (see 3.1), on the chat-room users do not face such issues as they all are, almost exclusively, fellow members of the Mitrovica Roma community. As my personal experience has shown me, it is not the case that users are not willing to accept individuals of different cultural backgrounds into their interactions. In fact, as I mentioned in 4.2.1 regarding the role of DJs as chat-room moderators, no one has ever been banned during my frequentation of the chat-room for not being Mitrovica Roma. However, only people able to understand RRM complex linguascape can actively participate in the interactions. Therefore, if hardly anyone other than Mitrovica Roma is on RRM chat-room, it is because the lack of the specific cultural competence required to interact on RRM prevents them from actively participate in the cosmopolitan practices characterizing the chat-room.

The almost exclusive participation of Mitrovica Roma in RRM chat-room, besides avoiding confrontations around racial issues, also allows for the recreation of forms of conviviality that were typical of the homeland, as for example the formalized shouting accompanying popular songs or the mutual teasing among users (see 4.2, 4.3.1 and 4.3.2). Since Sultano and Elvis clearly wanted to create an on-line memorial to the lost homeland, as shown by the inclusion in RRM homepage of pictures and videos of Mitrovica during and after the conflict and the Romani phrase ma bistren tumaro thand ‘don’t forget your country’ (see figure 3.2), the recreation of the homeland conviviality within this frame is recreating, even if just temporarily, the community itself.

The sense of shared belonging to this community is clearly flagged by users through the usage of two particular varieties of Romani (see chapter 6). However users are also flagging their own individual alignments to broader communities through the usage of other languages (as discussed in chapters 4 and 5). Again it is cultural competence in manoeuvring with the various languages in their repertoire that allows RRM users to do that. The composition of user-names (see 5.2) is the first visible instance of such hybrid identities. When Romani words are mixed with English and Net-talk elements, users are making a statement about simultaneously being Mitrovica Roma and participants in the global popular culture, when the languages of their country of residence are used alongside Romani they present themselves as both Mitrovica Roma and French, German or Italian. Similarly, Serbo/Croatian usage displays users’ awareness of their origin in former-Yugoslavia. Users continue to display these hybrid identities when switching between languages in their interactions and here, through usage of Qur’an Arabic expressions, some users also display their affiliation to Islam (see 4.3 and 5.3). This mixing and switching
between languages thus flags a multiplicity of belongings, a further characteristic of non-elite forms of cosmopolitanism.

The linguistic practices of RRM users are then confirming the characterization of cosmopolitan practices as instances of sociability in which the individuals’ cultural competence is displayed and a multiplicity of belongings is made manifest, even if no aesthetic and intellectual values underpin such practices or the embraces of a global political project emerges out of them. Furthermore, as the multiplicity of belongings each user is manifesting retains being Mitrovica Roma at their core, linguistic practices on RRM also show how engagement in cosmopolitan practices allow for both maintenance of a group identity and the simultaneous displaying of personal, individual alignments with other groups.

A further contribution to the definition of cosmopolitan practices comes from RRM users’ disregard of hierarchies of languages based on the socio-cultural values attributed to them I mentioned earlier (7.1). This disregard for value hierarchies attached to cultural features (languages, religious beliefs, dietary customs, dress codes, etc.) which I will refer to as flexibility, is what makes it possible for all languages in the speakers’ repertoires to be used as none is considered the ‘appropriate’ one. As a result the boundaries that stigma or prestige attached to linguistic resources (Milroy 1989) can create are disrupted. This creates an open and inclusive space in which anyone equipped with the required cultural competence can participate and freely display its multiple belongings. Contrary to what Hannerz (1990) suggested, it would seem that the quality people should posses as a precondition for engagement in cosmopolitan practices is not an attitude of openness but flexibility. Openness however, emerges from cosmopolitan practices and, as Kothari (2008) suggested, over time it can become embedded into people’s attitudes.

In the case of RRM, flexibility seems to stem from the Mitrovica Roma’s condition as diasporic subjects. Their late arrival and settlement in Europe, as it is the case for all Roma, put them in a position of marginality, from which they however had to carve themselves a niche. In the already multi-ethnic environment of the Balkans they were confronted with different yet, depending on the context, equally appropriate sets of socio-cultural norms. Flexibility, and therefore the possibility of engaging in cosmopolitan practices, allowed them both to maintain their specificity and to successfully engage with other ethnic groups in Kosovo, up to a certain extent as they however remained in a position of economic marginality. Following their dispersal from Mitrovica flexibility and cosmopolitan practices, as seen in RRM interactions still allows for maintenance of the group identity. Furthermore, thanks to their flexibility, RRM users have seized on the possibilities offered by the Internet (ease of communication, access to large repertoires of
cultural resources) and overcome its limitations (textual, disembodied nature of communication) and added a new domain of language activity which also required a transition from orality to literacy.

7.5 Concluding remarks

The case of RRM clearly shows that codification of languages does not need to be carried out within the ideological framework of the nation-state to be communicatively effective. Provided that all people involved share what I have termed flexibility, the disregard for value hierarchies attached to cultural features, language codification can be achieved through cosmopolitan practices. As engagement in cosmopolitan practices simultaneously allow for maintenance of a group identity and for individuals to display personal alignments with other identity facets, the codified language does not become a symbol of a monolithic ethnic identity. Rather it is yet another linguistic resource available in people’s repertoires. Furthermore, as cosmopolitan practices breed openness, variation in the written form, both in term of the varieties being written and in terms of the spellings employed, is not seen as a problem. Written language practice emerging from cosmopolitan practices thus closely mirrors spoken language practice. This seems to me an advantage over traditional attempts at language codification that in many cases have failed because speakers rejected the written variety as too distant from their everyday practice. Furthermore, if spaces for common development and interaction in the written variety are available any speaker can potentially take part in a codification process carried out through cosmopolitan practices. This could help avoid the feeling of imposition that can accompany traditional language codification efforts and the association of the written variety with (or appropriation by) particular sectors of a given population. Finally, if these spaces are provided speakers participate throughout the entire codification process largely reducing the effort required to implement the written variety.

The Internet, as an often easy to use and access technology, is offering a perfect space for common development of a written variety of a language. Furthermore, it gives access to examples of standard varieties of well established languages, of unofficial spellings for these same languages and of other languages in the process of developing their own written forms. All these resources can be drawn upon when people attempting to write their own language need to solve problems like the representation of sounds or to assist in the elaboration (lexical expansion) of the written variety. Because of these qualities the Internet also reduces the relevance of academies, committees of experts and other centralized authorities that have normally led language codification process. This further reduces the possibility of small groups taking over the whole process leading to imbalances
of power as they gain prestige from the written variety or producing a written variety that is rejected by the majority of speakers.

As far as I know, cases of language codification based on cosmopolitan practices helped by the Internet like that of RRM only occurred bottom-up among speakers sharing a diasporic background (Deuber & Hinrichs 2007, Rajah-Carrim 2009). Considering that diasporic communities across the world keep increasing and that they are often located in urbanized areas where the Internet is readily accessible, it seems logical to expect more cases like the one discussed in this work to take place.

The challenge for scholars will then be to keep documenting such process of language codification so that, since there certainly are practical advantages in achieving language codification through cosmopolitan practices, traditional actors involved in language planning would support them. For this to happen, however, the picture we should provide must be as accurate as possible. As in this kind of codification process strictly linguistic phenomena criss-cross issues of migration and diaspora, access to technology, representation and performance of identity, traditional sociolinguistic analysis alone will not be sufficient. We will need, as I tried to do in this work, to supplement our analysis with ethnographic methodologies and concepts grounded in anthropological theories. This will allow us not only to account for language internal processes, but also for those socio-cultural dynamics shaping them.

One of the socio-cultural dynamics involved in the emergence of a written variety of Romani on RRM is the absence from the chat-room of any culturally privileged subject like activists or intellectuals (see 7.3). Although of little relevance for the final outcome of language codification on the specific web-site, the fact resulted in a so far undocumented absence of identity representations in a diasporic web-space. Due to the scope of this work, I could not investigate this issue further, but it suggests that we should start reconsidering our descriptions of diasporic web-spaces to take into consideration the different behaviours of more and less privileged users.


