

**Language maintenance-attrition among generations of
the Venetian-Italian community in Anglophone
Canada.**

Submitted by Claudia Bortolato to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Italian
In September 2012

This thesis is available for Library use on the understanding that it is
copyright material and that no quotation from the thesis may be published
without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and
that no material has previously been submitted and approved for the award of a degree
by this or any other University.

Signature:

Abstract

This study reports on language contact phenomena among the Italian-Venetian communities of Anglophone Canada. The analysis perspective is twofold: on one hand it studies language maintenance/attrition comparing two cohorts of migrants, those already well researched who migrated during the period of mass migration (1945-1967) and those who did so in the following four decades (1970-2009). On the other, it investigates language maintenance/attrition taking an intergenerational perspective on three generations of speakers.

The corpus used in the analysis is composed of 56 interviews, collected during three months of fieldwork in Canada in 2009. These data were supplemented by 99 questionnaires, which set the background of the analysis, discussing in particular the linguistic habits and attitudes of the community investigated. Given the huge amount of data considered and the mainly quantitative approach taken in this research, two statistical software programs, Taltac and SPSS, were employed to help with the analysis. Another tool, meta-linguistic observation, is also used to broaden the general framework of the study and whenever possible support it with more evidence.

The literature on language maintenance/attrition among Italian migrant communities is sizeable; however, there remains room for further investigations. This work, in particular, addresses two major aspects still rarely explored: first, quantifying the decline in heritage language skills on a generational scale, and secondly, comparing the linguistic skills of post-Second World War migrants, on which research has mostly concentrated so far, with those of new waves of migrants.

Although this thesis is concerned with a particular geographical and historical framework and the findings are therefore representative of this specific context, the work aims to point to some observations from which generalisation may be possible. By setting side by side these two very distinct cohorts and discussing the new linguistic tendencies in language proficiency among the most recent groups of migrants, research is opened to the new scenarios evolving among Italian communities abroad.

Table of contents

Abstract	3
List of Figures	8
List of Tables	9
Acknowledgments	10
Introduction	13
Part 1: Literature review	19
Chapter 1: Bilingualism in a migration situation: maintenance and attrition	19
1.1 Definitions	20
1.2 Bilingualism among speakers of a heritage language	22
1.2.1 First versus second language: time and sequence of acquisition	22
1.2.2 Primary versus secondary language: prevalence of usage	24
1.2.3 Minority versus majority language: socio-political status	24
1.3 Competence of speakers of a heritage language	26
1.3.1 Baseline	26
1.3.2 Linguistic phenomena shaping the competence of speakers of heritage languages	27
1.3.2.1 Incomplete language acquisition	27
1.3.2.2 Language attrition	29
1.3.2.3 Transfer and Communal language	31
1.3.3 Classification of speakers of a heritage language	31
1.3.4 Sociolinguistic perspective	32
1.3.4.1 Sociolinguistic factors	32
1.3.4.2 Extra-linguistic factors	40
1.3.5 Linguistic perspective	47
1.3.5.1 Linguistic outcomes	47
1.4 Conclusion	54
Chapter 2: Italian: the language and its speakers	56
2.1 The Italian linguistic repertoire	56
2.1.1 Diglossia, Dialectal and Regional Italian	57
2.1.2 Italian language features	59
2.2.1.1 Lexicon	59

2.2.1.2 Verb morphology	60
2.2 Italian migration	61
2.2.1 Historical review	62
2.2.1.1 From the origins to the 1960s	63
2.2.1.2 From the 1970s onward	64
2.2.2 Italian and Venetian communities abroad	67
2.2.2.1 Italian and Venetian migration to Canada, the ‘last best west’ land	68
2.2.2.2 Italian and Venetian presence abroad and in Canada	70
2.2.2.3 Italian communities in Canada	71
2.2.3 Italo-Canadians between integration and isolation	71
2.2.3.1 From the beginning of the 20 th century to the 1960s	71
2.2.3.2 From the beginning of the 1970s onwards	73
2.2.4 Italian and Italese	74
2.2.4.1 Italian	74
2.2.4.2 Italese	76
2.3 Sociolinguistic research on language maintenance among Italians abroad	77
2.3.1 Generation	79
2.3.2 New environment	80
2.3.3 Domain	82
2.3.4 Gender roles	83
2.3.5 Attitude	84
2.3.6 Contacts	85
2.3.6.1 Associations	86
2.3.6.2 Religious services	86
2.3.6.3 Language courses	87
2.3.6.4 Mass media	87
2.4. Summary and concluding remarks	90
Part 2: Methodology	94
Chapter 3: Methodology	94
3.1 Methodology aspects	94
3.1.1 Sampling methods	94
3.1.2 Sampling population	95

3.1.2.1 Interviews methodology	99
3.1.2.2 Questionnaire methodology	104
3.2 Fieldwork in Canada	105
3.2.1 Place and time	105
3.2.2 Ethical issues	107
3.3 Analysis tools	108
3.3.1 Data collection methodology	108
3.3.1.1 Interview	109
3.3.1.2 Questionnaire	112
3.4 Corpus	113
3.4.1 Pre-fieldwork test	113
3.4.2 Interviews	114
3.4.2.1 Transcription issues	115
3.4.2.2 Lexicometric measurements	121
3.4.2.3 Tagging and manual disambiguation	122
3.4.2.2 Measures for the study of the vocabulary	124
3.4.3 Questionnaire	129
3.5 Final remarks	130
Part 3: Analysis	131
Chapter 4: Language attitudes and behaviours	133
4.1 Introduction	133
4.2 Social variables	134
4.3 Analysis and discussion of the questionnaires	136
4.3.1 Attitudinal factors	139
4.3.2 Linguistic habits	145
4.4 Results interpretation and analysis predictions	147
Chapter 5: Language maintenance and attrition among speakers of a heritage language	150
5.1 Introduction	150
5.2 Analysis tools	151
5.3 Textual statistical analysis	152
5.3.1 Tokens	153
5.3.2 The Guiraud Index	155

5.3.3 Average length of content words	158
5.3.4 High, medium and low frequency content words	159
5.3.5 Vocabolario di Base	163
5.3.6 Grammatical categories	165
5.3.7 Content and function words	170
5.3.8 Code-switching	173
5.3.9 Verb tenses	177
5.4 Content analysis	181
5.5 Partial correlation	183
5.5.1 Significant divergences in the language outcomes between native speakers	185
5.5.2 Significant divergences in the language outcomes along the generational scale	193
5.6 Conclusion	200
Conclusion	203
Appendices	209
Bibliography	248

List of Figures

- Figure 1.1 Typologies of bilingualism by age and sequence of acquisition
- Figure 4.1 Questionnaire informants distribution by generation
- Figure 4.2 Affiliation by generation
- Figure 4.3 Difficulties by generation
- Figure 4.4 'Feeling migrant' by generation
- Figure 4.5 Liking Italian by generation
- Figure 5.1 Tokens Z score by generation
- Figure 5.2 High/Medium/low frequency words by generation (%)
- Figure 5.3 High frequency words Z score by generation
- Figure 5.4 Content and Function words by generation (%)
- Figure 5.5 Content/Function words Z score by generation
- Figure 5.6 Code-switching Z score by generation
- Figure 5.7 Simple correlation between % of Adjectives and Linguistic habits for native speakers
- Figure 5.8 Simple correlation between % of Code-switching and Linguistic habits for native speakers
- Figure 5.9 Simple correlation between % of Indicative tenses and Linguistic habits for native speakers
- Figure 5.10 Simple correlation between Average Length of Content Words and Linguistic habits for native speakers
- Figure 5.11 Simple correlation between Tokens and Linguistic habits for first/second/third generation
- Figure 5.12 Simple correlation between Average Length of Content Words and Linguistic habits for first/second/third generation
- Figure 5.13 Simple correlation between % of Verbs and Linguistic habits for first/second/third generation
- Figure 5.14 Simple correlation between % of Pronouns and Linguistic habits for first/second/third generation
- Figure 5.15 Simple correlation between % of Low Frequency Content Words and Linguistic habits for first/second/third generation
- Figure 5.16 Simple correlation between % of Indicatives and Linguistic habits for first/second/third generation
- Figure 5.17 Simple correlation between % of Subjunctives and Linguistic habits for first/second/third generation
- Figure 5.18 Simple correlation between % of Conditionals and Linguistic habits for first/second/third generation
- Figure 5.19 Simple correlation between % of Non-finite forms and Linguistic habits for first/second/third generation

List of Tables

- Table 3.1 Sample of interviewees included in this study
- Table 3.2 Number of informants by decade of migration
- Table 3.3 Informants by age at time of migration
- Table 3.4 Informants by level of education
- Table 3.5 Sample of questionnaire respondents included in this study
- Table 3.6 List of topics covered in the interview
- Table 3.7 List of topics covered in the questionnaire
- Table 3.8 Lexicometric measurements
- Table 3.9 Examples of outcomes
- Table 4.1 Attitudinal factors
- Table 4.2 Linguistic habits factors
- Table 4.3 Linguistic habits outcomes by generation
- Table 5.1 Tokens by generation
- Table 5.2 Guiraud Index by generation
- Table 5.3 Average length of content words by generation
- Table 5.4 High, medium and low frequency content words by generation (%)
- Table 5.5 Vocabolario di Base by generation (%)
- Table 5.6 Nouns by generation (%)
- Table 5.7 Verbs by generation (%)
- Table 5.8 Adjectives by generation (%)
- Table 5.9 Adverbs by generation (%)
- Table 5.10 Pronouns by generation (%)
- Table 5.11 Prepositions by generation (%)
- Table 5.12 Conjunctions by generation (%)
- Table 5.13 Content and Function words by generation (%)
- Table 5.14 Code-switching by generation (%)
- Table 5.15 Verb tenses by generation (%)
- Table 5.16 List of statistically significant differences between the two groups of native speakers
- Table 5.17 List of statistically significant differences between the first - second - third generations
- Table 5.18 Comparison of Simple and Partial correlation outcomes for first - second - third generation

Acknowledgements

I would like to thank my supervisors, Dr. Francesco Goglia and Dr. Aidan Coveney, for enabling me to pursue my research; supporting and providing me guidance whenever it was needed. I would like to also thank the University of Exeter for granting me a scholarship to carry out this study.

I would like to thank Prof. Arjuna Tuzzi and Prof. Flavia Ursini for introducing me to content analysis and sociolinguistics during my studies at the University of Padua. I am also grateful to my high school teacher Sara Fazioni, who has always been more than a teacher to me. She shaped my personality and had a deep positive influence on whom I am now.

I would like to wholeheartedly thank all those who helped and supported me during my fieldwork. Firstly, I would like to thank Mon. Toso Canuto and the 'Trevisani nel mondo' Association. Without their help I would not have been able to carry out my fieldwork in Canada. I would like to also thank the Regione Veneto – Ufficio Flussi Migratori for providing me with the data for my research. I am very grateful to everyone who took part in my research; the many people who participated in the data collection, sharing their stories, thoughts and open wounds with me. Thanks must also go to the people who opened their homes and those who were very kind and hospitable to me. Their participation in my data collection is only a part of what they actually have done for me. What they did 'behind the scenes' can not be in this thesis, but it is something I will never forget. I would like to particularly thank the entire Perizzolo 'Big clan' in Vancouver. Particularly Settimo and Teresa, who were so kind to welcome me at the airport and offer me invaluable support throughout my entire stay in Vancouver. I would like to thank Manuela, who became a key person and such an indispensable help during my data collection. Those aforementioned worked to make my stay in Vancouver not only profitable but also enjoyable and unforgettable. I give my sincere thanks to Maria Teresa 'Resi' in Toronto. I thank you again for what you have done for me; I have been so lucky to meet you. I am very grateful to Giuliana, Imelda, Luciana and Valeria and their families in Guelph and Mississauga. They welcomed me in their homes, helped me in so many ways and I never felt alone. I am greatly thankful to Imelda. Without her help my experience in Canada would have definitely not turned out

the way it did. She followed and helped me through my first steps as a fieldwork researcher, encouraging me whenever I needed it and supporting me to mature as a researcher, but most of all as a person.

I want to thank all my friends in Italy and in the U.K. I am especially grateful to Antonella, Francesca and Chiara. I would like to thank Alessandra, for her sincere friendship and Daniele whom I treasure so many special and funny memories. I cherished lovely moments with my housemates Akiko, Yuan and Isaure. To Isabel, who is always willing to help me whenever I am in need, and to whom I owe the help of translating the questionnaire into English. To Sue, for our enjoyable tea conversations. To Elisa who helped me with SPSS. To Donald and Barbara, for providing me with a peaceful environment to stay whilst I have been in England. I want to thank all those in the Italian Unit at the University of Exeter - in particular Caterina, Elena, Giulia, Iole and Ilaria - for having offered me a constructive surrounding to work in.

Many thanks go to my extended family for their love and encouragement. I am deeply grateful to my parents and my sister Angela for their love and patience. Since the very start of this experience they have supported me in so many ways and they have spent countless hours listening to me whenever I felt down or to share my happiness whenever things were going well. They believed in me more than anyone else.

Finally, I would like to dedicate this work to the memory of my grandfather Franco Bortolato, who I never met. His legacy has nonetheless accompanied me throughout my life. With deep affection and gratitude, I dedicate this work to him.

Introduction

The phenomenon of Italian emigration is a well-researched topic, much investigation having been done within linguistic and sociolinguistic frameworks. What has clearly emerged from research on Italian communities abroad, as well as from studies of other heritage communities, is the importance of looking at each context separately: it is only in the original combination of each wave of migration and country of destination that the final scenarios are forged. Given the scale of Italian migration, there are still aspects to be uncovered, some of which are addressed and partly explored in this thesis.

The great majority of previous studies have dealt with the flow of post-Second World War Italian migrants, referred to in this study as ‘first generation’. Since the drastic reduction in migration from the beginning of the 1970s, research has only partially kept pace. Little attention has in fact been paid to the new waves of migration, henceforth ‘new migrants’. These recent waves do not simply represent a chronological extension of the previous cohorts, but have revealed a divergent development of the social, cultural and linguistic traits of Italian migration. In other words, they expose the need to revise all the frameworks within which past cohorts have been considered, and which were specifically designed for them.

At present, the co-existence of earlier generations of migrants alongside new ones offers an unrepeatably scenario: we can draw a line between two completely distinct groups of speakers. If, with the analysis of the first waves of migrants, we can link this research to previous studies, with the cohorts who followed we are introducing the new social and linguistic scenarios that the most recent waves of migrants are establishing.

In order to follow this line of research, it was necessary to select a country of migration where both old and new waves co-existed. Canada was identified among the possibilities. Even though it did not become a main destination until the beginning of the twentieth century, since the Second World War, Canada has become a migration destination for many Italians, although with great variability in the flow of migrants and a clear peak of arrivals during the 1950s and 1960s.

A second innovative aspect of this study is the adoption of a predominantly quantitative perspective, at times backed up by a more classical qualitative approach. With particular regard to the topic of intergenerational language maintenance among Italian communities abroad, a vast bibliography already exists; however, this mostly takes a qualitative approach, examining and discussing the features that describe this

phenomenon, but rarely addressing it by giving weight to the use of Italian. The current study seeks to provide an analysis that would explore this phenomenon quantitatively, specifying not simply *what* but *how much* has been maintained, on a generational scale. The descriptive perspective, mainly used so far in research among Italian migrant communities, has here been backed up with a second approach, intended to uncover the possible links between the linguistic habits of the community and their actual skills. Hence I will not dwell simply on describing the language abilities of my informants, but I will also try to further develop the analysis aiming at identifying, whenever possible, the causes that gave rise, or at least favoured, these linguistic outcomes.

One last original aspect is related to the data that are discussed in this thesis, which have been gathered for the study during three months of fieldwork in the metropolitan areas of Toronto and Vancouver. The decision to conduct fieldwork in Canada to gather material, in addition to providing originality to the research, also offered the opportunity to structure the data collection specifically to suit the aims of the present study. The data collection thus had a threefold focus. First, it was important to gather material from all three generations (migrants and their descendants) in sufficient quantity to allow statistical treatment. This has also been favoured by the fact that the time is now ripe to track the whole development of language skills among Italian-Venetians, as a substantial number of third-generation speakers are potentially available and old enough to be able to participate in a study which requires language skills as well as a certain level of awareness and self-reflection on their language usage as well.

Secondly, it was essential to base the analysis on present-day data, so that the linguistic phenomena analysed would be of real topicality and so that the study would paint an up-to-date picture of these evolving phenomena.

Thirdly, the sampling, with regard to the two groups of native speakers, namely 'new migrants' and the 'first generation', was structured in order to explore the two main phases into which Italian migration to Canada can be divided. This implied a need to ground this classification specifically in the historical and political conditions of Italy and Canada in recent decades, rather than relying on generic classifications based on a chronological subdivision, which would not account for the real specificity of each situation.

A final aspect that is worth emphasising is related to how I decided to label my informants. The choice of the terms 'migrant' and 'migration', instead of 'emigrant' and 'emigration' or 'immigrant' and 'immigration', has precise connotations, regarding both a more truthful description of this phenomenon and a more accurate prospective

analysis. The term ‘migration’, without a prefix, conveys a nuance of fluidity. People who move abroad, even permanently, occasionally return to the home country. Their status as ‘migrant’ is thus better definable in dynamic terms, as the bonds with their native country are never completely broken (Caltabiano & Gianturco, 2005; Cucchiarato, 2010). In this sense, the more neutral ‘migration’ seems to be preferable. The choice of ‘migrant’ and ‘migration’ also suggests a less rigid perspective to my analysis. My primary aim is to avoid suggesting an Italian viewpoint on the phenomenon, and to offer a more impartial angle; people leaving a country may be considered emigrants by their fellow countrymen and immigrants by their new countrymen. Moreover, during my fieldwork I came to realise that people who had been born in Italy and had lived there for many years might nonetheless feel more ‘at home’ in Canada. Defining them as emigrants or immigrants does not do justice to the wide variety of their feelings and affiliations, but runs the risk of implying a set of false identity perspectives.

This work is divided into three parts, beginning with a review of the literature on language maintenance/attrition and an historical, social and linguistic account of Italian communities abroad (chapters 1 & 2), moving through a methodological description of the data collection and its treatment (chapter 3), to conclude with the analysis of these data (chapters 4 & 5). In greater detail, the thesis is structured as follows.

The first chapter offers an overview of the literature on the continuum of language maintenance - attrition, referring, whenever discussed in the literature, to Italian communities abroad. After a general outline of the main characteristics of the heritage language speaker's proficiency, the discussion moves onto the two main aspects on which the analysis in this chapter is centred: one regarding the extra-linguistic factors impacting on language maintenance and attrition, and the second pertaining to the concrete language skills of speakers of Italian as a heritage language. In more detail, with regard to the extra-linguistic aspects I will explore which factors, related to attitude, the use of the heritage language and the environment, are conducive to attrition and which, on the contrary, favour maintenance. Then with regard to the linguistic factors, I will explore which linguistic phenomena have been shown to be more likely to attrite, and which are those more resilient, focusing in particular on the divergences seen among the generations taken into account in this research.

The second chapter offers a historical, social and linguistic overview of Italian migration. Starting with an outline of the linguistic situation in the Italian peninsula, which explains and supports the reason for this work being confined to informants from a particular area of Italy, the Veneto region, I pass on to offering a brief account of the history of Italian migration. Attention is particularly focused on Canada as a country of destination and on the two most recent phases of Italian-Venetian migration there: the first between the end of the War and 1967, when Canada introduced restrictive immigration policies, and the second after 1970. The political contingency of these restrictions effectively halted mass migration to Canada from Italy, as well as from other countries, drastically changing the flows of people to the country, replacing the mass phenomenon of manual workers with a small elite of qualified, highly-skilled and educated people. I conclude this chapter with a discussion of some features of the Italian language, which I will come back to in chapter 5. This last section is intended to provide the necessary background on the Italian language with which to frame and discuss the analysis of my data.

The second part comprises one single chapter and provides a report on the methodology adopted during the fieldwork. Every methodological choice involves the exercise of discretion, but should be subject to a certain scientific rigour, which has to be made explicit. The chapter begins with an overview of the population investigated in this study, followed by an account of how my three-months of fieldwork in Ontario and British Columbia was conducted. The data collected consists of 99 questionnaires and 56 interviews, with an equal number of informants, and was mostly gathered during my stay in Canada. I then pass on to explaining how the questionnaire and the interview were devised and conducted. In order to fit better the aims of this research and the particular population investigated, I decided to create my own questionnaire and grid of questions for the interviews. In the last part of this chapter I discuss the interviews in more detail, giving in particular a full critical account of the transcription process and highlighting the choices made in this phase while linking them with the contingencies and the aims of the research.

The third and last part of this thesis discusses the findings which emerge from the examination of my data. Two tools were used during the fieldwork, namely questionnaires and interviews, which allowed the widening of the spectrum of analysis to make it appropriately comprehensive for such a multifaceted phenomenon. In

particular, the interview, thanks to its great versatility, has been exploited with regard to both the content of the interviews (metalinguistic observation) and as evidence of linguistic ability.

Chapter 4 explores the findings of the questionnaire. I focus in particular on my informants' linguistic habits and attitudinal factors regarding the Italian language, and in general on Italianness, assessing whether - and if so how - these variables differ between the generations. Although these pages can be taken as an independent discussion of some characteristics of the Italian-Venetian communities in Canada, they are mainly intended to offer a key to the interpretation of the linguistic results discussed in the following chapter.

Lastly, chapter 5 explores the findings of the interviews, relating these to the outcomes of the questionnaire analysis in chapter 4. The perspective of this last chapter is intended to examine how linguistic variables pertaining to lexical richness and verb morphology vary among generations of the Venetian-Italian-speaking community. This examination is firstly aimed at describing my informants' linguistic skills, passing on later to assess whether any observed diversity or trend among the generations is statistically significant. I will eventually investigate whether the variables discussed in chapter 4, which conceivably have an impact on the language skills of my informants, can offer a key to interpretation of these linguistic findings.

Part 1: Literature review

Chapter 1: Bilingualism in a migration situation: maintenance and attrition

The term ‘bilingualism’ denotes a broad and multifaceted concept, applicable to a wide range of linguistic contact scenarios (see among others Weinreich, 1963; Grosjean, 1982; Romaine, 1995; Winford, 2003). In general terms, it can be defined as the “knowledge and command of two or more languages, albeit to different degrees” (Montrul 2008, p.17).

In this chapter the notion of bilingualism will be addressed in relation to what has been defined as ‘heritage language communities’. Although this latter term broadly refers to all people who speak a minority language, hence also ‘indigenous languages speakers’ (Valdés, 2005; Montrul, 2010), in this review it will be considered specifically in relation to migrant communities.

Before discussing the literature on language maintenance and attrition, it is important to highlight the perspective underlying the approach adopted in this research. One of the pivotal concepts in linguistics is that of *continuum*: this conveys the idea that linguistic phenomena cannot be grouped within clear-cut classifications, but they collocate along a scalar line, with two well-defined and prototypical poles, and a plethora of different internal actualizations based on the co-presence, although to different degrees, of both the concepts at the two poles. With regard to bilingualism among migrant communities, and in particular with the perspective adopted in this study, the approach is actualized in terms of a dichotomy: *maintenance versus attrition*. These two linguistic phenomena are then seen as two sides of a single phenomenon. We cannot say that a speaker attrites or maintains a language in absolute terms, but that s/he concurrently presents features of loss and features of maintenance. Some speakers may preserve their language to a remarkable extent even after a prolonged period in a foreign country with reduced inputs in the L1, and thus may be close to the pole of maintenance. Others may show clearer signs of loss, thus being closer to the pole of attrition. The position of each speaker, as will be discussed in this chapter, is the result of many different variables; none of them can be taken as fully explanatory by themselves: it is the specific combination of all the variables that ultimately determines the location of each speaker’s performance along the continuum *maintenance versus attrition*.

1.1 Definitions

When dealing with bilingual people, the literature broadly assumes two different perspectives: native and foreign, or L1 and L2 speakers. These have been discussed from the particular dichotomy of ‘complete’ versus ‘incomplete’ speakers (Benmamoun, Montrul & Polinsky, 2010). Native speakers have often been regarded (although this stance is somewhat questionable) as people who “have a fully developed system for the production and processing of the phonological, morphological, syntactic and discourse patterns of their languages” (Benmamoun, Montrul & Polinsky 2010, p.6), as they have had the opportunity to learn the language in a predominantly monolingual environment with uninterrupted exposure over a long period of time and usually the support of formal education. Thus they eventually display a high level of proficiency in terms of pronunciation, size of their vocabulary and grammatical competence (Benmamoun, Montrul & Polinsky, 2010). Conversely, non-native (foreign) speakers “tend to exhibit persistent signs of non-target acquisition, particularly in areas of phonology, inflectional morphology, and syntax-pragmatics” (Benmamoun, Montrul & Polinsky 2010, p.6). In fact, they tend to begin contact with the new language after puberty (after the critical period¹), when the acquisition of another language cannot, in the very great majority of cases, be attained to a native-speaker level. Moreover, if not migrating to the new country and attending classes, they – as learners of the L2 in their L1 country – can benefit from quantitatively reduced inputs in the L2, usually only from a formal education environment.

In recent years more attention has been devoted to a new type of speaker, labelled ‘heritage speaker’. This is a distinct type of bilingual speaker², who shares some of the linguistic traits of native speakers and others of non-native speakers, as s/he “straddle[s] the boundaries between first and second language acquisition [...]” (Benmamoun, Montrul & Polinsky 2010, p.14). From a merely social perspective, they are “the children of immigrants born in the host country or immigrant children who arrived in

¹ The *Critical Period Hypothesis*, which is widely recognised in L2 learning, suggests that there is an age threshold (about 10-12 years of age) after which the acquisition of a language to native-speaker level is in all probability not possible.

² The discussion on the definition of ‘heritage speakers’ centres on the threshold criterion of mastery of the heritage language. The broader perspective, taken up by Fishman, does not consider proficiency in the heritage language as a prerequisite, but stresses that it is only necessary to “have some affinity with the language which stems from my [sic] family background. So I’m [sic] emotionally attached to the language and I want to learn it” (Kagan, 2008). In the narrow definition – which is adopted in this study – the speaker has to put this heritage attitude into practice and also possess some level of proficiency (Kagan, 2008).

the host country some time in childhood” (Montrul 2012, p.2). Linguistically, their competence in their heritage/minority language(s) does not completely fit that of either of the two groups, native and non-native speakers, discussed above.

As we will discuss later, although heritage speakers differ significantly in terms of level of proficiency, literacy and ability to master different varieties, nevertheless they all share one trait: the acquisition of the heritage language, which starts taking place at birth - and thus distinguishes them from non-native speakers as it potentially may develop as a native language - is at some point interrupted before being fully completed, eventually resulting in their competence not developing to its native-like potential. Consequently, “by the time they reach adulthood the heritage language is the weaker language” (Montrul 2012, p.2). This does not mean that they may not keep on using it to some extent, maybe purely at a passive level within their family, but that their level of confidence in the majority language becomes higher over time, from childhood into adulthood, and that this happens at the expense of the heritage language (Benmamoun, Montrul & Polinsky, 2010).

One last term to be defined is that of ‘heritage languages’, which refers to the “languages spoken by immigrants and their children. Socio-politically, the languages spoken by the wider speech community in the host country are majority languages with official status while the heritage language is a minority language” (Montrul 2012, p.2)³. In this study, speakers of a heritage language are thus considered to be all the people who use this minority language as members of the minority community.

Having offered the basic definitions, as they will be used in this study, I pass on to examining the language development of speakers of a heritage language employing a generational perspective. The discussion will in particular examine the linguistic proficiency of people who migrated to a foreign country as adults (first generation) and their descendents (their children as the second generation, and their grandchildren as the third).

³ Other scholars have used the term ‘heritage language’ in a different way. Polinsky (2007) defined it as “a language which was first for an individual with respect to the order of acquisition but has not been completely acquired because of the switch to another dominant language” (p.149).

1.2 Bilingualism among speakers of a heritage language

The performance of speakers of a heritage language is the result of a bilingual condition. Therefore, if we aim at getting a comprehensive understanding of their level of proficiency, it is essential to consider all the languages involved, and their relation in terms of mutual and uneven balance. As underlined by Montrul (2012), aiming at obtaining the linguistic profile of speakers of a heritage language means that we need to “keep in mind the distinction between the two languages of these bilinguals in terms of order of acquisition of the languages (i.e. *first* vs. *second* language), the functional dimension of the languages (*primary* vs. *secondary* language), and the socio-political dimension (*minority* vs. *majority* language)” (Montrul 2012, p.2).

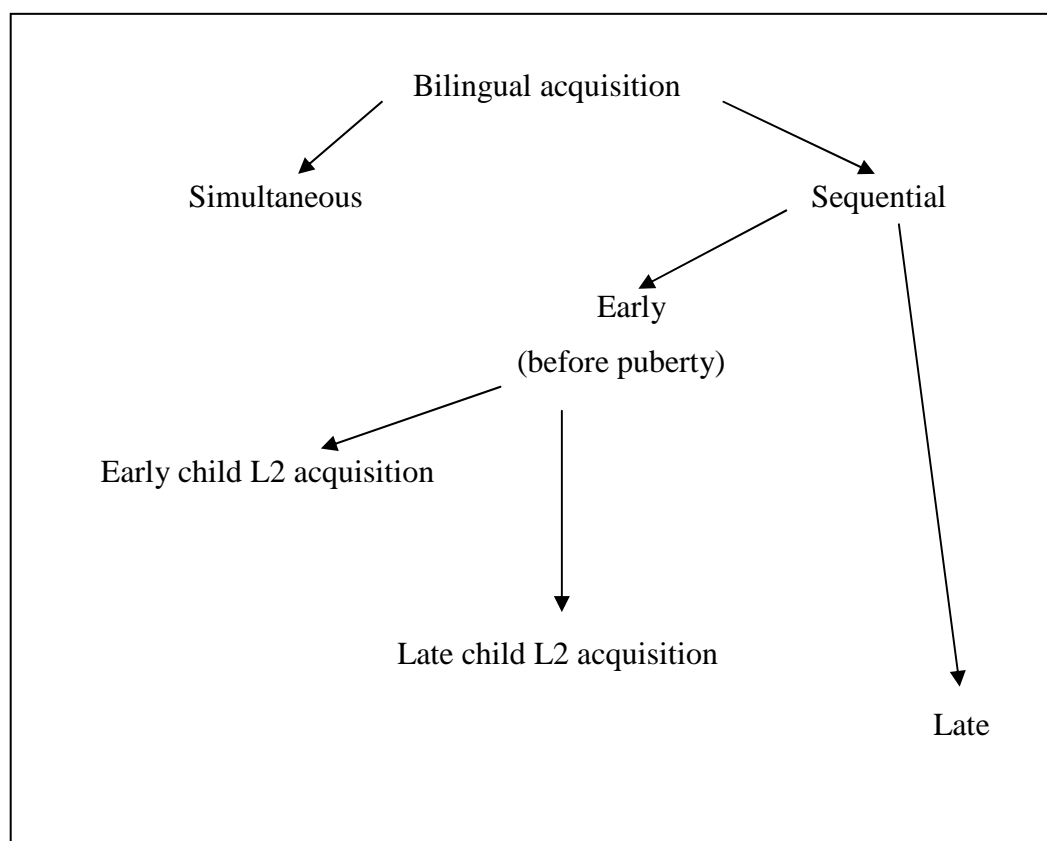


Figure 1.1 Typologies of bilingualism by age and sequence of acquisition

[Adapted from Montrul 2008, p.18]

1.2.1 First versus second language: time and sequence of acquisition

Acquisition of more than one language by a single speaker can occur through two different sequences: simultaneous, if the speaker acquires the two (or more) languages from birth and sequential, if the languages are acquired one after the other. Whereas in

the great majority of cases the third generation comes in contact with both the heritage language and the majority language concurrently (simultaneous bilinguals), the second and first generation normally present a different sequence of acquisition (sequential bilinguals), living their first years from birth in a monolingual environment, a span of time that can vary from birth till pre-school for early child L2 acquirers, to when they enter school for late child L2 acquirers, or to after puberty for late L2 acquirers. In other words, whereas the great majority of the second generation come into contact with their heritage language(s) within the family domain and the dominant language of the country where they live, through peer contacts and particularly through school, at a very early stage of their life (before the critical period), thus potentially allowing them to become perfect bilinguals, the first generation (those who migrated as adults) usually come into contact with the L2 at a later stage in life, when the L1 has already been consolidated and native-like attainment of the L2 is rare.

From this perspective, people who migrated as adults (sequential/late bilinguals) are the first generation, their children (sequential/early bilinguals) the second, and their grandchildren (simultaneous bilinguals) the third (Montrul, 2012). Yet, we must be aware that any attempt to group people together according to socio-demographic variables presents some flaws, as every case is unique and may develop divergently from the standard model. For instance, the pattern with regard to sequential and early bilinguals is even more variegated, in that birth order is particularly influential in the final attainment of second-generation speakers. Often, the status of the heritage language within the family domain is threatened by the introduction of the L2 by second-generation children who have been socialised with this language at school. While first and only children do not have siblings to talk with in the L2 at home, so often remain L1 monolingual within the family domain, those children who have older siblings usually have L2 interlocutors who are available (and usually willing) at home, so generally revert to L2 at an earlier age, depriving them of the opportunity to use and stabilize the L1. These internal differences between first/only children (sequential early/late child L2 acquisition) and their younger siblings (sequential early/early child L2 acquisition) in the second generation as regards their L1 competence are reported to be particularly significant (Bettoni, 1986). It may also be the case that some of the second generation fit in the simultaneous group, as they may experience the contact with the two or more more languages from birth. Equally the third generation, as the grandchildren of people who migrated to a foreign country at an adult age, may come in

contact with the majority language early in their life but at a later stage compared to the heritage/minority language (sequential early bilingual).

1.2.2 Primary versus secondary language: prevalence of usage

The two or more languages available to a speaker of a heritage language may be also regarded in terms of their functionality. A language is defined as primary when an individual “speaks it predominantly throughout the adult life. [..... But] if an individual dramatically reduces the use of his/her first language A and switches to using language B, then A is characterized as this person’s first/secondary language, and B becomes the second primary language” (Benmamoun, Montrul & Polinsky 2010, p.10). This switch appears particularly frequent among heritage speakers. However it can also take place in the first generation. With regard to heritage speakers, in the course of their childhood they show a functional shift of the two languages: the heritage language, the primary one, switches to become secondary in terms of language use. Heritage speakers, in fact, use the heritage language quantitatively less and less, to the advantage of the majority language. This runs parallel to, and is also triggered by, the development of their social networks: from a network based on their family and extended family (and possibly neighbourhood) in their early childhood, to one comprising also school and peer groups, in their late childhood. This shift in the prevalence of use of the majority language may also be linked to their sensitivity to the more ‘prestigious’ language (in most cases the majority language) and their literacy in the two languages. A higher level of literacy implies the use of the language, but heritage speakers “with few exceptions, [...] receive their formal education entirely in English [the language of the host country] [...] and, as a result, become literate only in English” (Valdés 2005, p.413). This condition then does not allow them to keep pace with the age-appropriate levels of monolingual development or to be equipped to employ their heritage language in a wide range of uses.

1.2.3 Minority versus majority language: socio-political status

In terms of a comparison of the socio-political status of the two or more languages, these can be either majority or minority languages. A minority language is the language of the migrant community (heritage language) and it usually enjoys a lower prominence and status within the society. The majority language is instead “typically the language spoken by an ethnolinguistically dominant group [...]. It has a standard, prestigious,

written variety used in government and the media, and it is the language imparted at school” (Benmamoun, Montrul & Polinsky 2010, p.10). On a collective level, this discrepancy triggers a different level of accessibility of these languages in school curricula and in public life. At an individual level, the socio-political status of a language also contributes to shaping of the attitude and feelings of its speakers, ultimately contributing to its higher or lower level of use (Montrul, 2012).

Before taking up again the discussion of the linguistic competence of speakers of a heritage language, I want to refer again to the notion of continuum with regard to heritage communities. As discussed above, the level of ability of these speakers in the heritage language can be placed between the two prototypical poles of maintenance and attrition. However, with the exception of a few speakers with a native proficiency level, their linguistic competence has to be recognised as distinct, even ‘defective’, in comparison with that of native speakers in a monolingual environment.

Comparing their linguistic performance with that of speakers in a monolingual environment, an interplay of different factors shaping their linguistic production emerges. The three key ones will be discussed below: incomplete language acquisition, language attrition, and transfers from an L2 or from a contact variety that has developed within the migrant community (Valdés, 2005; Rothman, 2007; Polinsky, 2011).

Whereas heritage speakers, namely the second/third generation, and possibly the succeeding ones, usually experience all three phenomena, albeit to very different degrees, late bilinguals, namely the first generation, who have migrated after the critical period and have had the chance to fully develop their competence in the L1 and to be literate in the L1, only experience the last two. The ‘incomplete language acquisition’ works as a threshold criterion between the first generation and the following ones. The three following sections address the competence of speakers of a heritage language, each analysing a different facet of the phenomenon. In the first section I treat the factors that shape the language competence of these speakers, namely incomplete language acquisition, language attrition and transfers. In the second, I discuss the factors, sociolinguistic and environmental, that in all probability impact on the language competence of the community investigated. Thirdly and lastly, the linguistic outcomes are presented, with particular focus on a comparison of fully competent speakers (first generation) and heritage speakers (second and possibly following ones).

1.3 Competence of speakers of a heritage language

1.3.1 Baseline

Before proceeding with the discussion, it is crucial to define the concept of baseline in relation to speakers of a heritage language. This can be broadly defined in terms of the actual language a person is exposed to (Polinsky & Kagan, 2007). When discussing language loss, both in terms of incomplete language acquisition and language attrition, we are assuming the existence of a yardstick and drawing a comparison between our informants' linguistic performance and this yardstick.

The definition of a yardstick is therefore essential to carrying out an accurate analysis. However, it is often not simple or straightforward to do. Certainly we cannot consider the standard language, that taught at school or used by the mass media, as a baseline (Polinsky & Kagan, 2007), particularly when dealing with heritage speakers. In fact, "the baseline language for a heritage speaker is the language that s/he was exposed to as a child. Since heritage speakers are typically not exposed to the language norm through formal schooling, the baseline should not be identified with the standard language available to fully competent speakers" (Polinsky 2008, p.41). On the contrary, heritage speakers often come into contact only with a spoken variety of their heritage language. Hence, if we want to establish our informants' actual baseline, we need to acquire a "knowledge of demographic patterns (who settled where and when) and a good understanding of dialectal and/or register differentiation in a given language" (Polinsky & Kagan 2007, p.373). A similar approach has also to be taken for the first generation: full speakers of the heritage language who moved to a foreign country as adults. Although they usually have received formal education in their home-country and had the opportunity to enjoy inputs in the standard language, it is misleading to set the comparison against the standard. This latter is in fact a prototypical concept, the language as it has been codified in grammars, but it is questionable if it is actually performed by speakers. Moreover, if we look at the real performance of monolinguals, we can note that all speakers, even fully competent ones, make errors while using their mother tongue (e.g. de Bot, 2004; Köpke & Nespoulous, 2001; Ribbert & Kuiken, 2010). Oral production, in particular, is affected, as the time for planning the sentence is significantly reduced and there is no opportunity to review the text once it has been uttered (Bazzanella, 1994).

1.3.2 Linguistic phenomena shaping the competence of speakers of heritage languages

1.3.2.1 Incomplete language acquisition

“Incomplete or interrupted acquisition [...] is used to refer to the case of bilinguals who never fully acquired one of the languages they were exposed to as children. [...] these are either simultaneous bilinguals or early child L2 learners who were exposed to the second or majority language early in childhood [...]” (Montrul 2005, p.203).

As discussed earlier in this chapter, heritage speakers begin using their heritage language from birth, and this can take place simultaneously with the majority language of the country (simultaneous bilinguals) or at an earlier stage (sequential bilinguals). Yet, their learning development, which at the beginning runs parallel with that of monolinguals, gets weaker and does not keep pace with their monolingual peers. In the course of time it also falls behind when compared to the majority language, which eventually becomes the stronger one. Their development in the heritage language is interrupted by a range of factors: from a temporal perspective, it usually only takes place during childhood (Montrul, 2012) and rarely stretches beyond it. Moreover, in terms of inputs, both the quantity and quality of inputs heritage speakers receive declines. Quantitatively, on a daily basis the heritage language begins to be used less than the majority one, with regard both to the frequency of exposure and to how much it is used by the same speaker (Montrul, 2010 & 2012). Qualitatively, the contexts in which the heritage speakers receive inputs are quite limited numerically, as usually these include only the family domain (and possibly the extended family/ neighbourhood) and input takes place only through the aural medium (Montrul, 2010). Moreover, the inputs that they receive from their parents are not only usually limited to a particular domain (familiar/domestic), but also further reduced through the inter-generational transmission process. In terms of each generation’s language skills and the inputs that they eventually pass on to the following one, there is a gap, with each generation transmitting only a part of its language skills in the heritage language (see Gonzo & Saltarelli (1983) with their *Cascade Model*). This translates into a decrease of language skills from one generation to another, eventually with the third generation as the ultimate heritage speakers. Moreover, even hypothesizing a first generation transmitting all of its proficiency to the second generation and the second to the third, the lack of formal education⁴ (Caruso, 2004), as well as the absence of different linguistic models

⁴ At this point, it is interesting to refer to the results of an analysis carried out by Clyne (reported by Bettoni & Rubino, 1996). He argues that in order to make improvements, it is important that children

outside their extended families (Scaglione, 2000), would deprive the second and third generations of the wide range of inputs in the language comparable to those received by a speaker in the L1 environment, and thus interfere with native-like acquisition. School, in particular, is recognized as playing a key role in language development and supporting heritage language speakers. In a scenario where heritage speakers already deal with the consequences of reduced inputs from their society and their family, the lack of formal education in the heritage language (through school/language courses) contributes to their falling behind compared to monolinguals (Montrul, 2010; Rothman, 2007), as they usually “miss the chance to learn formal registers along with the vocabulary and complex structures that are typical of written language” (Montrul 2010, p.9).

Hence, heritage speakers may end up acquiring native competence and show high-level proficiency in only a few domains, related to the home/informality, eventually becoming highly skilled mono-style speakers⁵ (Rubino, 2006).

A different view could be developed specifically for passive competence, as it has been shown that heritage speakers’ receptive skills are higher than their productive skills. As discussed earlier in this chapter, second-generation children tend to increase their use of L2 progressively. Over the years, after the L1 monolingualism of their early life, they become first bilingual and then, during their junior school years, more proficient in L2 (Grosjean, 1982). Their productive competence in L1 decreases drastically, while passive competence is still usually fostered by their family’s use of it (Kaufman, 2001). Conversely, the third generation usually cannot count on their parents’ active competence in L1 and so on regular exposure (even if only related to the family domain), which is often possible only when they talk with their older relatives. Their weakness in the heritage language will thus also spread to their passive competence, marking the end of transmission of the minority language.

studying their heritage language attend specific classes with other people of the same ethnic group. If this is not the case, they usually have the chance only to cover what they already know and thus they make no further progress.

⁵ In this vein, Andersen (1982) argues that “a second language learner will learn the language only to the degree to which he acculturates to the target language group. [I]f the learner remains socially and psychologically distant from the target language group, he will speak a linguistically-reduced variety of the target language” (p.88).

1.3.2.2 Language attrition

As discussed above, the level of proficiency of heritage speakers is the outcome of incomplete language acquisition. In fact, “[...] given that the variety of the language spoken by the parents of heritage speakers is the primary (and often only) source of linguistic input in heritage language acquisition, we can raise the question of whether some of the competence deficits in heritage speakers may be a categorical response to a quantitatively diminished and qualitatively reconfigured input” (Benmamoun, Montrul & Polinsky 2010, p.72).

However, their actual skills are also shaped by another phenomenon: language attrition over their life span. Whereas with incomplete acquisition, we are referring to a process of language acquisition that has not been completely accomplished compared to that of fully competent speakers, with language attrition we are considering skills that have been acquired but that for some reasons have also been lost. Although different, the two linguistic phenomena “are not mutually exclusive and can even co-exist with respect to the same or different grammatical phenomena [...]” (Montrul 2012, p.5).

Language attrition – defined as ‘first language attrition’ when it concerns the speaker’s L1 – can be described both as a process and a phenomenon⁶ (Schmid, 2008a).

In terms of *process*, language attrition is seen in terms of a decline in language skills previously mastered by the individual. The focus is thus placed on the decrease in linguistic competence during a span of time between non-attributed and attributed skills. From this perspective, language attrition has thus been defined, among other things, as a “progressive loss” of the ability to use a language (Schmid, 2002), “an overall decline of linguistic proficiency” (Schmid & de Bot 2004, p.215) or a “decrease in the level of proficiency” (Gardner 1982, p.24).⁷ This approach fits well with the idea of a continuum discussed above and the arguments of Andersen (1982) and Gardner (1982) on the subject. Attrition and retention are in fact the ends of a line that can be drawn, but there are several other points in between. Consequently, there are many gradations between a fully competent speaker and a full attriter (Andersen, 1982).

⁶ Another perspective is presented by Myers-Scotton (2002), who suggests looking at attrition primarily as an outcome, suggesting a more static view of this linguistic phenomenon, defined as “the state of any loss [at an individual level] at a point in time” (p.179).

⁷ In the same vein, Schmid (2011a) considers the term “loss” inaccurate in this type of approach, as it implies “a discrete all-or-nothing process [...] whereas language attrition] does allow for a more flexible and gradual interpretation of the forgetting process than the starkly dichotomous *language loss*” (p.3).

In terms of *phenomenon*, that is of loss of language skills previously mastered by the speaker, the perspective moves from a longitudinal towards a more static one. Despite this premise, the debate over the years has offered a spectrum of definitions and thus the discussion is again to be viewed in terms of a continuum. On one hand, it has been argued that if we want to claim that linguistic production is affected by attrition, it is crucial that the outcomes are in some way permanent. Attrition must also affect not only production, but also perception, comprehension and metalinguistic judgments (Pavlenko, 2004). Other perspectives present a more comprehensive view of this phenomenon, defining it as beginning “at a discourse level while the grammatical system remains intact” (Macevichius 2001, p.235)⁸ or suggesting how the starting point of attrition can be traced back to “an increase in the length of time needed for their [certain items] retrieval” (Paradis 2007, p.126), while linguistic production *per se* may be otherwise unaffected.

Whereas there is a broad body of research supporting a decrease in language skills among people who have at one time been fully competent (among others Schmid, 2002; Keijzer, 2007), it has been questioned whether the reduced language competence revealed by second and third-generation migrants (and possibly succeeding generations) can still be treated under the label of ‘language attrition’; as discussed earlier, their acquisition of the heritage language is indeed far from being complete and the lack of competence that we can detect as the result of attrition may in reality be learning that has never been acquired. However, by comparing heritage speakers and full speakers of different age ranges (child versus adult) with a matching linguistic background we can plausibly separate incomplete acquisition and attrition, and study the pattern of attrition among heritage speakers (Polinsky, 2011). So, “if a child performs as his or her age-matched baseline control but the adult does not, the feature can be assumed to have been acquired but may have subsequently been lost or reanalyzed” (Polinsky 2011, p. 306). In her research about relative clause use among Russian-speaking migrants, Polinsky (2011) found that the young heritage speakers performed similarly to the two control groups (full speakers, both adult and children), whereas adult heritage speakers stood out as a separate group. This shows that the language acquired can undergo restructuring since, whereas heritage children perform very close to their baseline, heritage adults' language showed divergent patterns. This outcome appears to relate to

⁸ In her study of referentiality features in two journals, one Lithuanian, the other American-Lithuanian, Macevichius (2001) discovered a mismatch in the use of registers. In particular, the American-Lithuanian journal did not maintain the register appropriate to a written journal, including traits of spoken language in a written text.

their language competence during their life. Their level of proficiency is in fact not only affected by incomplete language acquisition, but, as among fully competent speakers in a foreign environment, it is shaped by attrition as well. This gives support to the conclusion that “in cases when shift leads to loss, it is possible that these adults had more knowledge of their family language at some point during their early childhood, and then lost parts of it, stabilizing at a particular incomplete stage” (Montrul 2005, p.204).

1.3.2.3 Transfer and Communal language

As mentioned earlier in this section, the real competence of monolinguals differs from the standard language. So it is important to move “beyond the monolingual norm [which] must involve the rejection of the standard monolingual language (e.g., standard Spanish, standard Russian) as the norm against which the L1/L2 users are measured” (Valdés 2005, p.422).

The language used by speakers of a heritage language is also affected by other linguistic phenomena, related to the language contact with the majority language, namely transfers from the L2 but also to the influence of a communal language. The latter in particular needs to be briefly dealt with here. It is in fact crucial to be aware how the actual language which speakers of a heritage language are usually in contact with may also be a ‘contact variety’, their heritage language which “has undergone extensive changes through its contact with other varieties of the same language and with the dominant language” (Valdés 2005, p.418).

In this sense, when we are referring to a baseline for minority communities, and in particular for those (the second and third generations) who in most cases rarely receive quality inputs from outside the ethnic community, we are dealing with a different variety to that of L1 monolinguals, enriched through contact with other ‘contact varieties’, and this must be addressed when assessing the language competence of speakers of a heritage language.

1.3.3 Classification of speakers of a heritage language

As has been discussed, the presence of many different phenomena contributes to the shaping of the production of the speakers of heritage communities. The different weight of these factors and their combination help to create an extremely variegated linguistic

scenario among speakers of a heritage language. Hence, if even full speaker communities in a monolingual environment exhibit a variegated internal pattern, this is even more true for speakers of a minority community. In particular, heritage language speakers reveal a wide range (among others Montrul, 2005; Montrul & Polinsky, 2011; Polinsky & Kagan, 2007), as “heritage speakers are not a homogeneous group, but rather form a cline of those who may only understand the language [...] to the very advanced heritage speakers who may simply miss some registers in their language” (Polinsky 2008, p.41).

Although describable in terms of continuum, heritage speakers’ level of language proficiency can be placed along a continuum which is calculated on the distance of the speaker’s variety from a yardstick, namely the baseline language (Polinsky & Kagan, 2007). Speakers with lower levels of proficiency (basilectal) usually have restricted access to the heritage language, limited by the number of their contacts within the domestic sphere and only very marginally within the heritage community, the one which uses the heritage language dominantly (Polinsky & Kagan, 2007). Their competence in the heritage language is usually limited to the passive level with aural comprehension as is the level they are usually more proficient in as they often hear it spoken at home but then they do not use it. Moreover, they usually do not learn how to read and write in their heritage language (Polinsky & Kagan, 2007).

Speakers with a higher level of proficiency (acrolectal) have relatively few limitations in their use of the language and they may be even very close to fully competent native speakers (Polinsky & Kagan, 2007).

1.3.4 Sociolinguistic perspective

1.3.4.1 Sociolinguistic factors

In this section I move the focus of the discussion onto a sociolinguistic and extra-linguistic level, via an analysis of the impact of both the speakers’ background and the influence of the environment variables in preventing and/or favouring language loss among speakers of a heritage language. Despite the attempt to offer a comprehensive overview, the account cannot be exhaustive, as “there is a considerable individual variation among individuals in more or less similar attrition settings” (de Bot 2007, p.63).⁹ Drawing a parallel with the Dynamic Systems theory, de Bot (2007) underlines

⁹ Dynamic Systems Theory is an approach applicable to the analysis of the development of complex systems. It focuses on the study of not only every single variable that affects the system, but on how these variables interact, particularly over time. Each system—and thus also the language development of one’s

how variables in language attrition research (and thereby in that of language maintenance as well) are manifold and diverse, and that even though the same variables are at play, these can eventually interact differently and even change their patterns of interaction over time. This places the outcomes of attrition/maintenance in progress beyond complete prediction, and it also accounts for the differences in some of the research on this topic.

While discussing the evolution of the concept of ‘migrant groups’, Dabène and Moore (1995) suggest how this is settled a priori by two factors: generation and family/ kinship relations. In particular, one’s ethnic affiliation is determined by birth, therefore strictly depending on the generation one belongs to. In the course of one’s life, and in particular during early childhood, it is further reinforced, or conversely undermined, through interactions within the family and extended family. Later on in life, it increasingly comes to rely on one’s own attitudes, and therefore on one’s personal choices and actions. We can here draw a parallel between the evolution of the concept ‘migrant group’, as expressed by Dabène and Moore (1995), and the contribution of the different variables to shaping the development of language skills among speakers of a heritage language. We will start by recalling the impact of the variable ‘generation’, in particular in terms of ‘age at the onset of bilingualism’, expanding the discussion then to the other variables that have been revealed to play a role. We will secondly consider the role of the family in supporting the development of bilingualism among children, both in terms of offering children sufficient inputs in the heritage language and boosting positive feelings towards it. Thirdly we will consider the impact of the use of the heritage language, both in terms of quality and quantity of inputs, in shaping language proficiency. Lastly we will discuss the influence of the external-environmental variables in supporting language maintenance or conversely favouring language loss.

Before reviewing the different factors that contribute to shaping language proficiency among speakers of a heritage language, it should be underlined that this discussion is not intended as an exhaustive review of the factors that trigger language loss or maintenance among migrant communities. It is aimed at taking into account and exploring the factors that are considered significant and relevant to this analysis.

speaker, considered both in the sense of growth and decline— never completely settles, but it is prone to constant adjustments due to change in even a single variable.

Generation

As discussed earlier in this chapter, generation (age and type at the onset of bilingualism) has been considered the key predictive variable in language maintenance/attrition among speakers of a heritage language, having a deep impact in terms of the pace, depth and extent of its effects. Puberty, in particular, is considered to be a watershed: whereas in children the abrupt diminishing or halting of inputs in the heritage language, due for example to the move to another country, has dramatic consequences, eventually even leading to loss of their L1, as in the case of early adoptees to foreign families (Pallier, 2007), the effects among adults in terms of loss remain controversial (Köpke & Nespoulous, 2001). In the setting up of an age threshold, we can draw a parallel with research on L2 acquisition. Research in this field has not come to overall agreement on a precise age for the acquisition of a language as a native, depending also on the fact that different levels of the language require different spans of time for complete acquisition. Nevertheless, a “sensitive period” has been proposed, which starts around 6-7 years of age and may be considered completed by 12-13, in which it is possible to observe a progressive and domain-related decline in one’s chances of acquiring a language as a native speaker. A parallel, although reverse, pattern can be applied to language loss/maintenance among migrants (Schmid & Köpke, 2004). The chances of becoming and remaining fully proficient in the heritage language dramatically decrease if a child leaves the L1 environment well before puberty (Schmid, 2004a; Köpke, 2002a). His/her level of proficiency in the heritage language is also determined by cognitive skills: research has shown that the brain plasticity of young people leads to a faster and much more severe loss of the language following a reduction in the inputs (Köpke, 2004 & 2007; Pallier, 2007). In contrast, after having become a fully competent speaker, language retention is likely to occur even after many years of contact with the L2, although signs of attrition and other language contact phenomena may surface. For adult migrants there is no empirical evidence that the different time lapses since migration have particular effects on the heritage language (Schmid & de Bot, 2004).

Level of education

In language attrition/maintenance research, education is a variable that has so far received little investigation, probably due to its multifaceted nature which makes it difficult to define what counts as ‘education’ (Schmid & de Bot, 2004) and consequently difficult to actually deal with in research. Looking simply at school or

college qualifications has been reckoned somewhat reductive (Schmid, 2002). In fact, although this can be taken as a quite straightforward classification criterion, based on stable and discrete options, one's 'official' level of education may substantially diverge from one's effective erudition and literacy. Broadly speaking, a higher educational level has been suggested as preventing or mitigating attrition, and consequently favouring language maintenance (Waas, 1996; Yağmur, 1997). However, this field is still in need of further investigation (Schmid & de Bot, 2004). From the only two studies in language attrition that have so far considered education as an independent variable¹⁰(Schmid & de Bot, 2004), those by Jaspaert & Kroon (1991) and Köpke in 1999 (in Schmid & de Bot, 2004), apparently contradictory outcomes have emerged, showing an influential role for this variable in the first case but the opposite in the second. These divergent outcomes have been explained by noting that different tasks were used in the two pieces of research to elicit the data. The choice of instruments to use for data collection has in fact been found to have a significant impact on the final outcome. This variable should therefore be viewed from a methodological perspective, being aware that the conclusions that can be drawn are strictly grounded in the method adopted.¹¹

Gender

Gender is a difficult variable to handle as it is strictly correlated with women's and men's social roles, and is thus dependent on other variables such as social networks and education. It reflects a particular society in a specific historical period.

With regard to language maintenance, this variable has been found in general to have a certain influence, particularly along a generational scale, with women heritage speakers of the second and third generations maintaining/performing better than men (Fuller, 2013). With regard to fully competent speakers (first generation), research seems to be lacking, according to Schmid (2002). Yet, the results obtained so far show either no significant difference between men and women in terms of language maintenance/attrition (Leuner, 2008), or limited effects, with women, for example, revealing a tendency to make fewer errors in free speech (Schmid & Dusseldorp, 2010).

¹⁰ Other studies have considered educational level among other variables (i.e. Waas, 1996) without finding this variable qualitatively significant.

¹¹ Milroy & Gordon (2003) discuss the influence of the topic on the language produced. Some categories of people, with a quite similar level of education, may be more confident in using a particular type of test, while other categories may show a better capability in other types. This all has to be taken into account when setting up sampling.

Contact with L1

Contact with the L1 has been considered a key factor in language attrition/maintenance studies. It appears as self-evident that the more the L1 is used in an L2 environment by a speaker and by the whole community, the less attrition s/he will undergo as a single individual, and the higher the chances are of language maintenance along a generational scale, as heritage speakers are offered more chances to practise the language.

This variable has however a twofold nature which needs to be taken into account. While the overall role of social contacts in language attrition has been properly recognized, there seems to be a need to emphasize and single out the influence of quantitative versus qualitative contact. Criticism of an over-generalisation of these perspectives has been made, as they fail to distinguish different types and contexts of L1 use, lumping them all together under a single factor, namely ‘contact with’ and ‘use of’ the L1 (Schmid & Dusseldorp, 2010). If a simple overall view of this variable is taken, its difficulty in terms of quantifiability emerges clearly, as does its many-sidedness.

Moreover we must be aware that, although this variable is apparently concrete and thus easily detachable, it is on the other hand rather difficult to measure. Trying to ‘weigh’ contact is tricky, particularly as we do not normally have direct access to our informants’ contact experience and so it is usually assessed through their (subjective) self-reports.¹²

Language use among migrant communities has been widely studied both in relation to the domain of use and the interlocutor(s). With regard to the first aspect, analyses have focused on the use of the heritage language in various domains, which can be broadly categorised in the two groups of informal (including family, friends, neighbourhood) and formal (work, school, shops). With regard to the second aspect, the focus is moved to the interlocutor(s), and in particular to which language takes preference in use with peers (intra-generational) or with older/younger generations (inter-generational).

From a language maintenance/shift perspective, “if L1 use is reported in these more formal domains *outside* the immediate environment, there is a smaller chance of language shift and attrition” (Hulsen 2000, p.22). Therefore it is important to analyze not only how much a heritage language is spoken, but also when (domain) and with whom (interlocutor).

¹² The usefulness of self-reports is considered to be limited. If they do have some sort of methodological value, this can only be in relation to assessing purely sociolinguistic variables (Schmid & Köpke, 2004).

Social networks are key factors in influencing speakers' linguistic choices. Depending on the characteristics of these links, social networks can counteract language shift within the community. When the heritage community enjoys close-knit and strong social networks the chances of language shift are reduced but if on the other hand these weaken, language shift is expected to take place (Milroy & Gordon, 2003).

Participating in ethnic social networks beyond the sphere of family/friends and neighbourhood provides heritage speakers with more opportunities to use the language in different contexts, and therefore use a greater variety of registers and formality. When the community is numerically substantial and possibly settled in an ethnically dense area, such as in Chinatowns or Little Italies, it may also be capable of providing the community with the necessary resources, therefore reducing their need to interact with the mainstream or other ethnic communities. This contributes to the limitation of contacts with the majority language and subsequently to limiting language shift. Moreover, "network structure, in turn, is heavily related to language usage: as intergroup contact often involves two languages, the network structure will determine the language usage patterns. The more numerous are contacts with the dominant group, the more the dominant language is used" (Ehala 2010, p.372).

As has been discussed earlier in this chapter, heritage speakers often become proficient mono-style speakers. Although their level of proficiency may vary considerably, the great majority of them share a limited and domain-related competence. If not supported by formal education and/or a wide and variegated range of inputs, their heritage language skills will remain confined to the family/informal domain. Even the competence of the first generation may be hampered by restricted domain use, as they do not often have the opportunity to use their L1 in the formal domain. However, their competence may also be undermined by the same fact of using it in an L2 environment. Following the suggestions of Andersen (1982), recent studies have shown that generally defective input may undermine the stability of an L1. Schmid (2011b) contends that "both very frequent and very infrequent use of the L1 can accelerate attrition, either through contact-induced change within a bilingual migrant community, or through lack of rehearsal" (Schmid 2011b, p.155). In their everyday interactions, migrants may significantly differ in both the quantity and quality of their contacts with L1: if they mostly converse with other speakers of the heritage language, they probably receive impoverished L1 inputs. Subsequently, although contact with fellow expatriates may have positive effects on the overall proficiency in L1, this also favours an increase of

cross-linguistic influence phenomena (among others, Ammerlaan, 1996; Köpke, 2002b). In contrast, people who are exposed to means of communication¹³ from their heritage country, as well as contacts with fully competent L1 speakers, may preserve their L1 to a higher degree. Contact with competent native speakers ensures not only a higher quality, but also a wider range of inputs. Thus, although people who join social networks with other migrants usually experience fewer accessing or retrieval problems compared to isolated migrants, they appear conversely to suffer in terms of an impoverishment of their repertoire, in the complexity of both morphological and syntactical structures (Stolberg & Münch, 2010) and “in reduction in the variety of speech acts and discourse types realized through the native language” (Py 1986, p.166), showing a different typology of attrition, more prone to contact-induced change and convergence phenomena.

Attitude

Attitude toward the heritage language has been recognized as a key factor, being a good predictor of language maintenance among migrant communities. Among those discussed so far, attitude is probably the variable that deserves particular scrutiny, to explore its impact.

Literature on language attrition presented an early interest in external and ‘tangible’ variables, such as the number of contacts and time since onset of attrition (de Bot, Gommans & Rossing, 1991). Although recognising them as a certain influence, the literature later moved to pay more attention to internal variables, such as attitude, while gradually acknowledging the consequential role of such variables in preventing or favouring the onset of attrition (Schmid, 2002; Köpke, 2004)¹⁴ and pointing out how “emotional factors may outweigh even such well-established variables as age in attrition” (Köpke 2007, p.29).

In contrast with those discussed so far, attitude is not a clearly manifest variable and thus it presents more challenge when gauging it. However, it seems to be the one that better reflects maintenance/attrition in its continuous nature: all the other variables are

¹³ Thanks to the rapid development of communications technology, migrants’ language can keep abreast of changes, although they may not have frequent contact with people living in Italy. As shown by Marazzini (1994), Italian mass media like TV and radio keep up to date with the new linguistic trends and make use of different and new varieties of the language (as well as linguistic sub-codes) used within the whole society.

¹⁴ It has been shown how external factors (such as age or context of migration) are twofold factors, as they directly affect attrition, but they also function as modifiers of attitudinal factors (such as motivation), which in turn have effects on attrition themselves.

dichotomous or progressive, but they are all based on classifications or measurable clustering. Attitude, on the other hand, is a progressive phenomenon with two well-defined poles, but in between a scale of different stances.

We will not here go into a discussion of the factors that impact on attitude, which will be covered in the next section. Here we will limit the analysis to the role of this variable in terms of favouring or impeding language maintenance.

Research on second language acquisition has pointed to attitude as one of the most important factors in predicting the final attainment (among others, Gardner, 1985; Paradis, 2007; Schmid, 2004a). Gardner, in particular, puts attitude/motivation at the core of second language acquisition. In parallel with this stance, a positive attitude toward the heritage language has been pinpointed as favouring L1 maintenance among migrant communities.

In relation to this, a determinant role in supporting language maintenance among heritage speakers is played by parents, but, as recognised by Gardner (1985), in order to be effective in the long run, the parents' role has not to be that of imposing the study of their heritage language(s) on their children, or providing them with enough inputs, but to motivate and inspire them to acquire the language. If a positive attitude is rooted in them, then children will seek out and be willing to be engaged in activities related to the use of the heritage language.

Attitudes do seem to be strictly related to the notion of 'identity', which is an 'unstable' and 'fluid' concept (Prescher, 2007). The desire to preserve/lose one's mother tongue is in fact mirrored in the desire to maintain one's identity or detach oneself from one's origins (Prescher, 2007), and this is true both for native speakers (Schmid, 2002) and for their descendents (Cordero, 2008).

Probably one of the most relevant cases in point is the study by Schmid (2002), who focused her research on first language attrition among German Jewish emigrants to English-speaking countries during the three escalating phases of the onset of Nazi persecution. In contrast with the majority of the other studies, where minorities did not share the same language with the dominant population, the unique feature of this study is not only that German Jews used the national language of their country, but more that the circumstances in which this migration took place were highly distressing. The results support the hypothesis that the degree of persecution endured by different cohorts is highly correlated to the degree of attrition experienced. People who emigrated

in the late phase (when persecution of Jews was intensified) conserved their L1 language less than the other two groups and the (monolingual) control group. Emotions and attitudes have thus been found to have marked implications in language maintenance even among the same migrants, once fully competent speakers: from a cognitive perspective, L1 attrition can be viewed as the result of the combined effect of a strong emotional investment in L2, whatever¹⁵ the final level of proficiency may turn out to be, and a deep-rooted rejection of L1 (Köpke, 2007). Freud's theory [Repression Theory] on the memory effects of traumatic experiences and how people can, more or less consciously, remove data stored in their memory if linked to particularly negative experiences is an important neuro-linguistic predictor of language attrition (Ecke, 2004). Evidence supporting the role of emotions and attitudes in language attrition can be found whenever particularly traumatic events have taken place, as in the case discussed by Schmid (2002).

1.3.4.2 Extra-linguistic factors

In this section I briefly review non-linguistic factors that have an impact on language maintenance and those which are conversely conducive to attrition. This review is not intended to be exhaustive of all the scenarios or variables at play, but it aims at delineating the influence of the non-linguistic variables that are thought to be particularly relevant in this study.

Before proceeding, it is important to underline a few issues. Categorizing the variables that promote language maintenance or trigger language shift is complex and not always as straightforward as it may seem. For instance Clyne offered an overview of different models and taxonomies on language maintenance among migrant communities (see Clyne, 1991 & 2003), discussing objective/concrete factors and subjective/personal aspects that can favour or restrain the development of language shift. All these factors seem to carry some sort of explanatory power, although their capacity to predict linguistic processes and future developments are not singularly adequate. Moreover, not all factors have been recognised to be unequivocally conducive to maintenance or shift: some are instead ambivalent and they can favour either maintenance or shift according to their interplay with the other variables (Kloss, 1966).

This picture becomes even more complex and tangled if we consider subjective perceptions of objective factors. As Ammon remarks (2011),

¹⁵ Köpke (2004) stated L1 attrition not to be a direct consequence of L2 attainment, and this is usually why L2 attainment is not ascertained in L1 attrition research.

“thus, for example, ‘number of speakers’ or ‘intergenerational language transmission’ are mirrored as ‘belief about number of speakers’ or, respectively ‘belief about intergenerational language transmission’. Such beliefs can motivate individuals towards or against language maintenance efforts and thus function themselves as maintenance factors. [...] language maintenance indicators can multiply along the following lines: objective indicator (high number of speakers) → subjective indicator 1 (belief in high number of speakers) → subjective indicator 2 (positive evaluation of language maintenance)” (p.45).

The interdependence of several factors, including personal variables (subjective and objective), different migration environments and specific communicative situations, is a dynamic concept. Thus, every piece of research has to be read as an attempt to explore the exclusive singularity of the phenomenon investigated from a specific perspective, in that specific context and at that given time.

In this section the discussion will revolve around three main sets of variables: the geographical distance between the heritage country and the new one, the demographical characteristics of the heritage community, and finally its cultural specificities.

Geographical perspective

From a geographical perspective, the physical distance between the heritage country and the host country has an impact on language maintenance/shift. A greater distance from the home country tends to result in less frequent homecomings and also, particularly in the past when travel required more time and was relatively more expensive, it fostered the idea among migrants that they had left for good and that they were permanently settled in a new country (Sartor & Ursini, 1983; Auer, 1991). Moving to a far-away country has been found to usually (although not always) trigger a major language shift (Bettoni & Rubino, 1996). Migrants not only were more conscious that their choice was less reversible, but also had fewer chances to send their children back to their home-country to keep up their heritage culture and language by spending time with their families or even attending some years of schooling (Auer, 1991).

Demographical perspective

From a demographic perspective, the sheer number of people has been reckoned an ambiguous variable (Kloss, 1966), as it may favour the shift in some cases while in others it may bolster language maintenance. The literature suggests the value of looking at the relative concentration of the minority community instead (among others, Clyne, 1991; Bettoni & Rubino, 1996; Lindenfeld & Varro, 2008). A denser community makes the heritage group more visible to the whole community and it also favours a more regular use of the heritage language (Bettoni & Rubino, 1996). Moreover, “the geographical distribution in itself is not the causal factor in language maintenance and shift, but related communication patterns and the absence or presence of daily social pressure to use the prestigious language” (Appel & Muysken 2006, p.36).

Even the time elapsed since the waves of migration is a key factor. If the community has arrived more recently in the new country and if there are continuous arrivals of immigrants to support their community, maintenance is more likely to take place (Bettoni & Rubino, 1996).

The socio-economic position of the heritage community is, on the other hand, an ambivalent factor, as well as the level of education of its speakers (Kloss, 1966; Bettoni & Rubino, 1996). However with regard to the latter, a certain clear influence has been detected if considering the first and the second generation (and possibly the following) separately. The maintenance in the first generation is more likely with a lower level of education, whereas it is the other way around with the second (Bettoni & Rubino, 1996). With a focus on the family domain, the level of endogamy/exogamy of each community is reckoned to be an important factor. A higher level of endogamy favours language maintenance, as the heritage language has more opportunity to be used in the family domain and thus used by the parents with the children. Conversely a higher level of exogamy usually contributes to a faster shift. In particular, “in these marriages [mixed or inter-ethnic] the most prestigious language generally has the best chance to survive as the language of the home, and hence as the first language of the child” (Appel & Muysken 2006, p.35).

Cultural and social perspective

From a cultural perspective, the analysis revolves around the distance of the heritage group from the dominant. The linguistic distance is not very revealing (Clyne, 1992), as we have many examples of related languages whose communities eventually present

different degrees of maintenance. The cultural distance between the two communities seems to be more telling in fact. In particular with regard to Anglo-Saxon culture as the mainstream, according to Clyne (in Bettoni & Rubino, 1996; Appel & Muysken, 2006), the greater the gap between the two cultures, the higher are the chances of maintenance.

With regard to the types of contacts that a speaker of a heritage language may enjoy in a foreign country, revolving around the community as a whole, there are many facets, which are briefly addressed here.

The role of the family in language maintenance is reckoned to be highly significant in terms of language maintenance. When the heritage language is not used within the family, language shift is the natural outcome (Guiberson, Barrett, Jancosek & Yoshinaga-Itano, 2006). Therefore “if a language is not transmitted in the home, it is not likely to survive another generation” (Clyne 2003, p.22). Although many researchers have correctly emphasised the role of the family in the transmission of the mother tongue, family alone is not a sufficient source of support in language maintenance. Ethnic groups and communities outside the family are needed, particularly with regard to the second and the succeeding generations, in order to sustain or increase the level of proficiency of these speakers.

More opportunities to come into contact with and practise the heritage language take place through exchanges with other people of the same heritage group. Heritage associations can offer a vast range of activities, related for example to the cultural, social or sporting sphere. There is also the existence of ethnic neighbourhoods and the occasions to shop in the community shops. In particular the presence of a peer-group, in relation to the young generation, that supports and values the use of the heritage language is reckoned particularly important: it not only helps these young people to develop positive attitudes towards the heritage language, but also offers them opportunities to be exposed to a variety of activities in their heritage language (Li, 2007). Overall “a strong ethno-linguistic community [...] has a paramount importance in facilitating families’ efforts in HL [heritage language] maintenance” (Li 2007, p.19), as not only do they offer the opportunity to access more resources in the heritage language, but also to use the language outside the family domain (Li, 2007).

More specifically (although not simply) with language learning in mind, we have to mention school programmes, which can take place through educational courses in the

heritage language offered to students in school and also Saturday language schools organised by the communities themselves in order to promote language maintenance among the new generations. If in the past, public schools were not particularly committed to language maintenance among their heritage speaker students, or even not in favour of it (García, 2003), in the last few decades the less assimilatory policy taken up in many countries has led to more favourable attitudes towards language maintenance and its addition to school curricula.

There has been a debate about the effective contribution of such courses to language maintenance, particularly as these programmes, just offering a few hours of classes a week usually for a certain number of years, may not provide enough inputs to support literacy (García, 2003; Mah, 2005; Park, 2011), while they are also more restricted to conversational skills and thus do not provide opportunities to be engaged in writing activities (Mark, 2011). Moreover, it has been argued that the fact of attending these classes is not significant *per se* for a heritage speaker student. If the class is mixed, with both heritage speakers and students with no previous contact with the language, the effects on the heritage speakers eventually turn out to be very limited if not nil. The heritage speaker students are often studying areas of the language they already know and thus they do not get any concrete benefit from attending these classes (Bettoni & Rubino, 1996). However, it been shown that not all school programmes in the heritage language are ineffective. On the contrary, they are of real potential importance for language maintenance. What we however need to understand is that their contribution cannot be taken as effective independently from other sources of support for language maintenance (Fishman, 1985). School, in order to promote a language, needs the contribution and impetus of other players in order to bring about an effective and lasting impact.

Overall, the fact of attending classes and receiving education in the heritage language is “necessary to develop the child’s first language” (Appel & Muysken 2006, p.61) as s/he usually benefits from inputs that they can get only rarely within the family; thanks to these educational programmes these new generations can learn to read and write in the heritage language (Appel & Muysken, 2006), thus widening their opportunities for contacts with this language. They also become aware of the standard variety of the language (Appel & Muysken, 2006), as at home they are often exposed to non-standard varieties instead. Moreover, being involved with institutions that value minority language maintenance contributes to shaping heritage speakers’ positive attitude toward

the minority language (Li, 2007), valuing and validating minority languages and cultures in their eyes and somehow in those of the whole community.

A great number of parents, however, are against minority languages being taught to their children in school because of their negative feelings towards these languages (Appel & Muysken, 2006). The negative attitudes towards minority languages, frequently pervading the entire society, are often taken on by the same minority groups; the parents reinforce these feelings and eventually pass them on to the following generations (Appel & Muysken, 2006), significantly diminishing the chances of the new generations attending these courses and eventually the possibility of minority language maintenance.

Also relevant for promoting language maintenance are religious functions, when serving a particular ethnic community or in the case of a heritage language being a *sine qua non* for the practice of that religion (Appel & Muysken, 2006; Leuner, 2008; Bettoni & Rubino, 1996; Park, 2011), and other occasions of gathering promoted by the church, for example voluntary religious groups. However, if a particular language is not an essential aspect of a religion, it is often gradually replaced by the dominant language even if only partially. In fact, in order to promote religious institutions' pan-ethnic mission, services and meetings, once monolingual in the heritage language, usually become bilingual to some degree. This does not necessarily mean that the heritage language will cease to be used, but that specific measures have to be taken in order to partially maintain its use "because their underlying dynamics lead them ever so easily and naturally to the world of the unmarked language" (Fishman 1985, p.371).

The minority community may get administrative services in their mother tongue. This has a positive effect in language maintenance as not only does it validate the use of the heritage language, but also, because the number of interactions with these local or national state authorities are usually quite regular, it lessens the usefulness of the majority language (Appel & Muysken, 2006).

Mass media constitute another important opportunity to come into contact with the heritage language, particularly with the high level/standard language (Campanale, 2006). However, in contrast with contact with the L1 discussed earlier this chapter, in this case the language flow is unidirectional and the participation of the listener/reader is merely

passive. Mass media in general play a key role not only in sociological terms, helping people to remain in contact and be updated with what is happening in their heritage country, but also from a linguistic perspective (Leuner, 2008). Particularly in the case of broadcasts or programmes directly from the heritage country, migrants have the chance to continue receiving inputs, updating, maintaining, and possibly even increasing, their competence. The range of mass media available can be significant and range from TV and radio to periodicals. Nowadays a relatively new medium, the internet, has taken on particular relevance. The internet makes news from the home country easily available and it is more and more dominating or even replacing more traditional mass media (Leuner, 2008).

Practically, we know that the growth of mass media can affect language shift significantly as “broadcasting in minority languages [...] can boost these languages” (Appel & Muysken 2006, p.37). However, this seems to be particularly valid in the case of language maintenance among first generations, as the following ones are recognized as not usually taking advantage of the utilities they have at their disposal to support their language heritage (Schmid, 1993).

Very important also are the attitudes of the mainstream community. In the last few decades, in many countries, we have witnessed a progressive shift from assimilative to multicultural policies, favouring the recognition and development of linguistic and social realities different from the mainstream ones.

The attitudes of the mainstream community, however, lead to ambivalent outcomes (Kloss, 1966): if the mainstream society has a positive evaluation of the minority group it can favour the creation of ethnic institutions, but this at the same time may discourage maintenance as it gives the ethnic group a false sense of security and thus they reduce their efforts in maintenance. A negative evaluation attached to an ethnic group by the mainstream society, on the other hand, favours the development of negative attitudes. This can also spread among the heritage speakers themselves and undermine their self-perception, leading them into assimilation. However at the same time, this negative attitude may trigger phenomena of physical isolation both as individuals and as a community, leading for example to the creation of ethnic neighbourhoods, and thus consequently favouring language maintenance.

1.3.5 Linguistic perspective

1.3.5.1 Linguistic outcomes

In this last section I move on to a discussion of the linguistic phenomena that are considered to be more vulnerable, or conversely more resistant, to attrition. The focus will be centred, in particular, on a comparison between attrition among migrants, namely first generation and new migrants (first language attrition) and the following generations, labelled as the second and third generation.

Earlier in this chapter we noted how heritage speakers of a language do not represent a completely homogeneous group; they are instead very varied, much more so than full speakers. This diversity, as will be discussed in the next section, can be seen also in terms of their competence on different linguistic levels. In fact, “heritage speakers do not develop uniform native-like competence in all modules of the grammar” (Chang, Yao, Haynes & Rhodes 2009, p.81).

With regard to language attrition among native speakers, in parallel with the language acquisition process, the existence of a hierarchical scale in language attrition has been suggested, so “that the attrition process might not be an overall decline of linguistic proficiency, but certain levels or faculties might be affected earlier and more profoundly than others” (Schmid & de Bot 2004, p.215). Attrition is thus reckoned to be a selective process (Köpke, 2004; Selinger & Vago, 1991), affecting some domains more severely than others and following a different time pattern. As will be discussed, the lexicon has been largely proven to be the most strongly affected domain, but evidence of linguistic deterioration is also seen at levels considered more stable, such as phonology and semantics.

If, in fact, we consider linguistic levels separately, we discover that on the whole, “syntax and phonology seem to be the most resilient areas of grammar in heritage speakers, whereas syntax-discourse, semantics and inflectional morphology are quite vulnerable” (Montrul 2012, p.13).

As will be seen below, phonology in particular is a noteworthy aspect to be explored, a sort of threshold in language proficiency distinguishing heritage speakers from second language acquirers, as, though the two groups of speakers overlap with regard to many linguistic traits, “results on phonological competence indicate advantages for heritage speakers, who exhibit more native-like pronunciation than second language learners [...]” (Benmamoun, Montrul & Polinsky 2010, p.73).

Lexicon

Heritage speakers

The lexical competence of heritage speakers is, in the great majority of cases, weaker than that of native speakers (Benmamoun, Montrul & Polinsky, 2010). This applies not only in quantitative terms, but also in relation to the semantic sphere. As discussed by Montrul (2010), heritage speakers exhibit major gaps in their vocabulary. In particular, as we have explored earlier in this chapter, the acquisition of the language, and consequently of the vocabulary, among heritage speakers is often limited to the domestic sphere. Therefore, although they may potentially become sufficiently proficient in this domain, without the benefit of years of formal education in the heritage language and in general with the reduced inputs they get outside the family sphere, their skills are somewhat limited. Hence, the range of their lexical choices is relatively restricted compared to that of fully competent speakers, qualitatively comprising words often “related to common objects used in the home and childhood vocabulary” (Montrul 2010, p.3).

First Language Attriters

The lexicon has been the field in which research on language attrition has predominantly focused (Schmid & Köpke, 2009; Schmid, 2004b) and there is widespread agreement on this level being the most vulnerable to language loss¹⁶ (e.g. Andersen, 1982; Köpke & Nespoulous, 2001), “probably because vocabulary acquisition is a lifelong process” (Keijzer 2007, p.14) and the lexicon is an open-ended system. According to Schmid & Köpke (2009) “numerically, the lexicon is a much larger system than other areas of language knowledge. [...] Furthermore, the lexicon is a network of items that are far less densely connected and interdependent than, for example, the phonological inventory” (p.211).

In terms of reduction (due to disuse of the L1), it has been suggested that attriters tend to differ from non-attriters in their vocabulary, both in terms of quantity and quality: they frequently make use of a smaller portion of the vocabulary compared to that commonly employed by non-attriters, while this vocabulary also usually comprises highly frequent and unmarked items, particularly words linked to their current living

¹⁶ The lexicon is generally the level most likely to be affected by language contact. Lexical borrowing is very likely to take place in a multi-contact situation and lexical items are most likely to be code-switched.

conditions (Andersen, 1982). Moreover, words with common meanings are retained better than those related to specific concepts (Olshtain & Barzilay, 1991).¹⁷

One of the moot points in language attrition touches on the role of L2 influence. The contact with an L2 can naturally generate ‘deviant’ linguistic phenomena in L1, such as lexical borrowing, although this in itself does not imply that one can necessarily consider the speaker an attriter. In terms of influence from an L2, lexicon appears to be the most vulnerable domain.¹⁸ Schmid & de Bot (2004) propose a tri-partition of lexical interference from an L2 affecting an L1: code-switching, words altered in their meaning (semantic extensions, semantic transfers, loan-shifts etc) and inappropriate use of words that are homophonous in the two languages but differ in meaning. While the last two may be regarded as language attrition features, it is questionable whether code-switching can be taken as symptomatic of language loss. Switching into another language may in fact be intentionally driven by different communicative and emphatic purposes. Moreover, without knowing whether an informant has already come across a word (it may simply have entered the L1 since s/he left the home country), it is hazardous to state that the use of an L2 word in its place is evidence of attrition. In this sense, it is safer to take lexical interference as a correlated aspect of language attrition, rather than a diagnostic element.

Given these premises, a shift of focus towards disuse of L1 has been advocated in the empirical approach to the study of the lexicon in language attrition. A more reliable tool has been proposed by Schmid & de Bot (2004), emphasising the reduction of the L1 as a consistent clue to attrition. It has been suggested that “interferences of all these types are easily spotted and analysed in attrition studies, since they show up on the ‘surface level’ of utterances. Much more difficult is to find evidence for a predicted reduction of the vocabulary [...]” (Schmid & de Bot 2004, p.216). In the same vein, Schmid (2011a, p.33) argues that “while the L1 lexical system may be *changed* or *reconstructed* to some extent due to the L2 influence on L1, this change does not actually imply a *reduction* of vocabulary for the speaker” [my italics].

This implies that in order to get an accurate overall account of the real degree of attrition, it is essential to have more inclusive empirical data, which means that the researcher should not limit her/his attention and analysis to what is easily spotted as ‘deviant’ (what is there), as this is probably biased and inclusive of other language

¹⁷ A reduction in the vocabulary of attriters has been suggested by several authors, and evidence found in the study of German Jewish people (Schmid, 2002).

¹⁸ It has even been suggested that of the two factors (disuse and L2 influence), attrition in the lexicon is mostly due to the latter (Laleko, 2007).

contact phenomena, but to what is not immediately detectable (what is not there), in other words, a decreased range in the vocabulary employed.

Morphology

Heritage speakers

Morphology, in particular areas such as inflectional morphology (Montrul, 2010), is the level where heritage speakers appear to display more flaws (Benmamoun, Montrul & Polinsky, 2010). Among heritage speakers with a lower level of command of the language (basilectal) an over-regularization of morphological paradigms has been noted. In particular one finds the elimination of irregular and infrequent forms, the over-generalization of words' forms and meanings, and the high level of maintenance of fossilised forms within high-frequency items (Polinsky & Kagan, 2007). Overall, heritage speakers appear to struggle more with nominal morphology than verbal morphology (Benmamoun, Montrul & Polinsky, 2010).

With regard to the nominal domain, “heritage speakers of languages with overt gender, number, and case marking produce a significant number of errors as compared to native speakers or even their own parents (first generation immigrants)” (Montrul 2010, p.3). Other areas where the system is severely reduced and imperfectly mastered are case marking and agreement in noun phrases.

With regard to the verbal domain, parallel problems are noted. In particular, heritage speakers have been found to experience problems with subject-verb agreement, aspect and mood and, to a lesser extent, with tense paradigms (Benmamoun, Montrul & Polinsky, 2010). In particular, the subjunctive mood has been found to be poorly controlled as well as the conditional (Montrul, 2010).

First Language Attriters

Even with regard to morphology, an overall simplification and reduction of the linguistic system has been hypothesized and found in studies of first language attrition. Although it has been held that morphology is the level at which it is most difficult for a speaker to bypass gaps by using avoidance strategies, it seems that even within this level, “avoidance strategies to achieve an overall reduction of inflectional morphology can be developed in first language attrition” (Schmid & de Bot 2004, p.218). Thus, once more, the researcher must not be limited strictly to detecting deviant forms, since to

have a full picture we must also pay attention to the possibly limited morphological system of an attriter.

The effects of language attrition on morphology translate into a use of more analytical structures compared to non-attriters (Schmid & de Bot, 2004; Stolberg & Münch, 2010), but again the selective nature of attrition implies that some morphological features may undergo attrition whereas others may stay stable overall: some agreements (contextually driven NP inflection) have been found to stay substantially unaffected, while gender and plural agreements have shown more signs of attrition. Similarly, late system morphemes were found to be more affected than early system morphemes (Schmid & de Bot, 2004).

Syntax

Heritage speakers

With regard to syntax, this is a linguistic area that is more likely to be completely acquired by heritage speakers, although some areas may present deficiencies (Benmamoun, Montrul & Polinsky, 2010). Overall, the same trend of simplification discussed for other linguistic levels has been noted. In particular, “heritage speakers seem to develop some core aspects of the family language, but their grammatical systems show a marked tendency toward simplification and overregularization of complex morphological patterns and restricted word order” (Montrul 2010, p.5).

With regard to the length and structure of the sentences, if compared to full speakers, heritage speakers have been found to exhibit much shorter utterances while presenting a lower number of embedded clauses (Polinsky, 2007). Moreover, the surface word order in a sentence in a flexible word order language has been observed to freeze (Polinsky & Kagan, 2007)¹⁹.

Vulnerable domains in heritage languages are related also to long-distance dependencies (including pronominal references) and reflexive pronouns. Other problematic areas have emerged with more complex structures such as relative clauses (Montrul, 2010).

First Language Attriters

Syntax is deemed to be one of the most resistant levels, particularly in relation to L2 influence (e.g. Ecke, 2004; Keijzer, 2007 & 2010; Schmid, 2002; Stolberg & Münch,

¹⁹As Montrul (2010) observes, this can also be triggered by a language transfer from English, a language characterized by a SVO order and without overt case markers. The vast bulk of research on heritage speakers involves in fact people who live in Anglo-phone countries, historically characterized by a high level of immigration.

2010). This is an area of language which offers a wide range of options to express the same conceptual meaning, through linguistic structures that present different levels of complexity (Schmid & de Bot, 2004). Although the general expectation should be a preference among attriters for (simple) main clauses over (complex) subordinate clauses (Schmid, 2011a), the use of a syntactic style is often also the result of personal choice. There may be fully competent speakers with simple syntax, without this implying lack of competence. This suggests that taking simple syntax as a straightforward indication of attrition may yield biased results. Moreover, whenever this may be the case, it must be emphasized that “a trend away from more elaborate constructions [...] will often not result in ungrammatical utterances” (Schmid & de Bot 2004, p.218). This makes the search for clues to attrition even less straightforward.

Evidence of language attrition at a syntactical level is discussed by Yağmur (1997) and by Schmid (2002). Yağmur found complex forms of Turkish relative clauses (attained later in the process of language acquisition) to be most vulnerable to language attrition, while Schmid reports a tendency (albeit slight) among German attriters in an English environment to over-generalise the S-V-O structure of English, as well as forms of simplification such as using fewer long, hypothetical and embedded sentences, when compared to a monolingual control group (Schmid & de Bot, 2004).

Phonology and Prosody

Heritage speakers

With regard to phonetics, this is an areas where heritage speakers have been found to reach high levels of proficiency, being “better than late learners at approximating the phonetic norms of their two languages and maintaining cross-linguistic contrasts between similar categories” (Chang, Haynes, Yao & Rhodes 2009, p.14). Even those at a lower level (basilectal) may sound native-like and this advantage is often referred to as one of the main reasons for placing heritage and second language learners on different tracks. Although close to the pronunciation of the baseline, they usually also reveal a slight non-native speaker accent, which singles them out from full speakers (Polinsky & Kagan, 2007). Heritage speakers are usually described as having a good command of phonology, especially if compared to second language acquirers: they ‘sound’ in fact more native-like. Nonetheless, they also differ significantly from native-speaker control groups as they also display non-native phonological features, developing a sort of

‘heritage accent’ which distinguishes them from monolinguals (Montrul, 2010; Benmamoun, Montrul & Polinsky, 2010).

With regards to prosody, heritage speakers have also been found to speak at a much slower rate, compared to fully competent speakers (Polinsky, 2007). An explanation that has been proposed relates to the cognitive sphere and suggests that heritage speakers experience lexical retrieval difficulties, which result in their need for more time to “collect” the lexical items. (Polinsky, 2007).

First Language Attriters

At the top of a hierarchical scale of language attrition stand phonology and prosody. Overall, these levels have proved to be quite stable against the effects of attrition, although not wholly unshakable, since traces of attrition, albeit numerically less significant, have been detected in these two fields too.

Native pronunciation is one of the last attainments in a foreign language, often, but not always, a prerogative of early bilinguals and monolinguals. Although supported by much evidence, research suggests that this is not a categorical subdivision, as there are indications that even foreign speakers may attain a native pronunciation to the ears of native speakers (Hopp & Schmid, 2011), while native speakers may be perceived as foreigners (Schmid, 2002). The capacity to maintain a native pronunciation appears to be independent of external factors and mostly related to personal aptitude (Hopp & Schmid, 2011; Schmid, 2002). In fact, a “higher degree of language aptitude might mitigate the adverse effects of cross-linguistic influence on L1 speech production even after prolonged periods of non-use” (Hopp & Schmid 2011, p.38). However, external factors are not completely inconsequential: a prolonged holiday in the L1 country, as well as the quality of L1 contacts in the L2 country (mainly in a monolingual mode) may be beneficial and a deterrent to the possible onset of phonetic attrition (de Leew, Schmid & Mennen, 2010).

It has been suggested that prosody is affected by attrition to some extent as well, mostly related to the presence of self-interruptions (Stolberg & Münch, 2010) and dysfluency marker phenomena (Schmid & Beers Fägersten, 2010).²⁰

²⁰ An in-depth analysis suggests that empty pauses, repetitions and retractions were generally overused by attriters compared to monolinguals, underlying the accessibility difficulties of attriters in their L1. Conversely, filled pauses were overused only among the group of migrants whose languages were similar (such as German and Dutch), suggesting this to be more related to interlanguage effects.

Semantics

Heritage speakers

Research that has touched on the topic of the possible advantages of heritage speakers compared to L2 speakers, and conversely their deficiencies compared to full native speakers, with regard to semantics, (see among others Montrul, Foote & Perpiñán, 2008; Bolger & Zapata, 2011; Montrul & Perpiñán, 2011; Montrul, 2012) seems to show varied results, which would suggest a dependence on the specific semantic aspect investigated and on the type of heritage speakers, i.e. simultaneous versus sequential, considered.

First Language Attriters

From a semantic perspective, indications of what has been labelled ‘conceptual attrition’ have been found (Pavlenko, 2002; Skaaden, 2005). While not numerically as significant as in other domains (and not easily spotted), conceptual attrition has been seen to involve the L1 of attriters, with a change in their pragmatic and conceptual representations (Pallier, 2007), or with the extension of the meaning of a word in a way that would not usually be acceptable in the L1 monolingual community (Skaaden, 2005).

1.4 Conclusion

This first chapter of the thesis has offered an overview of two linguistic phenomena framed within the context of migrant communities: language maintenance and language attrition.

After a review of the main traits of bilingualism among migrant communities, analysing in particular their use of the heritage language, the focus has moved to a discussion of the particular phenomena that shape their linguistic competence in the heritage language, namely ‘incomplete language acquisition’, ‘language attrition’ and ‘transfer from the L2/communal language’.

The analysis has then passed on to a sociolinguistic perspective, exploring the factors that have an impact on language maintenance/shift among migrant communities. We have discussed in particular the impact of the main sociolinguistic variables, such as generation and attitude, and the influence of the external environment in supporting or conversely impeding language maintenance.

Lastly, the linguistic skills of the speakers of a heritage language have been explored, focusing in particular on a comparison between migrants (first generation) and their descendants (second and third) on different language levels.

In the following chapter the focus is moved towards an Italian perspective. The Italian presence in many communities outside Italy will be explored from a historical, social and in particular linguistic perspective. While in this first chapter, the bilingualisms of speakers of a heritage language have been explored without reference to any particular context, the next chapter will interpret and frame this discussion within the Italian migrant scenario.

Chapter 2: Italian: the language and its speakers

This second chapter moves the focus of the discussion of migrant communities on to an Italian-based perspective. In particular these pages are intended firstly to offer a review of the linguistic repertoire within the peninsula, and secondly to present an historical, linguistic and social overview of Italian communities abroad, with particular reference to Venetian migration and to Canada as a destination.

The chapter is divided into three main sections: in the first one I portray the complex Italian linguistic repertoire (diglossia), which constitutes the real linguistic background of the informants of this study, and the main features of the Italian language. In the second part I discuss the historical and social scenario of Italian migration and of the present-day Italian presence abroad, particularly that of Venetians. In the third and last part, I take up again the discussion developed in the first chapter about the main social variables having an impact on heritage language maintenance, here in the light of Italian communities.

2.1 The Italian linguistic repertoire

In this first section I briefly outline the Italian linguistic repertoire in the peninsula. Although in this study the focus of the analysis is on the use of Italian among heritage language migrants in Canada, and hence the data collection has been devised to assess this, a brief review of the linguistic patchwork of the Italian scenario in the peninsula is provided. This is aimed at accounting for and supporting my choice of dealing only with people who migrated from a particular geographical area of the country, namely the Veneto region. The linguistic patchwork that characterises Italy makes Italian in fact one of the most geographically diversified European languages (Hall, 1980; Dardano, 1996). Thus, when planning research where linguistic differences may have an impact on the outcomes, it is important to create a homogeneous sample. And as we will see, the Veneto represents a relatively favourable context.

In order to gain a better understanding of the linguistic situation in Italy, let us first briefly consider the historical background. Before National Unification (1861), Italy was politically, historically and linguistically fragmented. At this time it was assessed that only about 2.5% of the Italian population spoke what we can broadly define as 'Italian', while the rest of the population used a dialect. Each of the many dialects in

Italy has its own grammar and its own vocabulary, to the extent that if not geographically adjacent, they are not mutually intelligible (Maiden & Parry, 1997). Among these dialects is Venetian, which is characterised by a specific physiognomy, traced according to some major phonetic isoglosses (Devoto & Giacomelli, 1972) and rooted in the history of this region.

Although the percentage of Italian speakers in the population was gradually increasing, still in the middle of the 20th century Italian was not used by everyone, particularly as a spoken language. Although Italian had been the national tongue since the establishment of the National State, the process of the Italianisation of society became effective only at the end of the 1940s, after the Second World War. This was favoured by deep social and cultural changes that affected the whole society in those years, and which eventually promoted the diffusion of Italian on a larger scale. Among these were the increase of migration towards the large industrial centres of the north-western part of the country and in general towards the main cities, with the consequent need to use a common language, and the rise in the minimum level of education.²¹ During the 1950s, however, the linguistic competence of Italians gradually started shifting towards the standard language, thanks mostly to the prominent role of the mass media. Through the increase in mass media use, first of radio and later of TV, even less well educated Italians became acculturated to the use of this language.

Although from the middle of last century the use of Italian has steadily risen and that of dialects conversely been in constant decline, dialects have not completely disappeared in the Italian linguistic repertoire, neither as languages nor in their commingling with standard Italian (Rapporto Istat, 2006; Marcato, 2002).

2.1.1 Diglossia, Dialectal and Regional Italian

The sociolinguistic condition known as diglossia was discussed most famously by Ferguson (Berruto, 1995). It is described as the co-existence of varieties of the same language used by a speech community to fulfil different functions. The use of different codes is “dependent on each code’s serving functions distinct from those considered appropriate for the other” (Fishman 1967, p.29).²² This notion implies the existence of

²¹ The first milestone was the introduction of compulsory primary education during the Fascist regime. The extension to the secondary level was put into law in 1963.

²² This is the criterion that Berruto (1995) recognizes as crucial to distinguish diglossia from bilingualism. In the latter, the languages involved are not socio-functionally differentiated. Ferguson (1971) [in Grosjean (1982)] and Romaine (1995) suggest that the difference between diglossia and bilingualism lies in the fact that diglossia is a lasting societal arrangement, whereas bilingualism is a changeable condition.

social values connected with the linguistic system. In particular, Ferguson proposes a binary functional distinction between a high, prestigious variety and a lower one. Since the original definition of diglossia, several linguists, and Ferguson himself, have suggested revisions (Berruto, 1995).

In relation to the Italian repertoire, the notion of diglossia was applicable up to the second half of the 1960s, when the domains of use of dialect (low variety) and Italian (high variety) were kept separate by their speakers. However, the raising of the school leaving age, with younger generations becoming more and more Italophone, triggered major behavioural changes. As a consequence Italian entered the family domain, until then the exclusive realm of dialect, and became the language of family interactions. Thus, the compartmental use of dialect and Italian, which is a feature of diglossia, became blurred. The linguistic situation of the Italian peninsula has thus evolved into a new entity, significantly diverging from the notion of diglossia. In order to account for this peculiar linguistic repertoire a new term, *dilalia*, was introduced. Although sharing most of its properties with diglossia, dilalia differs in the occurrence of the high variety during informal conversation. This implies that the separation between the two varieties within the Italian repertoire is not entirely fixed. The notion of dilalia, however, assumes a specific connotation in the Veneto Region, where a generally blurred partition of the domains of usage of the two languages is detectable. Not only does the high language (standard Italian) enter the domains where the low language (dialect) is normally required, but vice versa (D'Agostino, 2007; Tomasin, 2010). This triggers a certain permeability in the use of the high and low languages, and an actual non-compartmentalisation of their use. The percentage of those who make use of both Italian and dialect in conversation with strangers (more formal contexts) is higher in the Veneto Region than the national average (Rapporto Istat, 2006) and it shows how the use of dialect is actually not completely limited to the family and informal domains.

From this it seems clear that what we can define as 'Italian' is a very variegated notion and the product of a specific linguistic context. Standard Italian²³ is in fact a purely conceptual entity: it defines an "official" and "abstract" language (Marazzini, 1994), as it has been fixed in the grammars and dictionaries. Since its codification (between the 16th and 17th centuries), Italian has been preserved as a conservative language and 'this

Grosjean (1982) stresses instead that in a diglossic situation, the speaker has little leeway in choosing which language to use.

²³ Other linguists employ a different terminology. Dardano (1996) labels it '*Comune*', to mark it as distinct from the locally-defined Regional Italian.

was possible because Italian was a literary rather than a spoken language' (Lepschy 2002, p.76). In recent years, the regulation deriving from its being based on a literary foundation has started to fade, as Italian has become a widely spoken language too. This has left room for a 'natural' and concrete evolution of the language into a new variety: Regional Italian. This notion identifies the group of local varieties of the standard language, the outcome of its diffusion into communities that are to some extent linguistically influenced by the local dialect. It is characterised by regional traits, particularly in pronunciation, but it is also partly influenced geographically in its lexical and syntactical choices (Dardano, 1996). Morphological divergences are conversely more rare. While standard Italian is the fixed language (although not the one used in most everyday conversations), Regional Italian is the language used by the very great majority of speakers in Italy, who can perceive the peculiarly regional phonetic and lexical choices of their interlocutors, although these do not impede mutual intelligibility.

2.1.2 Italian language features

In this second sub-section I introduce and discuss particular features of the language that will be explored in the following chapters with regard to the corpus gathered for this study.

2.1.2.1 Lexicon

Overall we can say that Italian is a rich language in terms of its vocabulary. Counting the words included in the encyclopaedias and dictionaries, its vocabulary consists of about 427,000. About 47,000 of these words are part of the *Lessico Comune* (Common Lexicon), that is words that are known and used by those who have a medium-high level of education, independent of their profession and personal hobbies. However, the words that everyone commonly uses are far fewer, although they can still cover all the necessities of everyday life. These words constitute the so called *Vocabolario di Base* (Basic Vocabulary) of our language and are about 6500 words, and 98% of our speech consists of these words. This latter group is further divisible into three parts: *Lessico fondamentale* (Fundamental lexicon), *Lessico di alto uso* (High usage lexicon), and *Lessico di alta disponibilità* (High availability lexicon).

Lessico fondamentale (also *VdB1*) is made up of about 2,000 words and includes words with a very high frequency, words that we learn while we are children. About 90% of

our speech is made using these words (i.e. most prepositions and some frequent conjunctions, nouns and verbs).

Lessico di alto uso (also VdB2) is made up of 2,500 words and these are words that we learn at school. They account for 6% of our speech.

Lessico di alta disponibilità (also VdB3) comprises about 1,900 words. These words are not so frequent as the words belonging to the first two groups, but they are all comprehensible to everyone and they constitute 1-2% our speech.²⁴

Another important feature that we need to briefly mention is that there are differences in the lexicon according to the regional variety (italiano regionale). We may in fact have different words to denote the same object, according to the region where the word is used; for example for the word *anguria*, which is typical of northern Italy, *cocomero* is the corresponding word in the central region and *melone/mellone* in the south (Marazzini, 1994). These variants are defined as ‘geosynonyms’. Hence, in Italian there are words that belong to the Italian language in all regional varieties (the very great majority), but there are also words that are typical of some areas and not widespread across the whole country. Although in the course of the last century many of these words have spread to other regions and become shared by more varieties, eventually even entering the Standard, we need to be aware that these lexical differences are still noteworthy and may also be words belonging to the core of the Italian vocabulary (Vocabolario di base) (Dardano, 1996).

2.1.2.2 Verb morphology

Italian presents five main moods/tenses: Indicative, Subjunctive, Conditional, Imperative²⁵ and Indefinite tenses (Infinitive/Participle/Gerund).

The indicative is the most used mood, and it serves to express a certainty or strong probability. The subjunctive express a hypothetical or uncertain state: it is known as the tense of ‘possibility’, used to express feelings, opinions and wishes. The Conditional is the mood that indicates uncertainty because the effective taking place of the action described is subject to particular conditions. The Infinitive/Participle/Gerund (defined

²⁴ Tullio de Mauro and his collaborators compiled the first two lists (VdB 1 & 2) using the most frequent lemmas of the Italian lexicon. While doing this, however, they realised that there were many lemmas known by everyone that did not eventually enter the list. This was due to the fact that they were not very common in everyday conversation (i.e. *forchetta* ‘fork’ or *pepe* ‘pepper’). Hence they decided to create a third list, containing these lemmas (VdB 3).

²⁵ The Imperative is simply mentioned here but not discussed, as it is not considered in the analysis.

also as ‘indefinite moods’) are verb moods used mainly in subordinate clauses. They indicate an action in itself (infinitive):

Sono solita sorridere quando vedo bambini giocare.

I usually smile when I see little children playing.

or describe an action in an adjectival form (participle):

La ragazza era sorridente per la gioia.

The girl was smiling in amusement.

or an action that is happening concurrently with another or the modality in which the main verb takes place (gerund):

Stavo sorridendo quando è entrato nella stanza.

I was smiling when he entered the room.

More interesting is the exploration of the actual use of these moods, in particular of the indicative, the subjunctive, and the indefinite moods. There is an interesting discussion among scholars (and beyond) about the growth in the use of the indicative at the expense of the subjunctive, observed also in other Romance languages. This tendency appears in the spoken language, whereas studies have shown a substantial stability of the subjunctive in the written language²⁶. Moreover, the occurrences of verbs in the subjunctive, as well as in indefinite moods, are a signal of subordinate sentences. The Subjunctive, in contrast with the indicative, is used only in subordinate clauses and it is therefore an indicator of a more complex syntax.

2.2 Italian migration

This second section is intended to provide a historical overview of Italian and Venetian migration. By offering a historical excursus, these pages will also highlight the social and cultural backgrounds of the different cohorts of migrants, in order to embed the linguistic analysis in a more comprehensive and explanatory framework.

The aim of these pages is thus twofold. First, they are intended as a historical overview of the Italian diaspora. Particular reference is made to Venetians and to Canada as a migrant destination. In order to account for the cultural, social and linguistic

²⁶ Overall, notwithstanding the real consideration that the indicative is used where the subjunctive should be used instead but not vice-versa, it seems that this development is proceeding more slowly than expected and that the subjunctive is far from disappearing.

development of Italian communities abroad, it is crucial to define the historical framework. When the extensive migration to Canada began, in the early 1950s, Italians found there a pre-existing community which had already established the grounds for the development of an Italo-Canadian identity and society, including the creation of predominantly Italian neighbourhoods (Little Italies) and diverse associations and mutual aid societies. The host community had already developed preconceptions (mostly negative) concerning these new waves of Italian migrants, leading to a certain degree of isolationism among these new cohorts and consequently to language maintenance.

Secondly, it is important to be aware that any linguistic analysis should be accompanied by a careful appraisal of the cultural, social and historical background, highlighting the resulting linguistic implications. Therefore, the discussion here aims at identifying phases and passages into which Italian migration can be divided. Particular attention is given to Canada and the dividing line of 1967, which signalled a break in the flow of Italian migration into the country, marking as it did the introduction of a new immigration policy which both dramatically reduced the number of arrivals and was conducive to the rise of a new type of highly qualified migration. This excursus is thus also intended to underline the differences in attitude, towards the host country as well as the motherland, among different cohorts of migrants, with specific reference to the two groups included in this study.

2.2.1 Historical review

The phenomenon of migration is considered an important chapter of Italian history, which helped to shape both the society and the culture of this country within its borders and beyond. It also marked the establishment of very substantial Italian communities abroad. Although migration has always characterised the history of this country, the phenomenon gained particular historical relevance after the unification of Italy in 1861. Conventionally, the onset of the period of migration is taken to have occurred in 1876, when departures began to be systematically recorded, thus providing a more precise account of this phenomenon. Between 1876 and 1973, about 27 million Italians officially left the country, although the real number is thought to be higher. This figure, however, refers to the total number of people who experienced migration, including about 14 million who moved permanently and about 13 million for whom migration was either seasonal or temporary.

The hundred years of Italian migration can be subdivided into three phases (with the exclusion of the new waves of migrants, which are usually considered separately), considering the two world wars as watersheds (Franzina, 1995). An investigation of the destinations reveals a clear preference for Europe, particularly for France, Switzerland and Germany (www.emigrati.it). Beyond Europe, the United States was chosen by 50% of those who moved to extra-European destinations, with Canada accounting for only 5.6% (www.emigrati.it). As for origins, while the migration drain hit southern Italy much more than the north, a regional analysis reveals that the Veneto was affected more than any other: about 3.3 million left this region, 12% of the total number of departures²⁷.

2.2.1.1 From the origins to the 1960s

Before the opening of the mass migration chapter (1876) there was an Italian presence abroad, although comprising just a few hundred people. This small 'elite' was composed of people with a certain level of education, or rather a specific professional skill, who offered their competence in various cultural and artistic fields. They were appreciated by European elites and aristocrats because of the reputation of Italian culture (Franzina, 1995).

Harney (1984) considers this early phase as particularly influential in shaping foreigners' perceptions of Italians. Since the Renaissance, people abroad had become used to identifying Italians as renowned and talented artists. Italy itself was associated with the image of a land of arts, a suggestion that was confirmed by these first cohorts of migrants. The impact of the ensuing masses of poor and backward peasants who started moving abroad at the end of the 19th century would dramatically change this picture, triggering the rise of indigenous hostility towards Italian migrants.

During the second half of the 19th century and the first two decades of the 20th, economic contingencies in Italy dramatically changed the migration phenomenon, in terms both of the number of people involved and of their educational background (Franzina, 2006; Vianello, 2006). A small elite group became a mass migration of

²⁷ In order to compare the percentage of departures from the Veneto region on a national scale (at 12%) with the percentage of population of the region (again on a national scale) I checked the data of four years, 1955, 1965, 1975 and 2013 (the first three chosen with an interval of ten years, during the period of mass migration). The results show a stable pattern, varying from 8.6% in 1965 and 1975, to 8% in 1955 and 2013 (www.istat.it). We can thus say that the number of people in the Veneto accounts for about 8% of the national population. However this region contributed 12% of the number of migrants.

millions of poor people.²⁸ It especially involved peasants from the countryside of the centre and south of Italy, but also from the north eastern areas of Triveneto and Emilia Romagna (Franzina, 1991; Lazzarini, 1981).

After about four decades, with the rise of the Fascist regime in the 1920s, Italian migration reached an effective standstill (Gallo, 2010). In 1928, Mussolini decided to make migration illegal. In addition, the world crisis of 1929 and the consequent closure of several countries to foreign immigration contributed to a temporary halt in the phenomenon.

During the third and final phase, which began after the Second World War, about 7.5 million people emigrated (De Clementi, 2010). The economic state of Italy after the conflict was critical, characterised by an exceedingly high rate of unemployment and a general condition of competitive deficit compared to other European countries. In order to boost the national economy, bilateral agreements were signed with European and South American countries,²⁹ offering labour in exchange for raw materials. The leading destinations were again mainly European (Switzerland, Germany and France), although new countries entered the array of potential destinations and Canada was amongst those. From the 1960s onward, the waves of migration gradually diminished. The combined effect of improved economic conditions in Italy and the reduced wealth of the countries to which Italians had migrated increased the number of repatriations significantly. The year 1972 is considered to be the closing stage of mass migration from Italy. From 1973, the balance of the flows in and out of Italy became positive in favour of the returns³⁰ (Fondazione Migrantes, 2009).

2.2.1.2 From the 1970s onward

Italian migration in recent decades has not received much attention from historians. Interestingly, books published in the last few years have seldom touched on the topic of new waves of migrants (from the 1970s onwards). This does not mean that this

²⁸ The causes of this workforce drain were many. Franzina (1995) states that it is only partially correct to point to endogenous factors, in particular the demographic increment and the subsequent need to reduce the population, as the grounds of this phenomenon. Other dynamics were simultaneously at play within the Italian situation, especially the difficult economic and agricultural conditions. But exogenous factors played a key role as well. Italian migration was completely dependent on the demand of countries in need of labour, of which Italy had a surplus. At that time, several states were seeking low-cost manual workers, promoting their countries abroad and financially supporting those interested in migrating in.

²⁹ The countries involved in these agreements were Argentina, Belgium, Brazil, Canada, France, Great Britain, the Netherlands, Luxembourg and Switzerland.

³⁰ De Clementi (2010) suggests connecting Italian migration waves to world economic phases and observes that 1973 represents the initial stage of the recession which hit western countries during the 1970s. Thus, this double correlation is replicated.

phenomenon has reached a complete end. On the contrary, in recent years it seems to have gained a new vigour and to be increasing, becoming a research field particularly for sociologists and journalists; and although during the last three decades of the 20th century these flows slowed down, young Italians have more recently reopened the routes of migration.

From a quantitative perspective, the sudden decrease in migration flows out of Italy in the 1970s is also to be situated in the particular economic conjuncture of that time. The process of ‘inverted migration’, with a significant number of people who had previously emigrated returning to the home country, was only partly the result of a personal (and collective) choice. It was mostly due to difficulties in finding work opportunities abroad. In particular, some European countries with a long history of Italian immigration tightened up their regulations regarding immigration in those years (Fondazione Migrantes, 2009). Canada was among those extra-European countries which introduced restrictive policies, in 1967, putting an end to new mass arrivals and preferring a highly skilled and selective immigration.

From a qualitative perspective, migration from the 1970s marks a clear break with the previous waves, now usually involving people with a high standard of education and/or high-level skills. Moreover, whereas migration in the previous decades was a mass phenomenon, involving entire families and villages (chain migration), what followed was an individual and numerically limited affair.³¹

Although, unlike those who came before, the new waves of migrants have had a high level of education, this is not the only difference. Particularly (but not solely) in the last two decades, along with a ‘brain drain’ in the strict sense, a new type of migration, called *nomadic migration*³², has begun, motivated more by a need for cultural and social change and involving people of different classes, though often graduates. Although these migrants, like their predecessors, may be pursuing better working opportunities abroad, the choice to migrate for some of them can be driven also, and even primarily,

³¹ Other factors in the migration flows were internal to Italy: the political and social conjuncture of the 1970s (terrorism and the Red Brigades) triggered ‘political migration’, not numerically significant, but still worth mentioning (Turchetta, 2005).

³² Although migrants have always had the possibility of returning to the home country, this was often not considered economically feasible. Furthermore, there was a sort of stigma of failure and decline of one’s self-esteem for those who returned to their home country or moved to another country without having reached a certain economic level. Their aim was in fact purely economic, that of ensuring better living conditions for themselves and their families. Nowadays, the attitude towards people who move through different countries, or who eventually return to their home country after a period spent abroad, and maybe leave soon after for a new destination, is more positive (Cucchiariato, 2010). An experience in a foreign country is seen as enriching, whatever the outcomes and the final relocation. The experience of migration has come to be seen from a nomadic perspective, where the move itself is considered an integral part, and a sign, of self-expressivity, autonomy and enterprise (Caltabiano & Gianturco, 2005).

by the rejection of a country perceived as static and inequitable towards younger people and by the search for values and different ways of life elsewhere (Cucchiurato, 2010):

[...] per alcuni neo-immigrati [...], la lontananza dall'Italia rappresenta non solo un atto di rottura, ma anche di denuncia verso una mentalità e un sistema di vita avvertiti come sintomi di una profonda arretratezza. (Caltabiano & Gianturco 2005, p.105)

[...] for some new immigrants [...], the distance from Italy represents not only an act of rupture, but also a denunciation of a mentality and a way of life perceived as symptomatic of profound backwardness. [My translation]

This does not mean that these feelings were not experienced by migrants of previous cohorts, nor that they are a particular trait of the majority of new migrants. Nevertheless, they seem to have turned the sentiments of a few individuals into a more generalized, although not collective, trend.

Although a sense of loss and detachment may still occur in the life of these new migrants, this does not have the force of a double absence, in terms of distance from the heritage of the home country and of only partial acquisition of that of the new country. On the contrary, migrants often experience a double presence, where both (and more) cultural references may interact and integrate (Caltabiano & Gianturco, 2005):

In particolare, la madrepatria non è più la terra mitica di un esodo senza ritorno. Nella vita d'ogni giorno, essa ricompare attraverso un flusso di notizie, immagini, messaggi interattivi e suoni che vengono incessantemente messi in circolo da Internet e dalla televisione. (Caltabiano & Gianturco 2005, p.24)

In particular, the motherland is no longer the mythical land of an exodus with no return. During everyday life it turns up again through a flow of news, images, interactive messages and sounds that are incessantly circulated by the Internet and television. [My translation]

In the globalised world in which they were raised, where the experience of being in contact with different cultures is often part of everyday life, and even sought and appreciated, their decision to move is also rooted in an enthusiasm for being in a cosmopolitan environment, as put into words by one of my new migration informants:

e mi sembra di essere stata in una città totalmente diversa. oppure vado a Richmond [...] e sembra di essere in Cina! oppure vado a Little India, che è giù

verso la quarantanovesima. io sono stato un paio di volte lì ed è veramente incredibile. è bellissimo. è bellissimo. non serve, non serve viaggiare una volta che si è in una città così. si vive, si viaggia in tutto il mondo all'interno della propria città.

And it seems to me that I was in a completely different city. Or I go to Richmond [...] and I might as well be in China! Or I go to Little India, which is down towards the 49th. I have been there a couple of times and it is really incredible. It is so beautiful. It is so beautiful. There is no need, there is no need to travel once you live in a city like this. You live, you travel all around the world within your own city. [New migrant living in Vancouver; my translation]

These new-wave migrants usually seem to have a positive attitude to multiculturalism and often choose to integrate into the host society, without establishing strong contacts with the Italian community, particularly that composed of first generation migrants, or even bypassing it completely. Hence, the very same forces that drove them out of the country propel them, once in the new environment, to integrate into the local community and to refuse (at least partially) contacts with the pre-existing Italian community. As one of my informants of the new migration waves maintained:

mi hanno fatto una lista degli indirizzi dei negozi italiani. pensando che io avessi molta nostalgia.[...] però non li ho un potuti deludere per cui sono stata zitta ho detto molto bene. ma in realtà non mi interessano. proprio per il discorso che non mi interessa frequentare la comunità italiana. perché non mi interessa.

They made me a list with the location of Italian shops, thinking that I was very homesick. [...] I didn't want to disappoint them, so I didn't speak my mind and just said OK, thanks. But I was actually not interested, just because I'm not interested in associating with the Italian community. Because it doesn't interest me. [New migrant living in Vancouver; my translation]

2.2.2 Italian and Venetian communities abroad

During the discussion of migration in recent decades, the data explored are taken from two censuses: the Canadian Census of 2006 and the Italian Aire³³ 2011. The first makes no reference to the regional origin of Italians, so the data are considered on a national

³³ Aire is an acronym for Associazione Italiani Residenti Estero, i.e. Association of Italians Resident Abroad.

scale, whereas when examining those from Aire, specific reference is made to Venetian migrants³⁴.

2.2.2.1 Italian and Venetian migration to Canada, the 'last best west' land

From a historical perspective, Canada was connected with Italy from early times, i.e. 15th century (Zampieri Pan, 2009). However, the phenomenon of Italian migration to Canada, as referred to in the present study, started during the 1880s. It was then characterised by a substantial increase in numbers throughout the decades, reaching its height after the Second World War (Harney, 1984). In particular, an Italian community began to develop only at the beginning of the last century, with the creation of Little Italies and of a sense of Italo-Canadian identity (Harney, 1984; Franzina, 1995).

Since the onset of migration to Canada, Italians perceived themselves to be the target of discrimination by the host society.³⁵ This became a key factor in promoting their partial segregation as a community and, notwithstanding local differences, promoting in the last resort the preservation of their cultural, social, and linguistic identity (Pautasso, 1978).³⁶

With the end of the Second World War, there was a sharp rise in the number of migrants from Italy. On one hand, tough conditions in Italy forced people to move abroad in order to seek economic opportunities. On the other, admission to Canada, in need of manual labour, was also regulated; this happened mainly through sponsorship, a legislative tool introduced by the Canadian authorities in 1948 and in force until 1967. It offered Italians (and people from other countries) the opportunity to enter Canada

³⁴ As De Clementi (2010) remarks, the investigation of the last 50 years of Italian migration is a particularly challenging task, since official archive documents are made available for consultation only after 50 years. Hence, the study of Italian migration during the last five decades is not particularly straightforward, given the scarcity both of official data and of books on the history of this topic.

³⁵ The tag used to refer to Italians in Canada was WOP, said to stand for 'Without Official Papers', in order to denote their arriving in the country without being fully approved. Over the decades the negative value has started to fade, however. As one of my second-generation informants told me: "noi italiani eravamo così tenaci così bravi [...] e la nostra scuola di [nome della scuola] [...] la maggior parte della popolazione erano figli di italiani. hanno messo una grande banner sul ginnasio. Wop are top. che vuol dire che i italiani sono i migliori. tanto per dirti come che è cambiato. da una cosa negativa è diventata una cosa positiva. *We Italians were so tough, so good [...] and at our school [name of the school] the great majority of the pupils were sons of Italians. They put a big banner on the school. Wop are top. Which means that Italians are the best. Just to tell you how much it has changed. From a negative thing it became a positive thing* [My translation].

³⁶ An emblematic episode is the arrest of about seven hundred Italians on the 10th of June 1940 simply because they were suspected of being a threat to the country. This shook the Italo-Canadian community and triggered a wave of resentment, because it was perceived as being against Italians because of their heritage, not because of any real danger.

legally on the condition of having someone already resident in the country willing to act as guarantor and to cover the expenses of the early period of their stay in the country.

Therefore, between the end of the Second World War and 1967, Italian migrants were mostly people with low professional qualifications, as the Canadian authorities initially set no minimum educational requirements to enter the country and, as has just been mentioned, they were particularly recruiting manual or blue-collar workers.

In 1967 the Canadian authorities introduced new immigrant acceptance criteria, based primarily on professional qualifications. Italian migration, as well as that from other countries, changed radically, first into skilled migration and later into a brain drain. The flows of Italian migrants to Canada thus turned into a selective phenomenon: Italians who emigrated in those years were mainly self-employed and/or qualified professionals. Then, in 2002, Canada passed into law the Immigration and Refugee Protection Act, which gave extra weight to certain selection criteria, particularly those relating to educational qualifications, working experience and knowledge of the two official languages of Canada. This favoured the arrival of even more highly qualified newcomers, and can perhaps be called 'high skilled migration'.

Today these communities of Italians are in decline, as they are not supported or reinforced by new arrivals. According to the current Consul General in Toronto, Gianni Bardini, about two or three hundred Italians migrate annually to Ontario³⁷, a very small number compared with the thousands of people who entered the country, particularly Toronto, in the years of mass migration. Italian migration to Canada can no longer be defined in terms of 'flows', as the numbers are very low, particularly if compared to the preceding cohorts. It can thus be better described as a phenomenon of 'individual migration'. The significance of the term is complex. Quantitatively, it marks a sharp decrease in numbers, while qualitatively the differences from the preceding cohort are two-sided: first, chain migration has been replaced by a pattern of single migrants or close family groups, who secondly tend to embrace their new country's culture, language and social life in such a way that their relocation can be better depicted as a single and isolated phenomenon, making them individual migrants as opposed to members of an Italian collectivity.

³⁷ Podcast at <https://fugadeitalenti.wordpress.com> of 17.09.2011

2.2.2.2 Italian and Venetian presence abroad and in Canada

Before proceeding with the discussion, it is important to note that the data to be examined are mainly taken from Aire.³⁸ This is the only available official register of Italians living outside the country and although it is known to be biased in some ways, as not all Italians abroad are registered, it can still be taken as an indicator of the phenomenon today.³⁹

In January 2011, the number of Italians registered at Aire was 4,115,235, a slight increase compared to the previous year, when there were 3,915,767.⁴⁰ This number represents about 6.6% of the whole Italian population. According to Aire (2011), Canada is in the 9th place for the number of Italians residing there.

As stated by the Canadian Census of 2006, the number of Italians residing in Canada but born in Italy was 299,965. People living in the country but with at least one parent born in Italy are 712,420 and are usually counted as second generation, while 432,945 are the members of the third generation (wholly or partly of Italian ethnic origin, half of these being of mixed ethnicity) (Aire, 2009). Overall there are almost 1.5 million Italo-Canadians.

Their distribution in Canada shows a great preponderance of Italians in the consular area of Toronto (the provinces of Ontario and Manitoba), which accounts for more than half of the whole Italian community in Canada, mostly located in the metropolitan area of Toronto (GTA), where they recently numbered more than 485,000⁴¹. Almost 30% of Italians in Canada have settled within the district of Montreal and about 10% in Vancouver, while Edmonton and Ottawa close the list with 5% and 4% respectively⁴².

In the list of countries in which Venetian-Italian citizens are registered, Canada lies ninth, with 3.1% of the total number of Venetian-Italians registered abroad. There are 9,272 Venetian-Italian citizens in Canada who are currently registered with Aire, a number which is far from including all who have actually migrated and who have to be added to the Italian migrants who switched their citizenship.

³⁸ According to Italian Foreign Ministry policy, Italian citizens who move to a foreign country for a period of more than 12 months and not for seasonal work must withdraw their residency in Italy and register with Aire (www.esteri.it/MAE). However, it has been estimated that only about 50% of Italians abroad are actually registered.

³⁹ It has been suggested that the figures are probably underestimated by 50% (<https://fugadeitalenti.wordpress.com>).

⁴⁰ It must be said that this number includes people with dual citizenship: of Italy and of another country.

⁴¹ Podcast of <https://fugadeitalenti.wordpress.com> of 17.09.2011

⁴² www.esteri.it/mae/doc_osservatorio/Rapporto_Paese_Canada.pdf

2.2.2.3 Italian communities in Canada

In this third and last section we reach the core of this chapter: having offered the linguistic, historical and social background of the community investigated in this study, we review the main topic and sociolinguistic variables that we have presented generally in the previous chapter, here specifically with regard to the Italian/Venetian community. In contrast with the nearby United States, Canada can be labelled a culturally and linguistically “fragmented nation” (De Maria Harney, 1998), characterised by a relatively low degree of inducement to assimilate into the mainstream culture. The multicultural patchwork environment provided by this country has had a positive impact on the cultural identities of the different ethnic groups who have moved to live there. A key concept coined by Fishman is that of ‘linguistic ecology’, which offers an interpretation to explain and possibly predict language maintenance among immigrant communities, by studying “the interactions between a language and its environment” (Edwards 2007, p.461). This perspective offers, for example, an explanation for the loss of ethnic language(s) among Italians in the USA: in contrast with other contexts, Italians in the USA have experienced high pressure to integrate into American culture (Bellù, 2009). Canada, Australia and several other countries, in contrast, represent ‘friendlier’ environments, in which Italians, although having sometimes been stigmatized to a certain extent, have been partially favoured in their cultural and linguistic maintenance (Bettoni & Rubino, 1996; Baldassar & Pesman, 2004).

2.2.3 Italo-Canadians between integration and isolation

2.2.3.1 From the beginning of the 20th century to the 1960s

Notwithstanding this relatively hospitable situation, as discussed in the previous section, the Italian presence in Canada was not always welcomed. Prejudices and stereotypes from the mainstream community contributed to the stigmatization of Italians and to their isolation.

After decades of lack of concern from the Italian political class, the Fascist regime became interested in Italian communities abroad (Pautasso, 1978; Franzina, 1995), aiming at inspiring in them a sense of pride and national identification with their home-country, but also promoting an anti-integrationist attitude among Italians, favouring the isolation of Italians from their mainstream and other ethnic communities. It must be underlined, beyond the effectiveness of these policies, that the real ability of Fascism and the concrete results it accomplished were the outcome of its exploiting the Italian

community's perception of a veiled intolerance on the part of the host society, rather than a tangible ability to make its policy have an impact. Particularly after the entrance of Italy into the Second World War⁴³, and in the following years, Italians suffered even more discrimination. This feeling of a perceived hostility was, among other things, at the root of the creation and growth of Italian ethnic neighbourhoods, known as Little Italies.

Little Italies

In parallel to what happened for other ethnic groups, the development of 'Little Italy', neighbourhoods with a high percentage of Italians, represents one of the more fascinating traits of the presence of Italians in Canada and of Italians abroad in general. They can be seen to embody the migrants' need to recreate themselves as a community in the new environment, a tendency that is also represented in a certain capacity of these Italian communities to preserve the heritage language. These communities indeed once acted as the first support centres for migrants arriving in the New World and now represent the legacy of this past within the cultural, social and architectural development of cities.

The tendency to set up Italian agglomerations in specific areas of cities started at the beginning of the 20th century in Canada and was maintained until the end of what have been labelled 'mass migration waves'. During recent decades, though, this trend has begun to fade as Italians have started moving to the outskirts of Canadian cities, while even the attitude of the mainstream culture towards Italians has changed significantly. The establishment of new suburban neighbourhoods has triggered social (and possibly linguistic) changes, both in the case of the new suburban Italian enclave of Woodbridge, in the Greater Toronto Area, and simply wherever Italians have melded with other communities. The true essence of the Little Italies was based not only on the physical proximity of Italians within an ethnic neighbourhood but also in their being settled in a limited area. De Maria Harney (1998) argues that the preference for moving into larger spaces and into bigger and detached houses has reduced the interactions within families and whole neighbourhoods, as people are increasingly likely to use cars and to drive out of the neighbourhood for their everyday needs. At the same time, the essence of these historical neighbourhoods changed, as "the charm of Mediterranean exotica in 'Little Italies' around the city creates opportunities for Italian Canadians to reap financial

⁴³ Italy entered the war in 1940, allying with Germany.

rewards by marketing their ‘Italian’ authenticity, their ‘practical knowledge’, and their ‘cultural vitality’ to the Canadian public; but it also limits and restricts those who wish to break out into different fields and new directions” (ibid., p.173). The true essence of Little Italies has substantially changed over recent decades, keeping pace with the new sets of identities of the Italian community and their level of integration into the Canadian culture, and these neighbourhoods have become an economic and fashionable reality. By moving along this path, Italians have given up language as an asset of their heritage identity, particularly among new generations.

2.2.3.2 From the beginning of the 1970s onwards

In recent decades the whole scenario has changed to a great extent, both with the introduction of new policies, a change in the attitude towards Italians from the mainstream Canadian society, and the more integrative attitude showed by Italians and above all by the new generations.

From a Canadian perspective, although at the beginning of Italian migration the hegemony of the English mainstream culture(s) was predominant, things changed after the Second World War. At the beginning of the 1970s, Canada launched new policies for a multicultural society with bilingual (English and French) foundations. Minority ethnic groups (including Italians) began pressing the Federal government to include their languages in educational curricula, arguing that it is not possible to pass on a cultural heritage without its language (De Maria Harney, 1998).

From an Italian perspective, from being poor and backward, Italy became a country with a growing economy and new fashion trends. As a result, the following years saw a profound transformation in how Italianness was perceived by Canadians and by the Italo-Canadian community as well. De Maria Harney (1998) argues that “new meanings entered the swirl of competing images to create further layers and greater complexity within the construction of Italianness. The image of Italians and Italianness was recast” (p.172). The second and third generations became aware that the nostalgia of their parents and grandparents for Italy was for a bygone age, a heritage that was neither part of Canadian history nor any longer a trait of Italian culture (Caltabiano & Gianturco, 2005). On the other hand, they began to appreciate aspects of Italian life that differed entirely from the local traditions of their parents and grandparents. The feeling of being Italian became ever more related to the new, trendy and high-status products and brands for which Italy is famous around the world (see Baldassar & Pesman, 2004 for a parallel

situation in Australia). However in those years the ability to speak Italian, at least for a certain percentage of Italo-Canadians, turned out no longer to be an asset for an Italian identity (Caltabiano & Gianturco, 2005) and different Italian sub-cultures were thus created, where Italian identity is individually negotiated and may appear in different forms and at various levels of intensity, both in terms of their affiliation and linguistic competence.

Notwithstanding these deep changes, prejudices against the Italian community did not completely disappear. Although these may still nowadays partly reunite the community, they mostly seem to undermine their heritage esteem, thus pushing them towards a complete assimilation. One critical point is the effect of stereotypes passed on by exposure to the mass media, where young (second and third-generation) Italo-Americans are still portrayed according to the more “picturesque” and belittling stereotypes (Fondazione migranti, 2011). It is clear that the effects of negative stereotyping of Italians were (and still are) weighty, both in the culturally dominant society and subsequently within the Italo-Canadian community as well.

A study by Giampapa (2001), investigating the continual self-definition of Italo-Canadians in different contextual and socially-situated practices, points to the fact that in more formal domains, young Italo-Canadians are facing exaggerated stereotypes passed on to all of society by the North American media. Particularly in the work domain, Italian is not widely used, even among competent speakers, as Italo-Canadians “feel they need to leave their ethnicity at the door in order to challenge these stereotypes and position themselves as legitimate players within this game” (Giampapa 2001, p.308).

2.2.4 Italian and Italese

In this subsection I briefly dwell on the variegated linguistic repertoire of Italians in Canada. I restrict the discussion to Italian and Italese, but note also the presence and influence of dialect in their repertoire, as well as in that of Italians in the peninsula, and consider, hence, the importance of creating a homogeneous sampling. My data collection is in fact based on my subjects' performance in Italian.

2.2.4.1 Italian

Vedovelli (2011) cites Rosoli as pointing out that the level of education of Italians in Canada (and in the other main countries of Italian migration) has not so far been

thoroughly studied and that research in this field is still inadequate. According to some authors, their education did not take place exclusively in dialect, but included some varieties of Italian (Turchetta, 2005; Sobrero & Miglietta, 2006; Vedovelli, 2011). Hence, their linguistic repertoire:

era [...] certamente eterogeneo e forse l'idea di un'emigrazione italiana connotata da alti livelli di analfabetismo andrebbe ridimensionata [...]. (Vedovelli 2011, p.419)

was [...] surely heterogeneous, and the idea of an Italian migration characterised by high levels of illiteracy should perhaps be reduced to its true proportions [...]. [My translation]

Among the cohorts of the end of the 19th century and the beginning of the 20th, the level of education and literacy of Italians who emigrated to Canada was relatively low. This meant a higher use of dialect compared to Italian. However, Haller (2006a), in a study of the Italian communities in the USA at the beginning of the 20th century, claims that along with predominant skills in dialect, they also had a certain capacity in spoken Popular Italian (*Italiano popolare*), at least on a passive level.⁴⁴

During the period of Italian mass migration, the great majority of migrants must, however, be considered active Italo-dialect speakers. They were also in fact capable of mastering at least one variety of Italian beyond their local dialect (Gobbi, 1994). Moreover, the very experience of migration itself facilitated and promoted Italian over dialect (Gobbi, 1994), first as a result of the increased opportunities to interact with Italians from other regions⁴⁵ and secondly as migrants became sensitive to the importance of literacy, which induced them to improve their ability to write and speak Italian (Gobbi, 1994; Haller, 2006b).⁴⁶

Vedovelli (2011) observes that there was a substantial development of published Italian media at the beginning of the 20th century, which the author takes as a sign of a general rise in the level of competence in Italian among migrants, adequate for them to read

⁴⁴Haller (2006a) analyses theatrical texts by Eduardo Migliaccio, an Italian-American writer of the early 20th century. As there was no corpus of the spoken language used by these migrants available to him, he had recourse to theatrical works in order to attempt “to hypothesize patterns of linguistic behaviour among immigrants in the early decades of the twentieth century” (p. 345).

⁴⁵ Before the experience of migration, the main opportunity to interact with people of different regions was (for young Italian males only) the one-year compulsory military service (Tempesta, 1978).

⁴⁶ This attitude is also conveyed in the letters that Italian migrants sent from America, advocating to their families who remained in Italy the importance of getting an education (among others Melillo, 1991; Turchetta, 2005; Sobrero & Miglietta, 2006; D'Agostino, 2007).

journalistic texts. Moreover, the number of readers grew considerably in those years, allowing the rise and development of several journals in Italian. Numerous journals were in fact launched between 1919 and 1940, considerable numbers surviving between 1940 and 2000, although their number has drastically decreased in the last decade, so that now only about ten newspapers are published in Italian in Canada (see also Turchetta, 2005). This is probably due to the combined effect of the generational turnover and the new modalities of information and entertainment available through new media.

2.2.4.2 Italese

The linguistic repertoire of Italians in Canada is enriched by a new language variety called Italese. Italese represents a facet of the linguistic presence of Italians in Canada and in other Anglophone countries⁴⁷ and it can be defined as:

una forma ibrida usata dalla prima generazione dei nostri migranti in Canada, derivante dalla dialettizzazione (o, per certi versi, dall'italianizzazione) di termini e di espressioni inglesi (Vedovelli 2011, p.426).

a hybrid form used by the first generation of our migrants in Canada, due to the dialectisation⁴⁸ (or, for some aspects, to the Italianization) of English words and expressions. [My translation]

The main linguistic feature of this koine⁴⁹ is the use of adapted loans from English. These loans were introduced for two different purposes: reflecting migrants' experience in a new linguistic, social and cultural environment and serving whenever words in English could not be easily substituted with a corresponding form in Italian (Haller, 2006b). These loans, taken from English, are morphologically and linguistically

⁴⁷ There are similar examples among Italians in other countries, including the cases of Chipilegno in Chipilo, Mexico and Cocoliche in Buenos Aires, Argentina, as well as with other languages: Daussie (Dutch in Australia), Franbreu (French in Israel) and Spanglish (spoken by Spanish-speaking migrants in English-speaking countries).

⁴⁸ The term "dialettizzazione" used by the author does not appear in the Italian dictionary. As for the English translation, it has been chosen to stay as close as possible to the original form, although I am aware that it is not in English dictionaries.

⁴⁹ Koine is a problematic concept, which has been used in linguistics with different meanings in different linguistic areas.

As suggested by Danesi (2011), I use the term 'koine' in referring to Italese. Although the applicability of the term to this specific phenomenon appears debatable, it seems, among those proposed by scholars, to best fit Italese. Other definitions, such as "the Italian/English contact language" (Giampapa 2001, p.280) or a "community language" (Vizmuller-Zocco 2007, p.355), involve the notion of language, which is itself open to debate.

adjusted to the migrant's first language (dialects or Italian). The great majority (more than 80%) of loanwords are nouns, which are usually given the masculine gender and its corresponding ending, according to the principle of markedness; following the same principle, verbs are assigned to the first conjugation.⁵⁰ Phonetically, Italese is characterised by an internal variegation. Loans are integrated into the phonetic systems of the different dialects of Italy, as speakers from different regions create their own versions of Italese. Thus, the linguistic patchwork of Italy is in a sense reflected in this idiom (Danesi, 2011; Haller, 2006b).⁵¹

Like other natural languages, Italese was subject to a process of evolution. In this regard, it has been proposed that it should be regarded as the language of the first generation (and linguistically as the dialectisation/Italianization of English words and expressions), while the variants used by the following generations should be termed *Nuovo Italese* (Clivio) or *Secondo Italese / Neoitalese* (Vizmuller-Zocco). According to Clivio (cited by Vedovelli, 2011), the new Italese is characterised no longer by the dialectization/Italianization of English words and expressions, but by the substitution of words and expressions in English and by various code-switching phenomena.

The future for Italese is however gloomy, and this koine is expected to disappear in the near future, as it is a spoken idiom rather than a written codified language (Vizmuller-Zocco, 2007), but mostly as its function has simply a phatic connotation, to be used with the older generations and the elderly (first generation) and it is openly disliked by the new generations.

2.3 Sociolinguistic research on language maintenance among Italians abroad

So far, I have presented an overview of Italian communities in Canada, with particular reference to their linguistic repertoires. What follows is an exploration of language maintenance among Italian communities from a sociolinguistic perspective, the core approach of this analysis. Empirical studies have provided further angles, shedding light on the roles of different factors. Striking results have been obtained over the years; nevertheless, this process is influenced by so many variables that it is difficult to present

⁵⁰ Among the three Italian verb conjugations, the first is regular in its inflection and represents the most numerous category in Italian (Danesi, 2011).

⁵¹ Both authors identify traits typical of Venetian-Italese, such as “degemiazione consonantica” (the use of a single consonant instead of the double required in standard Italian) and the dropping of the final vowel in a word, especially the /e/ (Danesi, 2011).

a simple picture. In the following pages I offer a summary of some of the most interesting analyses, aiming also at identifying areas of potential interest for future research.

All studies of Italians abroad have found that significant language shift has occurred. With the exception of a few cases, such as that of the secluded community from Segusino (in the province of Treviso) in Chipilo, in the south-eastern area of Mexico City, and that of the Venetian communities in the region of Santa Catarina (Brazil), the state of Italian among communities abroad is seen as supporting the idea that the third Italian-speaking generation is the last.⁵² Indeed, the competence of this last generation is often so dramatically limited that its members are unable to hold a proper conversation in their heritage language or even simply to understand it. In this respect, Bettoni (2007) argues that given the degree of shift among third generations, it would be better to look at their competence in heritage languages within the acquisitional linguistics paradigm. With a great shift among the third and following generations and a limited rate of migration from Italy, Italian communities abroad, and Canada is not an exception, are no longer numerically and linguistically supported by new arrivals, putting their future status as linguistic communities under threat.

For Clyne (2003), “language shift has emerged as a product of pre-migration and post-migration experiences mediated through culture” (p.69). In this respect, Italian communities experienced an Italian-dialect diglossia/dilalia in their home country, which is thought to have contributed substantially to the language shift among Italians abroad (Boyd, 1986; Rubino, 2006). Moreover, discrimination against Italians has induced them—and more so the following generations—to absorb stereotypes of and negative judgments against their culture. However, the same acts of discrimination have also contributed positively to language maintenance, favouring their physical proximity to other Italians (Little Italies) and cultural preservation as heritage communities, mainly through a relatively high level of endogamy, the formation of many diverse Italian clubs and societies, and the establishment of Italian mass media programmes.

⁵² These cases of language maintenance (Chipilo and Santa Catarina) are found in the countryside and remote environments. The circumstances within cities have been conversely less maintenance-supporting. In Canada, the Italian community showed a clear preference for cities, rather than the countryside (Edwards, 1998).

2.3.1 Generation

As discussed in the first chapter, generation is considered the most powerful predictive variable in language maintenance studies. Notwithstanding its crucial role and the fact that its impact seems plain and more easily predictable than other variables, there has been much discussion as to how to treat it concretely. Census data and macro-linguistic research have often considered place of birth as the discriminating variable between first and second generations. More qualitative research, on the other hand, has highlighted the importance of considering place of birth in conjunction with age of departure from the home country (and thus with the child's length of exposure to the heritage language in the L1 monolingual environment) (see also Bettoni, 1986).

With regard to the Italian communities abroad, the repertoire of first-generation Italian speakers includes at least one variety of Italian and one of dialect. The balance varies between different cohorts. Whereas at the beginning of the last century, migrants were mainly dialect-speaking, with few skills in Italian, over the years this relationship has reversed. However, although Italian migrants may have been mainly dialect speakers at the time of migration, they nevertheless found in their new environment the prerequisites to pass to a more frequent use of Italian (Gobbi, 1994), modifying their dialect in a pan-Italian direction: they had the chance to be exposed to the mass media and, more importantly, were able to interact with people of different Italian origins. Their Italian skills were thus strengthened, although not through formal education. With respect to their use of the host language(s), the physical proximity (Little Italies) and homogeneity of Italian communities abroad (particularly when in a rural context but even in a city setting), as well as a high level of endogamy, favoured a relatively high level of maintenance of their heritage language and conversely a lesser proficiency in the new language(s). New waves of migrants, however, are mainly closer to Standard Italian, although cases of dialect-speakers may still be found,⁵³ or more commonly of people with traces of Venetian dialect in their speech.

The second generation is linguistically and culturally more variegated (and presumably the third even more so) than their parents. Even in research with evidence of high levels of proficiency among the second generation (as in the case of Italian in the German-speaking part of Switzerland, discussed by Schmid, 1993), a weakened language performance is clearly observable. Differences with the first generation are noticed in

⁵³ During my fieldwork I interviewed a new migrant who had moved to Canada at the beginning of the 1990s. Although he was from a city, had a good level of education (up to 19 years of age) and was asked to use Italian during the interview, he spoke mainly in dialect.

terms of phonology (with a pronunciation less regionally marked compared to that of their parents) and generally with a greatly increased simplification (e.g. the generalisation of the auxiliary verb *have* or a reduction of vocabulary size (Schmid, 1993)).

In reality, the term ‘second generation’ *tout court* must be considered a general label, giving only a partial account of their internal variegation. For instance, Bettoni (1986), whose research addresses largely Italian communities in Australia, marks the importance of birth order, suggesting how second-generation levels of proficiency vary dramatically between first-born (including only children) and last-born Italians.

An appreciation of the boundaries between Italian and dialect is often particularly blurred for the second generation (Marcato, Haller, Meo Zilio & Ursini, 2002). This is related to the fact that they have lacked references to Italian other than their family and close community, in which the two languages are often mixed up, although, as Bettoni (1991) argues, this relative insecurity between Italian and dialect is also partly inherited from their parents. Moreover, although the second generation could potentially take advantage of both Italian language courses as linguistic models in the standard language and of mass media which broadcast in standard Italian (Schmid, 1993), “younger people – and especially the second generation – generally find radio programs and community language newspapers unappealing, rendering them ineffective as a resource for reversing LS [language shift]” (Clyne 2003, p.63).

2.3.2 New environment

The social, cultural and linguistic context into which Italian migrants settled played an important role as well. For example, the differences in social and cultural bases between Canada and the United States mentioned earlier in this chapter have caused these two geographically close countries to differ in the capacity of their migrant Italian communities to maintain their language. Comparisons between research in different environments is, however, made less viable by difficulties in comparing pieces of research, whether in terms of differences in homogeneity of sampling and in the time of the fieldwork, or of the tools used to gather data.

This subsection discusses two studies which, however, make direct comparisons between Italian communities in two different countries: Vanvolsem, Jaspaert and Kroon (1991) studied Italian communities in Belgium and the Netherlands, limited to the first generation, while Auer (1991) researched Italians in Canada and Germany, taking a

generational approach, both finding that this variable had a major effect on language maintenance.

Vanvolsem, Jaspaert and Kroon (1991) highlight the size and density of the minority community as key factors in predicting language maintenance, a result in line with other studies of language maintenance. While the Italian community in Belgium was numerically more substantial and cohesive, that in Holland was conversely more scattered and isolated. Moreover, although this is difficult to ascertain, it was suggested that a Romance language, such as French, may have had a positive influence on the maintenance of the first language among Italians in Belgium. The authors take these two criteria, density and linguistic relatedness, as descriptive variables which may explain the higher first language attrition among Italians in Holland.

The second piece of research highlighted the consequential role of the geographical distance between the country of origin and the host one, citing the fact that while Italians in Toronto “had a distinct feeling of leaving their homeland for good, and consequently quickly developed an Italo-Canadian identity, Italians in Germany continued to foster a wish of returning, although they may have lived in Germany as long as their compatriots in Canada” (Auer 1991, p.408). Geographical proximity offered Italo-Germans more opportunities to spend time back in Italy, also giving some young Italians the chance to spend some years at school in Italy before joining their parents in Germany (see also Campanale, 2006). Linguistic implications were accounted for along two lines: one quantitative and the other domain-related. The influence of L2 German, notwithstanding its membership of the same family as English, was considered to be less pervasive and quantitatively more limited in the speech of Italo-Germans than that of English in the case of the Italo-Canadian informants. Secondly, the competence of the second and third generations in heritage languages also differed in terms of domain use: whereas Italo-Germans employed their heritage languages both within their family and in peer conversation, paralleling in a way young Italians in Italy, the second and third generations in Canada conversely turned exclusively to English in their peer interactions. This difference may be due either to the fact that they feel their level of proficiency is so low that they are not confident enough to even code-switch in Italian, if not necessary, or it may be that the level of heritage affiliation among young Italo-Germans is higher than among their peers in Canada.

2.3.3 Domain

Divergence in Italian competence in different domains across generations is a key point to mark the passage from the first to the second generation. A pronounced reduction in the number of domains is often a clear sign of language shift in progress within the whole community.

Overall, although there are several factors at play in determining the most appropriate language(s) in each domain, some of them, without determining it completely, do nevertheless limit the speaker's range of choices (Bettoni & Rubino, 1996). Researchers have devoted more attention to interactions within the family, where it is easier to collect data from direct observation. In contrast, those who have treated other domains have mostly gathered data through self-reports or questionnaires. On the whole, two main results have emerged, the first being that the Italian diglossic pattern has been maintained in the new environment. Secondly, the shift to the L2 language appears usually to be stronger in formal and heterogeneous domains, such as work or church, but also (and quite unexpectedly) within the more private and informal ones, such as soliloquy or within the family (Bettoni & Rubino, 1996; Campanale, 2006).

According to the model of Smolicz (1981) on 'core values', Italians do not have language as the fundamental value to preserve and prove their heritage roots. Instead, this role is often played by the family, particularly the extended one. Hence among Italian communities abroad, "family cohesion [has taken] [...] precedence over language as a cultural core value" (Clyne 1991, p.92).

It is hardly surprising, observing the language shift that has occurred in the other domains (Bettoni & Rubino, 1996), that the literature on Italian communities abroad has seen the family as the last bulwark of Italian language in a foreign environment. But its relative strength should be described as such, in comparison with what has happened in other domains, rather than as a strength per se. The family domain has been one of the central points of the analyses by Bettoni and Rubino within the Italian communities in Australia over the last two decades, offering us a longitudinal overview. Their main finding is that the family has been erroneously considered the bulwark of language maintenance (Bettoni & Rubino, 1995).⁵⁴ In terms of quantity, English is in fact the language most used in the family (Bettoni & Rubino, 1995), despite a high degree of endogamy both in the first and in the second generation (Tosi, 1991). The heavy

⁵⁴ In the same vein is the result by Jaspaert & Kroon among Italians in the Netherlands and Flanders (1991). The domains of neighbourhood and church were conversely considered less resistant to language shift.

influence of the external society on the domestic environment (e.g. the use of mass media in English or the introduction of English through the schooling of the younger generations), as well as the increasing numerical presence of the third—more English-speaking—generation, has definitely favoured a shift from the use of ethnic languages. The family domain in itself does not seem to hold any explanatory power. It is in reality strictly related to other social variables, which all together control the pace of this shift. Among these, the number of children within the family is of key importance: the greater the number, the greater is their strength in choosing the language (Tosi, 1991). The dominant and preferred language among (third-generation) children within the family is English, with the occasional use of Italian and dialect, mainly limited to phatic and expressive functions (Rubino, 2000). Therefore, when in the family there is only one child, s/he is more likely to adopt to the language chosen by her/his parents as s/he is usually alone. If s/he has siblings, together they usually have more power in negotiating the language with their parents.

As suggested above, if maintenance is still occurring, this is related to the presence and influence of the extended family (Bettoni & Rubino, 1995), particularly in the case of family reunions, typical of Italian family life, when older people (the first generation, who usually do not master L2 well) are involved in the conversation.

2.3.4 Gender roles

In relation to gender, the literature has not always been unequivocal. Particular attention seems to have been given to the role of women as a subject of study, compared to the interest in that of men, underlining the view of the female role as being at times conservative and at others innovative (Milroy & Gordon, 2003).

In research on heritage languages, specifically among Italian migrants, outcomes seem to be more unidirectional, portraying women as more dedicated in the inter-generational maintenance of heritage languages (Clyne, 1991; Bettoni & Rubino, 1995; Marcato, Haller, Meo Zilio & Ursini, 2002; Campanale, 2006). This trend seems to be explicable in light of the different social roles of women and men that, particularly in the past but still nowadays, characterised Italian society:⁵⁵ predominantly those of child carer and housewife for the former, and of worker outside the domestic sphere for the latter. This

⁵⁵ The comparative study by Boyd (2001) of the effects of parents' gender on minority language maintenance among different cultures stressed that this influence depends on quantifiable variables (time spent with their children) which are related to educational roles, and thus socially defined and assigned to women and men in each heritage community.

implies more opportunities for men to come into contact with and use the language of the host country. Moreover, men usually show a greater drive for social integration and promotion. Thus, they are usually more inclined to learn the host language for social purposes.

However, attempts to generalize the compound variables at play rarely find complete corroboration and the situation for Italian communities has proved to be less straightforward and not always predictable. Italians were often employed by companies staffed by other Italians, which hindered their learning of the new language. Moreover, under some circumstances, Italian has even been taken as a *lingua franca* (thus with an expansive force) among Italians and people of other nationalities, particularly from southern Europe. When a group of Italian workers came to socialize with people from different communities, the latter may have even been inclined to learn and use Italian as the shared language of communication (Tempesta, 1978). Hence, as will be discussed in the metalinguistic observation (chapter 5), the condition of men working outside the home has not unequivocally meant that they were better speakers of the new language.

Despite these initial differences, however, gender divergences seem to have blurred within the Italian second generation (Clyne, 1991), as a result of new and more similar social conditions for women and men. This makes this variable possibly even less predictable than before, whenever it is considered separately from the array of other variables at play.

2.3.5 Attitude

As discussed in chapter 1, attitude is often a good predictor of heritage language maintenance/attrition among migrant communities. This section on the sociolinguistics of language maintenance concludes with a brief account of its key role in promoting language maintenance along a generational scale, an investigation of attitudes toward Italian and dialect, followed by some reflections on attitudes to the phenomena of code-mixing and code-switching among Italians abroad.

Overall, Italian is loved and appreciated as the language of cultured and well-educated people. Its prestige is well-attested in matched guise studies, which all confirm the high status that Standard Italian enjoys, particularly if compared with dialect (among others, Rubino, 2006; Marcato, Haller, Meo Zilio & Ursini, 2002; Bettoni & Gibbson, 1988; Turchetta, 2005; Haller, 2006b).

As in all language contact situations, the phenomena of code switching and code mixing are frequent among Italian communities, as is the use of local koinés. The situation of *dilalia* in Italy is usually entrenched in a new country with a new diglossia: the host country language as the higher language and the migrant language(s) as the lower. These diglossic borders are, however, permeable and mutable. Functionally, domains that have allowed in the past only one language may begin to gradually accept another. Thus, even code-switching and code-mixing usually increase consistently (Bettoni, 1991).

At the same time, attitudes toward these mixed forms are deeply negative (Bettoni, 1991) among both first and second generations.⁵⁶ Nevertheless, it is just the permeability of the diglossic boundaries and the subsequent use of ‘linguistic mixtures’ which allow second-generation migrants to use Italian, despite their competence in this language being insufficient to enable them to hold a conversation (Bettoni, 1991; Rubino, 1991). This mixing is thus not only a widespread practice and a sign of identity, but also a condition *sine qua non* for these speakers to continue partly using their heritage language(s) (Rubino, 1991). Interestingly, what they strongly dislike and value negatively is just what they have created and what lets them continue using their heritage language(s) to some extent (Bettoni & Gibson, 1988).

2.3.6 Contacts

In this last section I briefly consider the variety of contacts that Italians in Canada can enjoy in their mother tongue, looking in particular at their social networks and the mass-media in Italian available in the country. Although the number and range of this type of contacts is *per se* symptomatic of the vitality of Italo-Canadian communities and of the many ways to experience and use the Italian language, second and third generations usually do not take advantage of this situation. However, and this is valid for most of the aspects we will discuss in the next pages, the new generations have not completely neglected these types of contact, but partly moved to new forms of socialising. Their clear internet-oriented preference highlights the necessity to revisit the focus of the fruition among the youngest group, which comprises both new Italo-Canadian generations and new migrants as well.

⁵⁶ The generations do not seem to share exactly the same feeling. Whereas the first is more positively inclined toward standard Italian and purist forms, the second expresses more tolerance with regard to non-standard varieties (Marcato et al., 2002).

2.3.6.1 Associations

The scenario for the associations is extremely variegated (Zampieri Pan, 2011). From a geographical perspective, they range from very local (i.e. Associazione Selva del Montello– a little village in the province of Treviso - in Vancouver) to provincial (i.e. Associazione Trevisani, Associazione Vicentini and Associazione Bellunesi) to Italian or Italo-Canadian associations. However other typologies are well represented both in the Toronto and Vancouver areas, based on social groups (new generations or women), sports/outdoor activities or hobbies (e.g. bocce balls). Particularly lively also are the choirs, such as Cantitalia in Guelph and the many others in the churches offering services in Italian, generally involving people of Italian heritage but with the attendance also of other Italian speakers regardless of their heritage.

Particularly emblematic of the organizational capacities of many Italian regional groups is their success in grouping together while giving birth to complex associative realities, also in terms of economic sustainability. Regional centres, such as the *Veneto centre* or *Famee Furlane*, offer a wide range of services to the community and beyond; these range from summer camps to the opportunities to take part in sports, and to various cultural initiatives for people of all ages. Although bilingualism is widespread, with a prominence of English use in some activities, particularly when directed to the younger generation, these centres are still opportunities to meet other Italians and to use Italian language to some extent.

2.3.6.2 Religious services

Although decreasing, services in Italian in many Canadian cities are still offered, thanks to an Italian or Italian speaking clergy, though this is declining in number.

Overall a substantial shift to English has occurred, both as a natural process of shift, particularly but not solely among the new generations, but also favoured since the very beginning, among the same migrants, by the consistent presence of an Irish Catholic clergy, particularly active in North America at the time Italians migrated in large numbers. In contrast, Italians have partly benefitted from the presence of an Italian-speaking clergy. This is to some extent still possible nowadays thanks to the fact that Italian is the official language of the Vatican State and that a relatively significant number of non-Italian priests have learnt the language.

With regard to the effective participation in Italian services, apart from the weekly Sunday Mass in Italian, few other opportunities are taken by the Italian community.

Interestingly, although not quantitatively significant, Italian is still the preferred choice (maybe together with English to include the participation of the whole family) in the more emotive and personal occasions, such as the funerals of Italians.

2.3.6.3 Language courses

Language courses in Italian are available through a wide range of options. At a high level, there are undergraduate and master courses in Italian offered in many Canadian universities. According to the Italian Foreign Ministry, Italian is taught in thirteen Canadian Universities (www.esteri.it). Other possibilities are offered by state or private institutions, both Italian and Canadian: among the Italian ones particularly active are the Società Dante Alighieri and Italian Cultural Institutes, with headquarters all around the world. There are also local schools and language centres, such as the Columbus Centre in Toronto, which also promote activities related to Italian culture.

2.3.6.4 Mass media

Before proceeding to discuss some of the most prominent mass media and their programmes in Anglophone Canada, it is important to mention some general considerations.

Overall, Italian mass media are particularly vital in North America, and this is proof of a still significant Italian-speaking community. The fact that the media operate in a private market, thus supported economically by advertising, is symptomatic of a numerically substantial Italian audience (private conversation with P.R.⁵⁷). Moreover, both radio and television have played an important role in forging the Italian competence of migrants from the peninsula, helping them, notwithstanding their initially more limited skills in Italian, to have quality inputs in the standard language, despite local linguistic differences (private conversations with U.M.⁵⁸ and P.R.).

With regard to the content, Italian mass media in Canada may be delivered from Italy, thus proposing Italy-centred information, or they can be created specifically for the Italo-Canadian community; in this latter case they deliver local, international and Italian news (a '*glocal*' perspective), but always giving the information that affects the Italo-Canadian community.

⁵⁷ News Manager & Senior Editor at OMNI TV. Private conversation held on 17 July 2009.

⁵⁸ News and Programme Director at CHIN Radio. Private conversation held on 25 June 2009.

With regard to the language used, the variation is wide, ranging from the sole use of Standard Italian in Italian broadcast programmes or, in some Italo-Canadian programmes, to a mixed use of Italian, English and Italese, thus exploiting the whole linguistic repertoire of the audience. The choice is, however, also related to the type of programme and the audience.

The first generation, in particular, prefer more traditional Italian mass media, the radio and television, to periodicals and they use the web more rarely. Overall the first generation prefer oral mass media over those they have to read (private conversation with P.R.).

Among mass media, the process of language evolution has been different: television and radio are deemed to be less conservative than periodicals, which adhere to the more conservative model of the 1950s. The fact of including/mixing more languages, as happens in many TV programmes, is seen positively as the natural outcome of living in a foreign country: when you live abroad, hearing someone speaking in another language is in fact part of the your everyday life and the same in phenomenon in TV programmes is acceptable to the ear (private conversation with P.R.).

Radio and Television

Radio frequencies in Italian are not uniformly spread within the country, covering mostly Ontario, less of Quebec and existing only marginally in British Columbia (www.italiansinfonia.com).

The most important station is Chin Radio, a multicultural/multilingual radio station founded by a second generation Italian, which broadcasts in Italian for 80 hours a week in Ontario and from 2003 also in the area of Ottawa. Programmes range widely, from music and sport, to news from Italy and Canada and debates (Marchesin, 2011). The same variety is mirrored in the language used in these programmes: in order to satisfy all the needs of the different strata of Italians in Canada, with different levels of education and thus of linguistic competence, the language varies significantly, with some programmes closer to the Standard language and others characterised by a use of mixed codes, namely Italian with English and Italese (Chin, private conversation with U.M.).

With regard to television, a similar pattern to that just described for the radio is seen. In the area of Toronto there are two main channels: Telelatino (1984) and OMNI 1 (1978),

the latter on air also in other Canadian provinces, including British Columbia. Telelatino is a multilingual channel which broadcasts programmes in Italian as well. Since 2003 some programmes from an Italian commercial channel have been broadcast too (Marchesin, 2011).

OMNI 1 transmits various programmes in Italian, ranging from soap-operas to soccer reports to cultural programmes or interviews, during the afternoon and evening time. Once a week it also broadcasts a programme 'Noi Oggi' about Italian and North-American culture aimed at involving new generations of Italo-Canadians (Marchesin, 2011). In reality the level of competence in Italian among the second generation is often so limited that, if they do watch Italian programmes at all, their viewing is usually confined to cooking programmes (private conversation with P.R). Thus, for the very great majority of programmes, the audience is composed of first-generation Italians: a niche audience (private conversation with P.R) but one that the station is committed to. In the future the demand from the Italian audience will drop significantly in quantitative terms. Also qualitatively major changes are expected, as the audience will be more sensitive to the quality of the programmes and at the same time will have less time to watch television (private conversation with P.R).

From 2003 the Italian national public broadcasting company (RAI) launched RAI International in Canada, with transmissions of programs on air in Italy. Thanks to an agreement, Rai is available also in the afternoon/evening programmes of Omni TV.

Journals/Magazines

These days there are many publications in Italian, about forty in the area of Toronto alone (Marchesin, 2011). Corriere Canadese (founded in 1954) is a journal distributed in Ontario and some areas of Quebec (Marchesin, 2011). Today it is made available jointly with an Italian newspaper, La Repubblica. From 1995 also, once a week, Corriere Canadese is distributed with an insert in English, Tandem, which is addressed to second and third generation Italo-Canadians while discussing topics related to Italy and Italian communities. On a local level, there are other flourishing cases as well. Lo Specchio (1984), for example, is a weekly newspaper in Italian which is aimed at some suburban areas of the Greater Toronto area (municipality of Vaughan) where the presence of Italians is the highest. It discusses news from Italy as well as news regarding the local Italian community (Marchesin, 2011).

There are also many Italian periodicals and newspapers available in Canada, both those directed to readers in Italy but available also abroad, and those created specifically for a worldwide Italian-speaking audience, maybe with one version for Italians in the peninsula and another for Italians abroad (e.g. *Il Messaggero di Sant'Antonio*).

Internet

The development of this new media is more elusive and difficult to treat. It divides the generations dramatically with the first generations on one side, and the second and third generations and new migrants on the other. Whereas the first show a plain preference for more traditional media, the latter are regular users of the internet. In relation to the more traditional media, namely television/radio and papers/journals, the internet allows direct contact with Italian media, accessing an updated image of the country and offering the opportunity to receive a wide variety of linguistic inputs, ranging from the standard language of Italian journals or national radio stations, to more informal and locally diversified inputs.

The same lively associative reality that we have discussed above in Canada is mirrored on the internet among the new generations and particularly among new migrants, who are using it as a platform to connect specifically with Italians in the peninsula and with other migrants all around the world as well (*Il Messaggero di Sant'Antonio*, 2009).

2.4 Summary and concluding remarks

In this chapter we have analysed the scenario of Italian communities in Canada, starting with a linguistic review of the peninsula and passing on to analyse the historical and social phases that have characterised Italian migration to Canada, with the consequent linguistic outcomes. In particular the year 1967 has been highlighted as a threshold when new restrictive policies were introduced in Canada, putting to an end the mass migration of manual workers and opening the way to quantitatively reduced migration, but of well-educated people.

From a social perspective, Italian communities in Canada appear to be highly variegated entities, particularly with regard to the second and third generations. The idea of Italianness has undergone profound adjustments over the decades: from a negatively stereotyped and diminishing culture, it turned into a fashionable and renowned lifestyle model. Whereas the first generation appears to be culturally tied to the homeland and

linguistically conservative, new generations have detached their Italianness from the knowledge of their heritage language(s), preserving other, more societal, aspects. However, as soon as the first generation joins the conversation, heritage languages are generally used to some extent, so the second and third generations are exposed to them. In contrast, heritage languages seem not to be used outside the family environment. This reveals the key role of the extended family in helping to preserve heritage language(s) among Italian migrant communities, something that is made possible by the prominent role that extended families still play socially in Italian family life (Caltabiano & Gianturco, 2005).

The sociolinguistic variables that have been presented in the first chapter have been reviewed here with regard to the Italian community. All the variables have turned out to be of some relevance, with generation being the most pivotal. Favoured by a relatively linguistically conservative first generation and by the appreciation of the Italian culture and language among the broader Canadian society, heritage speakers may benefit from a good variety of contacts with Italian, ranging from TV programmes and periodicals to language courses. However they do not seem to take great advantage of these, and the family, particularly the extended family and the presence of the first generation, remains the main stronghold favouring language maintenance among Italians in Canada.

As appears clearly in this chapter, the future for Italian in Canada is gloomy. Italian has become irrelevant to the everyday lives of a large number of second- and third-generation people (Rubino, 2000), replaced by the language of the country where they live; it has eventually lost both its “practical” and “symbolic” functions (Bettoni, 1991), which would have helped to prevent its decline among new generations.

In this first part of the thesis I have offered a review of the bilingualism of migrant communities. I commenced in chapter one with a discussion of the attrition-maintenance continuum. I explored particularly two aspects of these linguistic phenomena. On one hand I discussed the factors, in terms of social variables, which favour attrition, or which promote maintenance. Generation, considered in terms of the onset of bilingualism and exposure to the heritage language, is a pivotal variable in accounting for migrants' skills. On the other hand I have explored the linguistic outcomes of these linguistic phenomena. Overall, although all language levels may be subject to attrition, the lexicon seems to be the most vulnerable.

Chapter two, on the other hand, moved the focus to an Italian-centred perspective, discussing the main marked traits of Italian heritage communities abroad, with more reference to Canada (and other Anglophone lands) as a country of destination. I paid particular attention to the social variables that are deemed to have an impact on the level of language proficiency of these heritage-speaking communities. Among all those that possibly play a role, generation has emerged as a key factor. Furthermore, the historical section of this chapter highlighted the presence of different cohorts of Italian migrants in Canada, which vary significantly from each other in terms of attitudinal, educational and social aspects. The threshold is marked by the year 1967, when Canada introduced new and more selective immigration policies.

These theoretical premises allow us to put forward some research questions that will drive the analysis in the following chapters.

Besides the explanatory function that the variable ‘generation’ has in accounting for the level of proficiency among these communities, we may suggest other variables as playing a role. In chapter 1 we have discussed how language attrition (and the reverse can be said for maintenance as well) is due to the combined effect of the influence of a new language and the disuse of the heritage one. Whereas the first is a natural and common condition for many migrant communities, the latter seems particularly to develop due to a relative decrease in the number of contacts a speaker has with whom s/he may use the heritage language, and to the speaker's affiliation towards her/his heritage culture and language. We may therefore envisage the influence of these two variables, labelled ‘linguistic habits’ and ‘attitudinal factors’ as potentially of impact. The first point that will thus be addressed in the analysis regards the study of the pattern of these two variables, specifically with respect to the community investigated in this research. In particular this will be carried out by analysing the results of the questionnaires gathered during my fieldwork in Canada and studying possible differences across the four groups (new migrants, first, second and third generation) of my Venetian-Italian informants (chapter 4).

A second research question regards the results of interview analysis. Although we have seen in the literature that historically Italian communities have turned out to be relatively conservative and capable of preserving their heritage and language to some extent, even this community is not immune to language shift and cultural assimilation. I would, therefore, foresee a plain decline in language skills across the generations, even considering the specific nature of the data on which my analysis will be based. In fact,

only people who wanted to participate in this study, and thus were good enough to carry out a conversation in Italian, entered the sample. The data from my interviews account only for people with a certain level of proficiency in Italian, and this may have potentially levelled their skills across generations. Therefore a linguistic decline, although plausible, is not an outcome to be taken for granted. Alongside this expected trend in the results, my research question is in particular focused on quantifying this pattern. This will be carried out by looking at the generational scale across three generations of Venetian-Italians in Canada. In parallel to this analysis I will also be comparing different cohorts of migrants, and in particular those which are considered in this study as first generation (1945 – 1967) and new migrants (1970 – onwards). This last comparison will examine two specific groups of Italian migrants who, although sharing some traits, in primis being Italian native speakers, also differ in other respects (chapter 5).

The last research aspect I am interested in links the two research questions discussed above. In particular I will be looking at the impact of the variables discussed in chapter 4 in accounting for any significant differences which have emerged between the two groups of native speakers on the one hand, and along a generational scale on the other (chapter 5).

Part 2: Methodology

Chapter 3: Methodology

With this third chapter we enter the core of this research, establishing the bases of the analysis while offering an overview of the methodological measures undertaken and the choices that have been made in order to prepare the corpus that is to be analysed in the following chapters.

In particular, this chapter first explores the preparatory phases of the fieldwork and the fieldwork itself, looking at the sampling methods adopted and offering a review of the population investigated. This section is followed by an account of the fieldwork in Canada and how I proceeded with the administering of the interviews and questionnaires. I will then pass on to discuss how the corpus has been constructed, specifically how I proceeded with the transcription of the interviews and the codification of the questionnaires. This section will in particular highlight the general characteristics of the corpus, and introduce the lexicometric measures that are to be discussed in the following chapters in relation to this corpus.

3.1 Methodology aspects

3.1.1 Sampling methods

The first aspect worth exploring here is that of the sampling methods used in sociolinguistic research. The bibliography on the topic is vast (among others Milroy, 1987; Milroy & Gordon, 2003). In this section the discussion will only briefly review the background in order to focus on the methodology used in my fieldwork in Canada.

Ideally, in order to provide a totally accurate depiction of the population investigated, a researcher should include every single member of the community; but in the very great majority of cases this is not a viable option, so some sort of sampling method has to be applied in order to select a group representative of the whole population. In order to draw conclusions about a defined group of people (the sampling universe), we need to be concerned with the notion of representativeness (Milroy & Gordon, 2003). Generally speaking, sampling methods can be categorised as either probabilistic or non-probabilistic. Within the first group are the methods to ensure that each member of a sampled population has an equal chance of being selected. Although subject to some unavoidable bias, probabilistic methods eventually permit the claim that the sample is

representative to some extent, thus allowing generalisation of the results of the study to the whole population. In contrast, non-probabilistic sampling methods select elements of a given population in non-random ways, targeting specific individuals according to the researcher's subjective judgment.

The method chosen is strictly correlated to the set of conditions in which the research is embedded and to the aims of the study. Although some researchers view non-probabilistic methods as deficient compared to probabilistic ones in terms of representativeness, their use is supported by solid methodological premises. Moreover, Milroy (1987) notes a change in the approach to fieldwork methodology starting from the early 1960s. She suggests that a "shift in attitude which comes with the maturing of sociolinguistics as a field of research enables researchers to select more freely than was once possible from a range of methods which, within a defensible theoretical framework, will best enable them to achieve their goals" (p.38). It is thus crucial first to define the aims of the research and have a clear idea of the whole scenario in which the study is embedded. At a later stage, it is possible to opt for the most appropriate method, whatever it may be.

Considering this background and the scope of this research, I chose to follow a non-probabilistic approach, namely judgment sampling. This choice was driven by the aims of the study, which are not to analyse the linguistic repertoire of all the Italian communities in Canada, but to specifically address and investigate maintenance skills across the Italian-speaking community of Venetian heritage in Anglophone Canada. This study thus does not aim to be proportionally driven or representative of the average skills of people of Venetian heritage in Canada, particularly with regard to second and third-generation speakers. Judgment sampling depends on the researcher's belief that some subjects are more fit for the research than other individuals, which is why they are chosen as subjects. When using this method, the researcher must be confident that the chosen sample is fairly representative of the entire population.

3.1.2 Sampling population

With reference to the selection of variables, generation is considered to be a classic choice in sociolinguistics (see chapter 1). In this research the variable is defined as follows: *first generation* includes people who migrated to Canada after the Second

World War and before the new immigration policies were introduced in Canada in 1967. *Second generation* refers to people born of first-generation parents in Canada or born in Italy but who migrated before the age of 12 (Bettoni, 1986). *Third-generation* migrants are the children of second-generation parents. Lastly, *new migrants* are people who migrated to Canada after the introduction of the above-mentioned immigration policies. We now consider two further criteria used to classify the informants, one related to the speakers' Venetian origin and the other to the definitions of third-generation and new migrants. In relation to the speaker's heritage, the main criterion they had to meet was to be of Venetian origin. In the light of what has already been said (chapter 2), this choice was necessary in order to even out their different Italian linguistic repertoires and to facilitate comparability. This eventually raised a series of methodological issues that will be addressed below.

In a quite stable social situation, where people tended to live their lives in the place where they were born (as in the Veneto until the Second World War), the incidence of people whose parents were both Venetian was quite high. Linguistic and cultural differences between Venetian provinces even made marriage outside the province rarer than today. Therefore, the criterion of having a Venetian heritage was fully covered by the first generation: they were all born and mainly raised in the Veneto by Venetian families.⁵⁹ This strict classification was also applied to new migrants. This strategy favoured internal homogeneity and, therefore, afforded the opportunity to make comparisons and to identify within the interviews possible evidence of first language attrition. Important social changes inside the Italian community necessitated a re-evaluation of this criterion for the second and third generations, however. The Canadian survey highlighted a trend similar to that demonstrated by Clyne (1991) in Australia, namely a greater tendency in the first cohorts of Italian migrants to marry individuals not only from their own Italian region but also from other parts of Italy⁶⁰. The fact that they all shared a new life in a foreign country was conducive to this new social trend, resulting in an inevitable redefinition of the concept of in-group. Marriages with non-

⁵⁹ It is essential to underline the word 'mainly'. In a migratory context, people may move to different places before settling permanently. This is also the case with some people who took part in this research, especially in relation to new migrants. Therefore, it was necessary to define 'mainly' as a basic prerequisite so that emigrants should not have spent a substantial amount of time in a third non-English-speaking country. As for these countries, only seasonal migrants were considered as fitting the criteria.

⁶⁰ As I was told by several Venetians I met in Canada, the new tendency to marry people from different regions was strictly related to different migration patterns among Italians. People from the north moved more as individuals than people from the south and only once they had settled did they become 'sponsors' for other family members to emigrate to Canada. In contrast, migration from central and southern Italy usually involved entire families. The smaller proportion of women among northern Italians favoured marriages between young northern Italian men and southern Italian women.

Italians were less frequent. Therefore, it was not uncommon to find second-generation people with parents from different Italian regions. This tendency was in part accentuated among the second generation, when the pressure for women's emancipation affected the whole of Canadian society, including the Italo-Canadian community. In particular, daughters no longer wanted to recognise the patriarchal structure of the family in which they were raised and at the same time they were fascinated by the more emancipated female figures of Canadian society, resulting in their greater willingness to marry men from different national backgrounds. In parallel, second-generation Italian boys moved on to look for more conservative wives, extending their marriage range beyond the Italo-Canadian community (from an interview with one of my informants). With reference to the second criterion, according to the main classification, it should have encompassed people both of whose parents were born into the second generation, but the Canadian situation turned out to be more complex. Among the Italo-Canadian community, marriages between members of the second generation and new migrants or the late cohorts of the first generation were not exceptional. Particularly frequent were marriages between women of the second generation and new migrant men. Two main reasons were given by my informants. First, women were still more oriented to choosing someone with the same background, both from personal choice and to please their families:

e lui (il suo futuro marito) ha iniziato a venire a casa. Anzi non mi piaceva! Ma quella è un'altra storia. A mia mamma li piaceva tanto [...]

And [her future husband] started to visit my family. In fact, I didn't like him! But that's another story. My mother liked him so much [...]. [Second-generation woman; my translation]

Secondly, Italians, especially recent emigrants, were highly valued by second-generation Italo-Canadian girls:

ti devo dire la verità. Quando che io ero teenager le ragazze da scuola dove che io andavo preferivano gli italiani. Motivo che. Erano uomini più maturi. erano più maturi. [...] dopo erano anche belli uomini. [...]

I must tell you the truth: when I was a teenager, my female schoolmates preferred Italians because they were more mature 'men'. And they were also even handsome men. [Second-generation woman; my translation]

This situation posed the question of how to classify the children of such marriages, who could be considered either second or third generation. According to surveys, the mother is usually in charge of her children's education and her presence in their lives is usually quantitatively greater than that of the father. Therefore, it was decided to track the maternal side of the family, considering people born of a second-generation mother and a new migrant or first-generation father as belonging to the third generation.

With reference to the definition of new migrants, the above historical account (chapter 2) has introduced this category, exposing the effects of the new immigration policies introduced in Canada in 1967. This new set of immigration criteria had consequences in terms both of the type of migrants received or more straightforwardly accepted and of the networks they eventually created in the new country. As already discussed, these new requirements favoured highly qualified professional migrants rather than manual workers, as had been the case till then. These new waves were usually composed of highly-educated people or of entrepreneurs.⁶¹ The sponsor system, on which mass migration was largely based, was suddenly brought to an end and consequently, so also was the Little Italy system. Migration became an individual experience. These new migrants, of a different social type and in smaller numbers, were more likely to spread and to settle mostly outside Italian neighbourhoods.

In relation to this thesis, the criteria discussed here applied fairly consistently to the sample. New migrants reflected these traits, although their distribution has thus to be perceived as a continuum, rather than a strict classification. Their education level ranged in fact from vocational school certificates to PhDs, making them more highly qualified than first-generation migrants. Also, they settled more widely, as all but two of them chose to live in non-Italian neighbourhoods.

Moreover, specific criteria have been followed for each different tool used—namely interviews and questionnaires. With regard to interviews, in fact, besides the general sampling criteria used to classify my informants into migration cohorts and generations, specific standards were applied in order to select informants who could also be studied in relation to first language attrition.

⁶¹ This is another category willingly accepted by the Canadian Government, which wanted new businesses, boosting the economy and possibly generating new jobs for Canadians.

3.1.2.1 Interviews methodology

Interviewees were classified according to two variables: generation (first generation, second generation, third generation and new migrants) and gender (female and male). Notwithstanding that the variable of gender has not been considered as an independent variable in this research, it seemed important to keep it under control as it may potentially impact on one's language skills (see chapter 1).

In order to achieve a good balance, each cell was composed of the same number of informants (seven). The final sample is presented in the table 3.1.

Table 3.1 Sample of interviewees included in this study

	Female	Male	Total
First generation	7	7	14
Second generation	7	7	14
Third generation	7	7	14
New migrants	7	7	14
Total	28	28	56

The corpus analysed in this study comprises a collection of 56 interviews, selected from the total of 80 recorded during three months of fieldwork as the most representative of the population under investigation.⁶² The selection criteria differed for each generation: for the first generation, among whom I recorded the highest number of interviews, the criterion was related to the social traits of my informants, chosen as using English significantly in their everyday lives and thus also qualifying for first language attrition studies. For the second generation, the criterion was purely linguistic: interviews where the use of dialect was consistent were discarded. The main selection criterion for the third generation was related to their level of skill in Italian: the interviews chosen were those best performed by this generation. Among new migrants, finally, the criterion was again related to the informants' social characteristics: by seeking a balance among migrants who moved to Canada from the 1970s onwards, chosen informants were spread over these four decades.

⁶² Two other interviews were recorded with key informants of the Italo-Canadian community: P. R. (News Manager & Senior Editor at OMNI TV) and U. M. (News and Programme Director, CHIN Radio). The last one was recorded with a minister of religion who had served the Italian community for some decades.

Particular attention was paid to the choice of the informants from the two cohorts of first-generation and new migrants. These are sharply definable as two separate groups, differing from each other in terms of several traits, the most salient of which are worth briefly recalling here. According to this sampling division, it is not possible to refer to a single independent variable: the boundary between the two groups comprises a cluster of variables, all connected with each other. Divergences between the two cohorts are thus not merely associable with the traditional parameter of ‘time since onset of attrition’; they also involve other factors, such as amount and type of contact as well as their degree of education. These variables have been discussed in chapters 1 and 2 in general terms, where their role in preserving the language or, on the contrary, favouring attrition has been highlighted. Here, they are briefly reviewed specifically with regard to the population investigated in this study.

In a migrant setting, the two requisites of L1 disuse and L2 interference are linked, interdependent and mixed to varying degrees. In particular:

Change and deterioration of the L1 which may be witnessed among migrant populations may be determined by two opposite poles: speakers who do not use their L1 at all may experience some degree of ‘atrophy’, while those who live in a bilingual migrant community where L1 and L2 are used frequently alongside each other and mixed to some degree may find themselves sharing in a language with accelerated signs of contact-induced changes (Schmid 2011b, p.171).

Disuse of L1

As regards the disuse of L1, it is important to recall how Italian communities abroad remain in some measure conservative with respect to their culture and language, particularly from an intergenerational perspective (see chapter 2). Social traits that help Italians to maintain their language to some extent in the following generations may also play a role in preventing language attrition among native speakers. Hence, with awareness of the relatively significant level of language maintenance—as a result of language use—among Italians in Canada, it was reckoned crucial to include only those informants who made significant use of their L2 (English) in everyday life, in the working sphere and/or within the family and social domains, with almost monolingual English-speaking offspring and multiethnic friends.

L2 interference

With regard to the second aspect conducive to language attrition, L2 interference, some conclusions may be drawn from the domains just examined. As discussed in chapter 2, first-generation Italians in Canada commonly allow themselves few contacts with L2 English compared to new migrants; their everyday lives are generally lived partly within the Italo-Canadian community. Bearing this in mind, during the setting up of the sample, informants of the first generation were chosen as being comparatively well integrated into Canadian society. New migrants, in contrast, use English or English and Italian at work, while activities and hobbies in their free time are pursued mainly within the multicultural Canadian community, which implies the possibly exclusive use of English. Moreover, all informants were either monolingual or minimally bilingual at the time of onset of attrition. Although the study of foreign language/s has become compulsory in Italy, the level taught and required is relatively basic. Moreover, although English has recently become compulsory, in the past another language, French, was sometimes the only foreign language available to students.

There are other weighty variables at play, which, for the purpose of this research, need to be made explicit.

Attitude and motivation

The roles of attitude and motivation have already been discussed, in chapter 1 for their repercussions on language maintenance/attrition and in chapter 2 with regard to the Italian community. The informants in this study confirmed the previous findings: whereas first-generation men had left the country for economic purposes and women mostly followed their spouses or partners, new migrants had more varied reasons. Thus, seven of the fourteen mentioned the search for a new lifestyle, whether cultural or, in most cases, social; five others had migrated to improve their working opportunities abroad and the last two, both women, claimed to have moved to follow their husbands.

Time since migration

Time since migration—and thus since the possible onset of attrition—varied noticeably between the two groups of informants: whereas the first generation migrated

predominantly in the 1950s (11 of 14), new migrants⁶³ were spread through four decades, as Table 3.2 shows. The mean time since migration thus differed considerably between the two groups: 50.1 years for the first generation and 20.4 years for new migrants.

Table 3.2 Number of informants by decade of migration

1950s	11
1960s	3
1970s	3
1980s	4
1990s	3
post-2000	4

Table 3.3 Informants by age at time of migration

	First generation	New migrants
15-19	6	
20-24	5	2
25-29	1	5
30-34	2	2
35-39		3
40-44		2

Table 3.3 lists new and first-generation migrants by age at time of migration. The informants chosen were all over 14 years old, in order for them to have had the time to learn and consolidate their knowledge of Italian (see chapter 1). Their distribution by cohort was naturally skewed: none of the new migrants had migrated in the youngest age range (15-19), while none of the first-generation migrants was included in the last two categories (35-44). The mean for each generation was in line with this, the mean age of departure being 21.6 years for the first generation and 29.9 for new migrants.

⁶³ People who migrated from the 1970s onwards are considered by me to be ‘new migrants’. Initially it seems strange that people who migrated forty years ago are labelled ‘new’, however this term refers to the new social characteristics of this cohort.

This divergence was mostly due to another key variable which significantly differentiated the first generation from the new migrants: their level of education.

Level of education

Respondents' level of education is presented in Table 3.4, which shows that divergences between the two groups were considerable, although overall their level of education covered all categories of the Italian education system.

Table 3.4 Informants by level of education

	First generation	New migrants
Elementary/ Middle school	14	2
High school/College		3
University/ Post-graduate studies		9

Although they overlapped to some extent, a demarcation line can be drawn between the two groups, with the first generation on the lower rungs of the education system and most new migrants having a university degree. In longitudinal studies, particularly of relatively recent cohorts of migrants, the education variable is quite a delicate one, as it may intertwine with time since the onset of attrition. Schmid (2011b, p.169) notes that “those attriters with a lower level of formal education had a longer average migration span. This composition of the experimental population is the outcome of the socio-historical conditions of migration from Germany during the second half of the 20th century”⁶⁴. This description is applicable to the Italian context and for historical reasons it matches particularly well the partition into cohorts, as overlapping with the Italian educational system. The introduction of the middle school diploma as a compulsory requirement for those born after 1950 implied an advance in the level of education for the whole of Italy, including those who, as adults, migrated to Canada from the 1970s onwards.

⁶⁴ Attempting to separate the effects of these two variables, Schmid (2011b) suggests taking educational level as the leading one.

These observations suggest that differences between the two groups of interviewees may not be strictly related to attrition, but rather to the effects of their different levels of education. All native speakers possess an implicit, unconscious knowledge of their L1, which they apply automatically, but they will obtain explicit knowledge (i.e. conscious and learned) only by attending classes.⁶⁵ Thus, in order to minimise the effect on the outcomes of these social variables, the tool chosen for data collection was the interview.

Language mode

An interview is usually a formal event, in which the roles of interviewer and interviewee are defined and fixed, although partially negotiable. In this fieldwork, the use of a recording device and the academic purpose of the interviews favoured a formal linguistic style, Italian-oriented and with relatively little interference from other languages, namely English and Venetian dialect. From a purely linguistic perspective, however, the presence of an interviewer with broadly the same language competence as the subjects may have favoured occurrences of switching and mixing involving these other languages. The situation in which these interviews were conducted during the fieldwork can thus be labelled as ‘intermediate mode’, where “if [...] two bilinguals interact in a more formal context, or if the speaker knows that her interlocutor does not like to mix languages, code-switching and interferences may be reduced, although language B will still remain active” (Schmid 2007, p.138).

3.1.2.2 Questionnaires methodology

The final sample of the questionnaires is presented in Table 3.5.

Table 3.5 Sample of questionnaire respondents included in this study

	Female	Male	Total
First generation	13	13	26
Second generation	16	14	30
Third generation	10	12	22
New migrants	8	7	15
Total	47	46	93

⁶⁵ Moreover, this subdivision implies an age effect: all first-generation informants were older than any of the new migrants.

Although there was no comprehensive attempt to gather an exactly equal number of questionnaires according to the two variables used for the selection of people to be interviewed, a sort of balance was sought in the questionnaire sampling.⁶⁶ At the beginning I followed no particular distribution. Thereafter, I started being more selective in choosing my informants. I tried to vary my sample, including more people of the second and third generations, who were more difficult to reach. The first generation is usually more accessible, having fewer family and work commitments. They are also much more involved in Italian associations. Moreover, it was supposed that they would have been more difficult to reach after my fieldwork, as they usually do not use the Internet (or at least they are not autonomous users), while using postal questionnaires would have been expensive.

My group of interviewees was included in the group of informants who filled in the questionnaire. After her or his interview, each informant was asked to complete the questionnaire, usually in my presence. In this way, it was possible to offer further explanations whenever an informant needed them. Moreover, this allowed me to speed up the completion, as I was able to tick the boxes corresponding to answers that had already been given during the interview. Another group of respondents who were asked to fill in the questionnaire did not take part in the interviews. Their completion of the questionnaire took place either in my presence (for example during a family dinner or a party) or not. Their participation was mostly the result of snowball sampling:⁶⁷ people who had already participated were usually willing to help me further by asking their acquaintances to take part.

This section has given some general methodological information about the population investigated and how the sampling was structured. The next section presents the data collection fieldwork that I conducted in Canada during the summer of 2009.

3.2 Fieldwork in Canada

3.2.1 Place and time

The previous chapter reviewed Italian migration to Canada with particular reference to its historical features. Census returns show Toronto and Vancouver to be the cities with

⁶⁶ While using the SPSS software in the analysis of the questionnaires, a balancing was possible at a later stage, as the software can weigh the results according to the size of the subgroup. Conversely, in the analysis of the interview data, no balancing could be carried out with software for the analysis of the texts; once the data were obtained, the researcher had to adjust the results to the dimensions of each subgroup.

⁶⁷ This is a non-probability method, used particularly when the target population is very difficult to reach. Snowball sampling relies on referrals from initial subjects to generate additional subjects.

respectively the largest and third largest Italian presence; they are also the largest and second largest among the Anglophone cities of Canada. The decision to conduct the fieldwork in these two areas is consistent with the presence of a large number of Italians and consequently of Venetians.⁶⁸

In Ontario, two main areas were selected: the Greater Toronto Area and Guelph. The GTA consists of a metropolitan district which incorporates Toronto city and some other neighbouring cities. The area housing the highest number of Italians is Woodbridge, part of the district of Vaughan, situated in the north-western area of the GTA.⁶⁹ Other areas of significant Italian presence are located mainly in the north and west parts of the GTA: Mississauga, Etobicoque, Richmond Hill and Maple. Another place of interest for my research in Ontario was Guelph, where I conducted part of my fieldwork. It is located roughly 100 kilometres from Toronto and is known to be the Canadian city with the highest percentage of Italians.⁷⁰

In British Columbia, the main presence of Italians is currently in the Vancouver area. Significant places where Italians are located are Vancouver (downtown, north Vancouver and west Vancouver) and Coquitlam. A city which is presently home to a substantial Italian community is Burnaby, just outside downtown Vancouver. All my interviews were conducted in these areas, with the exception of one individual who had moved to a town about 50 km south of Vancouver.

The fieldwork lasted about three months (June-August 2009) and was divided into two main parts, each lasting approximately one and a half months. The first part was carried out in Ontario and the second in British Columbia. The fine Canadian summers are conducive to the organisation of a variety of frequent parties, picnics and other kinds of social activities which are part of the Italian clubs' calendar. In particular, the Venetian associations in Canada celebrated two milestone anniversaries in 2009: the 30th anniversary of the foundation of the Associazione Trevisani in Guelph (13th June) and the 25th anniversary of the Federazione dei Club e delle Associazioni venete in Ontario

⁶⁸ The metropolitan areas of Toronto and Vancouver were chosen because their Italian communities share social and historical traits, thus conceivably also linguistic ones. Given the difficulty in finding third-generation migrants still speaking Italian and the fact that only a few hundred new Italian migrants enter Canada every year, this number being much smaller still if only those with a Venetian heritage were to be included in the sample, the possibility of conducting the fieldwork in two different areas offered the advantage of increasing my chances of meeting people who would fit the criteria. Lastly, each of the two areas could have become a back-up option if the fieldwork had come to a halt in the other.

⁶⁹ It is interesting to note that the progressive movement from downtown Toronto to the outskirts of the city and then to external residential areas has followed a precise direction. If we draw a line from downtown Toronto to Woodbridge we encounter first College Street, then St Claire West, Weston and Woodbridge in that order.

⁷⁰ Guelph is the twin city of a group of towns in the province of Treviso (Veneto) called 'Sette comuni della Castellana'. Numerous people from this Italian area migrated to Guelph during the 1950s.

in Woodbridge (27th June). On 26th July the annual picnics of the Associazione trevisani and the Associazione vicentini in Vancouver coincided with the start of my fieldwork in British Columbia. These and other meetings (e.g. the Italo-Canadian Festival in Guelph) were crucial appointments for my fieldwork, as they enabled me to make contact with Venetians and to make them aware of my research. The opportunity to spend the entire period of my fieldwork hosted by Venetian families was indeed a crucial source of data on the Italo-Canadian community. In particular, I learned important information about their social life, their traditions and their attitudes towards Italy and Canada. It also presented an opportunity to listen to them speaking in a more informal environment than that created during the interviews.⁷¹ Moreover, their willingness to help my research frequently made them intermediaries and guarantors within the Italo-Canadian community and in particular with people I wanted to interview.

Each interviewee was given the choice of venue, the only proviso being that it should be a quiet place where I could record them without any background noise. A total of fifty-six interviews were conducted: twenty-two interviewees were recorded at their homes, fourteen at their place of work or in the Italian consulate office, thirteen in my accommodation, four in a public park, two at the Veneto Centre in Woodbridge and one in a public library. The interviews were usually conducted on a one-to-one basis, although occasionally another person was present in the room but was asked not to intervene. In the case of minors, the presence of a family member was required during the interview. This person was also asked to sign the personal consent form.

3.2.2 Ethical issues

A critical ethical issue related to the use of research tools is the degree of informants' awareness of the aims of the research. It has been suggested that one should not let informants know about the real focus of a study, in order not to influence their answers, both in terms of content and in the way they use language. It must be said, however, that whether or not people being interviewed are paid, their participation still represents a favour in terms of cooperation (Nortier, 2008). Thus, informing them about the objects of the research they are participating in is an ethical choice which underlines a sense of respect for the informants, so it is usually preferred at least to outline the broad topics that the research is to cover. In doing this, it is still possible to preserve the authenticity of participants' answers, as well as the ethics of the research. The present study adopted

⁷¹ This information source was also key in investigating whether the data obtained from observation matched the fieldwork results, in other words whether interviewees actually practise what they say.

this policy, presenting itself in the broad terms of a cultural, linguistic and social analysis of the Venetian communities in Anglophone Canada⁷².

3.3 Analysis tools

The dichotomy of methodology versus technique is of primary importance when dealing with analysis of data. Whereas methodology is about discretionary choices, technique is a merely instrumental operation and thus executive (Tuzzi, 2003). But these categories—methodology and technique—are not to be seen as strictly separated. Notwithstanding the mediation of software, even instrumental measures are composed of discretionary choices, which are not only a part of the process, but also represent its true strength (Tuzzi, 2003). The next subsection therefore deals with instruments and techniques, but also includes alternative options carefully considered and chosen by the researcher.

3.3.1 Data collection methodology

In this study I decided to work with two well-used tools in sociolinguistic analysis: questionnaires and interviews. There is a broad body of research which examines and offers suggestions on the design and administration of questionnaires and interviews, and on subsequent treatment of the data obtained (e.g. Dautriat, 1979; Bernardi, 2005; Caselli, 2005; Milroy & Gordon, 2003; Tagliamonte, 2006 & 2012; Codó, 2008). What is important to remember here is that these are different tools, particularly in relation to the outcomes achievable from each. Broadly speaking, whereas the first adopts a quantitative-descriptive perspective, the second offers more a qualitative-interpretative view of the phenomena being investigated. This does not imply a different level of objectivity but different research goals: the questionnaire is more effective for extensive sampling and thus obtaining a sizeable amount of data, whereas the interview aims to obtain in-depth data from a relatively small number of informants.

Another significant divergence lies in the evidence we can gain from each. In the case of the linguistic analysis of an interview (rather than the analysis of its content) it is possible to examine the subject's linguistic behaviour. In contrast, the aim of the questionnaire in linguistic research is not to discuss respondents' linguistic skills but to

⁷² I attended a class organised by the University of Exeter on 'Ethical issues in Social Science Research and the Data Protection Act' on the 26th of January 2009. For this research I followed the guidelines of the University of Exeter on ethical issues (<http://intranet.exeter.ac.uk/socialsciences/ethicscommittee/>).

reveal their attitudes. Attitudes can to some extent suggest the likelihood of behaviour; while they may not correspond closely to actual behaviour, they nonetheless play a supportive role in directing it. These two tools are thus to be taken as complementary rather than contrasting, each revealing different facets of the same picture.

3.3.1.1 Interview

The interview is a flexible tool, which also allows the researcher to obtain a large amount of information, both as to the language competence of the informants as well as information about their background, attitude, linguistic behaviour and contacts with the heritage language. Moreover interviews seemed to be the most apt to level out the differences in education between my informants: whereas the first generation and new migrants differ sharply in terms of years of schooling (as seen above), second- and third- generation informants usually have learnt the language from their family (and sometimes attended classes), thus having an implicit knowledge.

The interview technique has some undoubted advantages: it allows the focus of language attrition to move from examining errors to a broader analysis of what is retained, which is the ‘new frontier’ of language attrition studies. It also is related to a relatively more spontaneous way of using language, allowing the researcher to obtain data in fairly natural conditions.⁷³ As Schmid and Köpke (2009) remark, “if the goal of an investigation is to judge to what degree language attrition is a ‘real’ phenomenon that might impact on people’s lives and their ability to communicate, then millisecond differences in [response times] in a picture-naming task may be of little relevance” (p.221). Particularly in the case of intimate and personal topics, the use of interviews in language attrition studies helps in measuring and analysing language attrition at the level at which it naturally occurs (Schmid & Köpke, 2004).

During the preparation of my fieldwork I worked on a set of topics to use with my informants. These were divided into five main categories, each subdivided in turn into subgroups. The set of topics used during the interviews is as follows, ranging from more general and background questions (usually easier to answer) to those where evaluations were required:

⁷³ Moreover, the level of literacy, which we have seen to be linked with both social and cognitive variables, is dependent also on the choice of the task. Its effect is more apparent in formal tasks, where a difference among the informants has a strong impact on the results, but it seems not to play a primary role in spontaneous speech (Schmid & Dusseldorp, 2010).

Table 3.6 List of topics covered in the interview

<ul style="list-style-type: none"> • Family migration history <ol style="list-style-type: none"> a. Who emigrated, when, from where. b. What he/she/they did in Italy (work). c. Memories of the journey d. What the main reason was for migration. Did he/she/they think they would stay in Canada for the rest of their life, or were they intending to going back to Italy? e. How migration to Canada was, compared to other countries which many Italians moved to.
--

<ul style="list-style-type: none"> • New Environment <ol style="list-style-type: none"> a. How the situation in the new country was when they first arrived (with particular reference to the Italian and Venetian community) and how it is today. b. If there was anyone who returned to Italy. If they wanted to do so in the future. c. If one/some of their sons/daughters did so, what they would think.

Family

<ol style="list-style-type: none"> d. The composition of his/her family. If he/she is married and the nationality of his/her partner. e. Presence of mixed marriages and the community's attitude to them⁷⁴.

3. Linguistic domains

<ol style="list-style-type: none"> a. Family b. In Veneto, people continue to speak dialect with family members. Comparison with Canadian linguistic behaviour. c. Neighbourhood d. Extra-family
--

4. Meta-linguistic evaluations

Heritage languages

<ol style="list-style-type: none"> a. When in Italy, if there is the opportunity to listen to dialect being spoken. On what occasions.

⁷⁴ Mixed is a generic term as it may refer to a joining of different ethnicities, religions, or languages. During the conversation the informant was able to explain which aspects of a mixed marriage s/he would be willing to accept and which not.

- b. If they do or do not use dialect, and the reason for this choice. Generally speaking, is it considered important among Italians to maintain dialect? (The value of language conservation versus language innovation).
- c. Potential differences between Italians (young people versus adults, women versus men).
- d. Words that are used in Canada but no longer in Italy.
- e. In Italy some people say that dialect is going to disappear. Personal considerations on the possible future scenario in Canada and also in Italy.
- f. Contacts with people in Italy (with whom, since when, what kind [means of contact: letters/emails, travel to Italy, visits from Italy]).
- g. What language people in Veneto mainly use.
- h. If Italian is being spoken among Italo-Canadians.
- i. Have they studied Italian? (If yes, when, where, for how long, and if they feel confident in it).

English

- j. When it was studied (if not a mother-tongue speaker, ability in using it and any particular difficulties).
- k. Eventual role of English in the disappearance of Italian and/or dialect.

5. Personal form to fill in

(Age, education, occupation, family status) and consent form.

Questions of a more personal nature were posed in the first and last sections. Whereas the historical background (participants' personal or family stories of migration from Italy and settlement in Canada) is usually an opening topic, as it helps in creating a friendly and collaborative atmosphere, specific personal information calls for more confidentiality, as it often concerns questions that may turn out to be sensitive to some informants (Bernardi, 2005), for example those related to their social status.

The aim of the second, third and fourth sections was to investigate the core topics of this research: identity and attitudes, languages and domains of use and metalinguistic evaluations. The key point was to introduce the object of the research gradually: within each grouping, the sequence of questions followed a 'funnel' technique, from the general to the more specific, with factual responses being elicited first, followed by questions requiring the expression of an opinion (Caselli, 2005).

3.3.1.2 Questionnaire

The second tool used is the questionnaire, which is composed of 52 questions (the Italian version) and 53 (the English one)⁷⁵ [see Appendices]. It is therefore quite long, although not all the questions had to be answered by everyone. My informants were in fact guided through and asked to skip a question/s if not relevant to them. Moreover, in order to facilitate the analysis they were required, if not expressly asked otherwise, to pick only one option. This request was made in order to facilitate the subsequent analysis of the correlation between these variables.

Considering the length of the questionnaire I decided to reduce as much as possible the number of questions that had to be answered by writing, as these tire the informants and are more likely to be left empty. They also create more problems during the analysis phase as they have to be categorised manually by the researcher.

The questionnaire grid paralleled that of the interviews. It was also subdivided into five sections as follows:

Table 3.7 List of topics covered in the questionnaire

Background	•Family and migration history
Affiliation	•Networking/ Identity and Attitudes (towards Italian and English). Environment (Isolation versus inclusion)
Language knowledge	•Knowledge of the languages and domains of use
Contacts with the Heritage Language	•Contacts with the heritage language (through travels to the home country, heritage associations, and mass media)
Personal Data	•Socio-demographic information

As seen above, both interviews and questionnaire touch upon the same themes. The choice of these topics was linked to the sociolinguistic variables that have been discussed in the previous chapters and that have been found to play a key role in language maintenance/attrition. Language usage, which language they use and with whom, their affiliation and sense of identity, the amount and quality of contacts with the heritage language, all have been found to play an important part in language

⁷⁵ The extra question is related to their knowledge of the Italian language.

maintenance, or on the contrary, to favour attrition. These variables were translated into the questions used in this study and served as the basis for creating my interviews and questionnaire grids.

3.4 Corpus

Every textual analysis is based on a collection of texts, a corpus. These texts are homogeneous according to one or more discretionary criteria and they also have to be coherent with the aim(s) of the research (Tuzzi, 2003). Given this definition, we can delineate the object of this analysis as the linguistic oral production in Italian of Venetian people who had emigrated to the selected areas of Anglophone Canada and of their descendants living there.

3.4.1 Pre-fieldwork test

Before going to Canada for the fieldwork, I conducted a pre-test to evaluate the reliability of the two instruments used to gather data, namely the questionnaire and the interview protocol.

While the interviews were all conducted in Italian, the questionnaire was prepared in both Italian and English versions (see Appendices), with participants being given the choice of which version to complete. Consequently, the English version of the questionnaire enabled the inclusion in the survey of those individuals who might have been reluctant to take part in an interview because of their lack of competence in Italian. The two versions were double-checked by native speakers of English and Italian, each having a background in the other language. This favoured a comparability between the two, avoiding any mismatches. A pre-test was then conducted on three people in order to evaluate the validity of the questions and to indicate possible changes. These people belonged to the second generation, third generation and new migrant groups and were selected for the pre-test as they were willing to participate in this research but were not suitable for the fieldwork because of their lack of Venetian regional heritage. This pre-test was also important in order to unify the choice of the vocabulary used in the Canadian context: words related to the education system or other specific fields, as well as the meaning Canadian people attached to that vocabulary, were investigated and adjusted when necessary.

A pre-test of the interviews was held two months before the fieldwork in order to evaluate the grid of questions. The test interviews were carried out in England on two

English people with an Italian background, although not a Venetian one. This helped to improve the interview structure in various ways: a reformulation of the questions to make them more precise and the addition of new points that emerged as significant. This pre-fieldwork was also an important phase in developing confidence with the instruments before the real fieldwork started: it helped in evaluating the suitability of recording in different environments and gave an average of the time needed to complete the grid of questions. An important additional benefit was that it helped me to become more confident in managing an interview.

3.4.2 Interviews

All the interviews were open-ended and conducted one-to-one. In order to improve the recording quality, a small microphone was used.

As we have just seen, the grid of questions was constructed before the fieldwork and tested in pre-fieldwork. The questions posed to each person did not follow a completely fixed order, although specific milestone topics had been identified and were included in every interview. The choice of a flexible grid was made in order to adapt it and make it appropriate for different generations. However, some questions were intentionally omitted for some informants because they were not relevant to a particular generation.

During the first part of the interviews, informants were asked to speak about their migration history or that of their family. By initiating an interview through talking about the interviewee's life, it is usually possible to gain two positive outcomes: first, the interviewer may stimulate changes in the speech of the interviewee to bring about a less formal style;⁷⁶ second, there is the opportunity to establish a more empathetic environment, where the interviewee will feel more at ease and where any barriers created by the formal roles of the interviewer and interviewee can be overcome more easily.

Another point related to the setting up of a more informal environment concerns the approach to specific types of questions. Usually, when there is a need to elicit particular types of information that may be considered intrusive or on which the interviewee does not wish to express a personal opinion, it can be useful to avoid a direct approach just by asking her/him to make a comparison between two things (in this case, for example, between the situation in Canada and that in Italy). In this way, it is possible that the

⁷⁶ However, we must be aware that although there are several expedients that can be put in place in order to elicit more informal speech, truly natural informal speech is by definition not attainable in an interview setting.

speaker will express her/his opinion and answer the question indirectly without the interviewer sounding too intrusive (see question 1.e). Another device that needs to be adopted is avoiding asking people directly about a topic that may be perceived as problematic. An example concerns from question 2.c, which asked (when applicable) how interviewees would have reacted if one of their offspring had gone to live in Italy. The aim of this question was not to assess their attachment to their family, but their attitude towards their home country and how they perceived it. In the case of a positive approach towards Italy (perceived as a pleasant, modern country to live in) their openness to the possibility of having children living there was assumed to be more affirmative. The opposing attitude would have been demonstrated by more disapproval of this hypothetical situation.⁷⁷

3.4.2.1 Transcription issues

Transcription poses for the researcher a series of problems and decisions to be made. I have hinted above at one first type, namely the problems related to the comprehension of the recorded materials and the importance of good quality recording as the first step. In this section I discuss those related to the actual production of a set of transcriptions. The fifty-six interviews included in this study were transcribed in their entirety. This section explains the process of transcription, and concludes by analysing some methodological choices adopted during the fieldwork and the subsequent phase of data handling. Turell and Moyer (2008) correctly assert that “transcription is already a first step in interpretation and analysis” (p.194); hence, as the selected methodological approaches have a great impact on the research outcomes, it is vital to incorporate a precise description of the stages in order to frame the analysis correctly and to interpret the results accurately.

The ability to speak Italian was a requirement for inclusion in the sample. However, it is well known that a bi/multilingual speaker may (more or less deliberately and consciously) use more than one language during a conversation. English, dialect, Italese and Italian were expected to enter into informants’ speech, regardless of the main language requested. Instances of code-switching⁷⁸ and passages in different codes were

⁷⁷ I noticed that this issue applies particularly to Canada, where people move quite frequently. Furthermore, the Italo-Canadian community is generally wealthy and therefore able to make trips to Italy quite frequently.

⁷⁸ I am using the term 'code-switching' in this work to signal the use of words in a language other than Italian. I am aware that this is an employment of the notion at its most simple level. Code-switching is a complex concept and in other circumstances would require in-depth consideration with regard to the theoretical aspects and in the phase of analysis, but such discussion is beyond the scope of this work.

possible; during the interviews they enabled the speaker to complete her/his discourse. As explained below, these phenomena were considered during the transcription phase. The work of transcription requires a sustained level of attention. It is therefore a demanding activity, particularly if it has to be done for many hours at a time. Although not all experience the same problems, there are some difficulties that recur for everyone. Apart from possible background noise, which was kept to a minimum thanks to a careful choice of the location and the use of a microphone, recorded voices do not 'sound' natural. In particular, while listening to a voice naturally our brain can select the voice as the only sound to pay attention to. While listening to a recording, on the other hand, all the sounds are registered as they actually occurred and we are not then capable of putting them into the background and devoting attention only to the voice. Background noise may therefore become distracting. Given these considerations, I decided to work by subdividing my transcription into small chunks during a relatively long period of time (around 9 months) and allowing myself about one month and a half more at the end for re-checking.

Before passing on to discussing the instrument used in the transcription process, I want to highlight one point regarding the choice of the type of transcription. There are in fact different kinds of transcription, broadly divisible into orthographic and phonetic. The choice of a particular type is dependent on the research questions pursued and on the possible use of software.

The choice adopted in this work is orthographic, a verbatim transcription of the recorded material. This choice fits the aims of this work, which pursues an analysis of the lexicon used among the Italian-speaking Venetian community in Canada. A phonetic analysis was therefore not suitable for the purposes of this work. Moreover my decision to use the software Dragon Naturally Speaking, an important choice in order to speed up the process of transcription, entailed an orthographic transcription.

Speech recognition software

In order to facilitate the transcription I decided to make use of speech recognition software. Among the programs available, I chose to use Dragon Naturally Speaking. This choice was made primarily as I was advised by other users in the first instance and because I already had a little experience with it. Before starting the real transcription of

my interviews I decided to take a week for practising, which allowed me to reach the competence I needed to use the software. Dragon Naturally Speaking is in fact a piece of software based on a voice recognition system and therefore training is necessary to allow the software to adjust to the specific voice it has to transcribe. Although in the more recent versions it is possible to operate automatically from a voice recorder to the software, the presence of many speakers in the corpus and the quality of the audio might compromise the outcome of the transcription. The best solution seemed to be direct use, with one person dictating to a microphone while listening to the recorded voice. This was my choice, and it allowed me to speed up the process of transcription.

Although the transcription improved as the process continued, some problems still persisted. First of all, the software can recognise only one language at a time and mine was the Italian version. This means that words in English, Dialect and Italese could not be recognised and had to be transcribed manually. This happened even with Italian words if not pronounced precisely, but a prompt check allowed me to correct these cases.

Transcription

Aside from more practical difficulties, the major problem of the transcription I undertook related to the transcription of words not in Italian. I first needed to distinguish between words in a foreign language (usually English), words in dialect and Italese, and words not found in any dictionary. Overall, my choice was to follow the transcription method that employs the spelling system of a standard language (either English for English words, or Italian for the remaining cases).

For English, in particular, I represented words articulated in non-standard pronunciation by the standard English form. By adopting this criterion, both intra- and inter-generational homogeneity were achieved, making lexical comparisons among migrants belonging to different cohorts easier. Doing this also made it possible to keep together on the one hand people belonging to the first generation and new migrants (for whom English was not the mother tongue and whose pronunciation may have differed both from the standard and from each other) and on the other hand the second and third generations, for whom English was the mother tongue or at least the language in which they were usually more confident.

As for dialect, the transcription was based upon the Italian spelling system, which enabled me to have one standard transcription scheme to follow. An exception to this

rule was made in order to facilitate the recognition of the common dialect form *se* /ze/ (meaning *is*, third person singular of the verb ‘to be’), which could have been mistaken for the Italian word *se* /se/ (meaning *if/whether*). A phonetic transcription would have made it possible to distinguish these two words, although it could not have been used with a software application for the analysis of texts. Therefore I decided to retrieve a Venetian writing tradition that renders an initial voiced ‘s’ with the letter ‘x’. Although this has not been common practice among other researchers, for the purposes of this study it was an expedient to facilitate the interpretation of the results.

In the case of words that were not found in any dictionary (being non-existent, in Italese or pronounced incorrectly), the rule that I followed was to transcribe them as they were pronounced, according to standard Italian orthography.

Another major problem, although quantitatively limited, is the occurrence of homonyms between the two languages, namely English and Italian. If within the same language homonymy is signalled by the software, as will be seen later in this chapter, that between two languages cannot be recognised by the software and thus is not signalled. Occurrences in English are thus put among those in Italian, and may be a problem when the software is tagging, particularly, as in the examples reported below, both forms have a high number of occurrences. The most relevant example, with all occurrences taken from the first generation, is reported below:

adesso se sbaglio xo⁷⁹ [so] what? non importa. comincio da capo e sennò mi faccio aiutare.

[now if I get it wrong so what? I doesn't matter. I start again or I ask for help]

il suo unico desiderio era di essere con noi e con il marito. non importa dove. xo [so] yeah quello che pensavo io era che una volta partita non poteva tornare più.

[her only wish was to stay with us and with her husband. it doesn't matter where. so yeah what I thought was that once she had left (Italy) she couldn't go back]

non so se lo ha mai sentito dire.

[I don't know if you have ever heard about it]

⁷⁹ This transcription choice is merely a technical expedient, as there is no such word in Italian, and it is not related in any way to the Italian phonological system.

oggi posso [tornare in Italia] e domani non so cosa che può succedere.

[these days I can (go back to Italy) and in the future I don't know what can happen]

in the first two examples the underlined word 'so' is used with the English meaning. In the following two examples it is used with the Italian meaning of '(I) know'. In order to have the software distinguish between the two forms it is important to choose a different orthographic format for the language which is not recognised by the software, in this case English. The choice is to a great extent personal (i.e. the adding of an accent or the changing of a letter, which I decided to do), although it has to be consistent throughout the corpus. The fact that we automatically infer the appropriate meaning of a word from the context means that we are usually unaware of such cases of homonymy. Also, as in the case reported above, although the two forms are written in the same way, they are pronounced differently and thus the difficulty may not be apparent before transcription. It can therefore be useful to make a list of these words while transcribing and make the relevant corrections during the revision.

Code-switching

Other important issues that were met during the transcription related to the phenomenon of code-switching, where interviewees briefly used an idiom other than Italian, the only language that they were asked to use in the interview. They did in fact make some use of English and dialect during the interview so that they could complete their discourse without interruption. The investigation of code switching is beyond the scope of this thesis, however, as it would have required a qualitative approach, in contrast to the quantitative approach followed for the analysis of the corpus. I therefore attempted to find a balanced qualitative and quantitative approach to dealing with switching.

A reading of the first transcription permitted the isolation of passages in which a language other than Italian was used. These were considered qualitatively, aiming to distinguish whether the language was chosen to cover a temporary or more permanent lack of competence in Italian, or to substitute entirely for Italian by expressing a concept or idea that the speaker was not able to articulate clearly in the language s/he was asked to use; in other words, whether the role of this other language was to support Italian, covering any linguistic gap in it, or whether the speaker made no attempt to express

her/himself in Italian. However, it could also be the case that words/sentences were uttered intentionally in a foreign language, without meaning that the speaker was not able to use the correct words in Italian.⁸⁰

Following reflections on the standard length of utterances belonging to the two types of switches mentioned above, a maximum length of 65 words was fixed on for strings of words in a language other than Italian.⁸¹ This quantitative choice was made upon a qualitative consideration: it appeared that this threshold acted as a boundary, highlighting the function that non-Italian sentences played in participants' speech. Below this limit were included all sentences in which a language other than Italian seemed to me to be used to cover a deficiency in the individual's Italian competence or because it was her/his deliberate choice to use another language, while above it were all sentences that reflected the speaker's lack of ability to master Italian at a basic level.

The final outcome was the exclusion of four pieces of discourse, each exceeding 65 words in length, spoken by a third-generation male and a male new migrant. Although probably motivated by different factors (a real inability to express himself properly in Italian in the former, and an overt preference to articulate those sentences in English in the latter), all four examples, two for each speaker, showed no attempt at conveying a message in Italian.

Limits of the transcriptions

One last point worth discussing is related to the transcription choices undertaken in this study and in particular to their limitations. The problem of devising the most appropriate transcription rules is linked to the necessity of finding the best compromise between the resources available, the potentialities of the software used and the aims of the study. This choice cannot be entirely fixed *a priori* and an individual path has to be devised for each study.

In order to obtain a full picture on the phenomenon investigated we should analyse different linguistic levels. The validity of this approach has been clearly corroborated in a study by Schmid, Verspoor and MacWhinney (2011), where it emerged that an informant (this time an L2 learner) may regress on one level (namely lexical sophistication) during the learning process because her/his cognitive resources are

⁸⁰ This was particularly true with dialect words. The speaker often wanted to prove that s/he was able to speak dialect, not just by telling me that s/he could, but by answering me in dialect or using some idiomatic or typical phrases in Venetian.

⁸¹ This was eventually applicable only to English, as dialect was always used in single words, idiomatic sentences or very short phrases.

diverted to implement other level(s), namely the complexity of the sentences produced. What is highlighted by the authors is that the informant runs a sort of trade-off of her/his cognitive resources, which only apparently shows as a decrease in language proficiency. If a full picture is drawn, progress on different levels may emerge, showing even a general improvement of the speakers' skills.

My choice not to transcribe certain phenomena typical of the spoken language (such as hesitations) is therefore debatable to some extent, as I am taking into account only some aspects of the linguistic competence of my informants while not considering others that have however been proved to be important in language attrition studies (such as Schmid & Beers Fägersten, 2010; Yilmaz & Schmid, forthcoming). However this choice has to be seen in terms of the need to transform a spoken text into a new version, easily usable with software for the analysis of the text and fitting the purposes of the research, which is aimed at studying lexical use and verb morphology among migrant communities. Other linguistic aspects, regarding, for example, pronunciation or hesitations, do not relate to the aims of this analysis and have, therefore, been excluded from the transcription. They are, of course, relevant to the study of skills among speakers of a heritage language and can offer a more complete picture of the phenomenon investigated; however, here they were outside the more narrow scope of this piece of research.

3.4.2.2 Lexicometric measurements

In order to evaluate whether the corpus had the features to allow a statistical study of its content, it was necessary to verify that the lexicometric measurements were below specific thresholds (Tuzzi, 2003). The lexicometric values for the whole corpus are given in Table 3.8.

Table 3.8 Lexicometric measurements

Word tokens (N)	342,411
Word types (V)	15,317
Type/token ratio (V/N * 100)	4.47%
Hapax legomena (*%)	47.3%
Average frequency (N/V) (*%)	22.3

The corpus comprised a total of 342,411 word tokens, which gives a measure of its size, and 15,317 word types, which represents the extent of the vocabulary.⁸² The type/token ratio (TTR) and the percentage of *hapax legomena*⁸³ are two indicators of lexical richness. As the table shows, in this study they were below the respective thresholds of 20% and 50%, which means that this corpus can be analysed statistically (Tuzzi, 2003).⁸⁴ Given the size of the corpus, these outcomes were as expected.

Before proceeding with the presentation of the results, it is essential to highlight a methodological choice that was made during the phase of preparation of the corpus. The choice of dealing with word tokens cannot be taken for granted, as other paths could have been followed. Working with lemmas, for instance, would have been another option (Cortelazzo & Tuzzi, 2007; Schmid, 2002).

3.4.2.3 Tagging and manual disambiguation

The last stage before proceeding with the analysis was that of manual disambiguation. Grammatical tagging is a tool offered by software for the texts' analysis, which helps to classify word types into their corresponding grammatical categories. Two examples are given in table 3.9.

Table 3.9 Examples of outcomes

Graphic forms	Occurrences	Length	Grammatical category	Imprinting	Lemma
è	8337	1	V	indic_pres_s_3	essere
dialetto	1248	8	N	s_m	dialetto

The two word types (*forma grafica*) are *è* (*is*) and *dialetto* (*dialect*). Their number of occurrences in the corpus (second column) is followed by the number of characters in each word type (third column). Grammatical category refers to the grammatical

⁸² By distinguishing word tokens vs. word types, we can introduce the concept of frequency referred to by these two terms. Type frequency is the frequency of an item pattern, while token frequency is the frequency of actual items. As it contains between 50,000 and 500,000 statistical tokens, this corpus can be described as large.

⁸³ *Hapax legomena* (or simply *hapaxes*) are word tokens that appear only once in a corpus.

⁸⁴ If we look at the high frequencies section, we notice that these occupy 31.46% of the total number of occurrences. The first four word types (*e*, *che*, *è*, *non*) account for 10.25% of the total. Since *e* is statistically the word most often used in spoken Italian, this result is in line with the suggestions of the literature. If the medium frequencies are added to the high frequencies, the result is 77%, which means that the 598 most common word types account for more than three-quarters of the number of word tokens in the entire corpus, whereas they constitute only 3.9% of the entire vocabulary (word types).

category to which each belongs. Here, V stands for verb, N for noun. Whenever possible, the result includes ‘imprinting’ (fifth column), which offers further detail of the grammatical tagging. In the examples given, the verbal form *è* is labelled as the 3rd person singular of the simple present, whereas *dialetto* is tagged as masculine singular. Following the sixth column, which allows different word types to be grouped under the same lemma, there can also be a column giving more information about the usage frequency of each word type/lemma.

This operation is not completely accurate, however, as there will be some ambiguous cases. Whenever a word token is not recognised as belonging to Italian, or where multiple entries are possible for the same word token (synonymy), software for the texts’ analysis labels the word type with an empty space or with the letter J respectively. Although the empty categorisation is variable, dependent on the nature of the text (more than one language used and the accuracy of the transcription), the J category is usually rather large, at around 1/3 of the entire vocabulary. This would prevent any analysis of these data, as it drastically reduces the proportion of the vocabulary it is possible to work on. The researcher must therefore carry out a manual disambiguation to reduce the non-tagged portion of the corpus to below an acceptable threshold.

Thus, in order to render the corpus treatable, I carried out a manual disambiguation, before which the level of ambiguity, comprising the percentage of non-tagged and category J items resulting from grammatical tagging with texts analysis software was 46.8% (7,177 word types from a total of 15,317). In terms of word tokens, ambiguity was higher still, at 56.1% (192,238 tokens in a corpus of 342,411). In particular, category J was rather large, comprising 3,783 word types and 176,643 tokens. Following manual disambiguation, the total number of non-tagged, and thus still ambiguous, entries was reduced to 3,098 types (20.2%) and to 42,204 tokens (12.3%). Category J was also much reduced, to 44 types (0.3%) and 28,365 tokens (8.3%).

Wherever the numerical condition allowed it, the disambiguation was done manually, by checking each single occurrence in the corpus. However, this was possible only where there were at most a few dozens of occurrences of a word type. Manual disambiguation is a time-consuming activity, as each occurrence in the corpus has to be checked, which is not feasible for word types with hundreds or thousands of occurrences. In order to make my final data more reliable to work on by reducing the proportion of ambiguous forms, I decided to adopt, whenever grammatically possible, a stricter criterion of disambiguation. Reading the list of word types with higher

frequencies, I noticed that there were some which, although rightly categorised in more than one group by software for the texts' analysis, seemed to have their occurrences greatly skewed towards one grammatical category, which was also the most commonly used in Italian. I therefore decided to classify these as if all their occurrences belonged to this category. By doing this I am aware that I will have misplaced a proportion, albeit very small, of the data in my corpus. However, the loss of this small amount of information was considered acceptable and not to have affected the final outcome significantly.

One final aspect to be discussed about the tagging phase regards articles. These represent a quite problematic grammatical category in Italian when operating a tagging assisted by software. In fact, articles in Italian are all homonymic with forms of other grammatical categories. In order to clarify this, I present an example: in this particular case, the same word type (*la*) takes up the function of an article in the first case, and of a pronoun in the second:

Questa è la mia nuova macchina.

[This is my new car]

La possiamo provare?

[Can we try it?]

Therefore, unlike other grammatical categories, with regard to articles all the occurrences fall automatically into the J category⁸⁵. In order to perform an analysis of articles it would, therefore, be necessary to conduct a manual disambiguation for all the occurrences, which, given their very high number, would be time-consuming. For this reason, and given the fact that articles do not constitute the focus of this research, I decided not to perform a manual disambiguation and to leave them in the J category. Consequently, as we will see in chapter 5, the grammatical categories to which the word types are assigned will not include articles.

3.4.2.4 Measures for the study of the vocabulary

Lexical diversity is an area that has been widely treated in linguistic research. However, it is only recently that we have seen the emergence of a comprehensive approach to this

⁸⁵ Articles, moreover, make up a substantial component of the ambiguous category of 'J' forms.

topic, more aware of the multifaceted nature of the concept of 'lexical diversity' (Jarvis, 2013b).

The measures that had been devised to calculate lexical diversity hitherto raised questions of validity. Although there have been improvements in their strength, there are still problematic aspects (Jarvis, 2013a). On a more technical level, the measures used are mostly dependent on text length. As stated in Schmid and Jarvis (forthcoming), "simple TTRs have been shown to be a problematic measure of lexical diversity in that they tend to vary a great deal as a function of text length, since the rate of word repetition inevitably increases as the text grows longer" (p. 14). Attempts have been made to overcome this problem, but even the most accurate indices in this respect (Johnson's MSTTR⁸⁶ and McCarthy's MTLT⁸⁷) have turned out not to be completely satisfactory, as "an important problem with both measures is that neither evaluates the text as a unified whole" (Jarvis 2013a, p.94).

On a more theoretical level, we need to understand that lexical diversity has to be approached from different perspectives. The indices created to determine lexical diversity can capture some valuable information, but they have been shown to be "blind" measures of lexical diversity" (Schmid, Verspoor & MacWhinney 2011, p. 44), unable to obtain all the information and to provide a full picture of the informants' actual level of lexical diversity. The importance has, therefore, been stressed of the need to go "beyond indices designed to reflect only the relationship between types and tokens, such as VOCD⁸⁸ or Guiraud⁸⁹, and seek ways in which lexical diversity and lexical access can be described more fully" (Schmid & Jarvis forthcoming, p.9). There are in fact other linguistic measures that together can give us a more complete depiction of the phenomenon of lexical diversity, such as word rarity and word dispersion, the occurrences of unique words and the incidence of relatively longer words, the occurrence of particular grammar categories (Schmid, Verspoor & MacWhinney, 2011; Schmid & Jarvis, forthcoming), all aspects that have not been 'caught' with the indices used so far.

This new approach implies a more rooted change in the perspective, highlighted by Jarvis (2013a) with the fact that:

it is recognized that indices of lexical diversity are useful, even though language researchers have neglected the question of what it is that they

⁸⁶ Acronym for Mean Segmental Type-Token Ratio.

⁸⁷ Acronym for Measure of Textual Lexical Diversity.

⁸⁸ Acronym for Vocabulary Diversity.

⁸⁹ It is an index of lexical richness.

are actually measuring. Unfortunately, lexical diversity indices tend to be validated in accordance with how well they avoid sample-size effects and/or how well they predict other constructs (e.g., proficiency) [...] rather than in accordance with how well they measure the construct they are intended to measure (i.e., lexical diversity). In other words, the problem is not that the existing measures fail to predict language proficiency [...]; the problem is that they lack construct validity because they have not been derived from a well-developed theoretical model of lexical diversity (pp. 94-95).

New research approaches in the field are, however, leading to promising findings and opening up new pathways, with the awareness that “lexical richness measures also do not take into account grammar, sentence structure or other textual features, such as cohesion, coherence or organization of the text” (Šišková 2012, p. 35). The concepts of lexical diversity and richness are thus more multifaceted aspects to investigate than is at first apparent.

Given this background outline, I now pass on to briefly discussing how this topic of lexical diversity has been approached in this study. Whereas a linguistic examination of these aspects of language is offered in the following chapters, when discussing the results of my corpus, what I give now is an introduction to the lexical tools that are available with the software application for the analysis of texts and that has been used to carry out the analysis.

As claimed by Šišková (2012),

the biggest drawback of lexical richness measures in general, however, is when looking at words used in isolation. The software available mostly recognizes a word as a group of letters separated by spaces, which means that it does not take into account compound words written separately, polywords, collocations, idioms, formulaic language or any other stretches of text which are often not further analysed into individual types and could be viewed as belonging together or having a single meaning (p. 34-35).

In the phase of corpus preparation with software for the texts' analysis, one of the measures that can be undertaken is 'normalization', which means the 'polishing' of the corpus that we are going to analyse by, for example, making uniform the use of the

accents or capital letters. Among the options available there is the recognition, through a list available in software for the texts' analysis, of the polywords. Although a word, taken as a string of letters separated by spaces, is usually taken as a single entity, it is possible to set the software to recognise strings of words that, in Italian, acquire a sense or change of meaning if considered together (e.g. *per esempio* 'for example', or *capo di dipartimento* 'head of department'). The software can recognise these forms as single entities, calculate their occurrences as in the case of all the other types and add them to the vocabulary. By doing this, we can retrieve part of the lexical information that would have otherwise been lost.

Guiraud Index

In this research I also use the Guiraud Index. Its formula is calculated by dividing the number of types (V) by the square root of the number of tokens (N), aiming in this way to compensate for variable text length. As has already been discussed above, the Guiraud Index does not completely overcome its dependency on corpus size. Therefore it cannot be considered a fully reliable index. Moreover, as we have seen, there are many linguistic aspects that operate together to shape the informants' lexical diversity, and it is only the combined analysis of these that can give us a more complete depiction of the phenomenon investigated, a multifaceted phenomenon that cannot be 'caught' with a single index. On these premises, I consider the Guiraud Index as a first indicator of the lexical diversity of my informants, a starting point for further lexical analysis that will go deeper and will unveil other aspects of their language skills.

Range

As we have seen in the previous pages, beyond the mere numerical outcomes of the quantitative treatment of the corpus, lexical statistical tools may provide "additional interpretative linguistic information such as part-of-speech annotation, grammatical parsing, and prosodic transcription" (McEnery & Wilson 2001, p.114). What I want to discuss now is the concept of 'range',⁹⁰ which is a valuable measure for corpora treated with statistical software. In fact, it allows us to isolate the (reduced) portion of the vocabulary that is of relevance to the content analysis. Moreover, within particular parameters, this notion of range can also be applied to perform purely linguistic

⁹⁰ This term is often associated with another, 'rank', which refers to the position that a word token occupies in the frequency table. This means that different word tokens which have the same number of occurrences are labelled with the same rank.

analyses. In fact, by keeping the content of corpora pretty much constant across all informants, as in the case of semi-structured interviews, we are somehow allowed to analyse this subdivision also in terms of linguistic skills.

With software for the texts' analysis, we are offered the possibility of subdividing our vocabulary list into high, medium and low ranges, from those forms which are most frequent in the corpus to hapaxes.⁹¹ This division is made automatically by the software. However, it can easily be done manually. Scrolling down the list of word types from the top, the first pair of word types with the same number of occurrences constitutes the limit between the high and medium frequency ranges, while from the bottom, the first gap in the number of occurrences signals the limit between the low and medium ranges. From a content perspective, the high range part of the vocabulary is characterised by word types with a very high number of occurrences which are not very meaningful (mainly prepositions, articles and conjunctions), while the low range is covered by the great majority of word types with a relatively low number of occurrences, including the substantial group of hapaxes. Words belonging to the medium range are usually more informative and it is among these that we have to look in order to find the most important topics covered in the corpus (Tuzzi, 2003).

From a language skills point of view, we can identify the most common word types, used by all our informants, from those forms positioned in the middle and used by the great majority, and finally from those that are the least used. In line with other research that has devised a categorisation of the vocabulary (Yilmaz & Schmid, forthcoming) we will analyse the outcomes in order to determine if tendencies and patterns emerge across our four groups.

Vocabolario di Base (VdB)

We have already introduced discussion about the Vocabolario di Base in chapter 2, dealing with it from a purely linguistic perspective.

Among the tools offered by some Italian software for the texts' analysis, the user has the opportunity to make use of the VdB list for the Italian language devised by De Mauro, to have tagged the occurrences shared with those of the VdB list. As seen in chapter 2, this list is in three parts (VdB1, VdB2 and VdB3); however, for the purposes of this study, it will be considered singly as VdB. This seems relevant with regard to the analysis of my informants' linguistic competence. A relatively high occurrence of VdB

⁹¹ Hapaxes usually represent the great majority of a small corpus and for a large one they usually cover 45-50% of the whole text.

forms may in fact be symptomatic of a use of the vocabulary made up of basic and common words. Again, it is important to stress that this analysis does not allow us to draw a conclusion about my informants' knowledge of the language. We are simply assessing their use of the vocabulary available to them, which is however indicative of their underlying language skills.

Before bringing this methodology chapter to an end, I am going to briefly discuss how the questionnaires have been treated and how the data have been entered in order for them to be analysed with the software SPSS. Compared to the choices and the methodological measures undertaken for the transcription, as discussed above, these are relatively straightforward, although it is important that they are stated clearly.

3.4.3 Questionnaire

With regard to the more practical aspects of data entry, the main features of the codification will be briefly introduced here. With the exception of questions regarding social variables (such as gender and age), the answers to the questionnaire have been coded according to a general rule: values would usually vary between 0 and 1. The first of these indicated an 'English pole', meaning either a low level of use of Italian or a low degree of affiliation to/liking for Italy and its language. With the opposite 'Italian pole', the degree of use of/affiliation to/liking for Italian was expressed with 1.

A great number of questions were devised according to a five point Likert-type scale. In this case, the data were assigned according to this scale: never = 0, seldom = .25, sometimes = .5, often = .75, very often = 1. When entering answers to the questionnaire, some questions were averaged together. For example, question 32 in the Italian questionnaire (and 33 in the corresponding English one), which is composed of 7 grids, was averaged so that a single result would appear. This helped to reduce the number of questions/variables by putting together those answers that could be considered jointly.

The questionnaire was composed of more questions than those actually analysed in the following chapters. The selection of which questions to treat was based mainly on two criteria: the first is related to the focus of my analysis, which has been further narrowed down during the review of the data collected during my fieldwork. The second, on the other hand, is related to the literature, and in particular I decided to focus my analysis on the variables that previous research suggested were more telling.

Firstly, although my informants were asked to rate their competence in, or affiliation towards, three languages/cultures, namely Italian, English and Dialect, only the first two were considered, and in particular Italian, as the main focus on this research. It was decided not to take Dialect into account in this study, while being aware of and recognising its influence in the Italian linguistic and cultural repertoire (as discussed in chapter 2). Whenever there was a question about affiliation to the Veneto region, or the origin of the people that an informant was used to spending time with, the choice of Venetian was considered equivalent to Italian. The affiliation to a particular region of Italy, in fact, is clearly related to a sense of affiliation to the country as a whole.

Secondly, according to the discussion of language attrition and maintenance as presented in chapter 1, two main sets of variables have turned out to be particularly significant among speakers of a heritage language, namely linguistic habits and attitudinal factors. Although recognising the plausible influence of a plethora of other factors, I decided to focus my analysis on these two. Therefore, only questions pertaining to these two topics were selected for analysis.

3.5 Final remarks

This chapter has completed the depiction of the framework of the study, with the methodological groundings of the analysis. Following a review of the literature on language maintenance and attrition among migrant communities, and an account of the linguistic, historical and social background to Italian migration (chapters 1 and 2 respectively), then a description of the fieldwork and data analysis methodologies in this chapter, there remains the third and last part of this thesis. The next chapter thus enters the analysis part, with a general study and discussion of the data gathered during the fieldwork, in particular through the questionnaires, highlighting if and how attitudinal and linguistic factors impact on the language skills of speakers of a heritage language across different generations and cohorts.

Part three: Analysis

Part two of this thesis provided information concerning the participants in this study, the methodological choices and instruments used during the fieldwork and the construction of the corpus. Here, in this third and last part, the outcomes of the sociolinguistic and linguistic analysis of the data collected during the fieldwork in Canada are discussed.

There is an overall ‘funnel’ structure to these remaining two chapters, from the more general background information presented in chapter 4 to the more linguistic discussion of the interviews in chapter 5. The structure of this part is thus bipartite: the fourth chapter, through an analysis of the data collected by questionnaire, examines the linguistic habits and attitudes of the communities investigated as self-reported in questionnaires. The fifth chapter involves sociolinguistic and linguistic analysis, exploring the results of interviews, first through a quantitative analysis of aspects of the language skills revealed in the interviews, then through an attempt to correlate these results with the sociolinguistic patterns which emerged in the fourth chapter.

Although treating a well-researched topic, particularly with regard to language maintenance among different generations, each chapter in this part presents and discusses novel elements: in chapter 4, while the main purpose of the investigation is to provide the grounds for the socio-linguistic analysis of the following chapter, it also offers an up-to-date depiction of the Italian language situation in Anglophone Canada, in particular pertaining to linguistic habits among speakers of Venetian-Italian heritage language. Chapter 5 is an attempt to move beyond the well studied field of the lexical variation that the Italian vocabulary of migrant speakers has undergone in contact with English (see the literature review of Italian bilingual communities abroad, and in particular the section on Italese), to propose a wider description of these communities’ linguistic skills, highlighting how much their linguistic abilities vary both among different cohorts of migrants and along a generational scale. In other words, I am not interested in framing and discussing language maintenance using a qualitative perspective as has been largely done so far. What is of more relevance in this study is quantifying these changes. Moreover, as far as I know, no study of first language maintenance/attrition has been carried out among the new waves of Italian migrants of recent decades. Although not numerically very substantial, they nonetheless embody a new phenomenon that should be of real interest for Italian society, as it represents a current and future social issue for this country and many others in Europe.

Overall, we can assert that variations within the first generation promote cascade effects on the following ones. It is thus safe to say that the pattern of language maintenance which we can now try to delineate for the old waves of migrants and their descendants may not be valid in the future. Different social and linguistic variables, both in the country of origin and in the host one, will in all probability lead towards different language maintenance paths and delineate new linguistic developments. The models that have been created to interpret and predict language maintenance and loss among minority communities seem to fit the first cohorts of migrants, but they will probably need to be revisited in the future. Moreover, I need to underline that at the moment the only possible longitudinal study is that of the descendants of the first cohorts of migrants: the relative recentness of the new migrant cohorts does not allow us to find enough second-generation speakers,⁹² while a third-generation requirement would be even more challenging. I thus hope that this research will stimulate more interest in this phenomenon, as well as opening up a discussion of possible future scenarios for the Italian language in Canada.

⁹² At the time of the fieldwork, the age of second generation speakers was on average under 16. With the exception of a few people aged about 30, the great majority were teenagers or children.

Chapter 4: Language attitudes and behaviours

4.1 Introduction

This fourth chapter is devoted to an analysis of the questionnaires, which were mainly completed and collected during the fieldwork in Canada. The choice of employing this particular tool was made in order to gather a great deal of information about the local Venetian community: questionnaires are less time-consuming than interviews and offer the opportunity to gather and treat large amounts of data using a quantitative and statistical approach.

The results of the questionnaire differ intrinsically from those of the interviews, which will be treated in the next chapter. The interview, if well designed and conducted, and notwithstanding some limits intrinsic to the very character of linguistic research, represents a relatively reliable and straightforward tool to gain access to speakers' linguistic behaviours. On the other hand, the reliability of the questionnaire as a research tool and of its outcomes, has given rise to more debate in the literature. Although recognising that the questionnaire is a useful tool when dealing with a substantial amount of data, and that it is relatively easy to handle with software, the literature advances some criticisms too. The questionnaire requires a categorical response, and therefore "respondents must be able to determine their own usage and, when given choices, compare that usage with alternatives presented in the question" (Milroy & Gordon 2003, p. 53). Moreover, it is not possible to be certain how truthful respondents have been as their answers may not be fully authentic. This may not be due to a deliberate choice to give an untruthful answer, but simply because of a personal misinterpretation of the question. The questionnaire, in fact, entails active participation by interviewees and can thus be biased by their role. However, if this tool may not be completely reliable in investigating interviewees' actual language behaviour, it can still have a certain legitimacy. Specifically, it is recognised as having the capacity to bring to the fore informants' attitudes regarding concrete behaviours, and, although attitudes may only partially correspond to linguistic practice, they are nevertheless linked to it and they often represent the lever for the real behaviour.

4.2 Social variables

In the first chapter of this thesis I presented and discussed the variables that play a role in language attrition/maintenance. The review highlighted the intricacy of the effects of many different variables on the linguistic skills of speakers of a heritage language, and the subsequent impossibility, in the very great majority of the studies presented, to single out one variable to explain the pattern of language skills outcomes among bilingual communities. However, whereas language attrition/maintenance may be dependent on many variables, some seem to emerge as producing a stronger impact.

Generation has been largely considered in the literature as a key variable, the threshold between native speakers of a language – those who came into contact with an L2 at a relatively later stage of their life, and who have received full schooling only in their L1 - and heritage language speakers. In this study generation will thus be considered as the independent variable.

In this work I will deal with two other variables discussed in the literature review, namely ‘contact with the L1’ and ‘affiliation/attitude towards the heritage country’, both of which are assumed to have an explanatory power with regard to the findings of this study.

As discussed in chapter 1, language attrition is the outcome of the combination of disuse of the heritage language and the introduction of a new language. Whereas the latter is, in the great majority of cases, intrinsic to the same experience of being part of a heritage-language speech community, the reduction in the use of the heritage language is an aspect that seems to be subject to more personal choices. As discussed in chapter 2, the modalities of use and in particular the opportunities to use the Italian language have changed considerably over time among Italian communities in Canada. Whereas in the past the occasions to use this language were mostly limited to interactions with other members of the community and within one’s family, today, although there is an undeniable reduction of use in these more traditional domains, the circumstances in which Italian can be used have multiplied, ranging from a wider access to different mass media, to Italian courses, to more opportunities to travel to Italy. The quantity of contacts made with Italian has become less dependent on the community and more a matter of personal choice. This has led to a more in-depth investigation and to the need to frame the analysis within the specific population investigated.

Hence, in this chapter I am going to explore whether any particular trend or difference regarding contacts with the Italian language across my population emerges.

With regard to the second variable, I will explore whether I can identify significant differences in the level of affiliation towards Italy among the four generations. This variable, as already discussed in the literature review, has been seen to play a significant role in language attrition/maintenance. In particular, the research by Schmid (2002) ascribes a high explanatory value to this variable with regard to the linguistic outcomes which emerged from her research with different cohorts of German Jewish migrants. We have, nevertheless, to be aware that the impact of this variable depends on the particular circumstances of migration: the author showed how the role of the attitude towards their home country, correlated to the escalation of the persecution, has played a key role in language attrition. Yet, in the very great majority of cases, circumstances that lead to migration are far less traumatic and therefore the level of attrition less sizeable.

The population investigated in this study is on the contrary less easily categorised according to migration cohorts. In chapter 2 I discussed the role of attitudes towards Italy, with particular reference to the two groups of migrants. Whereas a sense of nostalgia towards Italy is more common among people who migrated during the period of mass migration (first generation), as they were driven to leave almost solely by harsh economic conditions back in their home country or were following their families, the more recent cohorts of migrants are not exclusively in pursuit of better work opportunities abroad. On the contrary, there is a significant number of people who nowadays leave Italy seeking a new way of life, while looking at their heritage country with a disillusioned eye. The situation among heritage speakers is difficult to categorise as well. Their sense of affiliation towards Italianness, in addition to being linked to their personal background and up-bringing, is subject to external influence. In particular, the dominant image of Italy in Canada, and of the Italo-Canadian community itself, has changed considerably in recent decades. As discussed in chapter 2, notwithstanding still existing stereotypes, younger Italo-Canadians can benefit from a society that nowadays largely values Italianness and is less driven to assimilate its ethnic groups into its mainstream culture and language. New generations of Italo-Canadian are thus more likely to build a positive sense of Italianness and to feel proud of their heritage.

However, notwithstanding these general tendencies, the scenario of Italian migrant communities in Canada, with regard to their attitude towards Italy and the Italian

language, is variegated. There are not only major differences between different generations, but also within the same generation. Attitude is not a factor completely linked to one's group, but it is first and foremost the outcome of a personal inclination.

4.3 Analysis and discussion of the questionnaires

This section examines and discusses the results of the questionnaires administered during the fieldwork.

As mentioned earlier, a total of 124 questionnaires was returned, primarily during the fieldwork. Of these, 99 (45 in English and 54 in Italian) were selected at a later stage; questionnaires only partially completed or not signed⁹³ were not taken into consideration. As explained in chapter 3, a double version of the questionnaire was created, one in English and the other in Italian. Although 45 people opted for the English questionnaire,⁹⁴ only five of them reported not being able to speak Italian. All the others are assumed to have been more confident in English than Italian (second and third-generation informants), or more inclined to use this language despite being native speakers of Italian (first-generation and new migrants).

Although an attempt was made to create a balanced sample during the data collection in Canada, the final composition of the questionnaire sample was skewed in favour of the second generation (30 questionnaires), followed by the first (26), third (22) and new migrants (21). In terms of percentages the distribution is shown in figure 4.1.

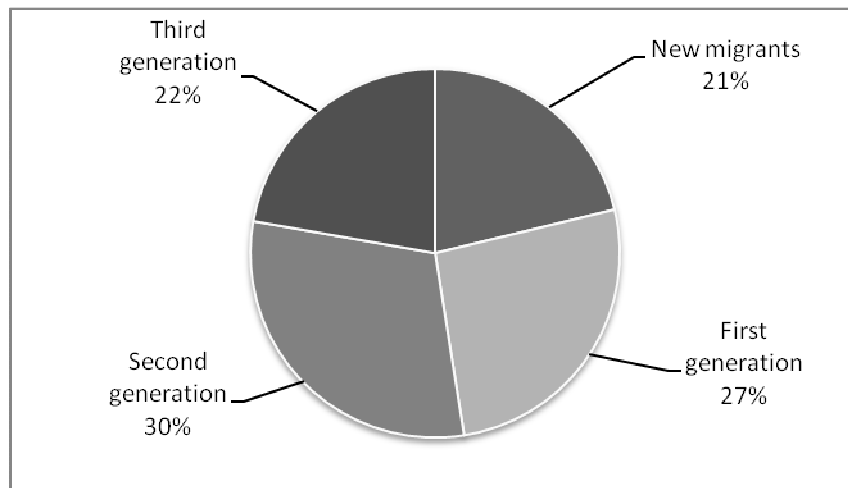


Figure 4.1 Questionnaire informants distribution by generation

⁹³ The signature (see Appendices) was relative to the protection of personal data.

⁹⁴ The composition of the informants who chose the English questionnaire was: two people of the first generation, 22 of the second, 18 of the third and three new migrants.

To a certain extent, this pattern was predictable, as it reflects the composition of the Venetian communities in Canada and the feasibility of contacting informants from each different group.

Each questionnaire was composed of 53 items (see Appendices), some comprising different questions. For the purposes of this study, only some of the questions were selected and treated statistically with SPSS. Their particular reference to the sphere of 'linguistic habits' and 'attitudinal factors' made them most relevant for this research.

In the following pages I present and discuss some data from the questionnaires. I will commence by giving a descriptive review of the most significant outcomes of the analysis with SPSS, to investigate if and how variables vary in the population investigated. Eventually some possible hypotheses will be formulated, hypotheses that will be tested in the following chapter against the results of the analysis of my informants' linguistic production.

Aspects pertaining to linguistic habits and attitude in the questionnaire are each seen in several items. Therefore, in order to reduce the number of variables to be analysed I decided to create two compound variables, one pertaining to linguistic habits and the other to attitudinal factors, following the work by Schmid and Dusseldorp (2010). Before combining them and creating one single compound variable, it is important to assess the reliability of this operation. I therefore performed a Chronbach's Alpha test, which enables us to obtain an index that measures the internal correlation and shows if we are able to join all the variables under one single compound variable. In fact, "if items within a scale are intended to measure aspects of the same construct, then they should all be fairly strongly correlated with each other" (Brace, Kemp & Snelgar 2009, p. 368).

One characteristic of Chronbach's Alpha test is that it is not robust with missing data. In my questionnaire I had in particular two questions, related to the language/s chosen by my respondents with given interlocutors (numbers 38/39), which were characterised by a high number of missing answers. This was due to the nature of the questions and the heterogeneity of the population investigated, which implied a substantial number of 'not applicable' answers. Notwithstanding the fact that they would have contributed to a depiction of a more comprehensive overview of the linguistic habits of my informants, I could not insert them into the analysis.

The attempt to create the two compound variables was however only partially successful. Questions pertaining to the Italian attitude turned out to have an outcome in the Chronbach's Alpha test well below the acceptable threshold. Having unsuccessfully

tried other suitable combinations, I decided to discuss all the questions pertaining to the concept of attitude towards Italy separately, to determine whether any particular trend would emerge.

Table 4.1 recaps the questions concerned (questions numbers are those of the English questionnaire in the appendices):

Table 4.1 Attitudinal factors

Attitudinal factors	<ul style="list-style-type: none"> • Do you feel Italian, Canadian or Italo-Canadian? (question 8) • Do you think you will go back/go to live in Italy? (question 9) • In Canada, has being Italian/with an Italian heritage put you in difficult situations? (question 13) • Do you still feel like an emigrant/son/grandson of emigrants? (question 14) • How much do you like Italian? (question 28)
---------------------	--

With regard to the second compound variable, the list of the combined questions consists of eight items, those with the highest Chronbach's Alpha test scores (α .885). These questions were all therefore placed under the label of 'linguistic habits'.

Table 4.2 Linguistic habits factors

Linguistic habits (8 questions)	<ul style="list-style-type: none"> • How frequently do you use Italian? (question 24) • How much do you understand Italian? (question 29) • How much do you speak Italian? (question 30) • How much do you read in Italian? (question 31) • How much do you write in Italian? (question 32) • How often and how are you in touch with people living in Italy? (question 33) • How often do you listen to/ watch Italian radio/ TV? (question 35) • How frequently do you read in Italian? (question 37)
------------------------------------	---

4.3.1 Attitudinal factors

As just seen, five questions pertaining to attitude were chosen from the questionnaire. They are discussed in the following pages, firstly by offering an overview of the outcomes, and then by assessing whether any of these features shows possible differences between the four groups which are statistically significant.

Affiliation

The first feature concerns the sense of affiliation among the four groups of informants, and in particular the comparison between the two groups of native speakers.

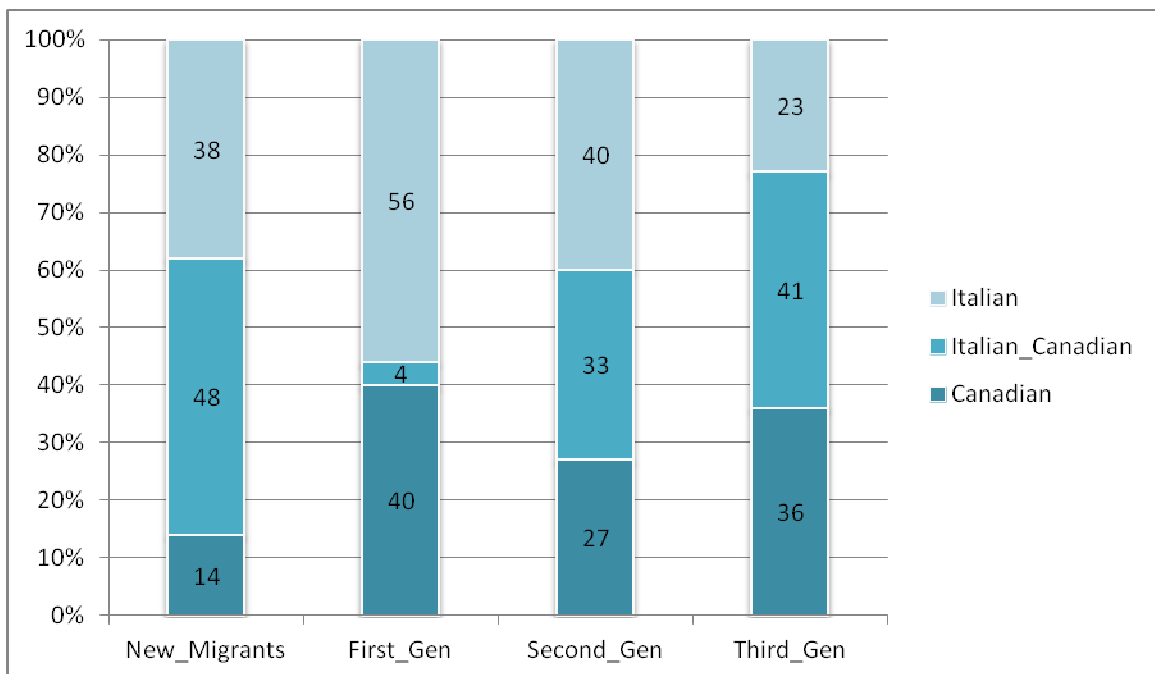


Figure 4.2 Affiliation by generation

In the cross-table displayed in figure 4.2, some interesting and rather unexpected results have emerged and they are worth highlighting here. Among the second and third generations, a sense of Italian affiliation seems not only to clearly exist, but even to be particularly widespread. If we look at the results for the (single) 'Italian' affiliation, these reach 40% and 23% for the second and the third generations respectively. On the same track are the results for Italo-Canadian affiliation, which reach 33% for the second generation and 41% for the third. Overall we may suggest that, among the second and third generations' informants in this study, the sense of Italianness and their affiliation towards their heritage roots are still remarkable and deep-seated.

The sense of affiliation among the two groups of native speakers seems to be particularly interesting, revealing a contrast between the two groups. Whereas new migrants display the highest result of people choosing an Italo-Canadian affiliation (48% of the respondents), this outcome is very slight (4%) among the first generation. Within this latter group, a feeling for an Italian identity is prevalent (56%), although followed by a relatively high occurrence of a Canadian identity (40%). The tendency to create a mixed identity, which, although not abandoning one's roots, embraces the culture of the new country, seems to be typical of new migrants. Among the first generation, on the other hand, the choice seems to have been more radical, avoiding blends in favour either of the maintenance of an Italian identity or the embracing of the Canadian one.

Going Back

The question regarding the possibility of a return/move to Italy shows a partial reverse trend compared to all the other features, pertaining to my informants' affiliation, that will be discussed in this section. Out of the 98 respondents, 92 (about 94%) answered negatively, that is that they do not have plans to move/go back to Italy, whereas only 6 (about 6%) answered positively.

The decision to relocate to another country is a quite radical choice that only those who are free from family bonds, or who have a particular family situation, can afford to make. My first generation informants have their close relatives in Canada, and therefore they usually do not have reasons to move.

As seen in chapter 2, for new migrants the choice to live in Canada has been a conscious decision, also with regard to the selection of this country. As one of my informants in Vancouver stated:

Anche perché c'abbiamo messo un sacco di sforzo e energia. E di tempo per ottenere questa permanent residence. Non ho assolutamente nessuna intenzione e voglia di tornare indietro.

[Also because it took us a lot of effort and energy. And of time to get the permanent residence card. I don't have the slightest intention and desire to go back.]

Siamo arrivati un po' per caso se vogliamo ma fino a un certo punto. [...] Allora escludiamo la zona est di Toronto e Montreal dove fa un freddo cane [...]. Non fa per noi. perché dobbiamo proprio fare i conti dell'oste. fa troppo freddo e troppo caldo d'estate quindi abbiamo escluso la zona est. perché non Victoria e Edmonton? allora Victoria è piccola. è molto più piccola di Vancouver. Ed è in un'isola.

E non ci piace l'idea di non essere liberi di muoverci. non che ci muoviamo questo granché ma ci piace l'idea della possibilità. però è più piccola e quindi è già molto più limitata. ed è molto turistica. quindi non è proprio l'obiettivo che noi cerchiamo. Edmonton? è fredda.

[We came here a bit by chance but only to a certain extent. [...] So we excluded the eastern part [of Canada] with Toronto and Montreal where it's freezing cold. [...] It's not for us. Because we wanted to be picky. It's too cold and too hot in the summer so we excluded the eastern part. Why not Victoria and Edmonton? Yeah Victoria is too small. It's much smaller than Vancouver. And it is situated on an island. And we don't like the idea that we are not free to move. It's not that we move that much but we like the idea of the possibility. But it is smaller so it is just more limiting. And it is very touristy. So it didn't match the target we had. Edmonton? It's cold.]

The two groups of heritage speakers may have developed a strong and positive affiliation towards Italy, but they are usually also bound to their native country and have family ties there. Among the six people that showed an interest in moving to Italy, five were third generation and one a new migrant. The third generation has a family situation that favours their relocation plans. The choice to move back to Italy appears to be closely linked to external conditions, in particular to family ties. Although affection towards the country is a necessary condition and a motivating factor favouring relocation, it is nevertheless not a sufficient factor on its own.

Difficulties

The question regarding possible difficulties encountered as being part of an ethnolinguistic minority community - which may also suggest the level of integration of my informants - shows quite predictable results, as seen in figure 4.3.

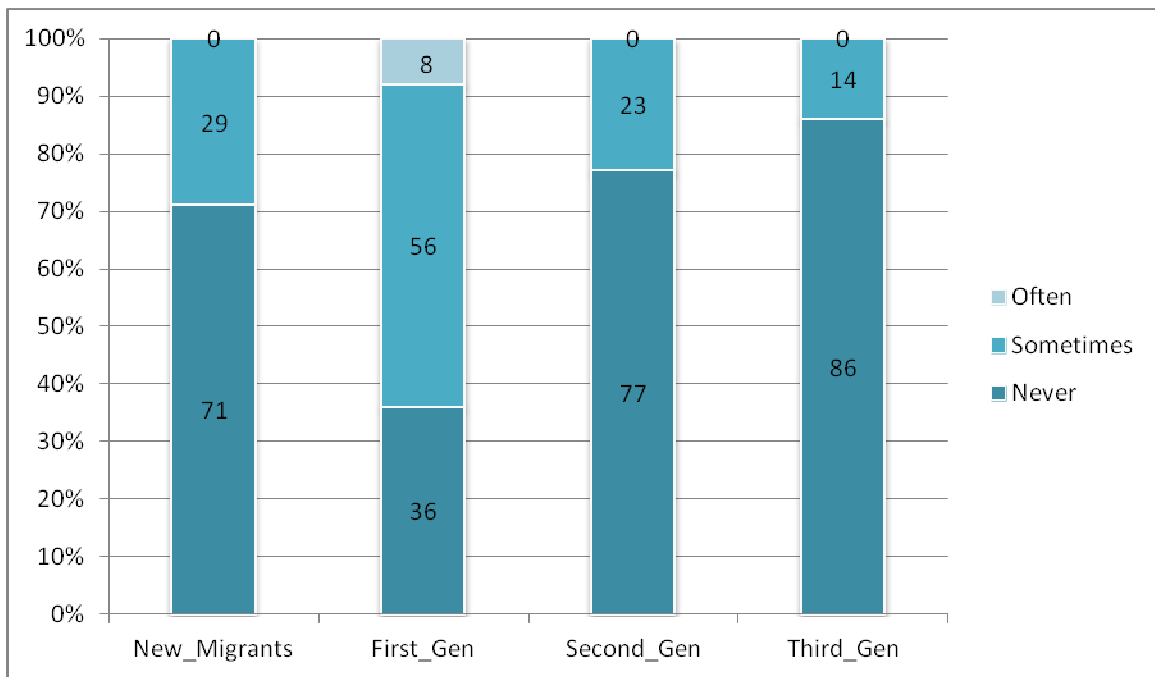


Figure 4.3 Difficulties by generation

In particular, a comparison between new migrants and the first generation reveals that the latter have run into more difficulties. There are even 8% of first-generation respondents who ticked ‘often’, although they were all referring to linguistic problems.

We can suggest two possible explanations for these results. Firstly, there is the different social scenario that the two groups have encountered in Canada, with the latest cohorts having benefitted from a more friendly, multicultural environment. Secondly, as seen in chapter 2, even the approach to the experience of migration differs for the two groups: whereas the first generation mostly moved to live in Italian neighbourhoods (Little Italies), new migrants were mostly seeking – or at least not rejecting the possibility of integration - with Canadian society or with other ethnic groups.

The increasing percentage of respondents along the generational scale that have not encountered any problem has to be related both to the characteristics of the new generations, now perfectly integrated into Canadian society, and again to the progressive acceptance of the Italo-Canadian community and its heritage by the mainstream community.

'Feeling Migrant'

The question pertaining to how much the respondents still feel themselves to be migrants, or descendants of migrants, shows, at first glance, puzzling results, particularly with regard to the second and third generation (figure 4.4).

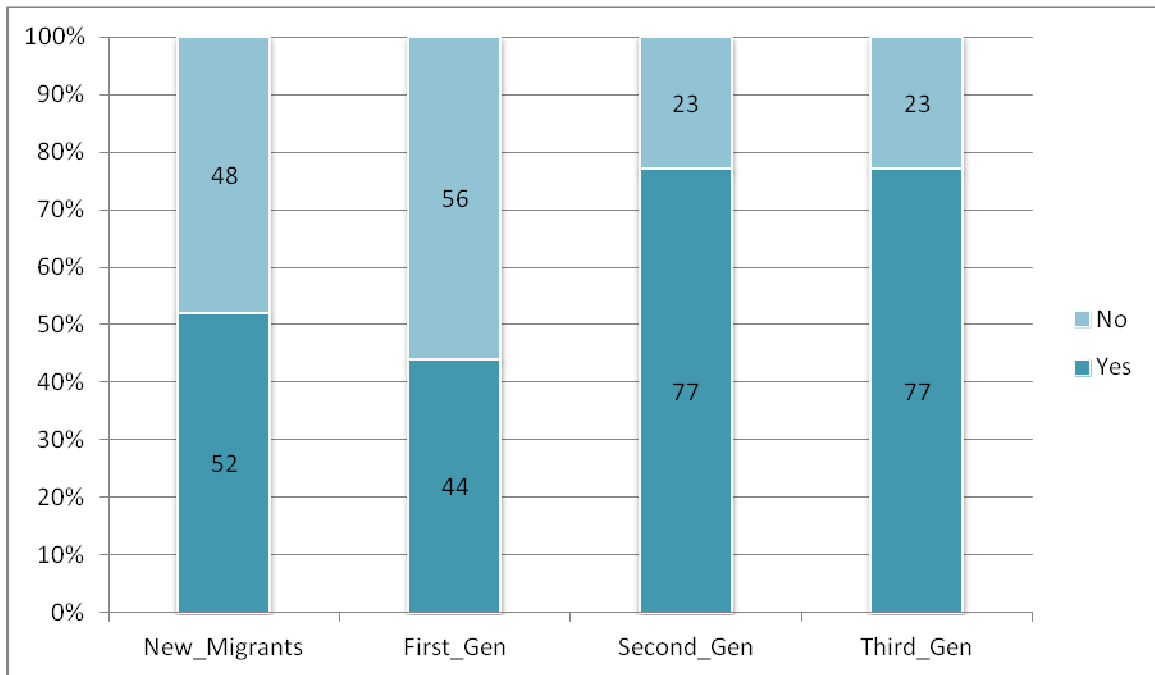


Figure 4.4 'Feeling migrant' by generation

The results for the two groups of migrants are in line with what I expected, with a percentage of respondents who still somehow feel migrants that touches 52% for new migrants and 44% for the first generation. Given the difference in the time elapsed since their moving to Canada, these results seem to suggest a correlation between years spent in this country and their self-description as migrant. In particular, a shorter stay in Canada is related to a higher rate of self-definition as 'migrants'.

The results for the second and third generations are difficult to explain. In both cases 77% of the respondents declared themselves as still feeling descendants of migrants, a result well above the corresponding one among the two groups of actual migrants. An explanation for these results seems not to be obvious. However we can propose a hypothesis. Earlier in this chapter we saw how people can interpret questions in different ways. Here we may suggest that different generations construct their own meaning of 'migrant'. Whereas the two groups of actual migrants may have read the question in terms of feeling not completely like the local Canadian-born population, the

second and third generation may have interpreted it not in this sense, but rather as being the bearers of a minority heritage that can even be seen as an added value to their identity.

Liking Italian

This last result indicates the appreciation that each generation displays for the Italian language. The overall outcome shows a great affection for the Italian language, with about 70% of the informants showing ‘very much’ appreciation, followed by 22% with ‘quite a bit’, and lastly ‘a bit’ and ‘little’ at about 4% each.

The results for the four groups separately are shown in figure 4.5, where other interesting patterns emerge.

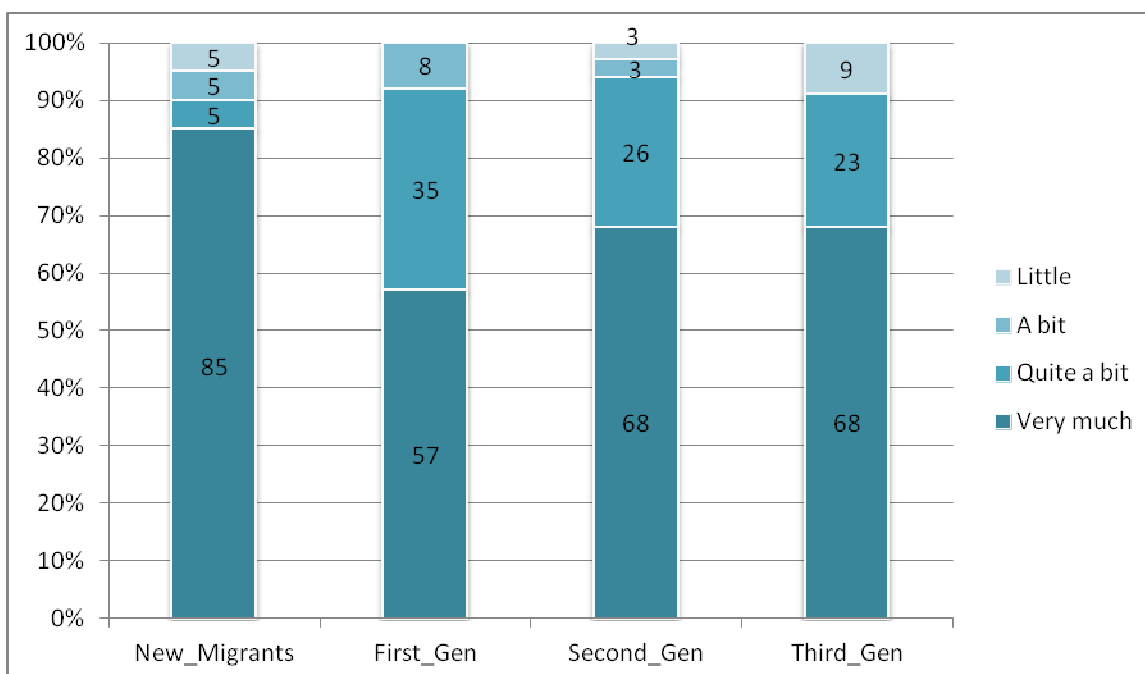


Figure 4.5 Liking Italian by generation

Whereas the great majority of new migrants (85%) highly value the Italian language, only a little more than half (57%) showed the same level of affection among the first generation. This latter result is also the lowest among the four groups, as the second and third generations both reach 68%.

A possible explanation can be found in the specific diglossic/dilalia situation in Italy (see chapter 2), which, mainly among the first cohorts of migrants, is reflected in a juxtaposition of Italian and dialect. Those speakers who were more dialect-oriented users (namely the first generation), although recognising the higher prestige of the national language, may have developed less attachment to the Italian language.

Along the generational scale, the most noteworthy result seems to be the growth of the category ‘little’. This trend seems to outline a decreasing attachment to the Italian language among the new generations and therefore less interest in the study of this language. However, given the low number of informants and the slight differences, we need to interpret these results cautiously. In view of this, on these five questions I ran a Pearson’s chi-square test, which is “used to test whether an observed pattern of events differs significantly from what would be expected by chance alone” (Brace, Kemp & Snelgar 2009, p.175). With regard to this analysis, the test has been used to assess whether the different results of these four generations discussed above are statistically significant.

The results of the test indicate that three out of the five results show statistically significant differences: the outcomes regarding a possible return to Italy (p-value .002), difficulties encountered in Canada (p-value .000) and my informants’ self-perception of still feeling migrant/descendant of migrants (p-value .025) all scored below the threshold of .05. On the other hand, the differences among these four groups turned out not to be statistically significant for the two other questions, regarding their identity (p-value .474) and love of the Italian language (p-value .081).

4.3.2 Linguistic habits

I now pass on to discussing the outcomes of the ‘linguistic habits’ index, which assesses my respondents’ use of the Italian language (the questions used for this assessment are in table 4.2). The results are displayed in table 4.3.

Table 4.3 Linguistic habits outcomes by generation

Group	Mean ⁹⁵	N ⁹⁶	Std. deviation
NewMig	.739	21	.157
FirstGen	.710	26	.125
SecGen	.482	30	.167
ThirdGen	.444	22	.191
Total	.588	99	.205

⁹⁵ Means are between 0 and 1.

⁹⁶ N stands for the number of speakers in each generation.

The first outcome that is worth discussing is the mean. The highest value is recorded for the new migrants (mean = .739), followed by the first generation (mean = .710), and then coming down to the second (mean = .482) and the third generation (mean = .444). These results are in line with what was expected. However the decreasing trend is not linear. In particular we can observe a major gap between the groups of native speakers on one hand, and the groups of heritage speakers on the other.

The new migrants and the first generation, those who are native speakers of Italian and usually still have close relatives back in Italy, are the ones that are quantitatively more in contact with the Italian language. The fact that new migrants scored the highest result may be due to the fact that they left Italy more recently, and thus may still have strong ties to the country, and to the fact that they usually have more close relatives back in Italy. This would lead them to travel more frequently to Italy, as clearly emerged both in their questionnaires and interviews. Moreover, while in Canada, although they have less preference for the more traditional mass media, such as radio and TV, they can usually enjoy a wider range of inputs in Italian, particularly from the web. With regard to the two groups of heritage speakers, both busy in their everyday life, in work activities or at school, we can assume a naturally lesser interest among the younger generation in spending time with other Italian-speaking people, who in Canada are usually the older generations, or in just being engaged in activities requiring the use of Italian.

With regard to the standard deviation, the results are relatively similar across the four groups. A lower score is however recorded for the first generation (std.deviation = .125), followed by new migrants (std.deviation = .157), and then by the second (std.deviation = .167) and third generation (std.deviation= .191). With regard to this index, the two groups of native speakers seem to be the more internally homogeneous in terms of their linguistic habits in Italian, a trait that appears to fade along the generational scale.

Lastly, I also ran the Pearson's chi-square test on these results. The outcome (p-value .000) is well below the threshold of .05, implying that the differences discussed among the four groups are statistically significant.

4.4 Results interpretation and analysis predictions

The results discussed in this chapter have firstly provided an overview of aspects of linguistic habits and attachment to Italy/Italian among different generations of Italo-Canadians. However, they can also serve as a base to support some predictions, which will be tested in the next chapter against the results of the linguistic skills in Italian of a group of interviewees, taken from those who filled in the questionnaire.

Before discussing working hypotheses, some observations on the attitudes of my informants are necessary. The non-homogeneous pattern of their answers in the above discussion, did not allow the creation of a compound variable for use in the analysis. Nevertheless, the results can still be used to support the formulation of the working hypotheses, as some general trends do seem to have emerged. Secondly, and most importantly, these results reveal the particularity of the Italian scenario and also validate the theoretical premises discussed in the second chapter as they mirror the very same multifaceted concept of Italian identity. The national language and feelings of nationality are co-present with different degrees of regional and local affiliation, a legacy of the fragmented history of Italy in terms of language and culture. A person can, therefore, experience different degrees of Italian affiliation, in diverse aspects of her/his identity, leading to the creation of individual identity sets, which may be compound and not internally homogeneous.

As was predictable, this chapter has shown differences between the two groups of native speakers on one hand, and the two groups of heritage speakers on the other. However this divergence mostly pertains to their linguistic habits. In contrast, the patterns regarding their attachment to Italy seem not to show clear trends. It also seems safe to say that their feeling towards their heritage country is quite dependent on external variables. As we have seen, family ties or the change in attitude of the mainstream society have an impact on how these communities relate to their heritage country and consequently to its language.

The results in this chapter seem to allow us to put forward a couple of interpretations. Among the two groups of native speakers, results that point to a higher proficiency level are expected for the new migrants. As discussed in chapter 3, new migrants and the first generation differ significantly in terms of their level of education, which, as we have seen in chapter 1, may have an impact on their language skills in terms of maintenance/attrition. New migrants have also more recently arrived in Canada, and

thus their Italian proficiency may have been preserved more. The results in this chapter both regarding contact with their heritage language and their level of affiliation towards Italy, show that new migrants had the highest score in both cases: they have on average the highest number of contacts with Italian, but they also revealed a stronger attachment to their Italian identity, for example by declaring a relatively higher degree of affection for Italian and also of affiliation towards an Italian identity. They are in fact the group which had the lowest percentage of people claiming a solely Canadian identity. All these observations point to a clear separation of these two groups. However, these differences may be partly levelled out and not as marked as one may expect, particularly when performing a spontaneous task, such as being interviewed about their life. Moreover, a migrant's lower level of education does not mean that they cannot keep on improving their level of language by reading, watching TV or other similar activities.

Among the two groups of heritage speakers, higher proficiency level results are expected for the second generation. As we have seen in chapter 1, people belonging to this generation may have even been raised as monolinguals in their heritage language up to entering school, in contrast with the third generation who are usually bilingual from birth. Notwithstanding these factors, the responses to the questionnaires show not very dissimilar results, both in terms of contacts with Italian and attitudes towards Italy. The question pertaining to love of Italian, or the percentage of those who feel Italian or Italo-Canadians, as opposed to solely Canadian, presented not dissimilar results for these two groups. On these grounds we can suggest that, although the language skills in Italian of the second generation may be higher than those of the third, given their usually greater exposure to this heritage language in their earlier years, these outcomes may not be as sizeable as one might have anticipated.

These results allow us to formulate predictions for the linguistic variables, which will be discussed in the next chapter. These predictions in reality simply follow and validate what has been discussed in the first part of this thesis, regarding in particular my informants' belonging to a migrant cohort and generation: linguistic abilities are expected to be higher for new migrants (native speakers, with a higher level of education) and then to follow a decreasing pattern from the first to the second and third generation. The results in this chapter thus allows us to hypothesise a reinforcing role for the variable of 'linguistic habits', the only one we were able to create a compound

variable from, in the linguistic abilities of my informants and in the differences between the four groups. In particular, the two expected patterns are outlined below:

1. In the two groups of native speakers, new migrants are expected to outperform the first generation. We have seen how they benefit from circumstances usually favouring language maintenance, such as a higher level of education and, as noted in this chapter, they also take advantage of a (relatively slightly) higher level of contacts with the heritage language, which can further emphasize the differences in terms of linguistic skills between these two groups.
2. The decreasing intergenerational patterns which are the common outcome among Italian migrant communities abroad are expected to be further reinforced by the level of contacts with the heritage language among the three groups. We would expect the first generation to stand out clearly from the other two, with a slightly decreasing trend in the passage from the second to the third generation.

Chapter 5: Language skills among speakers of a heritage language and the role of linguistic habits

5.1 Introduction

The core aim of this last chapter is to discuss language maintenance/attrition among Venetian-Italian communities in Toronto and Vancouver, taking a mostly quantitative and statistical approach to the phenomenon. As mentioned in the introduction to this thesis, language skills among Italian communities abroad is a well-researched field, but not much attention has been given to quantifying language skills trends across generations. This chapter thus attempts to address the topic from this less-studied perspective.

In order to enable a quantitative analysis, a considerable amount of data has been gathered, through the recording and transcription of 56 interviews, equally divided between four generational groups (new migrants, the first, second and third generations) with fourteen informants in each subgroup⁹⁷.

The number of interviews conducted is a significant outcome of the project *per se*. Although they do not represent the actual skills in Italian of the entire Venetian-Italian community in Anglophone Canada, since people who considered themselves not competent enough in Italian did not participate,⁹⁸ this number remains significant, particularly in relation to the third generation. The existence of a third generation still speaking the language to some extent and still able to carry on a conversation in Italian is indicative of the capacity of the community to support and value the use of the heritage languages, even among later generations.

This chapter is structured as follows: the first pages offer a descriptive analysis of the linguistic outcomes of my interviews. As introduced at the end of the first part, there are two perspectives of research taken up in this study; one regarding a comparison between the two groups of native speakers, namely new migrants and first generation, and the

⁹⁷ As discussed in chapters 1 & 3, the variable of gender is somehow relevant in this field, although its impact is not always straightforward. In order to try to create a sample as homogeneous as possible, a balance also with regard to gender was sought. Hence, each group was composed of an equal number of female and male participants.

⁹⁸ Although some first-generation and new migrants did not agree to participate, mostly suggesting that they did not feel at ease in sharing their experience as migrants and their life in general, the great majority of those who declined were from the second and (mostly) third generations.

other following an intergenerational perspective, from the first to the second and third generations. The discussion in the next pages will focus on these two perspectives.

In the second part of this chapter, while maintaining the same research perspectives, I will develop the analysis by assessing whether the variations observed among groups are statistically different, and whenever this condition is verified, trying to identify possible explanations. In particular I will try to assess whether the patterns relating to the linguistic habits across generations, as discussed in the previous chapter, are in some way linked to my informants' language skills and would (partially) account for their level of proficiency.

The data explored and discussed in this chapter have been analysed by means of two softwares: SPSS and secondly an application for the analysis of texts (which will be described in the next section). In particular, the outcomes generated with the text analysis software, which I then checked manually (see chapter 3), have been entered in SPSS and statistically analysed. Although statistical software is a powerful tool to analyse linguistic data, helping the researcher to obtain a more precise and objective picture of the phenomena investigated, it is important to remember its mere instrumental role: it can process a substantial amount of data and return them in an analysable format, but it does not offer explanations or interpretations, which remain the researcher's task (Carlioni, 2005).

5.2 Analysis tools

The two tools used in this analysis are SPSS and Taltac, a software for the analysis of texts. TaLTaC is an acronym for *Trattamento Automatico Lessicale e Testuale per l'Analisi del Contenuto* (Automatic Lexical and Textual Processing for the Analysis of Content). This software application, created in 1999, comprises a set of statistical and linguistic resources "highly integrated with each other and [which] can be customised by the user" (www.taltac.it). These tools allow one to extract relevant information (by text mining) from any type of linguistic data and to carry out textual analysis of both its language and its content (e.g. by selecting the relevant or specific language from the corpus under investigation). The main purpose of this software, while returning the text corpus in a different format, is to offer a set of tools to allow content and linguistic analysis of a large quantity of data (Bolasco, 2010).

Research in language maintenance/attrition has used different methods of data collection as well as different procedures and software to analyse these data. By choosing a particular piece of software we have to be aware of the resulting limitations. This choice, in fact, does not only have effects on the modalities of the corpus transcription, as these must be appropriate to the properties of the software, but also in terms of the decisions about which features we can analyse. Taltac does not offer all the instruments that might be relevant to the lexical analysis of my corpus (see Schmid & Jarvis, forthcoming). It has, however, the great advantage of having built-in tools, devised specifically for Italian, which permits more in-depth analysis of the lexicon in this language.

The use of Taltac, a well-established instrument among Italian language scholars in this field, helps in looking at the language abilities of speakers from a specific perspective, mostly focused on their lexical skills.

5.3 Textual statistical analysis

I now pass on to explore aspects of the language used by my informants in the interviews. In particular, I present and discuss the results of 9 lexical features, namely: 'Number of word tokens', 'Guiraud Index', 'Word length', 'High/Medium/Low-frequency words', 'Vocabolario di Base (VdB)', 'Grammatical categories', 'Code-switching', 'Content versus Function words' and 'Verb tenses'. Some of these features have already been introduced with regard to the Italian language (VdB and Verb tenses) in chapter 2, or with regard to the methodology in language attrition/maintenance studies (Guiraud Index, High/Medium/Low frequency-words) in chapter 3. The quantitative data which resulted from processing the interviews with Taltac, regarding the 9 features listed above, have been subsequently analysed with SPSS and will be discussed in the second part of this chapter.

Before beginning the discussion, it is important to make a preliminary point. Although the corpus investigated with Taltac is sufficiently sizeable to be treated statistically (see chapter 3), the value of a statistical analysis with 14 people in each generation may be debated⁹⁹. This is particularly the case with regard to the study of standard deviation: the low number of informants per cell, in fact, does not allow us to assess whether the

⁹⁹ However, many variationist studies have been based on similar numbers, or fewer (cf. Milroy & Gordon, 2003).

deviation is due to the informants being well scattered along the continuum or whether it mostly depends on one or a few informants whose scores differ significantly from the mean. Although this critical aspect (standard deviation analysis with a low number of informants) can not be encompassed by any statistical software, a compromise was found in calculating the standard normal distribution (or z score), and in particular by visually presenting the results, signalling in particular significantly different results (outside the non-significant range). The range of significance is visually marked in every figure by the four dashed lines. Statistically, the probability of finding any occurrence in the range -2 to 2 is 95% and in the range -3 to 3 it is 99%. This means that if I find occurrences that are outside these ranges it is because they are significantly distant from the average, particularly those outside the range -3 to 3. By doing this we can determine whether the overall outcome is the result of one or a few informants who scored very differently, or whether this is a less marked but more widespread trend, applying to a larger number of informants.

With regard to the following analysis, I will mainly report and discuss the outcomes of SPSS (the minimum and maximum score, mean and standard deviation). Whenever the outcomes turn out to need further investigation, I will also run a z-score analysis.

5.3.1 Tokens

The first index I am going to explore is the length of the interviews, analysed in terms of 'number of word tokens'. This has to be taken, however, as a quite crude indicator of language proficiency; the length of interviews in itself cannot be considered a completely reliable gauge of language proficiency, not even among native speakers. It must in fact be recognised that several other factors impact more strongly, the first among which is the inclination of each person to speak. Therefore, even in a wholly monolingual sample, there are in fact great differences in interview length. Despite this, it is reasonable to suppose that the number of words used in the sub-corpora, in particular across different generations, is in some way related to this group's level of proficiency in Italian.

Table 5.1 Tokens by generation

Group	Min.	Max	Mean	Std. De.
NewMig	4663	15324	8213	3326
FirstGen	3880	11291	7416	2016
SecGen	3810	9557	5788	1664
ThirdGen	1140	8228	4078	2023

Looking at the means reported in table 5.1, results replicated also for the minimum and maximum scores, a clear pattern emerges: as largely expected, a decreasing line is seen along a generational scale, starting with the new migrants. Notwithstanding the linearity of this trend, a closer look reveals an interesting feature: the gap between the means of the four groups increases, passing from about 800 units between new migrants and the first generation, to double this result between the first generation and the second generation (1628), and between this latter and the third generation (1710).

I now proceed to a further investigation of my data. I have in particular done a one-way Anova in order to assess whether the groups' means, which we have just considered in general terms, are statistically different. When the Sig value of an Anova test is lower than .05, we can conclude that there is at least one group in the sample that statistically differs from at least one other. However, this outcome is somewhat limited in its explanatory power because it does not point to which these groups are. In other words, it does not tell where the significant differences lay. If we want to assess this aspect we need to do a post-hoc test. In this analysis I have used the Tukey HSD. This is a multiple pairwise comparison test, which can signal which groups in our sample statistically differ from each other.

I therefore started by performing an Anova, which shows that there is a significant difference between the four groups ($F(3,52) = 8.55, p = .000$). Because of this result, I ran the Tukey test, which indicated that there are three pairs that statistically differ from each other: the new migrants and the second generation (p-value .041), the new migrants and the third generation (p-value .000), and lastly the first generation and the third (p-value .002). All the other pairwise comparisons were not statistically significant (at p-value $< .05$).

Given the combination of a low number of informants and the high number of occurrences for this feature, the standard deviation presents very high outcomes (as seen in table 5.1), also quite diversified across the generations. In order to have a clearer picture I ran a z-score test, the outcomes of which are visually reported in figure 5.1:

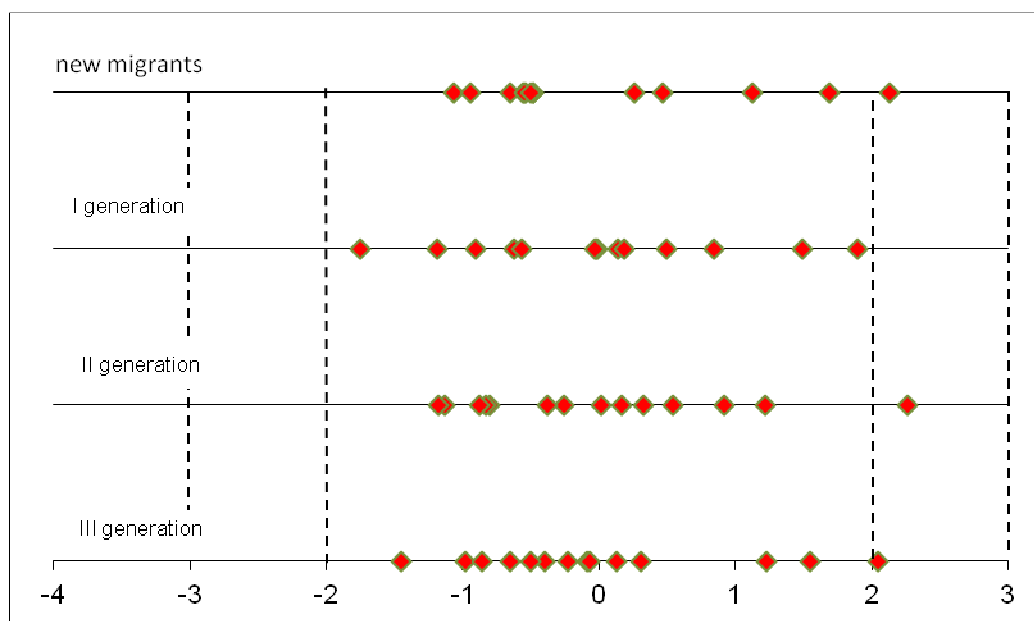


Figure 5.1 Tokens Z score by generation

In partial contrast with the results of the standard deviation we observe that three groups (new migrants, second generation and third generation) each have one informant that scores noticeably differently. However, if we do not take into account these three informants, and consider only the general pattern, we see that it is just the first generation which shows up as the most widely spread group. Moreover, there seems to be a slightly higher level of consistency among the second generation and also among the third, a tendency to cohere among the heritage speakers. Conversely, the two groups of native speakers seem to be the more variegated.

5.3.2 The Guiraud Index

The second feature that I am going to explore is the Guiraud Index. This is one of the indices employed to assess lexical diversity (as discussed in chapter 3), a measure “used by researchers in many fields as it has been found to be indicative of a wide variety of variables, such as writing quality, vocabulary knowledge, general characteristics of speaker competence, and even a speaker’s socioeconomic status” (McCarthy & Jarvis 2007, p.459). In this study, my aim is to apply the Guiraud Index. in order to examine the lexical diversity of my informants’ language performance. As discussed in chapter 1, in fact, a

reduction in the range of the vocabulary available for use is to be expected among native speakers, as a result of first language attrition being in progress (e.g. Andersen, 1982; Schmid, 2011a). From an intergenerational perspective, a general overview of the vocabulary used by informants and the wideness of its range may be taken as an indicator of their level of language skills, particularly because there is a large volume of data to consider, but also because these aspects of speakers' production are less closely controlled. Following the approach of Yilmaz and Schmid (forthcoming) I decided to calculate lexical diversity taking into account only content words. This choice will then be maintained throughout the major part of this analysis.

Before passing on to discussing the results, I need to establish certain points. First of all, we must be aware that, although the index of lexical diversity seems to be indicative of the speakers' lexical skills and can be taken as a valid measure when working with large corpora, a criticism should be considered in this regard. The richness of the vocabulary is a mere gauge of the use of the lexicon available to a speaker and therefore cannot be taken as proof of the actual extent of her/his own vocabulary. Indeed, it is impossible to specify even one's own vocabulary, as it is an abstract entity belonging to one's cognitive capacities. What can be assessed is simply one's use of it, which implies that this index cannot be directly indicative of the speaker's real capacity and may be biased by her/his lexical choices. One person may have a relatively small vocabulary but exploit it to its maximum capacity, allowing her/him to attain a higher vocabulary index than another speaker who, for whatever reason, does not make complete use of her/his larger vocabulary (Cortelazzo & Tuzzi, 2008). Therefore, if not completely indicative of this linguistic phenomenon, the vocabulary used can still be considered descriptive of the linguistic performance of my informants and to some extent also related to their language skills.

Secondly, a completely reliable instrument to assess lexical diversity has yet to be found. Several indices have been created for this use, with a progressive implementation of their reliability. However, as claimed by McCarthy and Jarvis (2007), all seem to present some flaws, although 'some are more obviously flawed than others.' (p.480). The presence of a large corpus (an average of 6500 word tokens for each of my 56 interviews), as well as the unbalanced size of the sub-corpora, would make the reliability of the lexical richness analysis complex and not completely satisfactory in any way. I eventually decided to use Guiraud Index, being aware that more accurate tools may be found. However, taking it as a mere indicator of a language pattern which

is explored also from other perspectives, I decided on its suitability for the scope of this study.

Thirdly, the presence among my informants of two groups of native speakers of Italian, new migrants and first-generation, offers the opportunity to consider their output in terms of first language attrition. Given the number of (interrelated) variables at play, including age and level of education, it would not seem justified to ascribe the outcomes directly to the effects of language attrition. However, if not completely indicative of this linguistic phenomenon, they can still be considered descriptive of the linguistic performance of my informants and to some extent also related to their language skills.

In table 5.2 the results of the Guiraud Index are displayed:

Table 5.2 Guiraud Index by generation

Group	Min.	Max	Mean	Std. De.
NewMig	12.79	19.76	16.60	2.02
FirstGen	12.40	17.51	14.76	1.38
SecGen	8.81	17.46	12.09	2.32
ThirdGen	8.35	13.36	11.05	1.54

The higher values are recorded for new migrants, both as the minimum and maximum scores (with 12.79 and 19.76 respectively). The lower values are registered for the third generation (8.35 and 13.36). The decreasing trend along a generational scale is reflected also with regard to the mean. The values in fact touch their peak with the new migrants, at 16.60, and then gradually decrease to 14.76 for the first generation and 12.09 for the second generation, and reaching the lowest point with the third generation at 11.05. Along this overall decreasing pattern it seems possible to pinpoint a major gap, that is between the two groups of native speakers on the one hand and the two groups of heritage speakers on the other. The difference in fact reaches 2.67, whereas between the two groups of native speakers it is at 1.84 and with the two heritage speaker groups at 1.04. The standard deviation is between 1.38 (for the first generation) and 2.32 (for the second generation), showing a relatively higher level of consistency among the first generation, but not significantly different from the other groups.

These results are in line with what would have been expected. However some aspects are worth briefly discussing here, first from a purely descriptive perspective, and then in relation to the results of the Anova. It seems that while there is a reduction in terms of lexical diversity from one generation to another, the decrease from the second to the third generation is not as great as one would have expected. It seems plausible to think that those of the third generation who still speak Italian, although considerably fewer than those of the second generation who do so, may be more highly motivated, and therefore may be trying to keep up with the second generation.

A noteworthy gap, both in term of its consistency and for the fact that it sharply divides the two groups of native speakers, is seen between the new migrants and the first generation. There may be more than one cause: we can suggest both the influence of a higher level of education among new migrants but also a possible decrease of language skills (language attrition) among the first generation.

It is lastly worth noting that the gap between the first two groups and the last two is greater, an outcome that once again separates native speakers and heritage speakers.

Despite this, the Anova showed a Sig value greater than .05 ($F(3,52) = 1.34, p = .271$), meaning that there is no statistically significant difference between any of the four groups considered in this sample.

5.3.3 Average length of content words

In Italian, as in many other Indo-European languages, there is a trend that relates the length of a word, calculated in the number of letters, and its frequency of use. In particular, the shorter a word, the more frequently it is used. Therefore, in a relatively sizeable corpus, the higher the average word length is, the higher is the percentage of use of relatively less frequent words, which is in some ways a sign of a more sophisticated level of language.

The average length of all the word types of a corpus in Italian is inevitably low due to the incidence of function words, which are short and very frequent in every type of text. Therefore, even with regard to this index, I ran the analysis only with content words. Table 5.3 shows the results for this.

Table 5.3 Average length of content words by generation

Group	Min.	Max	Mean	Std. De.
NewMig	5.53	6.29	5.87	.21
FirstGen	5.08	5.79	5.36	.23
SecGen	4.69	5.53	5.22	.20
ThirdGen	3.99	5.48	4.95	.40

These results are in line with what was expected: there is in fact a decreasing trend along a generational scale, starting with the new migrants, and this trend is replicated both with regard to the pattern of the minimum and maximum scores recorded in each group, as well as that of the means.

The standard deviation shows a similar outcome for the first three groups, while there is a higher result, therefore a lower level of internal homogeneity, for the third generation. The Anova performed showed a significant difference between the four groups ($F(3,52) = 27.17$, $p = .000$). The Tukey post-hoc comparison test was then carried out, indicating in particular that there are four groups that statistically differ from one another. The group of new migrants, above all, is highlighted: they differ significantly from all the other three groups; with the first generation (p -value .000), with the second (p -value .000) and with the third (p -value .000). The last pair that differs significantly is the first and the third generation (p -value .002).

I eventually ran a z-score test, the results of which show that only one informant out of the 56 (belonging to the second generation group) scored significantly differently. Hence the pattern was corroborated with the second test.

5.3.4 High, medium and low-frequency content words

In chapter 3 I introduced the subdivision by Taltac into High, Medium and Low frequency words. This classification is created automatically by Taltac and has a semantic value, helping us to isolate the key words (and therefore the key topics) of the corpus examined. However, when applied to sub-corpora that deal with the same topic(s), it seems appropriate to explore this classification also from a purely linguistic

perspective. This division into High, Medium and Low frequencies will thus be explored with the aim of determining if their percentage of use by different groups can be revealing of my informants' language skills. In particular High frequency lexis is associated with a more 'everyday' level of language, whereas the Medium and Low frequencies indicate a richer and more varied type of language.

In table 5.4 the results show the percentages of vocabulary in each of the three categories, High (HF), Medium (MF) and Low frequencies (LF), for the four groups of informants.

Table 5.4 High, medium and low frequency content words by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropHFContWords	19.74	31.49	24.63	3.63
	PropMFContWords	53.94	64.10	58.96	3.17
	PropLFContWords	13.84	22.42	16.40	2.26
FirstGen	PropHFContWords	16.30	27.12	20.79	2.86
	PropMFContWords	44.50	64.05	60.17	4.79
	PropLFContWords	14.16	38.04	19.03	5.64
SecGen	PropHFContWords	20.35	39.30	24.93	4.99
	PropMFContWords	53.83	68.55	62.52	3.68
	PropLFContWords	6.87	19.00	12.53	2.94
ThirdGen	PropHFContWords	18.01	48.54	31.48	9.43
	PropMFContWords	41.13	73.14	58.19	8.67
	PropLFContWords	5.86	15.02	10.31	3.13

In order to give a visual representation of the pattern of these results, the means are shown in figure 5.2.

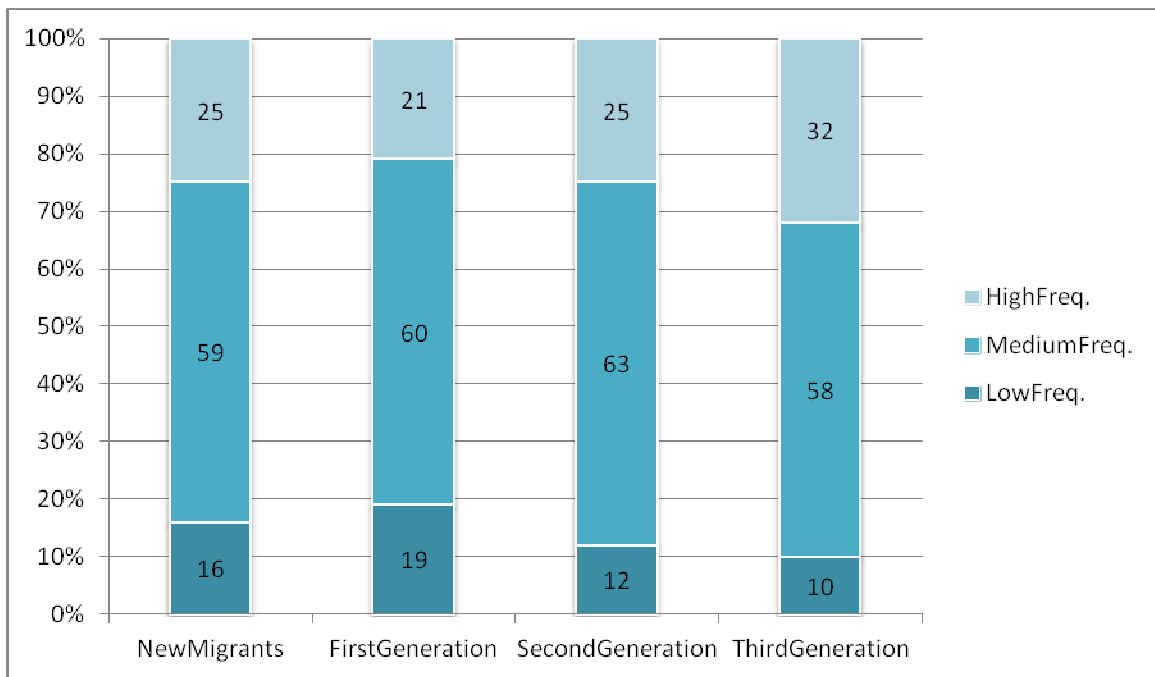


Figure 5.2 High/Medium/low frequency words by generation (%)

These findings partly confirm the pattern recorded with the preceding features. Along a generational scale an increased percentage of the category of High frequency words, and a corresponding decrease of Low frequencies, is observed. The percentage of Medium frequency, the most consistent one for all groups, varies between 58% for the third generation and 63% for the second.

Along the generational scale, this figure may reflect differences in language skills. In particular, the category of less frequent words is most used by the first generation, while the high frequency words are more used by the third generation.

The pattern for the new migrants places them between the first generation and the second generation, showing apparently lesser language skills compared to the other group of native speakers. This result stands in contrast with the preceding ones, which all showed the highest level of language abilities recorded for the new migrants.

The Anovas showed significant p-values, with the exception of the medium frequencies ($F(3,52) = 1.64$, $p = .190$), for which there is no statistically significant difference between any of the four groups investigated.

With regard to the high frequencies, the Anova showed a significant difference between the four groups ($F(3,52) = 8.14$, $p = .000$). The Tukey test, then, indicated three pairs that statistically differ from each other. Interestingly these combinations all include the third generation, which stands out as differing significantly from the new migrant group

(p-value .015), the first generation (p-value .000) and the second generation (p-value .022).

With regard to the low frequencies, the Anova presented again a significant difference ($F(3,52) = 15.30, p = .000$). In the analysis by Tukey, four pairs emerged as statistically different from each other, namely the new migrants and the second generation (p-value .040), the new migrants and the third generation (p-value .000), the first and the second generation (p-value .000), and lastly the first and the third generation (p-value .000).

With regard to the pattern of internal dispersion of these groups, table 5.4 shows quite diversified outcomes. I ran a z-score test to visually present the outcomes. Having obtained the same pattern for all three ranges, I present the results only of the High frequency range, which can be seen in figure 5.3. However, the discussion can apply also to the Medium and Low frequency.

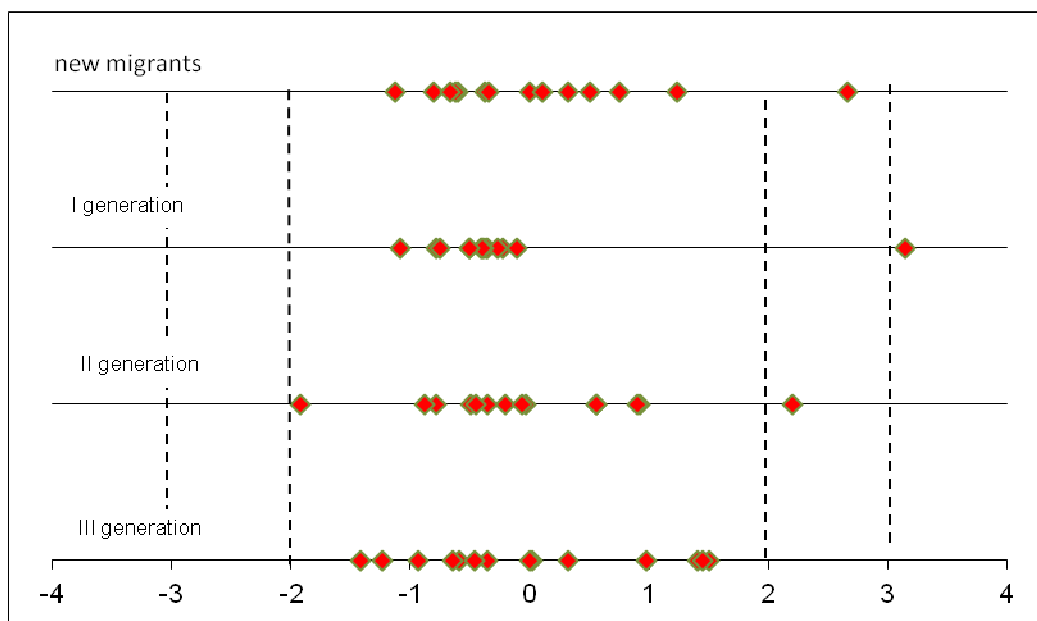


Figure 5.3 High frequency words Z score by generation

First of all, the figure highlights the occurrence in three groups, namely new migrants, first generation and second generation, of one informant in each group that scored significantly differently from the trend for the group. In particular, in the first generation, one informant goes even beyond the threshold of $|3|$. Beyond this result we can try to interpret these outcomes and delineate a general trend that seems to appear along a

generational scale, which suggests a tendency towards an increase of internal heterogeneity from the first to the third.

If we then try to interpret these outcomes and delineate a general trend, there seems to be an increase of internal heterogeneity along a generational scale, from the first generation to the third, and a major level of internal consistency among long-term migrants.

5.3.5 Vocabolario di Base

In chapter 2 we discussed the Vocabolario di Base (VdB) of the Italian language. This notion describes that section of the Italian lexicon that forms 98% of our everyday communication and which is known by almost the whole Italian population, regardless of their social class, level of education or regional origin. If analysed in terms of a comparison between different groups of migrants, and between different generations of Italo-Canadians, the percentage of VdB employed can be taken as an indicator of their linguistic skills¹⁰⁰.

A general decline of lexical abilities across generations is expected and well documented in literature on this topic. What is of interest here is to quantify this trend.

Table 5.5 Vocabolario di Base by generation (%)

Group	Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig VdB	71.67	78.90	74.88	2.11
FirstGen VdB	73.49	80.01	76.75	2.09
SecGen VdB	71.63	80.59	76.18	2.69
ThirdGen VdB	75.06	83.82	78.35	2.56

Table 5.5 shows the proportion of each groups' vocabulary that comes from VdB. Hence, a higher mean percentage seems to reflect a 'smaller' vocabulary. The outcomes

¹⁰⁰ However, it should be acknowledged that an interview of 1-hour represents a modest sample of an individual's speech.

present a clear but modest tendency emerging along a generational scale, once again commencing with the new migrants and going up through the three generations, although this is partly interrupted by a slightly lower result for the second generation compared to the first. Looking at the mean scores, in fact, we can see how the lowest score is obtained by the new migrants, followed by the second, first and lastly the third generation. This trend was more or less as expected and is in line with the majority of the previous findings discussed above: new migrants display the lowest percentage of use of the 'basic' language (74.88%), underlining their higher language skills compared to the other groups. Along a generational scale, my informants tended to rely more and more upon the basic vocabulary, with a percentage that eventually reached 78.35% with the third generation.

Moreover, differences between the groups' scores are worthy of mention as well. Two main gaps, in particular, are seen: one between the two groups of native speakers, and the other between the first/second and the third generations. With regard to the first gap we can again refer to a different level of education and also a longer period of time spent away from Italy. But with regard to the second gap, the trend is less easily explained. If, as it was predictable, the third generation has the highest score, with regard to the first and second generations the scores are the opposite of what was expected, with the second generation scoring a slightly lower percentage (but the difference is very slight).

In performing the Anova we found that there was a significant difference between the means' groups ($F(3,52) = 5.09, p = .004$). In the analysis by Tukey, however, just a single pair statistically differed: the new migrants and the third generation (p-value .002).

Lastly, with regard to the internal dispersion of these groups, we can consider the standard deviations, although with the same precaution mentioned earlier, that we are dealing with only 14 informants per cell. The two groups of native speakers are the most internally homogeneous, with very similar results. Between the two groups of heritage speakers, the third generation is the more homogeneous.

5.3.6 Grammatical categories

In this section I discuss the frequency of six grammatical categories: nouns, verbs, adjectives, adverbs, pronouns and prepositions. One last group will follow, i.e. code-switching (that is the use of words in a language other than Italian), which has been included as relevant for a study of bilingualism, although its composition is miscellaneous, in the sense that it includes words in English, dialect and Italese.¹⁰¹

As stated in chapter 3, I could not deal with articles¹⁰². The reason being that this is a problematic category to tag in Italian: the articles are in fact homonyms with pronouns and a manual disambiguation would have been necessary.

As far as I know, there is no previous research that deals with this topic among Italian communities abroad. Hence, without having any particular hypothesis or trend to validate or reject, I explore how these categories are distributed across my population, trying to see whether any trend emerge.

The results, as reported in tables 5.6 – 5.12, show patterns so diverse that it seems not possible to delineate an unequivocal trend.

Table 5.6 Nouns by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropNouns	16.28	21.85	19.04	1.97
FirstGen	PropNouns	15.63	19.79	17.56	1.23
SecGen	PropNouns	14.39	22.34	18.67	2.10
ThirdGen	PropNouns	16.12	37.11	22.75	5.56

¹⁰¹ At this stage we shall not discuss the structure of the code-switching category further.

¹⁰² In Italian there are many cases of ‘ambiguous’ words, that is homographic words in the written language. As discussed in chapter 3, a problematic category is that of articles, which cannot be categorised by Taltac because all its forms are ambiguous and shared with the category of pronouns (and therefore labelled as ‘J’). The reverse relationship does not exist: not all pronouns are ambiguous, because not all their forms are shared with the category of articles. Therefore, in this study, when dealing with pronouns, I am referring to those which are not ambiguous.

Table 5.7 Verbs by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropVerbs	22.11	28.95	25.43	1.76
FirstGen	PropVerbs	23.55	31.76	26.52	2.40
SecGen	PropVerbs	20.40	28.78	24.96	2.16
ThirdGen	PropVerbs	14.16	27.77	21.66	4.39

Table 5.8 Adjectives by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropAdj	5.92	9.75	7.71	1.13
FirstGen	PropAdj	5.47	7.52	6.31	.62
SecGen	PropAdj	5.62	8.77	6.82	.78
ThirdGen	PropAdj	4.74	7.98	6.50	1.06

Table 5.9 Adverbs by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropAdv	8.06	15.68	13.08	1.90
FirstGen	PropAdv	10.87	17.82	14.79	1.96
SecGen	PropAdv	9.26	15.87	12.52	1.85
ThirdGen	PropAdv	9.62	15.15	12.42	1.88

Table 5.10 Pronouns by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropPron	6.06	12.02	8.21	1.42
FirstGen	PropPron	7.97	11.77	9.19	1.02
SecGen	PropPron	4.30	10.32	7.03	1.53
ThirdGen	PropPron	2.22	8.77	6.18	2.16

Table 5.11 Prepositions by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropPrep	7.34	13.49	10.98	2.10
FirstGen	PropPrep	6.18	10.13	8.20	1.11
SecGen	PropPrep	6.15	12.57	9.48	1.84
ThirdGen	PropPrep	7.34	11.34	9.50	1.31

Table 5.12 Conjunctions by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropConj	11.03	17.72	13.56	1.79
FirstGen	PropConj	9.14	16.51	12.89	1.93
SecGen	PropConj	12.23	17.53	14.06	1.37
ThirdGen	PropConj	9.35	16.52	12.75	2.14

The only result that appears to reflect the researched literature is the pattern of use for the adjectives. Frequent use of this grammatical category is usually taken as a sign of a higher level of language. Therefore, a relatively higher percentage of adjectives, as seen among new migrants, could be taken as an indicator of a higher level of language.

However, the puzzling results among the other three groups, with the second generation followed by the third and eventually the first, makes this interpretation hard to sustain.

With regard to the Anovas, all but one of these variables showed a significant difference between the means' groups. The only exception is the variable 'conjunctions' ($F(3,52) = 1.52$, $p = .219$), for which there is no statistically significant difference between any of the four groups considered.

With regard to the variable of nouns, the one-way Anova showed a significant result ($F(3,52) = 6.96$, $p = .000$). The Tukey analysis that followed indicated that there are three pairs that statistically differ from each other, all including the third generation, which thus stands out from the others: the new migrants and the third generation (p -value .017), the first and the third generation (p -value .000) and the second and the third generation (p -value .007).

With regard to the variable of verbs, the Anova sig value was lower than .05 ($F(3,52) = 7.44$, $p = .000$). The Tukey analysis resembles much that has just been seen, regarding the category of nouns. Again, there are three pairs that statistically differ, which all include the third generation: the new migrants and the third generation (p -value .006), the first and the third generation (p -value .000) and the second and the third (p -value .018).

With regard to the variable of adjectives, the one-way Anova showed a significant result ($F(3,52) = 6.28$, $p = .001$). In the Tukey analysis, two pairs can be said to statistically differ from one another: new migrants are significantly different both from the first generation (p -value .001) and from the third (.006), thus partly supporting the descriptive discussion above.

With regard to the variable of adverbs, the Anova performed showed a p -value $< .05$ ($F(3,52) = 4.64$, $p = .006$). The Tukey post-hoc comparison test indicated then that there are two pairs that statistically differ from each other: the first generation and the second (p -value .014), and the first and the third (.009).

With regard to the variable of pronouns, the p -value is significant ($F(3,52) = 9.64$, $p = .000$). The Tukey test showed that there are three pairs that statistically differ from one another: the new migrants and the third generation (p -value .007), the first generation and the second (p -value .004), and the first generation and the third (.000).

Lastly, with regard to the variable of prepositions, the Anova outcomes are: $F(3,52) = 6.73$, $p = .001$. In the Tukey analysis, just a single pair statistically differed: the new

migrants and the first generation (p-value .000), hence showing a gap between the two groups of native speakers.

The results seen in the tables presented above, are hard to interpret. We can suggest that it is indeed not possible to find a clear trend or, more likely that there is a need for a larger corpus and/or a higher number of informants. Given these factors, in the next section I will group the results into two categories, content words and function words, to test whether a broader classification would reveal some clear generational trends.

5.3.7 Content and function words

Generally speaking, words can be divided into two categories,¹⁰³ usually with a skewed distribution in a corpus: content words and function words. The former, which includes nouns, verbs and adjectives, are open-class categories, meaning that words can be added more freely, either as new terminology or as derivatives. Furthermore, words belonging to this category are extremely informative of the content of a corpus.¹⁰⁴ Conversely, function words (adverbs,¹⁰⁵ pronouns and prepositions) form closed-class categories, usually with a high frequency of use. Their function is to bind lexical words; that is to say that they occur frequently in a text but are not very indicative of its content.

This classification is not completely water-tight: although auxiliary and modal verbs would appear to be classifiable as content words, being verbs, they are often considered as function words, because they form a closed class and have no independent meaning. However, some analysts point out that modal verbs express “obligation, permission and ability, and therefore convey content” (Stubbs 2004, p.40).¹⁰⁶ Additionally, adverbs, which in this study have been counted as function words, represent an undefined class whose items belong partially to the class of content and partially to that of function words.

¹⁰³ The present study takes no account of another possible group, labelled ‘insertions’, which is composed mainly of words intended to express feelings, gain attention or take time, but with no particular meaning in themselves.

¹⁰⁴ Spoken language is richer in function words than written language, yet its lexical density is less than that of written language, in which conversely there is a higher proportion of content words per clause.

¹⁰⁵ The association of adverbs with one of the two categories is a moot point and varies across different languages. In Italian there are adverbs that more clearly belong to one category and others that belong to another. Considering the size of my sample, and with the impossibility of disambiguating all the cases, I took the decision to treat them all as ‘function words’, which in my view is more suitable for the majority of the occurrences.

¹⁰⁶ Although I generally agree with auxiliary and modal verbs not being counted as content words (but being included in the category of function words), for the present study I include these two types of verbs in the category of content words, because, as discussed above, the individuation of their role is beyond the aims of this quantitative chapter.

Other key characteristics are that function words are usually shorter than content words and more frequent in a corpus (Osterhout, McLaughlin, Kim, Greenwald & Inoue, 2004).

In table 5.13 the outcomes of SPSS are reported for the proportion of Content vs Function words.

Table 5.13 Content and Function words by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropContentWords	39.25	56.45	52.66	4.41
	PropFunctionWords	43.55	60.75	47.33	4.41
FirstGen	PropContentWords	49.74	57.73	52.74	2.38
	PropFunctionWords	42.27	50.26	47.25	2.38
SecGen	PropContentWords	48.60	56,44	53.90	2.22
	PropFunctionWords	43.56	51.40	46.09	2.22
ThirdGen	PropContentWords	49.48	65.25	55.63	4.27
	PropFunctionWords	34.75	50.52	44.36	4.27

In order to get a better overview, these outcomes are visualised by way of two figures, figure 5.4 in which the mean of the percentages of content and function words for the four groups are reported, and figure 5.5 where the internal dispersion is visually expressed.

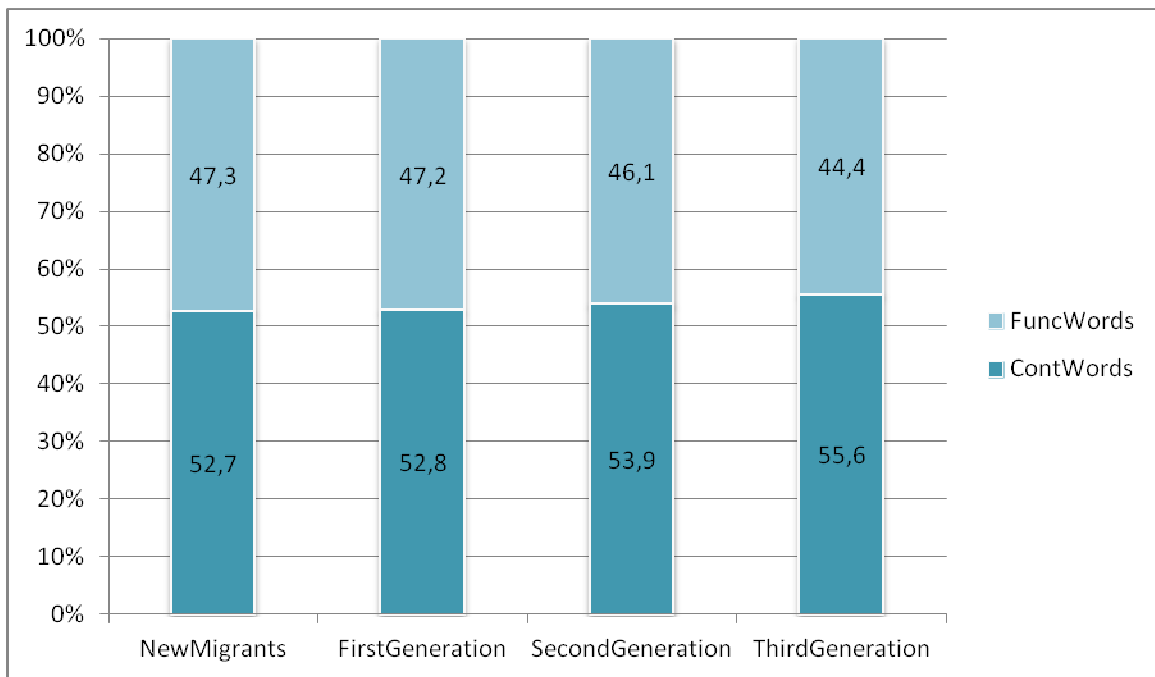


Figure 5.4 Content and Function words by generation (%)

The first outcome that appears from the bar chart is the very similar result seen for the two groups of native speakers: both their corpora, in fact, are split into about 53% of content words and 47% of function words. This associates these two groups, and at the same time marks them out from the other two groups. Although the percentages of the second and third generations differ by only a few points compared to those recorded by native speakers, this result seems to suggest an interesting result: along the generational scale, the percentage of content words increases, up to 53.9% for the second generation and 55.6% for the third generation while the pattern of function words, as mirroring that of content, decreases.

These outcomes seem to highlight two interesting patterns and suggest an interpretation of the development of the language among bilingual communities. With regard to the two groups of native speakers, we can hypothesise that the distribution between content and function words among different cohorts of migrants does not undergo variations over time; this balance represents a pattern that characterises native speakers and that seems to remain stable. This result, along a generational scale, seems to modify, and in particular the ‘structure’ (grammatical aspects) of the language used by second and third generations appears to weaken. Although we know from the literature and from the results discussed above that there is an indisputable decline in language skills among heritage speakers, this may not proceed uniformly, but seems to spread more severely across the functional aspects of the language. This pattern, however, shows that, despite

the decrease in language skills among heritage speakers, a relatively high presence of content words, those which carry the meaning of the language, make their speech still informative and linguistically dense.¹⁰⁷

From a purely statistical perspective, we found that the significant value which emerged from the Anova was greater than .05 ($F(3,52) = 2.4, p = .078$). This means that, with regard to these two categories, if there are differences between these groups, these are not statistically significant.

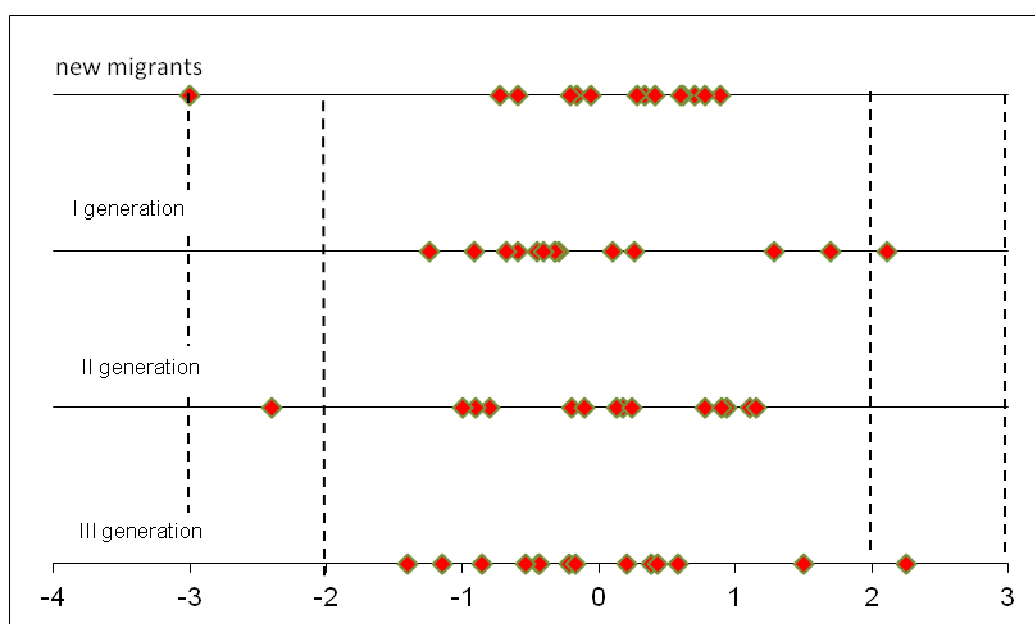


Figure 5.5 Content/Function words Z score by generation

With regard to the results of the z-score test, as shown in figure 5.5, all generations have one informant who deviates particularly from their group's mean. Apart from this, we can observe that new migrants is the most homogeneous group, followed by second generation, while first and third generation are very similar in this respect.

5.3.8 Code-switching

I shall now discuss the class of words categorised by Taltac as not belonging to Italian. Although we must be aware of the crudeness of this index as an indicator of language

¹⁰⁷ Lexical density is defined by Halliday as the ratio of content words to function words, calculated on the total number of words. Spoken language is characterised by low density, written language by high density (Berruto & Bescotti, 1995).

maintenance/attrition,¹⁰⁸ it seems worth considering the general pattern across the different generations and in particular making a comparison between new migrants and first generation, who were speaking in their native and strongest language.¹⁰⁹

In the language maintenance literature it is assumed that a substantial increase of interference in heritage speakers' performance will have taken place. What needs to be determined, in any particular case, is the extent of this increase. Therefore, in pursuit of a better understanding of this trend in this particular study, I have sought to quantify it, by comparing the mean percentages of non-Italian words produced by each of the generation groups (table 5.14).

Table 5.14 Code-switching by generation (%)

Group	Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig PropCS	.33	10.89	1.95	2.81
FirstGen PropCS	.86	11.43	4.51	2.81
SecGen PropCS	2.74	15.85	6.42	3.99
ThirdGen PropCS	1.13	20.38	8.19	6.73

The results follow a clear linear tendency, suggesting that the introduction of words not in Italian increased gradually across the generations, while new migrants clearly used a smaller percentage of such words. Particularly worthy of note is the difference between the two groups of native speakers: the percentage of these words used by first-generation informants was more than double that of new migrants.

With regard to the two groups of native speakers, in particular, two qualitative aspects of these results must be underlined. First of all, these figures provide only a crude index of language interference, as they tell us neither what language other than Italian has

¹⁰⁸ This category aims to give a general overview of the phenomenon of language contact, which is multi-faceted and includes phenomena such as code switching and borrowing.

¹⁰⁹ It is important to reiterate that informants were expressly asked to speak Italian, avoiding any other language. That said, we must be aware that these results have to be taken as crude indices of a linguistic tendency; as in all quantitative analysis, they do not aim to give a precise depiction of the phenomenon investigated. In the case of native speakers, they do not allow us to draw conclusions on the attrition of their Italian: as we have seen in chapter 1, we must proceed with caution before taking them as evidence of language attrition, although they may be considered symptomatic of a weakening of competence. The second and third generations had spent most of their lives in Canada. Therefore, although Italian or dialect may have been their mother tongue, English was certainly the language in which they received their schooling and, for the great majority, the one in which they felt most confident.

been chosen, nor in what context it has been used. It is fairly safe to assume that for new migrants most of these non-Italian items would be words in English, although not exclusively, whereas their composition would be more varied among the first generation (and indeed the second and third),¹¹⁰ including words of English, dialect and Italese. The presence of interference from more than one language may (partially) explain the higher percentage being recorded among the first generation. Another possible causal factor is the length of time since migration to Canada, much greater for the first generation than new migrants. However, while the first cohort of migrants had lived longer in Canada, this factor will have been counterbalanced by their membership of closer Italian networks compared with the new migrants, who tended to live in isolation from the Italian community.

If the increasing percentage of non-Italian words is predictable along the generational scale, the results are less straightforward for the two groups of native speakers. However, they do seem to be in line with the metalinguistic observations emerging from the interviews, which suggest a rather ‘purist’ attitude to Italian among new migrants. Native speakers all showed a great awareness of the changes occurring in Italian in Italy, along with a purist attitude. This tendency seems to have been most strongly manifested within the new migrant cohort, particularly with reference to the introduction of English loanwords into Italian:

Nella lingua italiana di oggi ci sono molte parole in inglese.

In today's Italian language there are many words of English [New migrant informant]

Sì questa forzatura dell'inglese [...] mettere parole inglesi dove l'italiano c'è l'ha benissimo. [...] e non son contrario a usarla la parola [...] per evitare una perifrasi magari oppure se è una parola originaria dell'inglese che esprime un concetto [...] quindi son d'accordo non è che io sia contrario assolutamente tout court però è l'esagerazione l'ostentazione di una parola.

Yes, this twisting of English. Putting English words where Italian has a complete synonym. And I am not against using this word, maybe to avoid a paraphrase. Or where it is an original word from English to express a concept. So I agree, I am not absolutely against it as such. But [what I am against] is exaggeration and showing off. [New migrant informant]

¹¹⁰ A key consideration here is that in order to avoid any communicative disruption and so maximise their performance in Italian, the fieldworker did not attempt to stop them using words or phrases in a language other than Italian.

Although confident in their new language, new migrants appeared not to like mixing Italian and English; in other words, they seemed particularly keen on maintaining a monolingual mode. In the long run, this may help both to prevent the attrition of their Italian and to maintain the competence of Italians with whom they interact, providing inputs of ‘unaltered’ Italian.

Metalinguistic observations offer another key to further depict the outcomes of code-switching phenomena among speakers of a heritage language, and in particular among the second and third generations. A relatively significant group of my heritage speaker informants claimed an inability sometimes to distinguish words in dialect from those in Standard Italian:

A scuola ho scritto la “putea” [“bambina” in Italian] perché ho difficoltà a capire se è italiano o dialetto

At school I wrote “putea” [dialect word for ‘baby girl’] because I have difficulty in knowing if it’s Italian or dialect. [Second-generation informant, my translation]

Younger Italo-Canadians largely lack the opportunity to receive substantial inputs in Italian outside the family. The distinction between Italian and dialect, which for them is already blurred by the non-compartmentalisation of *dilalia*, can thus only be inferred from other and more qualified inputs, such as that received from a substantial formal education. Their code-switching with dialect may thus become an involuntary and not always deliberate choice.

The one-way Anova outcomes presented a significant result ($F(3,52) = 5.19, p = .003$). The analysis with Tukey indicated that there are two pairs that statistically differ from each other: the new migrants and the second generation (p-value .046), and the new migrants and the third generation (p-value .002). These outcomes thus highlight the gap between one group of native speakers (new migrants) on the one hand, and the heritage speakers on the other.

The z-score test shows one final interesting outcome. Table 5.14 displayed the same result for the two groups of native speakers with regard to the standard deviation. However, the z-score in figure 5.6 shows how in reality these groups differ.

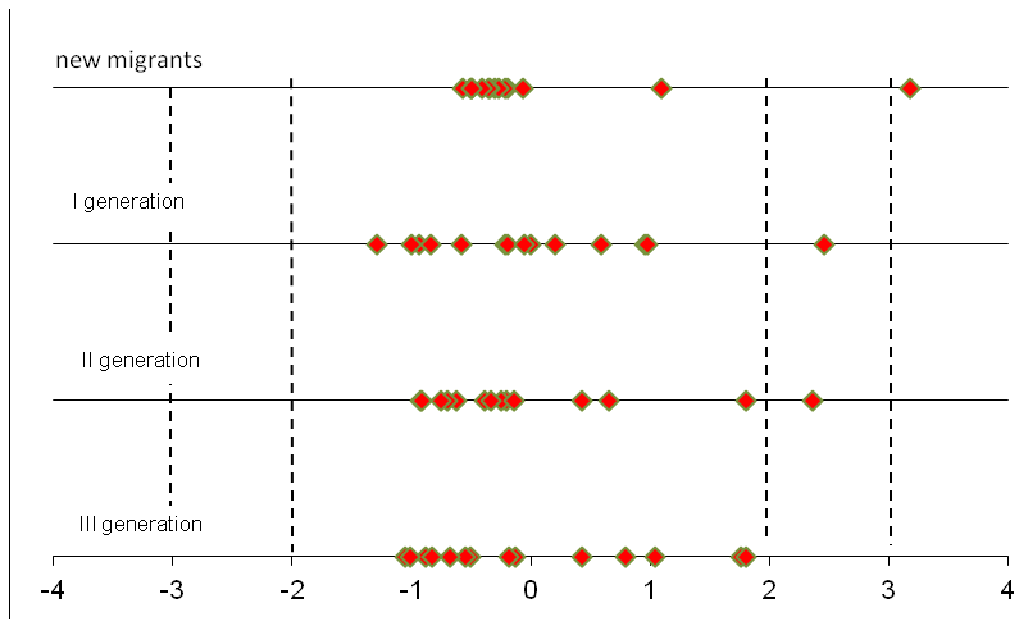


Figure 5.6 Code-switching Z score by generation

The results of the first generation are spread out more along the continuum. On the other hand, new migrants show a higher level of internal consistency, along with one single result that very significantly differs from the mean, being over the threshold of $|3|$. This latter in all probability explains the similar standard deviations of the two groups, although they have very different internal patterns.

5.3.9 Verb tenses

This section considers the morphology of the language used by interviewees, specifically the distribution of verb forms. The aim is to bring to the fore the pattern of verb choices among the speakers, intended mainly as a comparison between more complex (conditional and subjunctive moods) and more ‘everyday’ tenses (indicative). The analysis of verb forms in Italian in contact with English seems to have been little studied from a quantitative viewpoint. A comprehensive study of language attrition among Italians in Australia was conducted by Caruso (2004 & 2010),¹¹¹ whose analysis is framed by the concept of markedness and focuses on an internal comparison of attrition across the first and second generations.¹¹² Moreover, in order to bring her

¹¹¹ In order to offer a more balanced linguistic view and to take into account potential regional differences, Caruso restricted her sample to people from a defined area of the Calabria region in southern Italy.

¹¹² Caruso focused her data collection on family groups’ conversations. Therefore the second generation’s linguistic productions were compared to those of their parents. This helped to take into account

analysis closer to spoken Italian, she focused on a particular mood, the indicative, ignoring the other two finite moods, subjunctive and conditional, which are more often used in written production. Few areas of attrition were identified within the first generation,¹¹³ in contrast to the reduction and simplification of the second generation as well the increase in their use of deviant forms and in the number of forms they did not produce. Another significant aspect of her analysis, consistent with the findings of Bettoni in the same Italo-Australian linguistic context, was “a sharp drop in correct forms from the more to the less competent speakers of second generation, rather than a gradual decrease” (Caruso 2004, p.21). This result is also in line with the higher level of internal variability (distribution) among second-generation speakers in other indices, as discussed above.

The research carried out by Scaglione (2000) among the Lucchese (Italian Tuscan) community in San Francisco showed a stable incidence of indicative forms among the first and second generations. In contrast, the conditional and in particular the subjunctive were even rarer than would have been expected. What was deemed to be remarkable was that while occurring very rarely, they were mostly used correctly, in particular by first-generation speakers.

The following analysis of the occurrence of verb forms in this corpus takes account of the specific characteristics of the spoken language. A criterion of internal comparison is adopted: speakers in the four categories are compared with each other, according to their generation and their native language, in order to take due account of each group’s peculiar linguistic situation and to avoid making skewed assertions by comparing their production with the standard written language. Moreover, my use of a single grid of questions means that all informants were asked to talk on the same topics and so were directed towards using the same verb tenses. Divergences are thus ascribable to their language choices and perhaps to their skills.

incomplete language acquisition phenomena, as second-generation speakers were not compared with standard (oral) Italian, but with the language of their relatives.

¹¹³ Briefly, these were wrong use of the auxiliary in compound verbs and a lack of agreement, supposed by Caruso to be the result of the intrinsic nature of a spoken language (little advance planning), thus a characteristic of the spoken language of non-attriters as well. Another trait was the presence of English interference, “resulting in fully integrated lexical transfers formed on the basis of the most productive conjugation in Italian” (Caruso 2004, p.15).

Table 5.15 Verb tenses by generation (%)

Group		Min. (%)	Max (%)	Mean (%)	Std. De.
NewMig	PropIndicative	57.91	69.41	62.75	4.17
	PropSubjunctive	.64	3.58	2.08	.96
	PropConditional	.53	2.59	1.34	.60
	PropInfPartGer	26.68	37.89	33.81	3.67
FirstGen	PropIndicative	59.92	73.81	66.11	4.06
	PropSubjunctive	.37	2.65	1.02	.59
	PropConditional	.35	4.11	1.47	1.02
	PropInfPartGer	25.04	38.60	31.38	3.96
SecGen	PropIndicative	64.43	76.37	71.20	3.55
	PropSubjunctive	.00	1.07	.39	.29
	PropConditional	.00	1.06	.34	.31
	PropInfPartGer	22.59	34.03	28.05	3.55
ThirdGen	PropIndicative	63.12	83.16	71.77	5.80
	PropSubjunctive	.00	1.19	.46	.35
	PropConditional	.00	1.95	.70	.78
	PropInfPartGer	16.48	34.26	27.05	5.05

Although the indicative tenses appear widely in both spoken and written language, their relative frequency is considerably greater in the former (Berruto & Bescotti, 1995). Hence, a higher percentage of this mood in a sub-corpus may be indicative of an output closer to the spoken language and a lower percentage, conversely, of closeness to the written form. The non-indicative verb forms are associated with more complex syntax. With respect to the indicative tenses, the patterns indicate clearly three levels of frequency: the third generation scored the highest percentage, followed closely by the second. Then the first generation fell in the middle of this continuum, whereas new migrants showed the lowest percentage of indicative use. This finding is consistent with the research of Caruso (2004) and with the hypothesis of a rise in the use of the indicative tenses along the generational scale. The better educated and most recently arrived new migrants, on the other hand, showed a relatively lower level of indicative use.

I shall now discuss the use of the other forms, which seem worth examining despite the fact that the percentages —with the exception of the non-finite forms—were very low,

as consequently were the divergences among the four groups. With regard to the subjunctive mood, the pattern which emerged was the mirror-image of that for the indicative. The subjunctive forms are considered the most difficult to use and therefore mainly restricted to the written language (Berruto & Bescotti, 1995) or the speech of highly educated users; as expected, the second and third generations had the lowest percentages of use, again fairly close to one another, while the first generation was around the middle of the continuum and the highest percentage use was among the new migrants.

The pattern for the conditional partly differs from the preceding: it mostly shows a difference between native and second-generation speakers on one hand, while the third generation is placed in the middle. The gap between the two groups of native speakers, although slight, showed the reverse of the pattern obtained so far, with the first generation recording the highest percentage of use.

Finally, the pattern for the non-finite forms (infinitive, participle and gerund) was similar to that for the subjunctive, although the difference between the first generation and new migrants seemed to be slighter, as was that between the second and third generations.

These findings are in line with reports in the literature, although a few aspects are worth discussing briefly. First, the gap between the second and third generations was smaller than I might have expected, while there was also a slightly higher incidence of the more complex verb forms among the third generation, in particular of the subjunctive. Hence we may propose the hypothesis that the use of verbs is a complex aspect of the decrease in the language skills of heritage speakers from generation to generation, in that there is not a steady decline in their use but the major declines between the first and second generation then moderate or even undergo a reverse in the passage from the second to the third. Secondly, the divergence between the first generation and new migrants was marked, with the latter recording the highest percentage for the most difficult mood (subjunctive) and the lowest for the most common (indicative), a trend that was reversed among the first generation. Differences between these two groups were once again noteworthy.

The discussion above, however, was only partly supported by the Anova. With regard to the indicative, the Anova performed showed a significant difference between the four groups ($F(3,52) = 12.92$, $p = .000$). The Tukey test indicated in particular how the four

groups that statistically differ from one another: the new migrants differ both from the second (p-value .000) and the third generation (p-value .000). Similarly, the first generation differs both from the second (p-value .021) and the third (.008).

With regard to the subjunctive, the one-way Anova outcomes again showed a significant result ($F(3,52) = 22.7, p = .000$). In the Tukey analysis four groups emerged as statistically different from each other. In particular we have the new migrants who statistically differ from the first (p-value .000), second (p-value .000) and third generation (p-value .000). Also, the first generation statistically differs from the second (p-value .043).

With regard to the conditional, the one-way Anova outcomes displayed a significant result ($F(3,52) = 7.53, p = .000$). The Tukey analysis that followed showed that there are three pairs that statistically differ from each other: the new migrants and the second generation (p-value .004), and the first generation both with the second (p-value .001) and the third (p-value .034).

Lastly, with regard to the non-finite forms, the Anova again showed a significant difference between the groups ($F(3,52) = 7.99, p = .000$). In the Tukey analysis, three pairs turned out to statistically differ from each other: the new migrants from both the second (p-value .003) and the third generation (p-value .000), and the first from the third generation (p-value .036).

These outcomes suggest the presence of a noteworthy gap between the two groups of native speakers on the one hand, and the heritage speakers on the other. However, the existence of a gap between the new migrants and first generation, proposed above, turned out to be statistically significant only with regard to the subjunctive.

5.4 Content analysis

This final subsection moves the focus from a purely statistical perspective to a semantic study of the vocabulary. The aim here is to integrate the results discussed so far with a description of my informants' linguistic skills through an analysis of what their lexical usage can tell us about their language skills.

Another task that can be performed with Taltac is the analysis of specificities, a function which allows the extraction of lexical units that are typical of, or exclusive to, a given text. Once a corpus has been divided into sub-corpora, according to the variables the

researcher aims at investigating, it is possible, through a hypergeometric formula¹¹⁴, to isolate the specific language used in each of its sub-groups. In this study I divided my informants according to the variable of generation.

Here this discussion focuses on the positive specificities,¹¹⁵ namely those units that were used significantly in each of the four sub-corpora. Within the various sub-corpora, the number of specific forms varied greatly: there were 597 in the first-generation sub-corpus, 423 for the second generation, 458 for the third and 939 for the new migrants. In order to carry out a more qualitative analysis, I selected the first 100 of every subset, beginning with those having the highest levels of specificity.

An overview of these forms reveals some interesting trends. First-generation speakers typically used terms with a negative connotations, such as *sangue* (blood) and *fame* (hunger) or related to their personal lives, such as *in pensione* (retired) and *telefonare* (to phone). *Telefonare* was typical of their output because they would not usually make use of the internet, reference to which was conversely more common among the younger generations and new migrants. Interestingly, two expressions that were specific to the first generation were *montagna* (mountain) and *in montagna* (in the mountains), which were probably related to their early lives in Italy, as they mostly came from the mountainous northern part of the Veneto region.

From the speech of the second generation, two topics seem to emerge: one related to the content and the other more linguistic. There was a significant presence of terms related to the family sphere, commencing with *in famiglia* (in the family) and including *fratello* (brother), *suoceri* (parents-in-law) and *cognome* (surname). The more linguistic trend is that the terminology was more variegated than in the other sub-corpora, apparently because their repertoire typically comprised three different languages. There was a noteworthy presence of dialect words such as *pomo* (apple, rather than the Italian *mela*) and *darente* (close; *vicino* in Italian), while sometimes both an Italian word and its dialect equivalent were specific to this sub-corpus, such as in the case of the respective words for fork: *forchetta* (Italian) and *piron* (dialect).

The third-generation sub-corpus also revealed two trends. The first suggests a broader vision of the participants' country of origin, with the presence among their specificities of terms related to Italian locations far from the Veneto, such as Latina and Puglia/Bari (respectively in central and southern Italy). The second trend is more linguistic and

¹¹⁴ In statistics this formula is used to calculate the probability for a random selection (without repetition) of an element from a given collection of elements.

¹¹⁵ Negative specificities are units that have been significantly underused in one or more sub-corpora.

related to their potential skills. A large number of their specific items were of simple and everyday terminology, including nouns such as *strade* (streets), *edificio* (building) and *stazioni* (stations), and basic adjectives such as *bello* (beautiful), *triste* (sad) and *importante* (important).

Among new migrants, there were also a significant number of terms related to everyday life, such as *mutuo* (mortgage), *posto di lavoro* (workplace), *studiare/studi* (to study/studies), *scuola superiore* (high school) and *universitario* (university, adj), mostly highlighting their different lifestyle compared to that of the first generation. In the same vein, there were expressions related to travelling or moving, which seem to have assumed a specific connotation among new migrants. Among these, *viaggiare* (to travel), and *turisti* (tourists), can be said to be typical of the lifestyle and condition of migrants. Lastly, from a purely linguistic perspective, it is interesting to note, among terms specific to new migrants, two items which have entered Italian recently and which were thus not part of the Italian repertoire of the first-generation migrants when they left Italy: *background* and *software*.

While summarizing these outcomes, some interesting trends are worth highlighting briefly, in particular related to linguistic aspects. While the first generation and new migrants revealed interesting content choices, linked to their past life in Italy, but mostly regarding their present and everyday life, among the second and third generations two linguistic trends can be underlined. In the case of the second generation, the co-presence of dialect and Italian words draws attention to their multifaceted linguistic skills, which may be seen as more variegated than those of the other groups; but it may also be related to their difficulty in distinguishing what is Standard Italian and what dialect, favouring the co-presence of the two in their speech. The language of the third generation seems by contrast to be characterised by words belonging to the basic vocabulary, in line with their possibly lower level of ability in Italian as mentioned earlier.

5.5 Partial correlation analysis

In the first part of this chapter I investigated the linguistic outcomes of the interviews from a descriptive perspective. These pages served to examine some aspects that will be further explored and discussed in what follows. Over and above the individual results, two major trends seem to have emerged. On one hand, there appear to be differences

between the two groups of native speakers, with new migrants overall scoring better than the first generation. On the other, there seems to be a trend, although not always validated, of a decrease in terms of language skills along the generational scale, from the first generation to the third. These preliminary outcomes seem therefore to validate the working hypotheses proposed at the end of part one of this thesis. Hence, these divergences across groups of migrants and generations, which have been discussed so far only from a purely descriptive perspective, will now be verified with the support of a statistical analysis through the software SPSS.

Before proceeding I want to briefly recall a couple of aspects that are particularly important to consider while performing the next step of analysis. Firstly, as we have just seen, a point of interest in this study is a comparison between two types of migrants, namely new migrants and the first generation. As suggested at the beginning of this work, the scenario I had the chance to investigate was unique in the specific historical contingency which offered the opportunity for a direct comparison of two different types of migrants. Particular attention, however, has to be given to manage the level of education of my informants, which, as discussed in this thesis, is a relevant variable in language maintenance-attrition (chapter 1) and which exhibits a non-homogeneous trend in the population investigated in this study (chapter 3). The two groups of native speakers have in fact reached different levels of education although, in contrast with the two groups of heritage speakers, all received their education in Italian. Some cases among the new migrants form an exception to this, but they are marked by their high level qualifications (masters or PhD), obtained beyond a university degree, which for all of them was awarded in Italy.

Secondly, another point of interest lies in an intergenerational comparison, that is between what I labelled in this study as the first, second and third generations. In this part the new generation category is excluded; the inputs that the two groups of heritage speakers have received, in fact, derived mainly from their (grand)parents. Their variety of contacts could, of course, have been expanded through personal practices; in particular, my informants of the second and third generation showed appreciation for (contemporary) Italian music. However, an early, consistent and regular contact with the heritage language was primarily provided within the family domain.

The intergenerational pattern starting with new migrants is not viable yet due to the young age of the following generations which, in addition to their low numbers, does

not allow them to be tested in a relatively demanding data collection method, such as a sociolinguistic interview.

In the following pages I will further develop these two research questions in turn. If the descriptive part at the beginning of this chapter has highlighted some general patterns, we now pass on both to a more in-depth analysis, and also to investigate whether it is possible to determine the grounds of any significant divergence among these generations. This last aspect will be performed by linking the linguistic results of the interview with the extra-linguistic variables discussed in chapter 4.

Although some relevant results about the extra-linguistic factors impacting on language skills emerged in chapter 4, the attempt to create an indicator, grouping together all the questions pertaining to one single aspect, for the variable of affiliation was unsuccessful. This made the analysis of this chapter, and in particular my intention to link the linguistic results with the level of affiliation of my informants, more problematic. In view of all this I was left with the alternative of using one single question. This however would have implied that the phenomenon investigated was, to a high degree, not genuinely representative. If the answers describing my informants' affiliation did not go together, the choice of one single question as 'representative' would have skewed the results from the start. Therefore, although aware of the influence that affiliation can have on language maintenance/attrition, I eventually decided to confine the analysis of the impact of the extra-linguistic factors described in chapter 4 to the indicator of 'linguistic habits' alone.

5.5.1 Significant divergences in the language outcomes between native speakers

In this first section I explore the differences between the two groups of native speakers, firstly calculating if any detected divergence in their outcomes is also statistically significant, and if so, assessing the impact of the variable of linguistic habits on these outcomes.

As recalled above, the variable of education is quite influential on language maintenance/attrition. In chapter 1 we have discussed how its impact does, however, vary, depending on the task used to collect the data: it is major in tasks concerning explicit knowledge (such as grammatical exercises) and less so in those regarding implicit knowledge (such as interviews).

Although in this study I am working on the linguistic outcomes of the interviews, therefore hypothesising a relatively minor impact of education on the final outcomes, the two groups that I discuss here differ very considerably with regard to their level of education. Given these premises, controlling for this variable was regarded as necessary. As seen in chapter 3, the two groups of speakers almost completely diverge with regard to their educational level, while having only a couple of informants overlapping. In fact, all the first generation received the lower of the three levels of education into which the Italian school system has been partitioned, while new migrants are mostly spread between the other two. Only two of my new migrant informants share the same level of education as the first generation.

This indicates that to a great extent the generational affiliation mirrors the level of education. In other words, a given generational affiliation corresponds largely to a given level of education.

In order to confirm these findings I ran two Anovas, one keeping the generation as the independent variable and the other considering the level of education as the independent variable. The results show that the list of linguistic outcomes presenting a statistical difference between the two groups of informants is very similar both considering the generation and the level of education as the independent variable. The two differences noted pertain to the Vocabolario di Base (VdB) and the Adverbs, both of which scored significantly when the independent variable was generation (respectively p-value = .027 and p-value = .028), but the opposite when considering the level of education as the independent variable (respectively p-value = .078 and p-value = .065).

As we can infer from these outcomes, the variable of education goes along with generation in that they almost completely overlap. Therefore the list of the significant outcomes is thus drawn up, without considering those for the VdB and Adverbs (which are the two linguistic variables that are not shared in the two lists).

The final list is reported in table 5.16:

Table 5.16 List of statistically significant differences¹¹⁶ between the two groups of native speakers

Average length of content words	p-value .000 (.000)
Percentage of high freq. words	p-value .005 (.002)
Percentage of nouns	p-value .025 (.002)
Percentage of adjectives	p-value .000 (.000)
Percentage of pronouns	p-value .047 (.013)
Percentage of prepositions	p-value .000 (.000)
Percentage of code-switching	p-value .024 (.020)
Percentage of indicative_tenses	p-value .040 (.048)
Percentage of subjunctive_tenses	p-value .002 (.004)

I have already discussed these linguistic variables in the first part of this chapter. Therefore, I will just confine myself here to a very brief review of those, among these outcomes, that are more worthy of attention. The list above includes some linguistic variables that we have seen to be plainly related to a higher level of language proficiency, such as the average length of content words, the percentage of adjectives and the percentage use of the indicative and of the subjunctive. As discussed above, these all point to native speakers as those with the higher level of language skills. The percentage of high frequency words can also be noted. However it goes in the opposite direction: the first generation showed in fact a lower result in the use of more ‘common’ words compared to the new generation. The analysis of the z-score, however, indicated that with regard to this linguistic variable we need to take a particularly sensitive approach and consider these outcomes as a guideline.

Finally the result for code-switching also shows a significant difference between the two groups of native speakers. We have seen how the use of words not in Italian (the target language of my interviews) is higher among the first generation. We have, nevertheless,

¹¹⁶ The p-values reported in this table are the outcomes of the Anova considering 'Generation' as the independent variable. Those reported in brackets are instead those considering 'Level of education' as the independent variable.

to remember that it is a rough result about the ‘purity of use’ of the language, and it does not tell us which (and why) other languages are used instead.

I move now to the second part of this analysis, with the goal of assessing the impact of the linguistic habits on the outcomes listed above. In order to do that, I performed a partial correlation with SPSS. This statistical analysis is used to describe the relation between two variables (one dependent and one independent variable) while controlling for an additional variable, which we presume may be influencing the two variables of interest that we are trying to correlate. By controlling for the effect of this intervening variable, we can obtain a more accurate indication of the real relationship between the two variables we are interested in.

In this study I computed a partial correlation in which each linguistic outcome listed above (see tables) was related with the variable of linguistic habits, while controlling for generation. We have already discussed how this latter is linked, and has an impact, on both the linguistic habits (see chapter 4) and the linguistic outcomes of the interviews (earlier this chapter). Here, in studying the relationship between the language outcomes and the variable of linguistic habits, we will thus need to take into account the influence of generation, that is by removing its effects on the relationship we are studying.

The actual performance of a partial correlation, as will be discussed in the following pages, implies two steps in the analysis: firstly, the examination of the relationship between the two variables of interest, both without controlling for the effects of one or more other variable (bivariate/simple correlation), and secondly by controlling for the effects of a third variable (partial correlation). It is this comparison which provides us with a real picture of the true relationship because it gives us an understanding of the contribution of one or more variables in the relation of interest.

First of all it is important to remember that correlating two variables does not allow us to make cause-and-effect statements, because correlation *per se* is simply a measure of association between variables. In other words, it measures the tendency of a variable to change according to another variable.

Correlation, in particular, measures the direction and the strength of the relationship between two quantitative variables. The direction of the relation can be either positive or negative. In the first case the two variables move in the same direction. Therefore, if one increases so does the other, and vice-versa. In the second case, the two variables move in the opposite direction, so if one increases the other tends to decrease.

The strength of the relation is measured with a numerical value and it can range between -1 and 1. A result closer to the value $|1|$ means a stronger relationship between the two variables, whereas an outcome closer to the value 0 indicates a weaker relationship.

If we want to visually display the relationship between the two variables, we can do it by creating a graph (scatterplot). With regard to the strength of the relationship, the more the dots are grouped around the line, the stronger is the relation between the two variables. The more uniformly scattered they are, the weaker is the relation.

I discuss now the results. For five out of the nine linguistic variables¹¹⁷ from the list above no correlation has been found between these linguistic outcomes and the variable of linguistic habits (for all of them the p-value is $< .05$), whether controlling or not for the effects of the intervening variable of generation. We thus move on to investigate and discuss the other outcomes¹¹⁸.

For three variables, namely percentage of adjectives, code-switching and indicative tenses, there is a fair significance (either positive in the case of adjectives [$r = .392$], or negative in the case of code-switching [$r = -.375$] and indicative tenses [$r = -.447$]) showing a simple correlation between these linguistic outcomes and the variable of linguistic habits.

¹¹⁷ These variables are: % of HighFrequencyWords, % of Nouns, % of Pronouns, % of Prepositions and % of Conjunctions.

¹¹⁸ Neither with regard to the two groups of native speakers discussed here, nor I can anticipate with regard to the intergenerational perspective discussed later, have I noted cases where the relation when not allowing for the intervening variable was not significant but became significant when controlling for it.

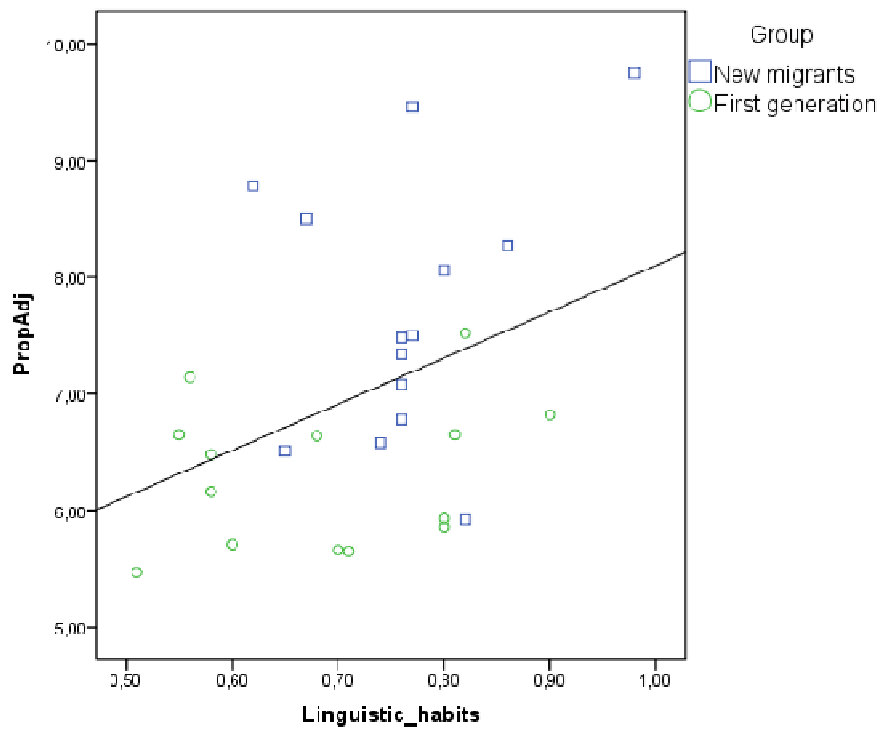


Figure 5.7 Simple correlation between % of Adjectives and Linguistic habits for native speakers

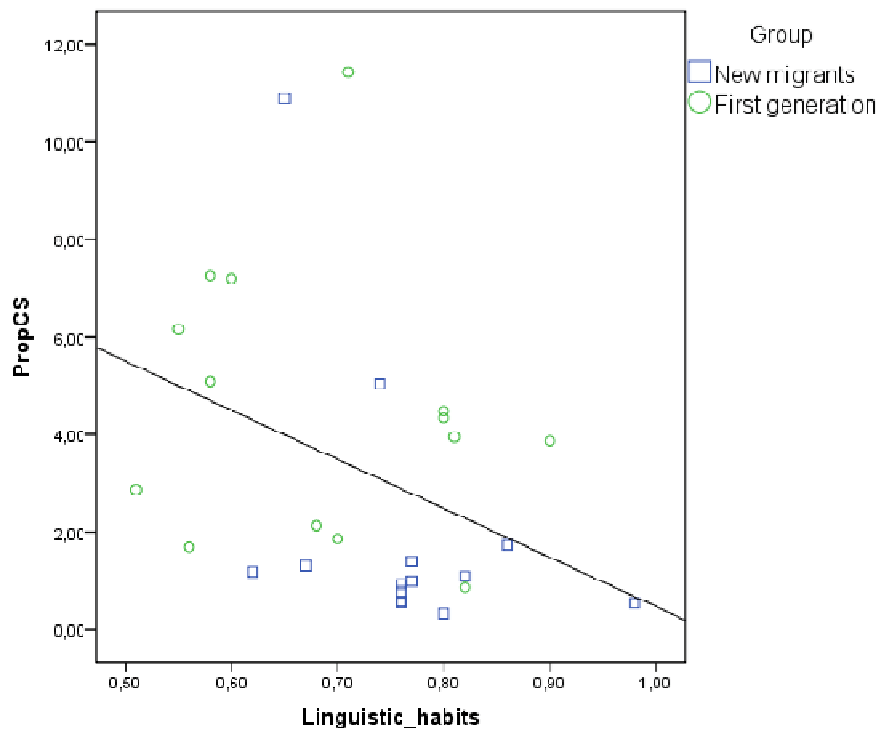


Figure 5.8 Simple correlation between % of Code-switching and Linguistic habits for native speakers

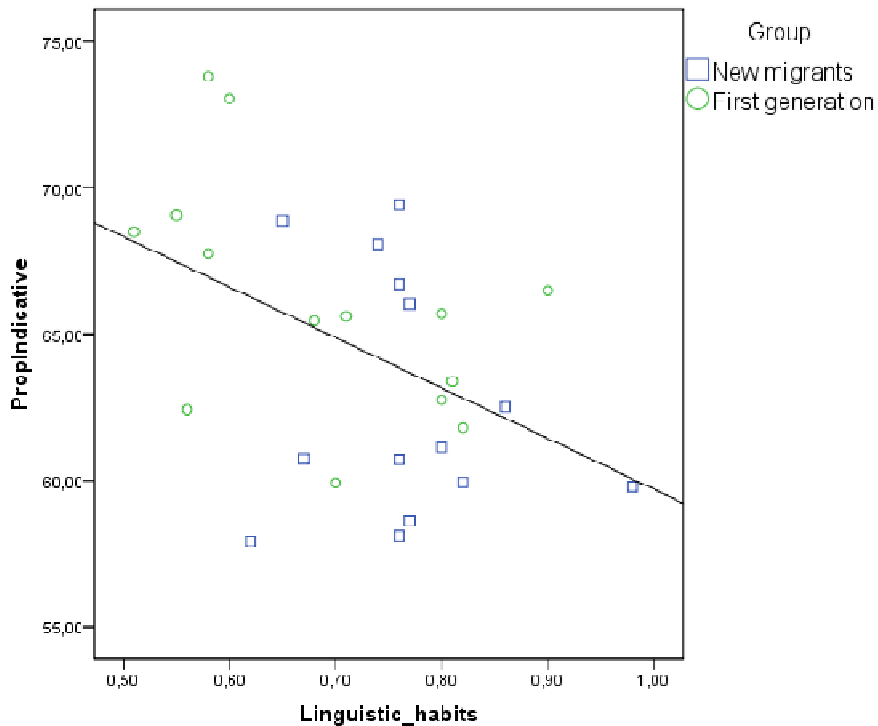


Figure 5.9 Simple correlation between % of Indicative tenses and Linguistic habits for native speakers

However, if we look at these correlations, while controlling for the effect of the intervening variable (generation), we see that they fade. In all of these three cases, against a significant simple correlation between these linguistic outcomes and the linguistic habits, we have a non-significant one when controlling for the effects of generation on the relationship. Therefore, by taking out the effects of the variable of generation from both of the two variables, we remove most of the apparent relationship between them. It is therefore the variable of generation which plays an influential role, determining this (apparent) relation between these linguistic outcomes and the variable of linguistic habits.

The last linguistic variable that remains to be discussed is the average length of content words, which presents a different outcome compared to those examined so far. I firstly performed a simple correlation, which shows a positive correlation between the two variables of interest, with a fair strength ($r = .514$).

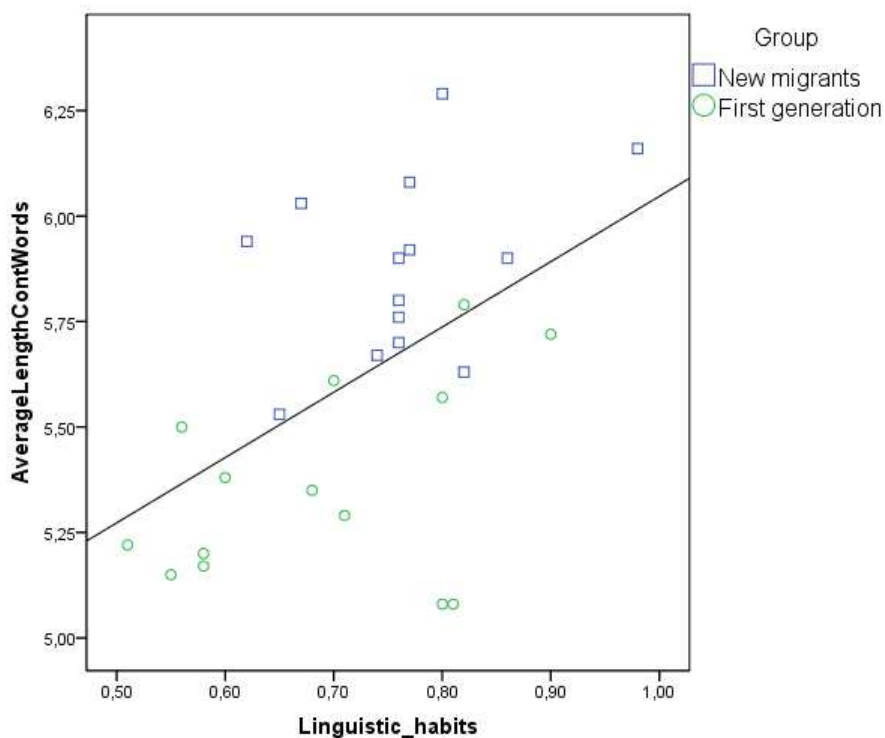


Figure 5.10 Simple correlation between Average Length of Content Words and Linguistic habits for native speakers

I subsequently performed a partial correlation. The relationship between the average length of content words and the linguistic variable remained statistically significant, even after removing the effects of generation. The strength of the correlation lowers to .398, a value that is, however, only moderately less than that when the effects of the variable of generation are not removed ($r = .514$). We can eventually deduce that the relationship between linguistic habits and the average length of content words remains even after controlling for the effect of the third variable. By comparing the two correlations we can see, however, that generation inflates the correlation, although playing a relatively marginal role: the strength of the relationship, when discounting the effects of the intervening variable upon both of the other two variables, is only marginally lower and the direction of the relationship remains unchanged. This suggests that relatively little of the relationship between the average length of content words and linguistic habits can be explained by generation.

5.5.2 Significant divergences in the language outcomes along the generational scale

In this second and last section I follow the same analysis path as above, this time exploring the results from an inter-generational perspective.

In the first step of this analysis the linguistic variables which present statistically significant differences between the generations were singled out. These outcomes were obtained by running an Anova test. The list of these results is reported in table 5.17:

Table 5.17 List of statistically significant differences between the first - second - third generations

Number of tokens	p-value .000
Average length of content words	p-value .003
Percentage of high freq. words	p-value .000
Percentage of low freq. words	p-value .000
VdB	p-value .050
Percentage of nouns	p-value .001
Percentage of verbs	p-value .001
Percentage of adverbs	p-value .003
Percentage of pronouns	p-value .000
Percentage of prepositions	p-value .034
Percentage of content/function words	p-value .048
Percentage of indicative_tenses	p-value .004
Percentage of subjunctive_tenses	p-value .001
Percentage of conditional_tenses	p-value .001
Percentage of non_finite forms	p-value .026

The first finding that our attention is drawn to is the higher number of items compared to those of the list of native speakers (15 against 9). Linguistic divergences among the groups at an inter-generational level thus seem to be more marked and statistically significant.

Again, many of the variables in this list are plainly linked with the level of language skills of my informants, as discussed in the first part of this chapter, where they are clearly shown to follow a decreasing pattern along the generational scale. There is, for instance, the decrease in the average content word length from the first to the third generation, as we have seen, suggesting a tendency to use more everyday words, or a higher percentage of use of the easier verb tenses (indicative). The differences that were identified in the first part of this chapter have turned out in many cases to be also statistically significant.

I move now to the second part of this analysis, assessing the impact of the variable linguistic habits on the outcomes displayed in the list above. I again computed a partial correlation in which each of the fifteen linguistic outcomes was correlated with linguistic habits, while controlling for generation.

For six out of fifteen variables¹¹⁹ there is no correlation with the variable of linguistic habits, all showing a p-value lower than .05. This result is seen when looking both at the simple correlation and at the partial correlation (after the third variable of generation has been taken into account).

Four variables, namely number of tokens, average length of content words, percentage of verbs and percentage of pronouns, show instead a positive, significant, simple correlation. For all of them the strength of the correlation (Pearson's *r*) is fair and the direction of the relationship is positive, meaning that the two variables mirror each other: as one increases so does the other.

¹¹⁹ These variables are: % of High Frequency Words, % of Nouns, % of Adverbs, % of Prepositions, % of Content and Function Words, and % of VdB.

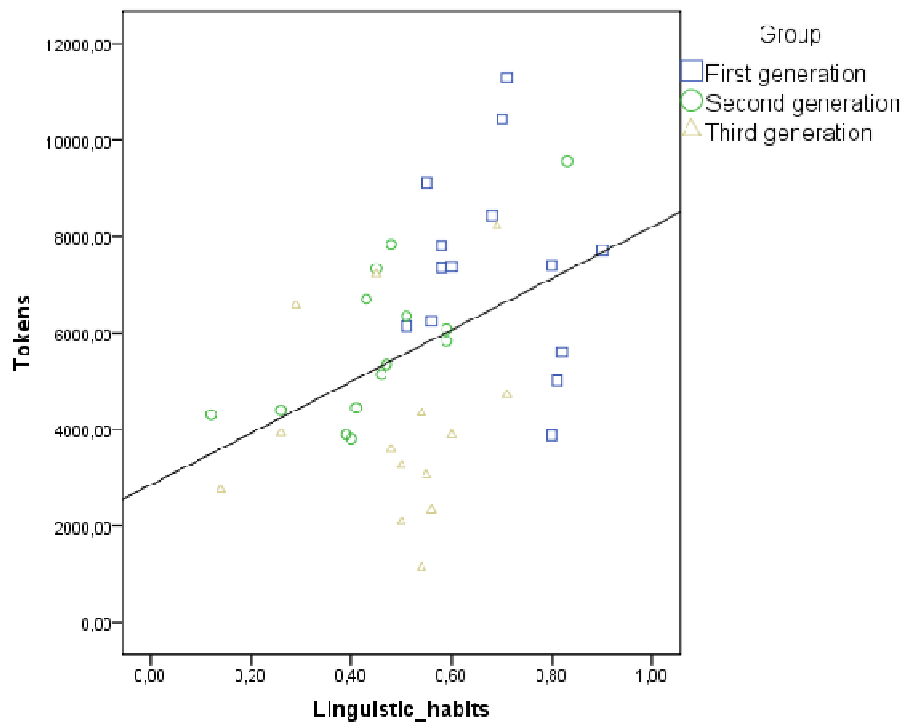


Figure 5.11 Simple correlation between Tokens and Linguistic habits for first/second/third generation

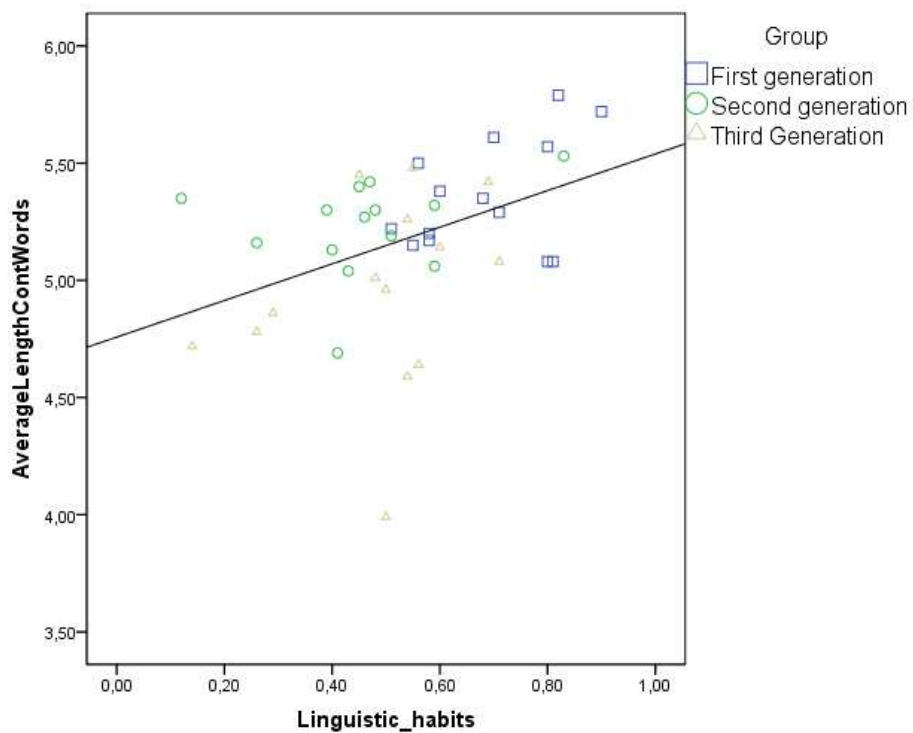


Figure 5.12 Simple correlation between Average Length of Content Words and Linguistic habits for first/second/third generation

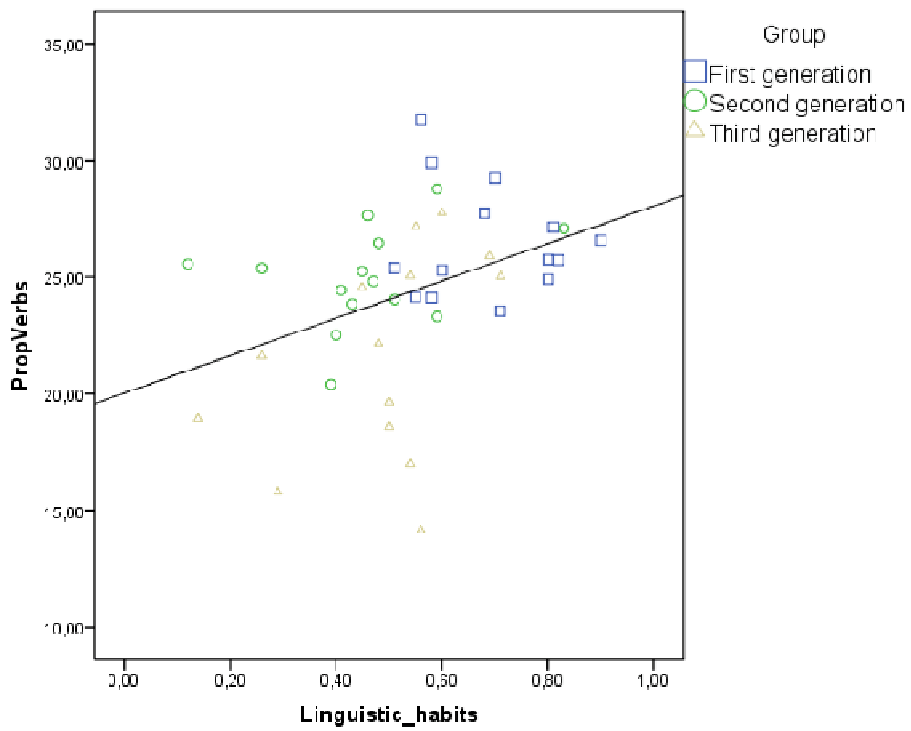


Figure 5.13 Simple correlation between % of Verbs and Linguistic habits for first/second/third generation

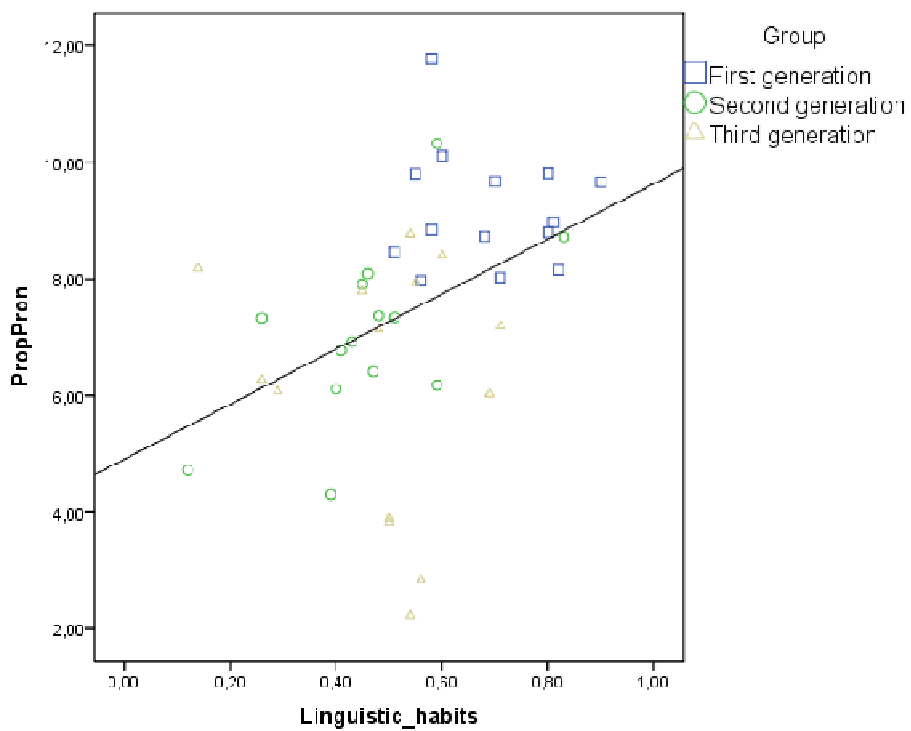


Figure 5.14 Simple correlation between % of Pronouns and Linguistic habits for first/second/third generation

However, if we perform a partial correlation, thus controlling for the effect of the generation variable, we find that the association between these linguistic outcomes and linguistic habits wanes. The significant bivariate correlations between the two variables of interest are not validated when controlling for the effects of the intervening variable on the relationship. Therefore, by taking out the effects of generation from both of the two variables, the relation ceases to exist.

The last five linguistic variables to be discussed are the percentage of low frequency content words, and the percentages of indicative, subjunctive, conditional and non-finite forms.

I firstly ran a simple correlation to assess whether these variables were correlated with linguistic habits. All of them turned out to have a p-value lower than .05. Also, these correlations were fairly strong for all of them - with values ranging from .481 to .579 - and all of them positive, with the exception of the percentage of indicatives which scored negatively.

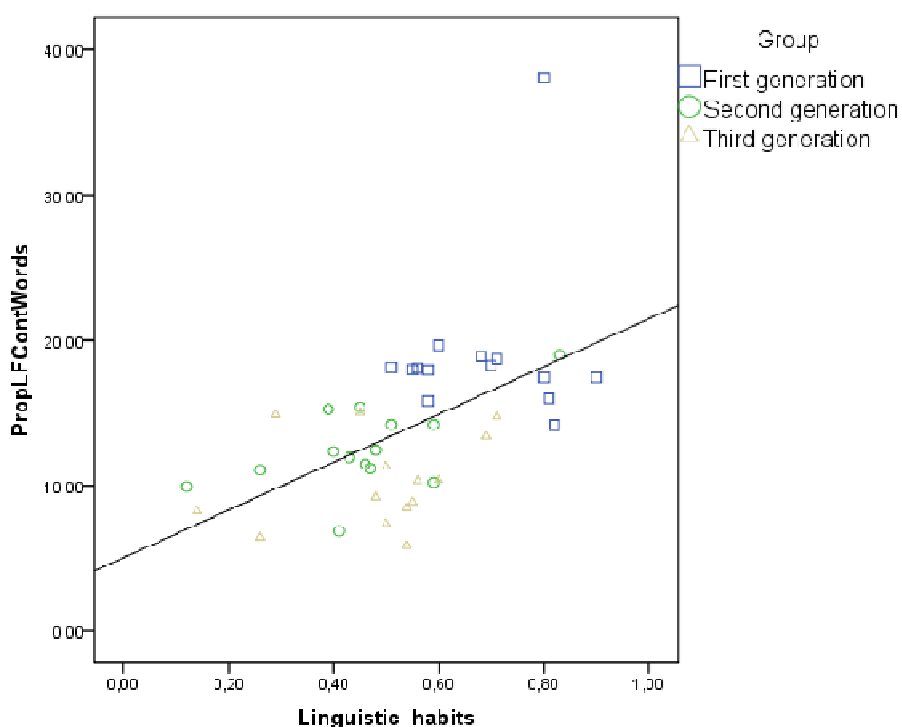


Figure 5.15 Simple correlation between % of Low Frequency Content Words and Linguistic habits for first/second/third generation

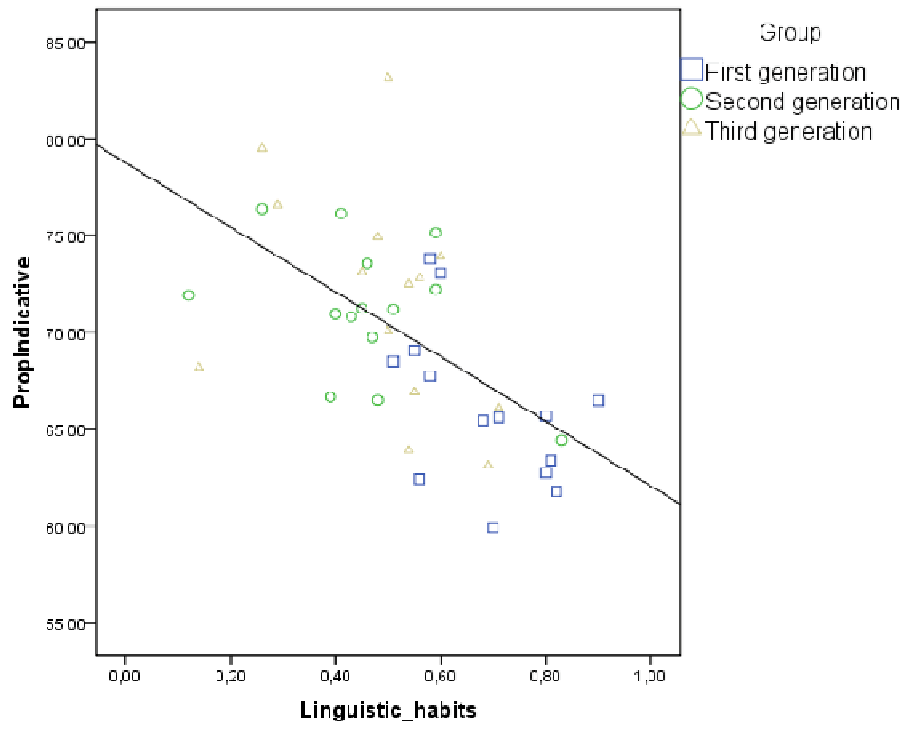


Figure 5.16 Simple correlation between % of Indicatives and Linguistic habits for first/second/third generation

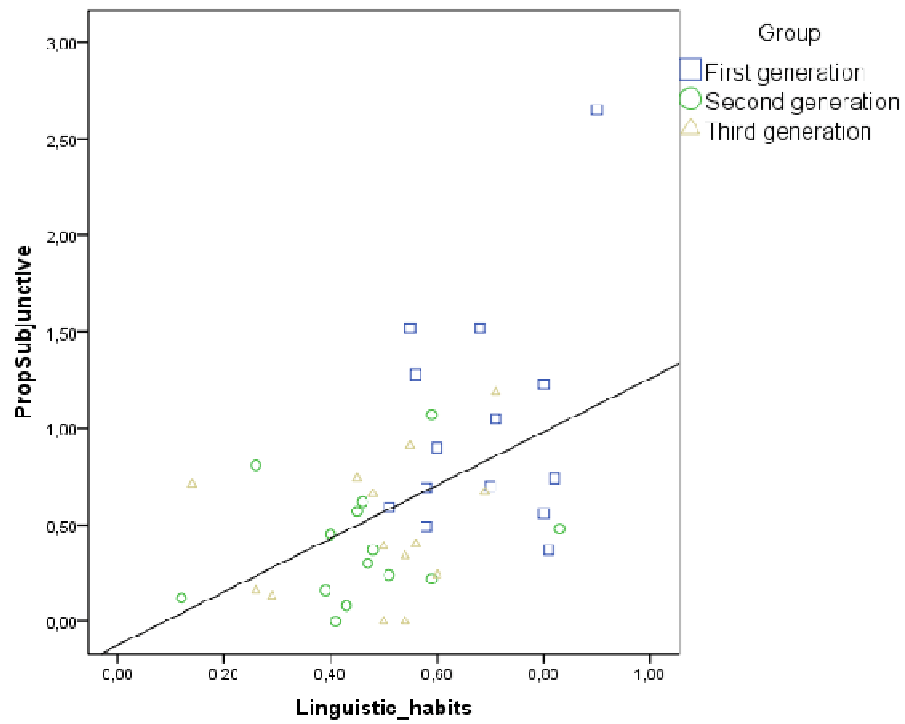


Figure 5.17 Simple correlation between % of Subjunctives and Linguistic habits for first/second/third generation

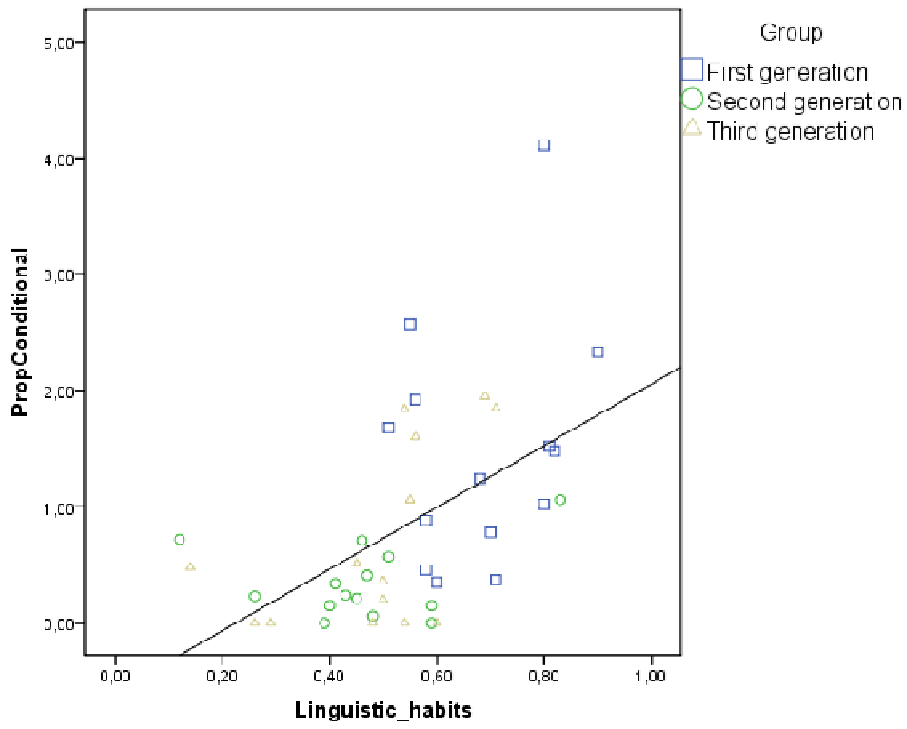


Figure 5.18 Simple correlation between % of Conditionals and Linguistic habits for first/second/third generation

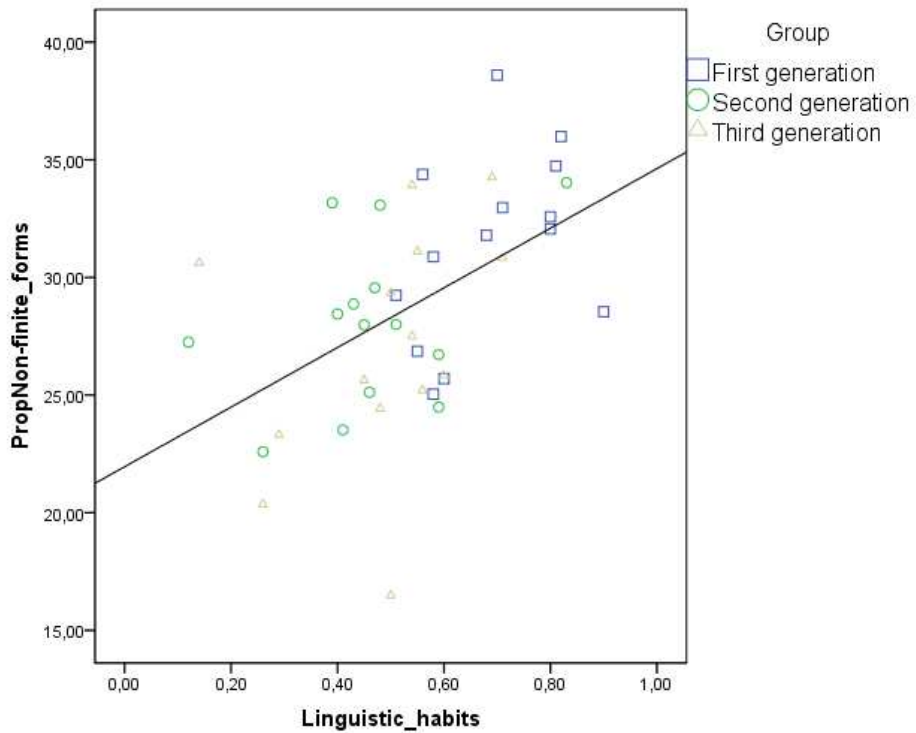


Figure 5.19 Simple correlation between % of Non-finite forms and Linguistic habits for first/second/third generation

As discussed above, if we want to have a real image of the correlation between the variables of interest we need to consider the partial correlation outcomes, and weigh them against the simple ones.

Table 5.18 Comparison of Simple and Partial correlation outcomes for first - second - third generation

		Simple correlation	Partial correlation
% LowFreqContentWords	Pearson Correlation	.537	.349
	p-value	.000	.025
% Indicative	Pearson Correlation	.579	.468
	p-value	.000	.002
% Subjunctive	Pearson Correlation	.481	.347
	p-value	.001	.026
% Conditional	Pearson Correlation	.532	.442
	p-value	.000	.004
% Non-finite_Forms	Pearson Correlation	.499	.389
	p-value	.001	.012

By comparing the correlation that has not been controlled against the one that has, we can gain insight into the contribution of different variables. We firstly notice that these correlations remain statistically significant (having their p-values below .05), although the relationships are somewhat weakened. In particular, the greater the difference of the two outcomes in the correlation, the more we can suggest that the two variables of interest are related to the intervening variable, and therefore that the strength of the observed (simple) correlation is due to their common relationship with the variable of generation. In this case, in particular, the gap for the low frequency words is bigger than those seen for verb tenses. This suggests that the variable of generation is influencing the simple relationship of this variable to a relatively higher degree than it does with verb tenses.

5.6 Conclusion

In this last section I summarise the outcomes of the analysis, reviewing the results of chapters 4 and 5.

In chapter 4 I discussed the extra-linguistic factors (attitude and language habits) impacting on language maintenance/attrition and how they vary across the population investigated. Attitude is a multifaceted variable and it has turned out to be even more so with regard to a very complex reality, such as the Italian one, where the creation of an identity set presents aspects which are incoherent to a certain extent. The linguistic habits showed a higher degree of internal consistency in the answers, which eventually allowed the creation of a compound variable. The result that clearly stood out is the disparity in quantity of language contact with the heritage language (Italian) between the two groups of native speakers on the one hand, and heritage speakers on the other. The first generation, and to a slightly greater extent the new migrants, enjoy quantitatively more contacts with Italian compared to the second and third-generation informants.

This second chapter opened with a descriptive analysis of the results of the interviews, highlighting and discussing language differences between the two groups of native speakers, and examining the expected generational decreasing language proficiency. The discussion was then backed up with a statistical analysis, assessing whether any differences between the two groups of native speakers and the three generations are also statistically significant. A first noteworthy point is the number of these results which have turned out to be higher in the case of the inter-generational analysis.

I eventually aimed at assessing whether these statistical differences were also correlated in some way with the linguistic habits of my informants, a variable that plays an important role with regard to the maintenance/attrition of a heritage language. This analysis was carried out by performing a partial correlation, aimed at studying the relation between each linguistic outcome and the linguistic habits while removing the effects of the variable of generation on the relation of interest.

It is important to remember that we are dealing with a correlation between two variables, and therefore we can not discuss this in terms of cause-and-effect. Nevertheless, and proceeding with caution, we can still suggest a role for the linguistic habits in shaping some of the language outcomes and the differences seen between the two groups of native speakers and among three generations of speakers.

The variable of generation has an impact on the correlation investigated in a relatively high number of variables for which, against a significant result for the single correlation, is opposed a non-correlation when discounting for the variable of generation. However there is room also for correlations between linguistic outcomes and language habits that

remain significant when discounting the intervening variable. In particular, it seems interesting to note the results for the verb tenses, among the three generations. All the four groups (indicative, subjunctive, conditional and non-finite forms) are significantly correlated with the linguistic habits of the informants, which seems to support the hypothesis that morphological aspects may be more dependent on the quantity of contacts with the heritage language compared to linguistic aspects, which in contrast are shown to be relatively highly linked to the variable of generation.

Conclusions

In this last chapter I briefly recap some of the main aspects of the present study, focusing in particular on the analysis discussed in chapters 4 and 5. Before reviewing the outcomes of the analysis, I shall briefly go over the main points of the data collection. The data analyzed and discussed in this piece of work were gathered for the specific purposes of this study during three months of fieldwork in 2009, in the areas of Toronto and Vancouver (see chapter 3). The focus of the data collection was twofold: the language skills and attitudes of three generations of Italo-Canadians were examined along a generational scale, and a more original intra-generational perspective was taken when comparing two different cohorts of Italian migrants, new migrants and the first generation. These groups differed consistently not only with regard to the time elapsed since they had left Italy, but mostly in respect of their educational background and their social interactions with other Italian migrants, as well as their attitudes towards Italy and Canada and the reasons that led them to migrate (see chapter 2). The same nature of these new waves of migration, rather variegated and at times nomadic, particularly if compared to the preceding waves - highlights their particularities but at the same time underlines the difficulties in managing this kind of sociolinguistic research.

Chapter 4 discussed some questions from the data collection through the questionnaires. My interest was in particular two types of factors, which were labelled as 'linguistic habits' and 'attitudinal factors'. With regard to the linguistic habits of my informants, the first point that we notice is a clear division between the two groups of native speakers on one hand, and the two groups of heritage speakers on the other. Quantitatively the native speaker groups are more 'in touch' with the Italian language. This result is in line with what was expected, suggesting that native speakers may enjoy more stimuli to use their native language, ranging from the fact that Italian is their mother tongue (and therefore they should not have any major difficulties of comprehension), to the fact that it usually comes naturally to us to retain an attachment to our motherland and to keep an interest, albeit sometimes a critical one, with one's roots. There is also the fact that this group probably have more close relatives in Italy to talk to and occasionally visit.

Internally, the two groups of native speakers on one hand, and heritage speakers on the other, seem to diverge to a relatively small extent, with new migrants and (in particular) the second generation showing the highest linguistic results within their group. The

result for new migrants may be due to two factors: they usually have more relatives and friends in Italy, both because they left more recently and because they may still have their parents, siblings and other close relatives there. They usually also are more wealthy, and therefore can spend more on Italian magazines, books, etc. We also need to consider the internet, and how this has transformed the ways to re-connect with one's native country when living abroad, widening the opportunities for contact and drastically reducing the cost of this.

With regard to the two groups of heritage speakers, the higher result for the second generation is in line with a decreasing generational attachment to the Italian language (and Italianness in general) as discussed in chapter 2. This is a natural and expected outcome for a heritage community that is going to be integrated almost completely into the Canadian multicultural patchwork, and that is now allowed to negotiate its Italian heritage without considering the use of the language as a necessary prerequisite of affiliation.

The attitudinal factors emerged as particularly difficult to handle. Grouping questions pertaining to this aspect were grouped together as with the linguistic habits, however this time the attempt was unsuccessful. Yet, beyond this somewhat unproductive outcome, a more conscious discussion on the inherent attributes of the concept of attitude seems worth having. The fact that I could not find a thread that would group all the questions together perhaps probably reflects the very nature of this concept. Another possible reason lies in the design of the questionnaire. In this respect, I found the influence of other 'more concrete' variables playing a major role, for example people free from family ties were more inclined to move to Italy. One last reason may lie in the population investigated. During my stay in Canada, I spent a lot of time within the Italian-Venetian community and I came across a certain inconsistency in their reflections and attitudes about their Italianness. The connection that they have developed with Italy, and in particular with what they consider their true motherland and culture, the Veneto region, is a feeling that for some of them has developed into a love-hate relationship; but even people who were more negative in their opinions of Italy and less active within the Italian community, never completely cut off all the connections with their mother/heritage country or with the Italian community in Canada. I may eventually suggest that it is the combined effect of all these factors that has led to some results that are not entirely clear-cut.

Before reviewing the linguistic outcomes of the interviews, I want to recall one point regarding my informants. The linguistic outcomes that I have discussed in this study, and that I am briefly going to review here, refer to the skills of the Italian-speaking community, and therefore they do not aim to be representative of the actual level of proficiency of the Italo-Canadian community. This clarification is intended especially for the two groups of heritage speakers, in particular for the third generation. During my fieldwork I met a great number of people who could potentially have been interviewed, but who were not. However the final number of interviewees was relatively small due to the fact that a substantial number, particularly of the third generation, could not hold a conversation in Italian. The inability to master the heritage language is in fact very common among the third generation of Italian heritage in Canada.

The linguistic outcomes of the interviews do not always produce the same picture. However, it seems safe to propose two general trends. The two groups of native speakers diverge, in the sense that the new migrants have shown better language skills in Italian than the first generation. This may be due to a cluster of factors, but we may suggest a major role for the variable of education. A higher level of education, which has to be interpreted in the broadest sense, can in fact help to prevent or mitigate attrition. People are more predisposed, thanks to their developed language skills, to widen their contacts or the variety of them.

Along a generational scale, my informants' skills in Italian tend to decrease. This fact is not surprising, and to a certain extent it is predictable. However it has not to be taken completely for granted without being subject to scrutiny. Even using a form of selection, where only those who were willing to participate because they were confident enough of their skills were part of the sample, a decrease is observed. Interestingly this decline seems not to proceed in a completely linear fashion, but the gap between the generations appears to reduce.

In this work I have looked in particular at the role played by two extra-linguistic variables in the language abilities of my informants: generation and linguistic habits, that is the amount of contact with the heritage language.

The variable of generation clearly plays a key role. The results of this analysis give further support to the literature discussed in chapter 1, in particular with regard to intergenerational language maintenance. In this study I also took into consideration two cohorts of migrants, both Italian native speakers but diverging in many respects, among others with regard to the span of time spent in Canada and their level of education. The

latter, in particular, we have seen can relate to a relatively higher number of contacts with the heritage language, favouring its maintenance. Despite the divergences between these two groups, the variable of generation has emerged as central.

It is important to remember that the language skills of my informants are placed along a continuum, and therefore that the boundaries between generations partly blur. There have been new migrant informants that scored more closely to some of the first generation (and vice-versa). This also applies to the divisions between the three generations. From a purely linguistic point of view, boundaries are not completely clear-cut.

The variable of linguistic habits seems to play overall a more marginal role compared to generation. However a certain impact of this variable on the linguistic outcomes is observable, in particular in the intergenerational analysis and, with regard to these three groups, especially pertaining to morphology (verb tenses). Although their impact is more marginal, we can nevertheless suggest that linguistic habits play a role in strengthening the effects of the variable generation on language skills.

This thesis has presented various empirical findings and has allowed problematic areas to emerge in parallel to these. I particularly want to highlight what I have referred to as 'new migration'. This label is rather crude, and as discussed in chapter 2, comprises a range of migration waves, which nonetheless seem to share certain sociolinguistic traits: a high level of education accompanied by frequent and direct contact with the L1, particularly via the new media. As suggested throughout this work, little attention has previously been devoted to these new waves. New cohorts of migrants are a phenomenon described by Bauman as extremely fluid, difficult to define and difficult to frame within the classical social variables that are in use in sociolinguistics (Caltabiano & Gianturco, 2005).

From a quantitative perspective, the number of Italian migrants to Canada in the new wave is smaller than that of the preceding cohort. However, more than their limited number, it is their isolation and their lack of involvement in Italian associations that makes them more difficult for researchers to meet. Particularly within Europe, but also elsewhere, their migration is a nomadic experience, with such a web of different variables as to make a sociolinguistic analysis difficult to perform.

Further potential investigations are several and diverse. A quantitative approach is time-consuming, particularly in the transcription phase. It is thus crucial to weigh up the

benefits and costs of such a choice. However, the time is now right for such an approach: the mainly qualitative perspective adopted so far by many researchers regarding Italian communities abroad could be supported by a more quantitative analysis, although, as we have also seen in this piece of work, this is not always straightforward. The intricate web of social and linguistic variables, which will doubtless become even more complex in the future, makes sampling particularly delicate. Moreover, research of this kind is difficult to develop fully in a single study, so joint projects would ideally become the basis of future research in this field.

As suggested in the introduction to this work, this analysis has revisited a well-beaten path, while adding elements of novelty. In particular I have worked both on developing a linguistic analysis along a generational scale, including three generations of speakers, and I have explored the relatively new phenomenon – at least from a linguistic perspective – of the new waves of Italian migration. These last have been linked to and investigated in relation to the well-studied Post-1945 cohorts.

New waves of migrants from Italy have exhibited a higher level of education, and in general more opportunities to use their heritage language, thanks to a higher standard of living and more opportunities for contacts with the Italian language, such as more frequent travel to Italy or the use of the internet. Finally, this study has shown how they emerged as displaying higher language skills compared to the preceding cohort.

Notwithstanding these promising features, the future of Italian as a heritage language in Canada has not to be taken for granted. Better inputs from their (grand)parents and more opportunities of contacts with the heritage language seem to suggest a brighter future. New appealing images of Italy are not always conducive to language maintenance and they may on the contrary lead the community to lack that ‘self-protection system’ that makes it more close-knit on a societal level and therefore more linguistically conservative.

The future for the young generations of Italo-Canadians is still uncertain. New research questions and analysis on what it will mean to be an Italian heritage language speaker will be addressed in future research.

Appendices

Appendix 1: Questionnaire in Italian

Questionario

La modalità di compilazione di questo questionario prevede che possa essere segnata solo una casella. Altri casi verranno specificati con una nota a fianco alla domanda.

1. Sei stato il primo della tua famiglia ad emigrare in Canada?

- No (vai alla domanda 1.1)
- Sì (vai alla domanda 2)

1.1. Se no, chi è emigrato per primo? (indicare l'anno)

- Bisnonni (_____)
- Nonni (_____)
- Genitori (_____)
- Zii/e (_____)
- Fratelli/Sorelle (_____)
- Altri parenti (specificare) _____ (_____)

2. Dove vivono gli altri componenti della tua famiglia?

	ITALIA	CANADA	ALTRO
Nonni paterni			
Nonni materni			
Genitori			
Fratelli/Sorelle			
Figli/e			
Altri parenti			

3. In quale anno sei emigrato? (indicare l'anno) _____

4. Per quale ragione principale hai scelto di vivere in Canada? (barrare solo una casella)

- Volevo ricongiungermi con i miei familiari o fidanzata/o
- Per motivi di lavoro/studio
- Per interesse verso il Canada
- È il Paese dove sono nato (vai alla domanda 7)
- Altro (specificare) _____

5. Quando sei partito pensavi di restare in Canada?

- No
- Sì

5.1 Se sì, per quanti anni? _____

6. Con chi sei partito? (puoi barrare più di una casella)

- Da solo
- Genitori
- Fratelli/Sorelle
- Marito/Moglie
- Fidanzata/o, Ragazzo/a
- Parenti
- Amici
- Compaesani
- Conoscenti
- Altro (specificare) _____

7. Chi si trovava già in Canada? (puoi barrare più di una casella)

- Nessuno
- Genitori
- Fratelli/Sorelle
- Marito/Moglie
- Fidanzata/o, Ragazzo/a
- Parenti
- Amici
- Compaesani
- Conoscenti
- Altro (specificare) _____

8. Ti senti : (barrare solo una casella)

- Canadese
- Italiano/a
- Veneto
- Altro (specificare) _____

9. Pensi di rientrare/andare a vivere in Italia?

- No (*se la risposta è No, salta alla domanda 10*)
- Sì

9.1. Se sì, vorresti tornare nel paese d'origine della tua famiglia?

- Sì, nello stesso paese
- Sì, nella stessa provincia
- Sì, nella stessa regione
- No, in una diversa regione
- Indifferente

9.2. Se sì, quando vorresti tornare in Italia?

- Prima dell'inizio dell'età lavorativa
- In età lavorativa
- In pensione

10. Trascorri parte dell'anno in Italia?

- No (*se la risposta è No, salta alla domanda 11*)
- Sì

10.1 Se sì, per quanto tempo? (specificare) _____

10.2 Per quale motivo? (puoi barrare più di una casella)

- Visita a parenti
- Visita ad amici
- Turismo
- Lavoro
- Altro (specificare) _____

10.3 Con chi vai in Italia?

- Parenti
- Amici italiani
- Amici canadesi
- Altro (specificare) _____

10.4 Con chi trascorri il tuo tempo quando sei in Italia? (puoi barrare più di una casella)

- Parenti
- Amici
- Compaesani
- Associazioni
- Gruppi parrocchiali
- Altro (specificare) _____

10.5 Quando sei in Italia ti trovi in situazioni di difficoltà?

- Mai (*se la risposta è No, salta alla domanda 11*)
- Qualche volta
- Spesso

10.5.1 Se ti trovi in difficoltà, per quali motivi? (puoi barrare più di una casella)

- Difficoltà linguistiche
- Diverse abitudini di vita
- Difficili relazioni con le persone
- Altro (specificare) _____

11. In Canada fai parte di qualche associazione italiana?

- No (*se la risposta è No, salta alla domanda 12*)
- Sì

11.1 Se sì, di quale tipo? (puoi barrare più di una casella)

- Assistenziale/Sanitaria
- Culturale
- Patriottica
- Regionale
- Religiosa
- Ricreativa
- Sportiva
- Turistica
- Altro (specificare) _____

12. Chi frequenti principalmente nel tempo libero? (barrare solo una casella)

- Canadesi
- Italiani di origine veneta
- Italiani originari della tua stessa provincia veneta
- Italiani di altre provenienze regionali
- Altro (specificare) _____

13. In Canada il fatto di essere italiano/origini italiane ti ha messo in situazioni di difficoltà?

- Mai (*se la risposta è Mai, salta alla domanda 14*)
- Qualche volta
- Spesso

13.1 Se ti trovi in difficoltà per il fatto di essere italiano, in quale ambito prevalentemente ti capita? (barrare solo una casella)

- Familiare
- Sociale
- Lavorativo
- Altro (specificare) _____

13.2 Se ti trovi in difficoltà, per quali motivi? (puoi barrare più di una casella)

- Difficoltà linguistiche
- Diverse abitudini di vita
- Difficili relazioni con le persone
- Altro (specificare) _____

14. Ti consideri ancora un emigrato/figlio/nipote di emigrati?

- No
- Sì

15. In Canada l'immagine dell'italiano e dell'Italia è :

- Sostanzialmente positiva
- Sostanzialmente negativa
- Migliorata rispetto al passato

15.1. Se migliorata, perché _____

16. Quali sono gli aspetti che ami dell'Italia? (puoi barrare più di una casella)

- Storia/Arte/Musica
- Cucina
- Stile di vita
- Calcio
- Lingua italiana
- Dialetto
- Attaccamento alla famiglia
- Tradizioni regionali
- Altro (specificare) _____

17. Quali sono gli aspetti dell'Italia che non ti piacciono? (specificare)

18. Secondo te i veneti sono diversi dagli emigranti delle altre regioni?

- No
- Sì

18.1 Se sì, perché _____

19. Trovi i veneti più integrati rispetto agli italiani di altra provenienza?

- No
- Sì

19.1 Se sì, quale è il motivo? _____

20. Quali lingue conosci?

21. Se conosci la lingua italiana è perché: (puoi barrare più di una casella)

- è la tua lingua madre
- la parli in famiglia
- la parli con gli amici
- ami l'Italia
- frequenti associazioni italiane
- la usi al lavoro
- l'hai studiata
- Altro (specificare) _____

22. Conosci il dialetto?

- No
- Sì

23. Quanto parli le seguenti lingue?

	Mai	Raramente	Qualche volta	Spesso	Molto spesso
Dialetto					
Italiano					
Inglese					

24. Che cosa rappresenta per te l'italiano? (barrare solo una casella)

- Affetti familiari
- Radici
- Cultura
- Strumento di comunicazione
- Strumento di lavoro
- Altro (specificare) _____

30. Lingue lette:

	per niente				molto
	1	2	3	4	5
italiano	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
inglese	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dialetto	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
altra lingua (specif.) _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
altra lingua (specif.) _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

31. Lingue scritte:

	per niente				molto
	1	2	3	4	5
italiano	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
inglese	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dialetto	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
altra lingua (specif.) _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
altra lingua (specif.) _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

32. Quanto spesso e come sei in contatto con le persone che risiedono in Italia?

Genitori

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)
Telefonate					
Biglietti/cartoline					
Lettere					
E-mail					

Fratelli/Sorelle

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)
Telefonate					
Biglietti/cartoline					
Lettere					
E-mail					

Figli/e

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)i
Telefonate					
Biglietti/cartoline					
Lettere					
E-mail					

Nipoti

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)
Telefonate					
Biglietti/cartoline					
Lettere					
E-mail					

Altri parenti

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)
Telefonate					
Biglietti/cartoline					
Lettere					
E-mail					

Amici

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)
Telefonate					
Biglietti/cartoline					
Lettere					
E-mail					

Altri (specificare _____)

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)
Telefonate					
Biglietti/cartoline					
Lettere					
E-mail					

33. Con quale frequenza ascolti/vedi radio e televisione canadesi?

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)	Tutti i giorni
Radio						
Televisione						

34. Con quale frequenza ascolti/vedi radio e televisione italiane?

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)	Tutti i giorni
Radio						
Televisione						

35. Quali sono i tuoi programmi preferiti? (nome della trasmissione, argomento, lingua):

televisione: _____

radio: _____

36. Con quale frequenza leggi in italiano e in inglese?

	Mai	Raramente (2/3 volte all'anno)	Qualche volta (6 volte l'anno)	Spesso (1 volta al mese)	Molto spesso (1 volta alla settimana)	Tutti i giorni
inglese						
italiano						

36.1 Se leggi in italiano, leggi prevalentemente

- Quotidiani
 Riviste
 Libri
 Altro (specificare)_____

36.2 Se leggi in italiano, i temi prevalenti sono

- Attualità
 Cultura
 Sport
 Narrativa
 Altro (specificare)_____

36.3 Se leggi in inglese, leggi prevalentemente

- Quotidiani
 Riviste
 Libri
 Altro (specificare)_____

36.4 Se leggi in inglese, i temi prevalenti sono

- Attualità
 Cultura
 Sport
 Narrativa
 Altro (specificare)_____

37. Quali lingua usi con le persone che frequenti?

	Solo Inglese	Solo Italiano	Solo Dialetto	Inglese/italiano	Inglese/dialetto	Italiano/dialetto	Italiano/dialetto/inglese
Moglie/Marito Compagna/o							
Figli/e							
Genitori							
Nonni							
Colleghi di lavoro/scuola							
Amici nel tempo libero							
Negozianti							

38. Con quale lingua si rivolgono a te le persone che frequenti?

	Solo Inglese	Solo Italiano	Solo Dialetto	Inglese/italiano	Inglese/dialetto	Italiano/dialetto	Italiano/dialetto/inglese
Moglie/Marito Compagna/o							
Figli/e							
Genitori							
Nonni							
Colleghi di lavoro/scuola							
Amici nel tempo libero							
Negozianti							

39. Quando parli una lingua, ti capita di inserire parole di altre lingue?

- No
 Sì

39.1. Se sì:

	Inserisci inglese	Inserisci italiano	Inserisci dialetto
Quando parli inglese			
Quando parli italiano			
Quando parli dialetto			

40. Sesso

- Maschio
 Femmina

41. Luogo di nascita

Canada (indicare località e provincia) _____

Italia (indicare comune e provincia) _____

Altro (indicare stato e anno trasferimento in Canada) _____

42. Anno di nascita _____

43. Luogo di residenza attuale (indicare località e provincia) _____

44. Altro/i luogo/i di residenza in Canada (specificare gli anni di permanenza)

45. Dove hai studiato? (se scuola superiore e/o università specificare quale)

	ITALIA	CANADA	SCUOLA ITALIANA IN CANADA	ALTRO
Asilo Nursery school				
Scuola elementare / Elementary school				
Scuola media / Middle school				
Scuola superiore/ High school				
Corsi post maturità/ College				
Università / University				

46. Hai seguito corsi di lingue?

- Inglese per quanti anni? _____
- Italiano per quanti anni? _____

47. Cittadinanza

- Canadese
- Italiana
- Entrambe
- Altra (specificare) _____

47.1 Se hai solo una delle due cittadinanze, vorresti acquisire anche l'altra?

- No
- Sì

48. Stato civile

- Celibe/Nubile
- Convivente
- Sposato/a
- Divorziato/a
- Vedovo/a

48.1 Se sei o sei stato/a sposato/a, tua moglie/marito/compagno/a è/era:

- Veneto/a
 Italiano/a
 Canadese
 Altro (specificare) _____

49. Hai figli?

- No
 Sì

49.1 Se sì quanti e di che età? (specificare) _____

50. Qual è attualmente la tua attività?

- Lavoro (*se la risposta è Lavoro, salta alle note finali del questionario*)
 Studio
 Altro (specificare) _____

51. In Italia lavoravi?

- No (*se la risposta è No, salta alla domanda 52*)
 Sì

51.1 Se sì, in quale settore svolgevi la tua attività lavorativa?

- Agricolo
 Industriale
 Commercio
 Impiego pubblico
 Servizi
 Altro (specificare) _____

51.2 Se sì, la tua occupazione era di quale tipo?

- Imprenditore
 Libero professionista
 Lavoro autonomo
 Dipendente
 Altro (specificare) _____

52. Attualmente in Canada lavori?

- No *(se la risposta è No, salta alle note finali del questionario)*
- Sì

52.1 Se sì, in quale settore svolgi la tua attività lavorativa?

- Agricolo
- Industriale
- Commercio
- Impiego pubblico
- Servizi
- Altro (specificare) _____

52.2 Se sì, la tua occupazione è di quale tipo?

- Imprenditore
- Libero professionista
- Lavoro autonomo
- Dipendente
- Altro (specificare) _____

52.3 Lavori con qualche membro della tua famiglia? (se sì, specificare chi)

52.4 Con quante persone lavori e quale è la loro nazionalità?

Nazionalità _____ numero _____
Nazionalità _____ numero _____
Nazionalità _____ numero _____
Nazionalità _____ numero _____

Conclusione

Nell'eventualità di aver bisogno di contattarti per chiarire alcuni punti relativi al questionario, ti sarei grata se volessi lasciarmi un recapito (indirizzo e/o indirizzo e-mail e/o numero di telefono), dove ti possa raggiungere:

Indirizzo _____

Indirizzo e-mail _____

Numero di telefono _____

Ti pregherei inoltre di indicarmi se sei interessato/a a collaborare ad un eventuale prosieguito di questo studio:

SÌ NO

Per qualsiasi informazione puoi contattarmi a uno dei seguenti indirizzi:

Claudia Bortolato
Room 242, Department of Modern Languages - Italian Unit
University of Exeter
The Queen's Drive
Exeter, Devon, EX4 4QH
England

Email: XXX

Tel.: XXX

Ti ringrazio per la disponibilità e l'attenzione che hai dedicato a questo progetto di ricerca.

Claudia Bortolato

Protezione dei dati personali

Questo questionario si inserisce nell'ambito di un progetto di ricerca riguardante lo studio delle comunità venete in Canada. Il fenomeno dell'emigrazione italiana all'estero è stato un ambito studiato da diverse prospettive e in differenti contesti geografici. Ricerche sulle caratteristiche delle comunità italiane in Canada sono state già condotte da alcuni studiosi. Sono però necessari studi più approfonditi per indagare alcuni aspetti tralasciati e per studiarne il divenire culturale, sociale e linguistico.

La raccolta di dati in Canada verrà seguita da Claudia Bortolato, dottoranda all'University of Exeter, attraverso la somministrazione di un questionario ed eventualmente di una intervista in profondità. Ti chiediamo la disponibilità a partecipare al progetto, ritenendo molto importante la tua esperienza personale e competenza.

In conformità con la legge sulla privacy, ti assicuriamo che il questionario e l'eventuale intervista verranno trattati in forma anonima e confidenziale, e che verranno usati solo per i fini di questo studio e per eventuali altre ricerche che da quest'ultimo potrebbero scaturire.

_____, ____/____/2009

Firma

Appendix 2: Questionnaire in English

Questionnaire

On completing the questionnaire it is necessary to tick one box at a time. A different option will be specified with a note next to the question.

1. Were you the first in your family to emigrate to Canada?

- No (go to question 1.1)
- Yes (go to question 2)

1.1. If no, who did emigrate first? (write the year)

- Great grandparents (_____)
- Grandparents (_____)
- Parents (_____)
- Uncle/Aunts (_____)
- Siblings (_____)
- Other relatives (specify) _____ (_____)

2. Where do the rest of your relatives live?

	ITALY	CANADA	OTHER
Paternal grandparents			
Maternal grandparents			
Parents			
Siblings			
Sons/Daughters			
Other relatives			

3. In which year did you emigrate? (write the year) _____

4. For which main reason did you choose to live in Canada? (you can tick only one box)

- I wanted to join my relatives or partner
- For work/study
- For interest towards Canada
- It is the country where I was born (go to question 7)
- Other (specify) _____

5. When you left Italy, did you think you would stay in Canada?

- No
- Yes

5.1 If yes, how long (years)? _____

6. Who did you leave with? (you can tick more than one box)

- Alone
- Parents
- Siblings
- Husband/Wife
- Fiancé/e, Boyfriend/Girlfriend
- Relatives
- Friends
- Fellow countrymen/women
- Acquaintances
- Others (specify) _____

7. Who was already in Canada? (you can tick more than one box)

- Nobody
- Parents
- Siblings
- Husband/Wife
- Fiancé/e, Boyfriend/Girlfriend
- Relatives
- Friends
- Fellow countrymen/women
- Acquaintances
- Others (specify) _____

8. Do you feel: (you can tick only one box)

- Canadian
- Italian
- Veneto
- Other (specify) _____

9. Do you think you will go back/go to live in Italy?

- No (*if the answer is No, go to question 10*)
- Yes

9.1. If yes, would you want to go back to your family's town of origin?

- Yes, to the same town/city
- Yes, to the same province
- Yes, to the same region
- No, to a different region
- Indifferent

9.2. If yes, when would you like to go to Italy?

- Before working age
- During working age
- Retirement

10. Do you spend part of the year in Italy?

- No (*if the answer is No, go to question 11*)
- Yes

10.1 If yes, how long? (specify)_____

10.2 What's the reason? (you can tick more than a box)

- Visit relatives
- Visit friends
- Tourism
- Work
- Other (specify) _____

10.3 Who do you go to Italy with?

- Relatives
- Italian friends
- Canadian friends
- Other (specify)_____

10.4 Who do you spend your time with when you are in Italy? (you can tick more than a box)

- Relatives
- Friends
- Fellow countryman/woman
- Associations
- Parish groups
- Other (specify) _____

10.5 When you are in Italy, do you ever find yourself in difficult situations?

- Never (*if the answer is No, go to question 11*)
- Sometimes
- Often

10.5.1 If you find yourself in difficulties, what are the reasons? (you can tick more than a box)

- Linguistic difficulties
- Different habits
- Difficult relations with people
- Other (specify) _____

11. In Canada are you a member of any Italian association?

- No (*if the answer is No, go to question 12*)
- Yes

11.1 If yes, of which kind? (you can tick more than a box)

- Welfare/Medical
- Cultural
- Patriotic
- Regional
- Religious
- Recreational
- Sporting
- Tourist
- Other (specify) _____

12. Who do you mainly associate with in your spare time? (you can tick only one box)

- Canadians
- Italians of Venetian origin
- Italians of your same Venetian province
- Italians of other regional origins
- Other (specify) _____

13. In Canada has being Italian/having an Italian heritage forced you into difficult situations?

- Never (*if you have answered Never go to question 14*)
- Sometimes
- Often

13.1 If you are having/have been in difficulties for being Italian, in which sphere does this mainly happen to you? (you can tick only one box)

- Family
- Social
- Working
- Other (specify) _____

13.2 What are the reasons? (you can tick more than a box)

- Linguistic difficulties
- Different habits
- Difficult relations with people
- Other (specify) _____

14. Do you still feel like an emigrant/son/grandson of emigrants?

- No
- Yes

15. In Canada the image of Italians and of Italy is:

- Essentially positive
- Essentially negative
- Improved, if compared to the past

15.1. If improved, why: _____

16. Which are the aspects that you love most about Italy? (you can tick more than one box)

- History/Arts/Music
- Cooking
- Lifestyle
- Soccer
- Italian language
- Dialect
- Bond to the family
- Regional customs
- Other (specify) _____

17. Which are the aspects of Italy that you don't like? (specify)

18. In your opinion, are Veneti different from other regions' emigrants?

- No
- Yes

18.1 If yes, why: _____

19. Do you think that Veneti are more integrated than Italians of a different origin?

- No
- Yes

19.1 If yes, what's the reason? _____

20. Which languages do you know?

21. If you know Italian language, how: (you can tick more than a box)

- it's your mother language
- it's spoken in your family
- you speak it with your friends
- you love Italy
- you attend Italian associations
- you use it at work
- you have studied it
- other (specify) _____

22. Do you know Italian?

- No
- Yes

23. Do you know Dialect?

- No
- Yes

24. How much do you speak these languages?

	Never	Seldom	Sometimes	Often	Very often
Dialect					
Italian					
English					

25. What does Italian mean to you? (you can tick only one box)

- Family ties
- Roots
- Culture
- Communication instrument
- Work instrument
- Other (specify) _____

31. Read languages:

	Not really				Very much
	1	2	3	4	5
Italian	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dialect	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other language (specif.) _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other language (specif.) _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

32. Written languages:

	Not really				Very much
	1	2	3	4	5
Italian	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
English	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Dialect	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other language (specif.) _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Other language (specif.) _____	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

33. How often and how are you in touch with people living in Italy?

Parents

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)
Phone calls					
Cards/postcards					
Letters					
E-mails					

Siblings

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)
Phone calls					
Cards/postcards					
Letters					
E-mails					

Sons/Daughters

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)
Phone calls					
Cards/postcards					
Letters					
E-mail					

Grandsons/daughters

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)
Phone calls					
Cards/postcards					
Letters					
E-mail					

Other relatives

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)
Phone calls					
Cards/postcards					
Letters					
E-mail					

Friends

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)
Phone calls					
Cards/postcards					
Letters					
E-mail					

Others (specify) _____

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)
Phone calls					
Cards/postcards					
Letters					
E-mail					

34. How often do you listen to/watch Canadian radio/TV?

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)	Everyday
Radio						
Television						

35. How often do you listen to/watch Italian radio/ TV?

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)	Everyday
Radio						
Television						

36. Which are your favourite TV programmes? (name of the programme, topic and language):

television: _____

radio: _____

37. How often do you read in Italian and English?

	Never	Seldom (2/3 times a year)	Sometimes (6 times a year)	Often (once a month)	Very often (once a week)	Everyday
Italian						
English						

37.1 If you read in Italian, what do you read mainly?

- Newspapers
- Magazines
- Books
- Other (specify)_____

37.2 If you read in Italian, the main topics are:

- Current events
- Culture
- Sport
- Fiction/narrative
- Other (specify)_____

37.3 If you read in English, what do you read mainly?

- Newspapers
- Magazines
- Books
- Other (specify)_____

37.4 If you read in English, the main topics are:

- Current events
- Culture
- Sport
- Fiction/narrative
- Other (specify)_____

38. What language do you speak with people that you see often?

	Only English	Only Italian	Only Dialect	English/ Italian	English/ Dialect	Italian/ Dialect	Italian/Dialect/ English
Wife/Husband Partner							
Sons/Daughters							
Parents							
Grandparents							
Colleagues/ Classmates							
Friends in free time							
Shopkeepers							

39. In which language do the people that you see often speak to you?

	Only English	Only Italian	Only Dialect	English/ Italian	English/ Dialect	Italian/ Dialect	Italian/Dialect/ English
Wife/Husband Partner							
Sons/ Daughters							
Parents							
Grandparents							
Colleagues/ Classmates							
Friends in free time							
Shopkeepers							

40. When you speak, do you to insert words from other languages?

- No
 Yes

40.1. If yes:

	Insert English	Insert Italian	Insert Dialect
When you speak in English			
When you speak in Italian			
When you speak in Dialect			

41. Sex

- Male
 Female

42. Place of birth

Canada (place and the province) _____

Italy (place and the province) _____

Other (place and the year you moved to Canada) _____

43. Year of birth _____

44. Place of current home residence (place and the province) _____

45. Other(s) place(s) of residence in Canada (specify for how many years)

46. Where have you studied? (if college and/or university specify what kind)

	ITALY	CANADA	ITALIAN SCHOOL IN CANADA	OTHER
Asilo Nursery school				
Scuola elementare / Elementary school				
Scuola media / Middle school				
Scuola superiore/ High school				
Corsi post maturità/ College				
Università / University				

47. Have you attended language courses?

- English for how long? _____
- Italian for how long? _____

48. Citizenship

- Canadian
- Italian
- Dual nationality
- Other (specify) _____

48.1 If you have only one of the two, would you like to gain the other?

- No
- Yes

49. Marital status

- Unmarried
- Co-habiting
- Married
- Divorced
- Widow(er)

49.1 If you are/have been married, is/was your wife/husband/fiancé/e:

- Veneto/a
- Italian
- Canadian
- Other (specify) _____

50. Do you have children?

- No
- Yes

50.1 If yes, how many and of what age? (specify) _____

51. What's your current state of employment?

- Study (*if you answer Study, go to the end of the questionnaire*)
- Work

52. Did you work while you were in Italy?

- No (*if you answer No, go to the question 53*)
- Yes

52.1 If yes, in which sector did you work?

- Agricultural
- Industrial
- Trade
- Public work
- Services
- Other (specify) _____

52.2 If yes, what type of employment?

- Entrepreneur
- Free lancer
- Self employer
- Employee
- Other (specify) _____

53. Do you currently work in Canada?

- No (*if you answer No, go to the Conclusion of the questionnaire*)
- Yes

53.1 If yes, in which sector do you work?

- Agricultural
- Industrial
- Trade
- Public work
- Services
- Other (specify) _____

53.2 If yes, which kind of employment?

- Entrepreneur
- Free lancer
- Self employer
- Employee
- Other (specify)_____

53.3 Do you work with anyone from your family? (if yes, specify who)

53.4 With how many people and of which nationality do you work?

Nationality_____ number_____

Nationality_____ number_____

Nationality_____ number_____

Nationality_____ number_____

Conclusion

If the occasion arises that I need to clarify some of the data collected in the questionnaire, it would be very much appreciated if you could leave me your contact details (address and/or email address and/or telephone number) so that I can contact you:

Address _____

Email address _____

Telephone number _____

Please indicate whether you would be interested in helping with any further stages of this study:

YES

NO

For any further information please do not hesitate to contact me at the following addresses:

Claudia Bortolato
Room 242, Department of Modern Languages - Italian Unit
University of Exeter
The Queen's Drive
Exeter, Devon, EX4 4QH
England
Email : XXX
Tel. : XXX

Thank you very much for your help, it has been greatly appreciated.

Claudia Bortolato

Personal data protection

This questionnaire is part of a research project about Venetian communities in Canada. The phenomenon of Italian emigration has been a field studied from different perspectives and in different geographical contexts.

Research about Italian communities in Canada has already been carried out, however more studies are necessary in order to analyse further aspects and investigate their cultural, social and linguistic development.

The data collection in Canada will be carried out by Claudia Bortolato, PhD student at the University of Exeter, through the circulation of a questionnaire and an eventual interview. I ask for your availability to participate in the project, valuing your personal experience and competence as very important.

In accordance with privacy law, I assure you that this questionnaire and the possible interview will be treated anonymously and confidentially, and that the data collection will be used only for the aims of this research and for possible developments that could arise from this study.

_____, ____/____/2009

Signature

Bibliography

- Ammerlaan, T. (1996) "You get a bit wobbly." *Exploring bilingual lexical retrieval processes in the context of first language attrition*. Ph.D. dissertation, Katholieke Universiteit Nijmegen.
- Ammon, U. (2011) A checklist of Sociolinguistic Language Maintenance Indicators for Diaspora Minorities (with a focus on German examples). In Moretti B., Pandolfi E. M. and Casoni M. (eds.) *Vitalità di una lingua minoritaria. Aperti e prospettive metodologiche / Vitality of a minority language. Aspects and methodological issues: Proceeding of the conference*. Bellinzona: Osservatorio Linguistico della Svizzera italiana, pp.43-64.
- Andersen, R. W. (1982) Determining the linguistic attributes of language attrition. In R. D. Lambert & B. F. Freed (eds.) *The loss of language skills*. Rowley, MA: Newbury House, pp.83-118.
- Appel, R. & Muysken, P. (2006) *Language contact and bilingualism*. Amsterdam: Amsterdam University Press.
- Auer, P. (1991) Italian in Toronto: a preliminary comparative study on language use and language maintenance. *Multilingua*, 10 (4) pp.403-440.
- Baldassar, L. & Pesman, R. (2004) *I Veneti in Australia: sfide di storia contemporanea*. Padova: ANEA.
- Bazzanella, C. (1994) *Le facce del parlare. Un approccio pragmatico all'italiano parlato*. Firenze: La Nuova Italia.
- Bellù F., Perché gli italo-americani non parlano italiano? (www.redacon.radionova.it, accessed on 25.05.2009)
- Benmamoun E., Montrul S. & Polinsky M. (2010) *White Paper: Prolegomena to Heritage Linguistics*. Harvard: Harvard University.
- Bernardi, L. (ed.) (2005) *Percorsi di ricerca sociale*. Roma: Carocci.
- Berruto, G. (1995) *Fondamenti di sociolinguistica*. Roma/Bari: Laterza.
- Berruto, G & Bescotti, K. (1995) Sulla complessità/semplificata sintattica dell'italiano parlato. *Studi Italiani di Linguistica Teorica e Applicata*, 24(3) pp.461-478.
- Bettoni, C. (1986) Italian language attrition in Sydney: the role of birth order. In C. Bettoni (ed.) *Altro Polo - Italian abroad: Studies on language contact in English-speaking countries*. Sydney: University of Sydney Press, pp.61-85.
- Bettoni, C. (1991) L'Italiano in Australia: tra innovazione e purismo. In I. Baldelli & B.M. Da Rif (eds.) *Lingua e letteratura italiana nel mondo oggi. XIII Congresso internazionale dell'Associazione Internazionale per gli Studi di Lingua e Letteratura Italiana*. Firenze: Olschki Vol.2, pp. 381-394.
- Bettoni, C. (2007) Un profilo sociolinguistico della comunità Italo-australiana. In A. Ciliberti (ed.) *La costruzione interazionale di identità*. Milano: Franco Angeli, pp.37-67.
- Bettoni, C. & Gibbson, J. (1988) Linguistic purism and language shift: a guise-voice study of the Italian community in Sydney. *International Journal of the Sociology of Language*, 72(1) pp.15-35.
- Bettoni, C. & Rubino, A. (1995) Lingua e dialetto a confronto tra gli italiani di Sydney (Australia). In M.T. Romanello & I. Tempesta (eds.) *Dialetti e lingue nazionali: Proceedings of the XXVII Congresso della società di linguistica Italiana*. Roma: Bulzoni, pp. 361-395.

- Bettoni, C. & Rubino, A. (1996) *Emigrazione e comportamento linguistico: un'indagine sul trilinguismo dei siciliani e dei veneti in Australia*. Galatina: Congedo.
- Bolasco, S. (2010) *TaLTaC 2.10*. Milano: Led.
- Bolger, P. A. & Zapata, G. C. (2011) Psycholinguistic Approaches to Language Processing in Heritage Speakers. *Heritage Language Journal*, 8(1) <http://www.international.ucla.edu/media/files/bolg-zap-eight-one-g3-1dr.pdf> (retrieved on 28th of October 2013).
- de Bot, K. (2004) Special Issue on language attrition. *International Journal of Bilingualism*, 8(3) pp.233–237.
- de Bot, K. (2007) Dynamic systems theory, lifespan development and language attrition. In B. Köpcke, M.S. Schmid, M. Keijzer & S. Dostert (eds.) *Language attrition: Theoretical Perspectives*. Amsterdam/Philadelphia: John Benjamins, pp.53-68.
- de Bot K., Gommans P. & Rossing C. (1991) L1 loss in an L2 environment: Dutch immigrants in France. In H.W. Seliger & R.M. Vago (eds.) *First Language Attrition*. Cambridge: Cambridge University Press, pp. 87-98.
- Boyd, S. (1986) Using the Present to predict the Future in Language contact: the case of immigrant minority languages in Sweden. In B. Weltens, K. de Bot & T. van Els (eds.) *Language attrition in progress*. Dordrecht: Foris Publications, pp.99-115.
- Boyd, S. (2001) A minority language as mother tongue or father tongue. Does it make a difference? In T. Ammerlaan, M. Hulsen, H. Strating & K. Yağmur (eds.) *Sociolinguistic and psycholinguistic perspectives on maintenance and loss of minority languages*. Münster: Waxmann, pp.33-45.
- Brace N., Kemp R. & Snelgar R. (2009) *SPSS for psychologists*. Houndmills: Palgrave Macmillian.
- Caltabiano, C. & Gianturco, G. (eds.) (2005) *Giovani oltre confine*. Roma: Carocci.
- Campanale, L. (2006) *I gelatieri veneti in Germania*. Frankfurt am Main: Peter Lang.
- Carlioni, F. (2005) La legge di Zipf sul numero dei significati in Italiano e Inglese. In I. Chiari & T. De Mauro (eds.) *Parole e numeri. Analisi quantitative dei fatti di lingua*. Roma: Aracne, pp.355-372.
- Caruso, M. (2004) Attrition in the verb system of Italian in Australia. *Australian review of applied linguistics*, (18) pp.9-24.
- Caruso, M. (2010) *Italian language attrition in Australia. The verb system*. Milano: Franco Angeli.
- Caselli, M. (2005) *Indagare col questionario*. Milano: Vita & Pensiero.
- Chang, C B., Yao Y., Haynes E. F. & Rhodes R (2009) Production of phonetic and phonological contrast by heritage speakers of Mandarin http://linguistics.berkeley.edu/phonlab/annual_report/documents/2009/changetal_JASA_submitted.pdf (retrieved on 20th of February 2013)
- Clyne, M. (1991) *Community Languages: the Australian experience*. Cambridge: Cambridge University Press.

- Clyne, M. (1992) Linguistic and sociolinguistic aspects of language contact, maintenance and loss: Towards a multifacet theory. In W. Fase, K. Jaspaert, S. Kroon (eds.) *Maintenance and loss of minority languages*. Amsterdam/Philadelphia: John Benjamins, pp.17-36.
- Clyne, M. (2003) *Dynamics of language contact*. Cambridge: Cambridge University Press.
- Codó, E. (2008) Interviews and questionnaires. In L. Wei & M. G. Moyer (eds.) *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Oxford: Blackwell Publishing, pp.158-176.
- Cordero, M. (2008) *Integrating Reading, Writing, and Talk in the Spanish for Native Speakers Classroom*. PhD Thesis, University of Illinois pp.5-16.
- Cortelazzo, M. A. & Tuzzi, A. (2007) (eds.) *Messaggi dal Colle. I discorsi di fine anno dei Presidenti della Repubblica*. Venezia: Marsilio.
- Cortelazzo, M. A. & Tuzzi, A. (2008) *Metodi statistici applicati all'italiano*. Bologna: Zanichelli.
- Cucchiariato, C. (2010) *Vivo altrove. Giovani e senza radici: gli emigranti italiani di oggi*. Milano: Mondadori.
- Dabène, L. & Moore, D. (1995) Bilingual speech of migrant people. In L. Milroy & P. Muysken (eds.) *One speaker, two languages*. Cambridge: Cambridge University Press, pp.17-44.
- D'Agostino, M. (2007) *Sociolinguistica dell'Italia contemporanea*. Bologna: Il Mulino, pp.23-147.
- Danesi, M. (2011) Il Veneto parlato in Canada. In G. Romanato (ed.) *Veneti in Canada*. Ravenna: Longo, pp.143-155.
- Dardano, M. (1996) *Manualetto di linguistica italiana*. Milano: Zanichelli.
- Dautriat, H. (1979) *Il questionario: guida per la preparazione e l'impiego nelle ricerche sociali, di psicologia sociale e di mercato*. Milano: Franco Angeli.
- De Clementi, A. (2010) *Il prezzo della ricostruzione*. Roma/Bari: Laterza.
- De Maria Harney, N. (1998) Eh, paesan!: *Being Italian in Toronto*. Toronto: University of Toronto Press.
- Devoto, G. & Giacomelli, G. (1972) *I dialetti delle regioni d'Italia*. Firenze: Sansoni, pp.30-40.
- Ecke, P. (2004) Language attrition and theories of forgetting: A cross-disciplinary review. *International Journal of Bilingualism*, 8(3) pp.321–354.
- Edwards, J. (1998) (ed.) *Languages in Canada*. Cambridge: Cambridge University Press, pp.1-124, 400-414, 461-469.
- Edwards, J. (2007) Societal multilingualism: reality, recognition and response. In P. Auer & L. Wei (eds.) *Handbook of Multilingualism and Multilingual Communication*. Berlin: Mouton de Gruyter, pp.447-467.
- Ehala, M. (2010) Refining the notion of ethnolinguistic vitality. *International Journal of Multilingualism*, 7(4) pp.363-378.
- Fishman, J. (1967) Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues*, 23(2) pp.29-38.

- Fishman, J. A. (1985) *The Rise and Fall of the Ethnic Revival: Perspectives on Language and Ethnicity*. Berlin: Mouton, pp.363-376.
- Fondazione Migrantes (2009) *Rapporto italiani nel mondo 2009*. Roma:Idos.
- Fondazione Migrantes (2011) *Rapporto italiani nel mondo 2011*. Roma:Idos.
- Franzina, E. (1991) *Storia dell'emigrazione veneta. Dall'Unità al Fascismo*. Verona: Cierre.
- Franzina, E. (1995) *Gli Italiani al nuovo mondo*. Milano: Mondadori, pp. 1-337.
- Franzina, E. (2006) *Una patria espatriata*. Viterbo: Sette città.
- Fuller, J.M. (2013) *Spanish speakers in the USA*. Bristol: Multilingual Matters, pp.1951-1967.
- Gallo, S. (2010) Emigrare da fascisti, tra bonifiche, guerre coloniali e l'alleato tedesco. In M. Colucci (eds.) *Archivio storico dell'emigrazione italiana*, 6(1) pp.53-76.
- García, M.E. (2003) Recent research on language maintenance. *Annual Review of Applied Linguistics*, (23) pp.22-43.
- Gardner, R.C. (1982) Social factors in language retention. In R. D. Lambert & B. F. Freed (eds.) *The loss of language skills*. Rowley, MA: Newbury House, pp.24-43.
- Gardner, R. C. (1985) *Social Psychology and Second Language Learning. The role of attitudes and motivation*. London: Edward Arnold Publishers, pp.1-15, 108-123.
- Giampapa, F. (2001) Hyphenated identities: Italian-Canadian youth and the negotiation of ethnic identities in Toronto. *International Journal of Bilingualism*, 5(3) pp.279-315.
- Gobbi, S. (1994) Italiano e dialetto in situazione di emigrazione. *Quaderni patavini di linguistica. Monografie* (13) Padova: Unipress.
- Gonzo, S. & Saltarelli, M. (1983) Pidginization and linguistic change in emigrant languages. In R.W. Andersen (ed.) *Pidginization and creolization as language acquisition*. Rowley, MA: Newbury House, pp.181-197.
- Grosjean, F. (1982) *Life with two languages: An Introduction to Bilingualism*. Cambridge, Mass: Harvard University Press.
- Guiberson, M., Barrett, K.C., Jancosek, E.G., & Yoshinaga-Itano, C. (2006) Language maintenance and loss in young preschool age children of Mexican immigrants. Longitudinal study. *Communication Disorders Quarterly*, (28) pp.4-14.
- Hall, R. A. (1980) Language, Dialect and 'Regional Italian'. *International Journal of the Sociology of Language*, 25(1) pp. 95-106.
- Haller, H. (2006a) Tra italiano, dialetto e anglo-americano: riflessi scritti della lingua parlata dagli emigranti italiani negli Stati Uniti nel primo novecento. In E. Banfi, L. Gavioli, C. Guardino, M. Vedovelli (eds.) *Problemi e fenomeni di mediazione linguistica e culturale*. Perugia: Guerra, pp.345-358.
- Haller, H. (2006b) Lingue degli emigranti e degli esiliati: italiano. In G. Ernst, M. D. Gleßgen, C. Schmitt & W. Schweickard (eds.) *Romanische Sprachgeschichte Historie linguistique de la Romania* (vol.2). Berlin: Walter de Gruyter, pp. 1886-1892.
- Harney, R. (1984) *Dalla frontiera alle Little Italies*. Roma: Bonacci.

- Hopp, H. & Schmid, M. S. (2013) Perceived foreign accent in L1 attrition and L2 acquisition: The impact of age of acquisition and bilingualism. *Applied Psycholinguistics*, 34(2) pp.361-394.
- Hulsen, M. (2000) *Language Loss and Language Processing. Three generations of Dutch migrants in New Zealand*. PhD dissertation. Katholieke Universiteit Nijmegen.
- Jarvis, S. (2013a) Capturing the diversity in lexical diversity. *Language Learning*, (63, Issue supplement 1) pp.87-106.
- Jarvis, S. (2013b) Defining and measuring lexical diversity. In S. Jarvis and M. Daller (eds.) *Vocabulary knowledge: Human ratings and automated measures*. Amsterdam: Benjamins, pp.13-44.
- Jaspaert, K. & Kroon, S. (1991) Social determinants of language shift by Italians in The Netherlands and Flanders. *International Journal of the Sociology of Language*, (90) pp.77-96.
- Kagan, O. (2008) What is a heritage language. <http://www.international.ucla.edu/languages/news/article.asp?parentid=93215> (retrieved on 7.02.2013)
- Kaufman, D. (2001) Tales of L1 Attrition - Evidence From Pre-Puberty Children. In T. Ammerlaan, M. Hulsen, H. Strating & K. Yağmur (eds.) *Sociolinguistic and psycholinguistic perspectives on maintenance and loss of minority languages*. Münster: Waxmann, pp.185-202.
- Keijzer, M. (2007) *Last in first out? An investigation of the regression hypothesis in Dutch emigrants in Anglophone Canada*. Utrecht: LOT publications.
- Keijzer, M. (2010) The regression hypothesis as a framework for first language attrition. *Bilingualism: Language and Cognition*, 13(1) pp.9-18.
- Kloss, H. (1966) German-American language maintenance efforts. In J.A. Fishman (ed.) *Language loyalty in the United States: the maintenance and perpetuation of non-English mother tongues by American ethnic and religious groups*. The Hague: Mouton, pp.206-52.
- Köpke, B. (2002a) Activation thresholds and non-pathological first language attrition. In F. Fabbro (ed.) *Advances in the Neurolinguistics of Bilingualism*. Udine: Forum, pp.119-142.
- Köpke, B. (2002b) Attrition is not a unitary phenomenon: on different possible outcomes of language contact situations. In A.M.L. Suarez, F. Ramallo & X. Rodriguez-Yanez (eds.) *Bilingual socialization and bilingual language acquisition: Proceedings of the Second International Symposium on Bilingualism*. Vigo: Servizo de publicaciones da Universidade de Vigo, pp.1331-1347.
- Köpke, B. (2004) Neurolinguistic aspects of attrition. *Journal of Neurolinguistics*, 17(1) pp. 3-30.
- Köpke, B. (2007) Language attrition at the crossroads of brain, mind and society. In B. Köpke, M.S. Schmid, M. Keijzer & S. Dostert (eds.) *Language attrition: Theoretical Perspectives*. Amsterdam/Philadelphia: John Benjamins, pp.9-37.
- Köpke, B. & Nespoulous, J. (2001) First language attrition in production skills and metalinguistic abilities in German-English and German-French bilinguals. In T. Ammerlaan, M. Hulsen, H. Strating & K. Yağmur (eds.) *Sociolinguistic and psycholinguistic perspectives on Maintenance and Loss of minority languages*. Münster: Waxmann, pp.221-235.

- Laleko, O. (2007) A Study of Language Attrition: American Russian in Minnesota. In K. Shaw, S. Wagner & E. Yasui (eds.) *Proceedings of the Fifteenth Annual Symposium About Language and Society*. Austin: Texas Linguistic Forum (51), pp.103-112.
- Lazzarini, A. (1981) *Campagne venete ed emigrazione di massa (1866-1900)*. Vicenza: Istituto per le ricerche di storia sociale e di storia religiosa.
- de Leeuw E., Schmid M.S. & I. Mennen (2010) The effects of contact on native language pronunciation in a L2 migrants setting. *Bilingualism: Language and Cognition*, 13(1) pp.33-40.
- Lepschy, G. (2002) What is the Standard? In A.L. Lepschy & A. Tosi (eds.) *Multilingualism in Italy, Past and Present*. Oxford: Legenda, pp.74-78.
- Leuner, B. (2008) *Migration, multiculturalism and language maintenance in Australia*. Frankfurt am Main: Peter Lang, pp.184-239.
- Li, G. (2007) The role of parents in heritage language maintenance and development: Case studies of Chinese immigrant children's home practices. In K. Kondo-Brown (Ed.) *Multiple factors and contexts promoting heritage language: Focus on East Asian immigrants*. Amsterdam: John Benjamin, pp.15-32.
- Lindenfeld, J. & Varro, G. (2008) Language maintenance among "fortunate immigrants": the French in the United States and Americans in France. *International Journal of the Sociology of Language*, (189) pp.115-131.
- Macevichius, I. (2001) The beginnings of language loss in discourse: A case study of referentiality in American Lithuanian. In T. Ammerlaan, M. Hulsen, H. Strating & K. Yağmur (eds.) *Sociolinguistic and psycholinguistic perspectives on maintenance and loss of minority languages*. Münster: Waxmann, pp.235-248.
- Mah, B. (2005) *Ethnic Identity and Heritage Language Ability in Second Generation Canadians in Toronto*. Master Thesis, Ryerson University pp.2-13, 42-54.
- Maiden, M. & Parry, M. (eds.) (1997) *The Dialects of Italy*. New York: Routledge.
- Marazzini, C. (1994) *La lingua italiana. Profilo storico*. Bologna: il Mulino.
- Marcato, C. (2002) *Dialetto, dialetti e italiano*. Bologna: Il Mulino, pp.11-128.
- Marcato C., Haller H., Meo Zilio C. & Ursini F. (2002), I dialetti italiani nel mondo. In M. Cortelazzo, C. Marcato, N. De Blasi & G.P. Clivio (eds.) *I dialetti Italiani. Storia struttura uso*. Torino: Utet, pp.1073-1096.
- Marchesin, A. (2011) Media italiani a Toronto e nella Great Toronto area. In G. Romanato (ed.) *Veneti in Canada*. Ravenna: Longo, pp.183-188.
- McCarthy, P. & Jarvis S. (2007) Vocd: A theoretical and empirical evaluation. *Language Testing*, (24) pp.459-488.
- McEnery, T. & Wilson, A. (2001) *Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Melillo, A. (ed.) (1991) *Lettere dalla Merica*. Bari: Adriatica.
- Milroy, L. (1987) *Language and social Networks*. Oxford: Blackwell Publishing.

- Milroy, L. & Gordon, M. (2003) *Sociolinguistics. Method and interpretation*. Oxford: Blackwell Publishing.
- Montrul, S. (2005) Second language acquisition and first language loss in adult early bilinguals: exploring some differences and similarities. *Second Language research*, 21(3) pp.199-249.
- Montrul, S. (2008) *Incomplete acquisition in bilingualism*. Amsterdam/Philadelphia: John Benjamins.
- Montrul, S. (2010) Current issues in heritage language acquisition. *Annual review of Applied Linguistics*, (30) pp.3-23.
- Montrul, S. (2012) Is the heritage language like a second language? *EUROSLA Yearbook 12*, pp.1-29.
- Montrul S., Foote R. & Perpiñán S. (2008) Gender Agreement in Adult Second Language Learners and Spanish Heritage Speakers: The Effects of Age and Context of Acquisition. *Language Learning*, 58(3) pp.503–553.
- Montrul, S. & Perpiñán, S. (2011) Assessing Differences and Similarities between Instructed Heritage Language Learners and L2 Learners in Their Knowledge of Spanish Tense-Aspect and Mood (TAM) Morphology. *Heritage Language Journal*, 8(1) <http://www.international.ucla.edu/media/files/mont-and-perpi-eight-one-0i-5il.pdf> (retrieved on 28th of October 2013).
- Montrul, S. & Polinsky, M. (2011) Why not heritage speakers? *Linguistic approaches to bilingualism*, 1(1).
- Myers-Scotton, C. (2002) *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford: Oxford University Press, pp.164-232.
- Nortier, J. (2008). Types and sources of bilingual data. In L. Wei & M. Moyer (eds.) *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Oxford: Blackwell Publishing, pp.35-53.
- Olshtain, E. & Barzilay, M. (1991) Lexical retrieval difficulties in adult language attrition. In H. Seliger & R. Vager (eds.) *First language attrition*. Cambridge: Cambridge University Press, pp.139-150.
- Osterhout L., McLaughlin J., Kim A., Greenwald R. & Inoue K. (2004) Sentences in the Brain: Event-Related Potentials as Real-Time Reflections of Sentence Comprehension and Language Learning. In M. Carreira & C. Clifton (eds.) *The on-line study of sentence comprehension*. New York: Psychology Press, pp.271-308.
- Pallier, C. (2007) Critical periods in language acquisition and language attrition. In B. Köpke, M.S. Schmid, M. Keijzer & S. Dostert (eds.) *Language attrition: Theoretical Perspectives*. Amsterdam/Philadelphia: John Benjamins, pp.155-168.
- Paradis, M. (2007) L1 attrition features predicted by a neurolinguistic theory of bilingualism. In B. Köpke, M.S. Schmid, M. Keijzer & S. Dostert (eds.) *Language attrition: Theoretical Perspectives*. Amsterdam/Philadelphia: John Benjamins, pp.121-133.

- Park, S.M. (2011) The role of ethnic religious community institutions in the intergenerational transmission of Korean among immigrant students in Montreal. *Language, Culture and Curriculum*, 24(2) pp.195-206.
- Pautasso, L. (1978) La donna italiana durante il periodo fascista in Toronto, 1930-1940. In B. Boyd Caroli, R.F. Harney & L.S. Tomasi (eds.) *The Italian immigrant woman in North America*. Toronto: Multicultural History Society of Ontario, pp.168-186.
- Pavlenko, A. (2002) Conceptual change in bilingual memory: a neo-whorfian approach. In F. Fabbro (ed.) *Advances in the Neurolinguistics of Bilingualism*. Udine: Forum, pp.69-94.
- Pavlenko, A. (2004) L2 influence and L1 attrition in adult bilingualism. In M. S. Schmid, B. Köpke, M. Keijzer & L. Weilemar (eds.) *First Language Attrition: Interdisciplinary perspectives on methodological issues*. Amsterdam/Philadelphia: John Benjamins, pp.47-59.
- Polinsky, M. (2007) Heritage Language Narratives. In D.M. Brinton, O. Kagan & S. Bauckus (eds.) *Heritage Language education*. New York/London: Routledge, pp.149-164.
- Polinsky, M. (2008) Gender under incomplete acquisition: Heritage speakers' knowledge of noun categorization. *Heritage language journal*, (6) pp.40-71.
- Polinsky, M. (2011) Reanalysis in adult heritage language. *Studies in second language acquisition*, (33) pp.305-328.
- Polinsky, M. & Kagan, O. (2007) Heritage languages: in the 'wild' and in the classroom. *Language and Linguistics Compass*, 1(5) pp.368-395.
- Prescher, P. (2007) Identity, immigration and first language attrition. In B. Köpke, M.S. Schmid, M. Keijzer, & S. Dostert (eds.) *Language attrition: Theoretical Perspectives*. Amsterdam/Philadelphia: John Benjamins, pp.189-204.
- Py, B. (1986) Native language attrition amongst migrant workers: towards an extension of the concept of interlanguage. In E. Kellerman & M. Sharwood Smith (eds.) *Crosslinguistic influence in second language acquisition*. New York: Pergamon, pp.163-172.
- Ribbert, A. & Kuiken, F. (2010) L2-induced changes in the L1 of Germans living in the Netherlands. *Bilingualism: Language and cognition*, 13(1) pp.41-48.
- Romaine, S. (1995) *Bilingualism*. Oxford: Blackwell Publishing, pp.1-119, 181-240, 288-319.
- Rothman, J. (2007) Heritage speaker competence differences, language change and input type: Inflected infinitives in heritage Brazilian Portuguese. *International Journal of Bilingualism*, (11) pp.359-389.
- Rubino, A. (1991) Commutazione di stile e strategie comunicative degli scolari Italo-australiani. In I. Baldelli & B.M. Da Rif (eds.) *Lingua e letteratura italiana nel mondo oggi. XIII Congresso internazionale dell'Associazione Internazionale per gli Studi di Lingua e Letteratura Italiana*. Firenze: Olschki Vol.2, pp.587-603.
- Rubino, A. (2000) Playing with languages: language alternation in Sicilian-Australian children's conversation. *Rassegna Italiana di Linguistica Applicata*, 32(2) pp.89-108.
- Rubino, A. (2006) Linguistic practices and language attitudes of second-generation Italo-Australians. *International Journal of the Sociology of Language*, (180) pp.71-88.

- Sartor, M. & Ursini, F. (1983) *Cent'anni di emigrazione: una comunità veneta sugli altipiani del Messico*. Cornuda (TV): Grafiche Antiga.
- Scaglione, S. (2000) *Attrition. Mutamenti sociolinguistici nel lucchese di San Francisco*. Milano: Franco Angeli.
- Schmid, M.S. (2002) *First Language Attrition, Use, and Maintenance: The case of German Jews in Anglophone countries*. Amsterdam/Philadelphia: John Benjamins.
- Schmid, M. S. (2004a) Identity and first language attrition: A historical approach. *Estudios de Sociolingüística*, 5(1) pp.41-58.
- Schmid, M. S. (2004b) First Language attrition: The methodology revised. *International Journal of Bilingualism*, 8 (3) pp.239-255.
- Schmid, M. S. (2007) The role of L1 use for L1 attrition. In B. Köpke, M.S. Schmid, M. Keijzer & S. Dostert (eds.) *Language attrition: Theoretical Perspectives*. Amsterdam/Philadelphia: John Benjamins, pp.135-153.
- Schmid, M. S. (2008a) Defining language attrition. *Babylonia*, 2(8) pp.9-12.
- Schmid, M.S. (2011a) *Language attrition. Keytopic in sociolinguistics*. Cambridge: Cambridge University Press.
- Schmid, M. S. (2011b) Contact x time: External factors and variability in L1 attrition. In M.S. Schmid & W. Lowie (eds.) *Modeling Bilingualism: From structure to chaos*. Amsterdam/Philadelphia: John Benjamins, pp.155-176.
- Schmid, M.S. & Beers Fägersten, K. (2010) Disfluency markers in L1 attrition. *Language Learning*, 60(4) pp.753-791.
- Schmid, M. S. & de Bot, K. (2004) Language attrition. In A. Davies & C. Elder (eds.) *The Handbook of Applied Linguistics*. Oxford: Blackwell Publishing, pp.210-234.
- Schmid, M.S. & Dusseldorp, E. (2010) Quantitative analyses in a multivariate study of language attrition: the impact of extralinguistic factors. *Second language research*, 26(1) pp.125-160.
- Schmid, M.S. & Jarvis, S. (forthcoming) Lexical access and lexical diversity in first language attrition. Submitted to *Bilingualism: Language and Cognition*.
- Schmid, M. S. & Köpke, B. (2004) Language attrition: the next phase. In M. S. Schmid, B. Köpke, M. Keijzer & L. Weilemar (eds.) *First Language Attrition: Interdisciplinary perspectives on methodological issues*. Amsterdam/Philadelphia: John Benjamins, pp.1-43.
- Schmid, M. S. & Köpke, B. (2007) Bilingualism and attrition. In B. Köpke, M.S. Schmid, M. Keijzer & S. Dostert (eds.) *Language attrition: Theoretical Perspectives*. Amsterdam/Philadelphia: John Benjamins, pp.1-7.
- Schmid, M.S. & Köpke, B. (2009) L1 attrition and the mental lexicon. In A. Pavlenko (ed.) *The bilingual mental lexicon: Interdisciplinary approaches*. Clevedon: Multilingual Matters, pp.209–238.
- Schmid M.S., Verspoor M. & MacWhinney B. (2011) Coding and extracting data. In M. Verspoor, K. de Bot and W. Lowie (eds.) *A dynamic approach to second language development: methods and techniques*. Amsterdam: John Benjamins, pp.39-54.

- Schmid, S. (1993) Lingua madre e commutazione di codice in immigrati italiani di seconda generazione nella Svizzera tedesca. *Multilingua*, 12 (3) pp.265-289.
- Selinger, H. & Vago, R. (1991) The study of first language attrition: an overview. In H. Selinger & R. Vago (eds.) *First language attrition*. Cambridge: Cambridge University Press, pp.3-15.
- Šišková, Z. (2012) Lexical Richness in EFL Students' Narratives. In C. Ciarlo and D.S. Giannoni (eds.) *Language studies working papers*. Reading: University of Reading (4), pp.26-36.
- Skaaden, H. (2005) First language attrition and linguistic creativity. *International Journal of Bilingualism*, 9(3&4) pp.435-452.
- Smolicz, J. J. (1981) Core values and cultural identity. *Ethnic and Racial Studies*, 4(1) pp.75-90.
- Sobrero, A.A. & Miglietta, A. (2006) *Introduzione alla linguistica italiana*. Roma: Laterza, pp.57-112, 212-223.
- Stolberg, D. & Münch, A. (2010) "Die Muttersprache vergisst man nicht" – or do you? A case study in L1 attrition and its (partial) reversal. *Bilingualism: language and cognition*, 13(1) pp.19-31.
- Stubbs, M. (2004) Language Corpora. In: A. Davies & C. Elder (eds.) *The Handbook of Applied Linguistics*. Oxford: Blackwell Publishing, pp.106-132.
- Tagliamonte, S. A. (2006) *Analysing sociolinguistic variation*. Cambridge: Cambridge University Press, pp.37-65.
- Tagliamonte, S. A. (2012) *Variationist Sociolinguistics: Change, Observation, Interpretation*. Chichester: Wiley-Blackwell Press, pp.120-157.
- Tempesta, I. (1978) *Lingua ed emigrazione: indagine sul comportamento sociolinguistico degli emigranti salentini*. Lecce: Milella.
- Tomasin, L. (2010) *Storia linguistica di Venezia*. Roma: Carocci.
- Tosi, A. (1991) *L'italiano d'oltremare. La lingua delle comunità italiane nei paesi anglofoni*. Firenze: Giunti.
- Turchetta, B. (2005) *Il mondo in italiano. Varietà e usi internazionali della lingua*. Roma/Bari: Laterza.
- Turell, M.T. & Moyer, M.G. (2008) Transcription. In L. Wei & M. Moyer (eds.) *The Blackwell Guide to Research Methods in Bilingualism and Multilingualism*. Oxford: Blackwell Publishing, pp. 192-213.
- Tuzzi, A. (2003) *L'analisi del contenuto. Introduzione ai metodi e alle tecniche di ricerca*. Roma: Carocci.
- Valdés, G. (2005) Bilingualism, heritage language learners, and SLA research: opportunities lost or seized? *The Modern Language Journal*, 89(3) pp.410-426.
- Vanvolsem S., Jaspaert K. & Kroon S. (1991) Erosione e perdita di lingua presso emigrati Italiani in Belgio e in Olanda: primi risultati. In I. Baldelli & B. M. Da Rif (eds.) *Lingua e letteratura Italiana nel mondo oggi* (vol.II) Firenze: Olschki Editore, pp.677-691.
- Vedovelli, M. (2011) (ed.) *Storia linguistica dell'emigrazione italiana nel mondo*. Roma: Carocci.

- Vianello, F. (2006) Una vita altrove. L'emigrazione italiana dal 1876 al 1976. In M. Santipolo (ed.) *L'Italiano*. Novara: Utet, pp.157-184.
- Vizmuller-Zocco, J. (2007) Language, ethnicity, post-modernity: the Italian Canadian case. *Studi Emigrazione/Migration Studies*, 44 (199) pp.355-368.
- Waas, M. (1996) *Language attrition downunder*. Frankfurt: Peter Lang.
- Weinreich, U. (1963) *Languages in contact: findings and problems*. The Hague: Mouton.
- Winford, D. (2003) *An introduction to Contact Linguistics*. Oxford: Blackwell Publishing, pp.1-28, 235-267.
- Yağmur, K. (1997) *First language attrition among Turkish speakers in Sydney*. Tilburg: Tilburg University Press.
- Yilmaz, G. & Schmid, M.S. (forthcoming) L1 lexical accessibility among Turkish speakers in the Netherlands. Accepted for publication in *The Mental Lexicon*.
- Zampieri Pan, A. (2009) *Presenze Italiane in British Columbia*. Vancouver: Ital Press.
- Zampieri Pan, A.M. (2011) Italiani e Veneti in British Columbia. In G. Romanato (ed.) *Veneti in Canada*. Ravenna: Longo, pp.255-265.

Internet sources

- <http://www.emigrati.it> (accessed on 27.07.2011)
- http://www.esteri.it/mae/doc_osservatorio/Rapporto_Paese_Canada.pdf (accessed on 27.02.2009)
- Aire: www.esteri.it (accessed on 01.03.2013)
- www.istat.it (accessed on 20.02.2013)
- <http://www.italiansinfonia.com> (accessed on 27.03.2013)
- Radio 24 - Il sole 24 ore, *La fuga dei talenti* (on air on 17.09.2011). Podcast on <https://fugadeitalenti.wordpress.com> (accessed on 17.09.2011)
- <https://fugadeitalenti.wordpress.com> (accessed on 03.01.2012)
- <http://www.taltac.it> (Sergio Bolasco, University of Rome – La Sapienza) (accessed on 14.09.2010)