

JAMES WILLIAM HAWKEY

PhD Thesis

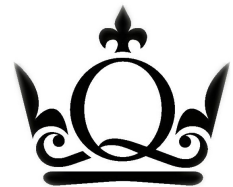
2012

---

**Language policy and  
language contact in Barcelona:  
a contemporary perspective.**

---

Centre for Catalan Studies  
Department of Iberian and Latin American Studies  
Queen Mary, University of London



# Declaration

I hereby declare that the work presented in this thesis, 'Language policy and language contact in Barcelona: a contemporary perspective', is my own.

Signed:

A handwritten signature in black ink, appearing to read 'J. Hawkey', written in a cursive style.

James Hawkey

Centre for Catalan Studies

Department of Iberian and Latin American Studies

Queen Mary, University of London

# Abstract

The present thesis provides an analysis of language policy and language contact in present-day Barcelona. On the one hand, the effectiveness of the implementation of Catalan-medium education in Catalonia will be critically evaluated. On the other hand, Catalan-Castilian bilinguals' awareness of non-normative instances of language contact will be assessed. This thesis brings these two strands together, in order to paint an accurate picture of the current Catalan sociolinguistic situation.

The 1983 *Llei de Normalització Lingüística* had numerous consequences, including the introduction of Catalan as the chief medium of education in Catalonia. Such legislation had many aims, not least to 'ensure Catalan language competence'. But to what extent has this been successful? Furthermore, the varieties of Catalan and Castilian in Barcelona are characterised by centuries of language contact, which has resulted in the incursion of numerous, non-normative linguistic items and constructions in both languages. To what extent are speakers aware of such non-normative language? Moreover, how do these questions concerning language policy and language contact bear upon one another?

This thesis is therefore looking at the consequences of language policy and of language contact. With the aid of an innovative, three-dimensional model of sociolinguistic phenomena, it will be shown that these are maximally differentiated, yet clearly related. These will be termed *top-down phenomena* and *bottom-up phenomena* respectively.

This is to be tested using a unique fieldwork experiment whereby fifty bilingual Catalan-Castilian *barcelonins* of two different generations (one educated in Catalan, the other in Castilian, due to different linguistic policies) are asked to identify instances of non-normative language contact in two prepared texts.

This work will allow conclusions to be drawn regarding the effectiveness of Catalan-medium education, as well as offering insight into the nature of how we examine issues of language policy and language contact.

# Acknowledgements

Much like this thesis, my academic background to date has consisted of two complementary strands. An undergraduate degree in modern languages inspired me to pursue further research into Hispanic and Catalan studies, while a master's program in theoretical linguistics introduced me to the discipline of sociolinguistics. The next step on my academic career, the present work, fuses these two components into one cohesive entity which truly reflects both my experience thus far and my prospective direction for future research. Given that this doctoral thesis draws so heavily on my personal academic background, many individuals and organisations must be thanked, for without their support, this project would never have come to fruition.

This PhD thesis was fully funded (maintenance and tuition fees) by Queen Mary, University of London's Catalan Research Studentship in association with the Institut Ramon Llull. Without this generous financial support, the thesis could not have been undertaken, and so all persons involved in the selection process, as well as the Centre for Catalan Studies at QMUL and the Institut Ramon Llull, are gratefully acknowledged.

Many thanks to my two supervisors, Professor Christopher Pountain and Dr Jordi Larios, for all of their guidance and support. Chris, your continued encouragement has kept me on track for the duration of the project and your encyclopaedic knowledge of Hispanic Linguistics has been greatly useful. Jordi, your input as a Catalan studies scholar has been both enlightening and beneficial to the project as a whole, not to mention your help in providing contacts and participants in the field and your attentive consideration making sure I remain motivated. Thank you both. Special thanks must also go to Professor David Adger, for his invaluable views on how this thesis should ultimately take shape. Many thanks also to my two examiners for their time and attention, and whose comments have made this thesis a work of which I am truly proud.

The fieldwork experiment is an undoubtedly crucial part of this whole project,

and so a great debt of gratitude is owed to all seventy-four participants, as well as to those who procured participants on my behalf. Marina Roman Castells, Naira Tortajada, Elena Artigas, Lurdes Franquesa and Ester Pou: si no fos per vosaltres, no hauria trobat ningú, i aquest projecte no hauria estat possible. Moltes gràcies, de debò.

QMUL has offered me an entire support network within both the Department of Iberian and Latin American Studies and the Department of Linguistics. I would like to give heartfelt thanks to all academic members of staff in both departments for the interest they have shown in my project, and for making me feel very much at home. Special mentions must go to Erez Levon and Linnaea Stockall for bolstering my meagre knowledge of statistics, and Devyani Sharma for her constructive comments after my upgrade exam. I must also mention the Linguistics Lab (and its occupants). ‘The Lab’ has provided me with a work space day-in, day-out, and allowed me the privilege of spending three years with a group of intellectually stimulating, brilliant young academics. Such a community would not have existed without the combined vision of David Adger and Jenny Cheshire. Their dream of a Linguistics department that encompasses both formal and sociolinguistics, in which ideas are freely exchanged between students and staff of all areas of the discipline, has resulted in an academic environment that I feel honoured to be a part of. So a big, heartfelt thank you must go out to Professors Adger and Cheshire. On a personal level, the support offered to me by the Lab has helped immensely through what have been some difficult and trying times, and for this I will be forever grateful. So thank you, Labbers past and present : Fryni Panayidou, Ruth Kircher, Eva Klingvall, Ollie Brownlow, Rachelle Freake, Philippa Law, Michèle Pettinato, John Weston, Agnieszka Knaś, Maria Sečová, Barbara Clark, Ahmad El-Sharif, Chiara Ciarlo, Álvaro Recio, Nee Rhurakvit and Jad. Fellow PhD students and colleagues in the Department of Iberian and Latin American Studies must also be mentioned, namely Ignacio Gregorio, Rocío Díaz and Ángel Osle. Muchísimas gracias por todo. I have also been lucky enough that my support network at QMUL does not begin and end with the departments in which I work: my PhD ‘History Boys’ (Andrew Smith, Paul Davidson, Phil Baldwin

and Cory Santos) have accompanied me on this doctoral journey, and their support has been worth its weight in beer! Thanks guys.

To the non-QMUL friends who have kept me sane throughout the last three years, and who I am sure will continue to do so for many more to come, thank you. To Katherine Brown and Katy Knight, who have been my enduring housemates throughout this entire experience, thank you. To Alexandra Wilby, my best friend and minibreak companion, thank you. To Ghislain Graziani, quand tout n'allait pas bien, tu étais là, et je t'en remercie. To Maja Jaworska, Judith Arkle, Natalie Ellis, Catherine Beckford, Simonetta Sarogni, Claire Grant, Sophie Parrott, Jennifer Head, Mara Lockowandt, Sam Rawlings, Kate Willman and the whole host of people I am lucky enough to share my life with, thank you. I am very blessed that there are too many of you to mention individually, but were it not for you all, I could never have reached this point.

Finally, to my family, for whose unconditional love I will be forever thankful. To my parents, Fay and Phil, who have always encouraged me and supported my decisions, thank you so much for all your guidance and love. To my sisters Joanne and Rachel, and to their wonderful children Megan, Mia, Lewis and Carey, thank you. To my grandmother Diana whose weekly bulletins have kept me cheerful and motivated, thank you. To my 'second mother' Cheryl, thank you for all your support. And to my younger sister Melissa: the start of this PhD coincided with a very difficult time for you, but you bounced back stronger than ever, and never fail to inspire me. For that reason, I dedicate this thesis to you. Love love.

# Contents

<b>Abstract</b>	<b>3</b>
<b>Acknowledgements</b>	<b>4</b>
<b>List of Tables</b>	<b>11</b>
<b>List of Figures</b>	<b>15</b>
<b>List of Terms and Abbreviations</b>	<b>16</b>
<b>1 Introduction</b>	<b>19</b>
1.1 Barcelona: historical and sociolinguistic background . . . . .	21
1.2 The importance of language policy and language contact in Barcelona	41
1.2.1 Policy and contact: Existing studies of Catalonia. . . . .	42
1.2.2 Policy and contact: Theoretical implications. ‘Top-down phe- nomena’ and ‘bottom-up phenomena’. . . . .	47
1.3 Ways in which to study language policy and language contact in Barcelona	64
1.4 Thesis plan . . . . .	66
<b>2 Language Policy in Barcelona</b>	<b>67</b>
2.1 Language policy and resultant phenomena (TDP): Theory . . . . .	69
2.1.1 Language policy and planning . . . . .	69
2.1.2 Acquisition planning and its relation to TDP . . . . .	83
2.1.3 Summary . . . . .	88
2.2 Language policy and resultant phenomena (TDP) in Barcelona . . . . .	89

2.2.1	Language policy in Barcelona . . . . .	89
2.2.1.1	Present-day language-in-education policies in Barcelona	100
2.2.2	TDP in Barcelona: Details of phenomena in language form and function . . . . .	102
2.2.3	Summary . . . . .	111
2.3	Language policy in Barcelona: Overview and research question . . . .	111
<b>3</b>	<b>Language Contact in Barcelona</b>	<b>114</b>
3.1	Language contact and resultant phenomena (BUP): Theory . . . . .	116
3.1.1	Language contact . . . . .	116
3.1.2	Language contact and its relation to BUP . . . . .	126
3.1.3	Testing the theory: BUP in the city . . . . .	132
3.1.4	Summary . . . . .	138
3.2	Language contact and resultant phenomena (BUP) in Barcelona . . .	139
3.2.1	Language contact in Barcelona . . . . .	140
3.2.2	BUP in Barcelona: Details of phenomena in language form and function . . . . .	143
3.2.3	Summary . . . . .	149
3.3	Language contact in Barcelona: Overview and research question . . .	150
<b>4</b>	<b>Experiment Methodology</b>	<b>154</b>
4.1	Methodological Context . . . . .	154
4.1.1	Addressing the research questions . . . . .	155
4.1.2	Designing the experiment . . . . .	156
4.1.2.1	Language policy concerns : Apparent time studies . .	157
4.1.2.2	Language contact concerns : The pre-pilot questionnaire	158
4.1.2.3	Summary . . . . .	168
4.2	Pilot study . . . . .	168
4.2.1	Procedure . . . . .	169
4.2.2	Stimuli . . . . .	171
4.2.2.1	Pilot study part one: Reading passages . . . . .	171



4.2.2.2	Pilot study part two: Conversation stimulus image . . . . .	177
4.2.3	Subjects . . . . .	177
4.2.4	Results . . . . .	178
4.2.5	Discussion . . . . .	179
4.3	Fieldwork experiment methodology . . . . .	181
4.3.1	Procedure . . . . .	181
4.3.2	Stimuli . . . . .	183
4.3.2.1	The Catalan text . . . . .	184
4.3.2.2	The Castilian text . . . . .	191
4.3.3	Subjects . . . . .	197
<b>5</b>	<b>Experiment Results</b>	<b>200</b>
5.1	Exercise Results . . . . .	200
5.1.1	Participant details: Overall frequencies . . . . .	202
5.1.2	Overall test scores . . . . .	204
5.1.3	Age group (younger / older) . . . . .	205
5.1.4	Sex (male / female) . . . . .	210
5.1.5	Language dominance (Catalan / Castilian) . . . . .	214
5.1.6	Language of interview (Catalan / Castilian) . . . . .	219
5.1.7	Level of education . . . . .	223
5.1.8	Social class . . . . .	228
5.1.9	Multiple interactions . . . . .	238
5.1.10	Summary . . . . .	241
5.2	Types and tokens . . . . .	241
5.2.1	Type-by-type analysis . . . . .	242
5.2.2	Token-by-token analysis . . . . .	246
5.2.3	Misidentified tokens . . . . .	253
5.3	Language competence: Participant self-report . . . . .	256
5.4	Overview . . . . .	261

<b>6 Discussion</b>	<b>263</b>
6.1 Answering the language contact research question . . . . .	264
6.2 Answering the language policy research question . . . . .	271
6.3 Related issues and directions for further research . . . . .	287
6.4 Final conclusions . . . . .	293
<b>Bibliography</b>	<b>295</b>
<b>Appendices</b>	<b>308</b>
Appendix A: Pre-Pilot Questionnaire . . . . .	308
Appendix B: Pilot Study Materials . . . . .	310
Appendix C: Final Fieldwork Experiment Texts . . . . .	314
Appendix D: Participant Personal Details Form . . . . .	319
Appendix E: Participant Consent Form . . . . .	325
Appendix F: Network of participants . . . . .	327

# List of Tables

1.1	Catalan and Castilian language competence in the population of the autonomous community of Catalonia .....	35
1.2	Language of identification of the population of the autonomous community of Catalonia .....	36
1.3	Language of identification of the population of the autonomous community of Catalonia, classified according to place of birth .....	36
1.4	Language of use with parents, classified by participant age group ....	38
1.5	Tabular representation of how my definitions of TDP and BUP relate to the dimensions of <i>metacommentary</i> , <i>overt prestige</i> and <i>planning</i> ...	60
1.6	Examples of TDP and BUP .....	62
2.1	Partial reproduction of table 1.6, pertaining to TDP .....	68
3.1	Partial reproduction of table 1.6, pertaining to BUP .....	115
4.1	Proposed findings of pre-pilot questionnaire .....	159
4.2	Questionnaire respondent details, categorised by sex and place of origin / residence .....	160
4.3	Linguistic features specific to Barcelona, according to questionnaire respondents .....	160
4.4	Pilot study results .....	178
4.5	Third-person direct and indirect object pronouns in normative Catalan and normative Castilian .....	189
4.6	Coding system for experiment participants .....	199
5.1	Overall test results. Scores of all participants and independent t-test results.....	204

5.2a	Test results divided by <i>language of exercise</i> , then <i>age group</i> . Scores and independent t-test results .....	207
5.2b	Test results divided by <i>age group</i> , then <i>language of exercise</i> . Scores and independent t-test results .....	208
5.3	Interaction of <i>language of exercise</i> and <i>age group</i> . Repeated Measures test .....	209
5.4a	Test results divided by <i>language of exercise</i> , then <i>sex</i> . Scores and independent t-test results .....	211
5.4b	Test results divided by <i>sex</i> , then <i>language of exercise</i> . Scores and independent t-test results .....	212
5.5	Interaction of <i>language of exercise</i> and <i>sex</i> . Repeated Measures test	213
5.6a	Test results divided by <i>language of exercise</i> , then <i>participant language dominance</i> . Scores and independent t-test results .....	216
5.6b	Test results divided by <i>participant language dominance</i> , then <i>language of exercise</i> . Scores and independent t-test results .....	217
5.7	Interaction of <i>language of exercise</i> and <i>participant language dominance</i> . Repeated Measures test .....	218
5.8a	Test results divided by <i>language of exercise</i> , then <i>language of interview</i> . Scores and independent t-test results .....	220
5.8b	Test results divided by <i>language of interview</i> , then <i>language of exercise</i> . Scores and independent t-test results .....	221
5.9	Interaction of <i>language of exercise</i> and <i>language of interview</i> . Repeated Measures test .....	222
5.10	Test results divided by <i>language of exercise</i> , then <i>level of education</i> . Scores and independent ANOVA results .....	225
5.11	Test results divided by <i>language of exercise</i> , then <i>level of education</i> . Duncan's homogeneous subsets .....	226
5.12	Interaction of <i>language of exercise</i> and <i>level of education</i> . Repeated Measures test .....	227
5.13	Composition of social class score .....	229

5.14	Participant scores / Participant parent scores and corresponding NS-SEC values .....	230
5.15	Participant scores and corresponding levels of education .....	231
5.16	Participant scores and corresponding area of residence, with mean household per capita income figures .....	232
5.17	Participant total scores and corresponding social class .....	233
5.18	Test results divided by <i>language of exercise</i> , then <i>social class</i> . Scores and independent ANOVA results .....	234
5.19	Test results divided by <i>language of exercise</i> , then <i>social class</i> . Duncan's homogeneous subsets .....	235
5.20	Interaction of <i>language of exercise</i> and <i>social class</i> . Repeated Measures test .....	235
5.21	Interaction of <i>language of exercise</i> and <i>social class as a scalar variable</i> . Repeated Measures test .....	236
5.22	Interaction of <i>language of exercise</i> with <i>different social class subvariables</i> . Repeated Measures test. Sphericity-assumed results only .....	237
5.23	Statistically significant Pearson correlations between pairs of variables	239
5.24	Catalan exercise test results divided by <i>type</i> , then <i>participant language dominance</i> . Scores and independent t-test results .....	244
5.25	Frequency of identification for each token .....	247

# List of Figures

1.1	Map situating <i>Catalunya Vella</i> , tenth to eleventh century (red), within current Catalan-speaking areas (light grey) . . . . .	23
1.2	Labov’s stereotypes, markers and indicators as CFAs / CFBs . . . . .	51
1.3	Graphical representation of how TDP and BUP need to relate to the dimensions of <i>metacommentary</i> , <i>overt prestige</i> and <i>planning</i> . . . . .	58
2.1	Schiffman’s taxonomical language policy model applied to India . . . . .	76
3.1	Non-exhaustive typology of the consequences of language contact . . . . .	131
4.1	Vowel charts showing (from left to right): the seven-vowel system of Catalan in stressed position, the three-vowel system of Catalan in unstressed position, and the five-vowel system of Castilian in stressed or unstressed position . . . . .	162
5.1	Token-by-token analysis : Catalan exercise divided by <i>age group</i> . . . . .	249
5.2	Token-by-token analysis : Castilian exercise divided by <i>age group</i> . . . . .	250
6.1	The links established between <i>MOI choice</i> , <i>Recent Catalan language policy developments</i> and <i>Awareness of non-normative language use</i> . . . . .	277
6.2	The validity of conclusion 4 . . . . .	282
6.3	Reproduction of figure 1.3, showing where TDP and BUP appear on the cube . . . . .	288
6.4	The positioning of Castilianisms on the three-axis cube for older participants (blue dot) and younger participants (green dot). . . . .	289

6.5 Examples of potential configurations of simultaneous, distinct sociolin-  
guistic phenomena in hypothetical linguistic communities . . . . . 290

# List of Terms and Abbreviations

This brief glossary serves as a convenient point of reference, should at any juncture the reader require easy access to the definition of a particular term, as it is used in this thesis. At the end of most entries, there is a page or section reference indicating where the concept is explored in depth or ultimately defined.

\* \* \*

**Acquisition planning** is used to refer to instances of language planning which are ‘directed toward increasing the number of [language] users’ (Cooper 1989, 33). See also *corpus planning*, *language planning*, *language-in-education policy* and *status planning*. See page 72.

**BUP (bottom-up phenomena)** encompass situations which give rise to the use of linguistic items or structures. This language use is implemented spontaneously and below the level of consciousness. See page 59.

**CILP (contact-induced linguistic phenomena)** refer to a collection of speaker-driven phenomena, including simplification, overgeneralisation, analysis, transfer, convergence and divergence, though not code-switching. See pages 127–129 for a full definition and explanation of all of the practices mentioned here.

**Castilianisms** are linguistic elements of Castilian origin used in other languages. In this thesis, the term most often refers to Castilian lexical, syntactic or phonological elements found in Catalan speech. See page 144.



**Catalanisms** are linguistic elements of Catalan origin used in other languages. In this thesis, the term most often refers to Catalan lexical, syntactic or phonological elements found in Castilian speech. See page 146.

**Corpus planning** is used to refer to instances of language planning which ‘modify the nature of the language itself’ (Kloss 1969, 81). See also *acquisition planning*, *language planning* and *status planning*. See page 71.

**LNL** refers to the *Llei de Normalització Lingüística* (language normalisation law), passed in 1983 by the *Generalitat de Catalunya*.

**LPL** refers to the *Llei de Política Lingüística* (language policy law), passed in 1998 by the *Generalitat de Catalunya*.

**Language contact** is a situation whereby two or more distinct linguistic varieties co-occur, either on a microsociolinguistic level between speakers, or on a macrosociolinguistic level within in the linguistic community as a whole. In this thesis, it exemplifies a background situation to BUP. See also *BUP*. See section 3.1.1 for an in-depth discussion of language contact.

**Language planning** is ‘the ways in which... communities... consciously attempt to influence the language(s) their members use, the languages used in education, or the ways in which Academies, publishers or journalists [do so]’ (Ager 2001, 5). In this thesis, it exemplifies a background situation to TDP. See also *TDP*. See page 70.

**Language policy** is ‘official planning, carried out by those in official authority’ (Ager 2001, 5). In this thesis, it exemplifies a background situation to TDP. See also *language planning* and *TDP*. See page 70.

**Language-in-education policy** is a type of language policy, and one of several manifestations of acquisition planning in action. It chiefly addresses issues surrounding the selection of a language as medium of instruction in a given education system, as well as other questions concerning language teaching in schools. See also *acquisition planning*, *language policy* and *MOI*. See page 84.

**Linguistic community** refers to ‘a social group which may be either monolingual or multilingual, held together by frequency of social interaction patterns and set off from the surrounding areas by weaknesses in the lines of communication’ (Gumperz 1962, 32). It is generally employed in this thesis in the context of the Barcelona bilingual Catalan-Castilian linguistic community. See page 19.

**MOI (medium of instruction)** refers to the language(s) used for the delivery of classes in the education system of a given polity. See also *language-in-education policy*. See page 84.

**Status planning** is used to refer to instances of language planning which address ‘[a language’s] standing alongside other languages or vis-à-vis national government’ (Kloss 1969, 81). See also *acquisition planning*, *corpus planning* and *language planning*. See page 71.

**TDP (top-down phenomena)** encompass situations which give rise to the use of linguistic items or structures. Such language use is implemented above the level of consciousness, as the result of actions by an officially sanctioned organisation or body, which are driven by conscious decision-making and aim to approximate language usage rich in overt prestige. See page 59.

# Chapter 1

## Introduction

The vitality of civil society has turned [Catalan] into one of the most creative non-official languages in Europe. Today it is used in all domains, it is a scientific language with a continuously expanding output of neologisms, it has a rich literature, it is the main language of theatre, arts, universities and television, and it is making great inroads in the movies, as well as in the daily and weekly press. All of this was unthinkable not many years ago (Conversi 1997, 266).

[Catalan] is widespread and is in formal official daily use; in effect, [it] is the majority language within its regional area (Ager 1997, 59).

For a large part of the twentieth century, Catalan language users in Spain were subject to an extensive degree of oppression, and the above citations serve to highlight the fact that a corner has undoubtedly been turned since those dark days. However, the dawn of a new era for Catalan brings with it fresh challenges and new questions to be answered. The central aim of this thesis is to study language policy and language contact<sup>1</sup> in order to paint an accurate portrait of the contemporary sociolinguistic situation in Catalonia, and more specifically the city of Barcelona, providing an in-depth view of some of the issues faced by the linguistic community.<sup>2</sup>

---

<sup>1</sup>For full definitions of these terms, please consult the list of terms and abbreviations starting on page 16. In this chapter, discussion of ‘language policy’ does not necessarily exclude instances of ‘language planning’, but the term ‘policy’ is generally used as it is the clearest, simplest term available. Further discussion on the differences between policy and planning will appear in chapter 2.

<sup>2</sup>The term *linguistic community* will be employed throughout the thesis, and is to be interpreted

As we shall see in this chapter, the linguistic situation in present-day Barcelona is characterised by developments in language policy as well as an extensive degree of language contact between Catalan and Spanish. This project aims to provide a more complete picture of the current sociolinguistic reality of Barcelona than existing studies by addressing these two facets: policy and contact. This will be supported by a fieldwork experiment which should offer empirical evidence concerning the impact of language policy and language contact in this linguistic community. The experiment will answer a set of research questions, to be given at the end of chapter 3, which in turn will be based on the theoretical and practical findings concerning policy and contact given in the first three chapters. Specifically, the fieldwork will aim to ascertain the extent to which non-normative contact-induced linguistic items in both Catalan and Spanish figure in the conscious awareness of native bilingual speakers. This thesis will show that any cross-generational development in this awareness will allow conclusions to be drawn as to the effectiveness of certain language policy measures. All of this will be fully explained over the course of the first three chapters. In short, this thesis will tie together findings concerning language policy and language contact to allow researchers greater insight into the contemporary sociolinguistic situation in Barcelona. Any related issues (such as language change), while interesting, are beyond the remit of this work, which for practical considerations must limit itself to the examination of policy and contact.

This introductory chapter will set the scene for the thesis. Section 1.1 will provide an overview of the historical and sociolinguistic background of Barcelona, which will allow us to ascertain what exactly brought the Catalan language to the brink of such

---

as Gumperz's notion of 'a social group which may be either monolingual or multilingual, held together by frequency of social interaction patterns and set off from the surrounding areas by weaknesses in the lines of communication... [Linguistic communities] may consist of small groups bound together by face-to-face contact or may cover large regions, depending on the level of abstraction we wish to achieve' (Gumperz 1962, 32). This definition suits the Barcelona situation since the community can be multilingual, geographically demarcated and most importantly, defined by the languages used by its members. Moreover, the related term *speech community* is avoided, since this is often interpreted in a Labovian sense of a group which is 'not defined by any marked agreement in the use of language elements, so much as by participation in a set of shared norms' (Labov 1972, 120-1). This view of speech communities as not necessarily being determined by language use has been adopted in subsequent literature (cf. Romaine 2000, 22), and does not suit our purposes. *Linguistic community* will therefore be used.

despair, and reveal how speakers are attempting to fight back. It will conclude with a brief overview of current demographic linguistic data, highlighting any current issues that researchers may face. Leading on from this, section 1.2 will pose the question of *why* it is necessary to study both language policy and language contact in Barcelona. This will be done firstly through analysis of current literature which addresses Catalan sociolinguistics (in section 1.2.1), which will reveal that this work fills a key gap in the market. I will support this with theoretical claims (in section 1.2.2) that policy and contact comprise key elements of maximally differentiated phenomena – which will be termed *top-down phenomena* and *bottom-up phenomena* respectively – and that the study of a community where simultaneous, distinct linguistic forces are at work should prove of interest to future researchers. Section 1.3 briefly raises some practical questions concerning the requirements to be fulfilled by a fieldwork experiment that would test all of this, before section 1.4 offers a plan of the whole thesis.

## 1.1 Barcelona: historical and sociolinguistic background

Barcelona... és el centre cultural i polític més important [de Catalunya] i això fa que tots els avenços en aquests dos terrenys es notin més que a d'altres zones de Catalunya (Calsamiglia and Tusón 1980, 38).<sup>3</sup>

Aquesta és la idea que tenc dels valors innegociables. Els vull recordar: llibertat, cultura, terra i llengua. Si tenim l'actitud ferma de no renunciar a cap d'aquests valors i combinar-los amb una dosi suficient de senzillesa i creativitat, és possible sentir-te digne de viure en català (Majoral 2001, 114-5).<sup>4</sup>

---

<sup>3</sup>TRANSLATION: Barcelona... is the most important cultural and political centre [of Catalonia], meaning that all advances in these two fields are more readily perceived there than in other areas of Catalonia.

<sup>4</sup>TRANSLATION: This is my view on what the deal breakers are. Let me remind you of them: freedom, culture, land and language. If we remain firm in our resolve to not renounce any of these values and to combine them using the right amount of simplicity and creativity, it is possible to feel proud of living in Catalan.

For Catalan nationalists, language has always been the most important form of national distinctiveness (McRoberts 2001, 139).

The above citations reflect the commonly held view of Barcelona as the centre of Catalan culture and politics, and demonstrate how these are inextricably linked to the Catalan language. However, Barcelona finds itself within the borders of the Spanish state, and is Spain's second largest city: as such, the Spanish language (henceforth Castilian) also holds official status, placing Barcelona at the heart of a Catalan-Castilian bilingual community. The following historical and sociolinguistic overview will reveal how this state of bilingualism was reached, providing a detailed description of the cultural, political and social backdrop to the study. The section will be divided chronologically. The different historical periods will address diachronic sociolinguistic developments that are applicable to Catalonia as a whole, with a separate subsection accorded to Barcelona-specific facts. Finally, the present state of bilingualism will be accurately detailed through analysis of current demographic data which shows the sociolinguistic make-up of the city, and highlights some contemporary issues for discussion.

**Ninth to fourteenth centuries.** The Catalan language as we know it descended from the varieties of Vulgar Latin spoken in an area referred to retrospectively as *Catalunya Vella* (see figure 1.1, adapted from Ferrando Francés and Amorós 2005, 94). This area approximately corresponds to the eastern part of the *Marca Hispànica*, a buffer zone established in the late eighth and early ninth centuries by the Carolingians to the north in order to keep out invading Arab forces from the south. This led to the Catalan language developing relatively free of Arabisms (when compared to Castilian), and acquiring certain lexical and syntactic elements more akin to Gallo-Romance varieties, such as the *langues d'oc* and the *langues d'oïl*. This period saw the first instances of the endemic vernacular variety shifting from a purely oral to written medium, with the *Homilies d'Organyà*, composed around the end of the twelfth to the beginning of the thirteenth centuries, being one of the first written Catalan texts, as well as being among the earliest translations between Romance languages (from



Figure 1.1: Map situating *Catalunya Vella*, tenth to eleventh century (red), within current Catalan-speaking areas (light grey)

Provençal to Catalan) (Bastardas 1995, 100).

The thirteenth and fourteenth centuries saw yet further diffusion of the Catalan language and culture across many parts of the Mediterranean, from Sardinia – this is still evident in the Catalan-speaking community of l’Alguer – to Greece. This Golden Age of cultural expansion was later the source of great Catalanist pride inspiring such emotive discourse as:

[La llengua catalana] s’anticipà a ses germanes, les famoses castellana y portuguesa, en sa expansió ultramarina y s’extengué per totes les mars y penínsules del que era allavors nostre llach mediterrani, portada triomfalment per les ascones dels almogàvers, o les fitores dels almiralls (Rubió y Lluch 1908, 81).<sup>5</sup>

Moreover, this period saw the growth of a world-class Catalan literary tradition, with the works of great thinkers and writers such as Ramon Llull, Joanot Martorell and Ausiàs March establishing Catalan as ‘one of the great languages of medieval Europe’ (Costa Carreras 2009, 7). However, the fourteenth century brought with it a number of plagues and famines which decimated the Catalan population and coincided

<sup>5</sup>TRANSLATION: [The Catalan language] predated its sisters, the famous Castilian and Portuguese languages, in its overseas expansion and extended across all the seas and peninsulas which were to become the Mediterranean, brought triumphantly by the warriors’ lances or the admirals’ tridents.

with the end of the period of prosperity, exacerbated by the *Compromís de Casp* of 1412, which placed a member of the Castilian Trastámara dynasty on the Aragonese throne. The fifteenth century then bore witness to a generalised social and economic downturn, from a serious devaluation of the Catalan currency to outright civil war (1462-72), the outcome of which prompted the betrothal of Ferdinand of Aragon to Isabella of Castile, and as a result ‘Catalonia fell into Castile’s orbit’ (McRoberts 2001, 14). The linguistic consequences of all these events would prove understandably detrimental to Catalan: as stated, the ruling Trastámara dynasty were Castilian, and the unification of Aragon and Castile further cemented the presence of Castilian as the language of administration and government.

Barcelona-specific developments. Historically, this period was characterised by the annexation of other neighbouring territories (Tarragona 1116, Mallorca 1229-32, Ibiza 1235, Valencia 1229-45) and their subsumption under the County of Barcelona, and later the Crown of Aragon, with its chancellery seat at Barcelona (Nadal and Prats 1982, 127). Barcelona’s history as a focal point of Catalan activity is thus grounded in nearly a thousand years of tradition. Then, in the latter part of this period, Barcelona would come to feel a shift from Aragon to Castile, with the centre of power moving away from Barcelona to Toledo. There was a resultant increase in the importance of Castilian, although inhabitants of Barcelona continued to use Catalan as their sole, or at least primary, language

**The fifteenth to eighteenth centuries.** The subsequent centuries are often referred to as the *decadència*, a period of unprecedented decline in Catalan linguistic and cultural output. The increasing power of Castile and the progressive Castilianisation of the ruling Catalan bourgeoisie led to a decrease in prestige of the Catalan language, and its association with the lower classes (Amelang 1986, 190-3). This resulted in a concomitant rise in prestige of Castilian since ‘if the lower classes were restricted to the use of Catalan, those wishing to distinguish themselves as patri- cians were perforce obliged to use Castilian and/or Latin as well’ (Amelang 1986, 193-4). As a result, Catalan-Castilian bilingualism started to spread through the



general population, and by the mid-sixteenth century it was claimed that all men apart from those living in more isolated areas possessed some communicative skill in Castilian (Amelang 1986, 190). Moreover, the shift in the domains of usage of Catalan caused the firmly-established Catalan koiné of the fifteenth century to break apart, and awareness of the normative Catalan of the period rapidly decreased (Costa Carerras 2009, 8). Moreover, in this period, the Castilian language began a process of standardisation. In 1492, Nebrija's *Gramática de la lengua castellana* was released, this work usually considered the first grammar of a Romance vernacular. At this time, the importance and overt prestige of Castilian increased, its new codification going hand in hand with expansion into the New World; indeed, Nebrija described language as 'la compañera del imperio' (Pountain 2001, 122-3). Within Spain, 1492 was also the year of the final defeat of the Moors, and the Castilian language was being diffused to all corners of newly unified country in a nationalising wave. Castilian thus became the primary language of literature (here adhering to the prestige norm of Toledo rather than that of Seville) and education throughout the peninsula, including Catalonia (Penny 2002, 212). Arguably the nadir in this era of Catalan history came in 1714, when the end of the War of the Spanish Succession led the Bourbon Felipe V to pass the *Decrets de Nova Planta*, which abolished all Catalan institutions and public usage of the language. However, despite all these hardships and impositions, the Catalan language continued to be the primary vehicle of social interaction in Catalonia : 'Catalan remained, not as some antiquarian fossil, but as a living vernacular' (Hughes 1992, 189).

Barcelona-specific developments. As a city, Barcelona suffered from a demographic stagnation analogous to its linguistic and cultural decadence, its population only increasing by one fifth between the fifteenth and early eighteenth centuries (Amelang 1986, 8), and the city's inhabitants were the first to be touched by the incipient bilingualism spreading through the linguistic community.

**The nineteenth century into the twentieth century.** The early nineteenth century saw the arrival of a long-awaited cultural and literary Catalan reawakening,

known as the *renaixença* (Renaissance). Early inspirations for the movement were Bonaventura-Carles Aribau's 1833 poem *La Pàtria* (Fatherland), as well as Joaquim Rubió i Ors' 1841 poetry anthology, often viewed as the manifesto for the movement as it advocates a complete literary emancipation from Spain, arguably as a potential precursor to a more political independence (Conversi 1997, 13). The movement was influenced by the medieval Catalan Golden Age with poets referring to the language as *llemosí*, harking back to earlier Troubadouresque traditions, and in 1859 the medieval literary festival *els Jocs Florals* was revived in Barcelona after a hiatus of over four hundred years. However, by the late nineteenth century, it was widely appreciated that reminiscing on bygone glory days might not be effective in propelling the Catalan society into the modern age, and that the hitherto 'archaic, pedestrian and provincial' Catalan culture needed to be radically updated through 'national modes of creativity' (Costa Carreras 2009, 12). So, in 1886, Valentí Almirall published *Lo Catalanisme*, an early manifesto of a more forward-looking, political Catalanism, while literary, cultural and artistic output in this period also underwent transformations with the birth of the *Modernisme* movement. *Modernista* output changed the cultural landscape on an unprecedented scale, as exemplified by the architectural innovations of Antoni Gaudí, as well as the literature of Joan Maragall and Víctor Català. Catalan thus, for the first time in four centuries, returned to the status of literary language. Ultimately, however, *modernisme* was to provide no more than a springboard for the evolution of Catalanism, and indeed its ethos of change and upheaval proved 'unsusceptible to political channelling' (Costa Carreras 2009, 14); and thus out of *modernisme* was born the more abstemious *noucentisme* (1900s-ism), characterised by order and discipline, in an attempt to forge a political and nationalist Catalanism. Indeed, this period saw the creation of early Catalanist groups such as the *Unió Catalanista* and the *Lliga Regionalista* and in 1914, for the first time since the *Decrets de Nova Planta*, the inauguration of a Catalan governmental institution, the *Mancomunitat Catalana*. Linguistically, an important consequence of the cultural developments of the nineteenth century was the aforementioned advancement of Catalan from purely oral to literary language, and all concomitant increases in

prestige, however marginal. However, despite the ravages of the *decadència*, Catalan had always been the primary communicative vehicle for the inhabitants of Catalonia. What the new, literary language served to highlight was the lack of a standardised form of Catalan, with certain writers preferring varieties that approximated urban speech, and others adopting a rural style. Diatopic and diastratic variation in Catalan were still very much apparent, and no standard (or at least koiné) Catalan had existed since the Middle Ages. A further development, of particular importance to this study, was the increased attestation of Castilianisms in Catalan speech as an example of contact-induced language phenomena.<sup>6</sup>

Barcelona-specific developments. During this period, the previous economic and demographic downturn felt by the city of Barcelona began to be reversed in earnest. Between 1834 and the mid 1850s, the population of Barcelona increased by forty per cent, reaching a total of 189,000, a growth significantly larger than that felt in other towns and cities of the region (Hughes 1992, 254), this being linked to the development of (the mainly textile) industry. This unprecedented increase in population, along with the fact that Barcelona was fast becoming a lively cultural centre thanks to *renaixença*, *modernista* and *noucentista* artistic production, as well as the growth of political Catalanism, cemented the city's position as an increasingly important hub of activity. Moreover, the aforementioned contact-induced language phenomena (i.e. the increased attestation of Castilianisms in Catalan) in the context of nineteenth century Barcelona was heavily imbued with social meaning and notions of language prestige (see section 1.2). Wealthy *barcelonins* were aware that the market for their industrial output could only be accessed through the rest of Spain. They thus spent increasing amounts of time in Madrid, speaking Castilian, and upon returning to Catalonia 'larded their Catalan with *castellanismes* to show that they were people of the better sort' (Hughes 1992, 375). The status of Barcelona as an industrial powerhouse of increasing importance in this period thus had linguistic consequences, namely the ever-growing presence and prestige of Castilian and the numerous ways in which this

---

<sup>6</sup>See section 3.1 for discussion of contact-induced language phenomena, and section 3.2.2 for specific discussion of Castilianisms in Catalan. Moreover, definitions of these terms can be found in the list of terms and abbreviations starting on page 16.

manifested itself in the language practices of members of the linguistic community.

**The early twentieth century: language.** The cultural metamorphosis, inclination towards nationalist sentiment and growth of industry, all witnessed in Barcelona in the nineteenth and early twentieth centuries allowed for a gradual stirring up of the Catalan intelligentsia, which in turn impacted upon the linguistic situation. In 1906, the First International Congress on the Catalan Language took place in Barcelona: academics from around the world gave speeches and presentations on a wide variety of topics linked to the Catalan language. Among their number was Pompeu Fabra, a chair of chemistry at the School of Engineering, Bilbao, a man who would go on to standardise and revolutionise the Catalan language, largely creating normative Catalan as we know it. In 1907 the *Institut d'Estudis Catalans* was founded, and promptly set about the standardisation of the Catalan language, with Fabra as their chief proponent.<sup>7</sup> Fabra's output took many forms, namely:

Letters, dictionaries, speeches, interviews, responses to others' contributions to the 1906 Congress, grammars, manifestos, miscellanies on grammatical correctness, prologues, reviews of lectures, postcards and transcriptions of lecture courses. (Costa Carreras 2009, 33).

Of particular interest are Fabra's *converses filològiques*, a total of 915 short articles addressing numerous issues pertaining to Catalan language standardisation, many of which were published in the 1920s in newspapers such as *La Publicitat* (Rafel 2010, 77-8). These remain of great interest today and will be examined in more depth on page 91 when we study the specific history of Catalan language policy. Fabra's codification work proposed standards for Catalan syntactic constructions, morphology and lexis, such as the *Diccionari ortogràfic* of 1917, which was essentially the full list of lemmas of the *Diccionari general de la llengua catalana*, published fifteen years later. Fabra and colleagues investigated the etymology of these items while also taking into account diatopic variation in their pronunciation, in order to arrive at a prescriptive norm for each item, thus resolving problems of whether Catalan words

---

<sup>7</sup>For a full discussion of the theory behind language standardisation, see section 2.1. For a more in-depth discussion of Catalan language standardisation, see section 2.2.

were to be spelt with *b* or *v*, with *ç*, *s* or *ss* (Costa Carreras 2009, 141-2). These norms received immediate backing from Catalan regional administration and could soon be witnessed in the literary output of writers and journalists (Costa Carreras 2009, xviii). In 1932, Fabra's aforementioned *Diccionari general de la llengua catalana* was published and indeed, 'within this period of just twenty years the victory of Fabra's idea of a modernised and standardised Norm for Catalan was almost total' (Costa Carreras 2009, xxvi). Fabra's importance is undoubted and his work shall receive greater attention in chapter 2.

Barcelona-specific developments. While taking into account regional and dialectal variation, when developing a normative standard, Fabra's work evidenced a certain *barcelonisme*, preferring the vernacular of Barcelona to that of other regions. Costa Carreras justifies this by making reference to Barcelona's unequivocal position at the centre of the Catalan linguistic universe (Costa Carreras 2009, xxvi). So throughout this period, Barcelona was playing a central role by offering Fabra and colleagues a linguistic model which served as the foundation for their newly-standardised Catalan. Thus, Barcelona's role at the heart of Catalan cultural, linguistic and political matters continued apace at an important time of growth, standardisation and normativisation of the Catalan language.

**The early twentieth century: politics.** While the Catalan language flourished and progressed at the turn of the twentieth century, the political situation both in Catalonia and Spain as a whole was tumultuous. Decades of unrest and instability characterised the first third of the twentieth century throughout Spain, and during the first twenty-one years of Alfonso XIII's majority (1902 to 1923), sixteen different Prime Ministers (*Presidentes del Consejo de Ministros de España*) served thirty-two different terms of office. Continued social and political volatility resulted in a military dictatorship (1923-30) headed by General Miguel Primo de Rivera under the auspices of the monarchy of Alfonso XIII. While this dictatorship was encouraged by the more right-wing of Catalanist organisations at the time, such as the *Lliga Regionalista*, it soon became apparent that Primo de Rivera was motivated by centralist ideologies

intent on diminishing Catalan autonomy. Within days of the declaration of the dictatorship, laws were passed which banned the use of the Catalan language in public settings, all Catalan language publication was subject to government censorship and within two years, the *Mancomunitat Catalana* had been abolished altogether. However, such oppression may not have been as detrimental to the Catalanist cause as one may think: in the face of a common enemy, the Catalan intelligentsia was galvanised into action, and literary output in Catalan increased, making extensive use of the new Fabrian standard. Most importantly, the dictatorship ultimately resulted in the declaration of the liberal Second Spanish Republic (1931-36), under which the Catalanist cause flourished, particularly in the initial *bienni reformista* period. The Catalan *Generalitat* (regional government) was re-established, an Autonomy Statute (1931) was ratified and approved, and progress was made to Catalanise and modernise the education system. On a linguistic-cultural level, Catalan book and newspaper production more than doubled in a period of six years, the presence of Catalan on the radio increased and Fabra's *Diccionari General de la Llengua Catalana* appeared (Costa Carreras 2009, 20-1). However, the Republic was beset with problems: the *bienni reformista* was followed by the *bienni negre*, a period of right-wing austerity, and then finally in 1936, power swung radically back to the left with the victory of the *Frente Popular*. Such constant political vacillation culminated in the Spanish Civil War (1936-9), sparked by the revolt of the Spanish garrison in Morocco headed by General Francisco Franco. In Catalonia itself, the mantle was taken up by the CNT-FAI (*Confederación Nacional del Trabajo - Federación Anárquica Ibérica*) anarchist confederation, in response to the political upheaval of the time. The war raged for just under three years, claiming approximately half a million lives and resulting in a Nationalist victory and ultimately, the Franco dictatorship.

Barcelona-specific developments. Throughout this period of instability, the city remained the nerve centre of Catalonia. Unrest was felt particularly acutely in the summer of 1909, during the *Setmana Tràgica*, a week of chaos, rioting and widespread violence, chiefly in response to extensive conscription of citizens to fight in the unpopular war in Morocco, and resulting in the death of at least seventy-five civilians.

The city's status as an industrial powerhouse arguably contributed to the unrest, with a highly developed working class who could be mobilised into left-wing protest (Conversi 1997, 29). In the first days of the Second Spanish Republic, the Catalan leader Francesc Macià hastily declared an independent Catalan republic as part of an Iberian federation from the balcony of the Palau de la Generalitat de Catalunya in Barcelona's Plaça de Sant Jaume, while at the end of the Civil War, the 'conquest' of Barcelona led to the dissolution of the Generalitat and the execution or exile of all Catalanist leaders (Conversi 1997, 41-2). Moreover, this was a period of exponential demographic growth for the city. According to the Instituto Nacional de Estadística, the city's population rose from 544,137 to 1,081,175 between 1900 and 1940. When viewed alongside demographic figures for the whole of Catalonia, this increase represents a gradual move towards urbanisation, since in 1900, 27.68% of Catalonia's population lived in Barcelona, shifting to 37.41% forty years later. Barcelona thus maintained its position as the focal point for Catalan culture, and politics, and as the pole towards which increasing numbers of Catalans gravitated.

**The Franco dictatorship (1939-75).** This marked a nadir in the history of the Catalan language. The regime was characterised by the brutal oppression of all those viewed as enemies of the state. Franco's slogan of *España: una, grande y libre* dismissed as anti-Spanish anyone who was deemed to promote any values of regions such as Catalonia. People with anti-Spanish or anti-regime political or cultural views were seen as a threat to Franco's monocentric concept of Spanish national unity, and thus were viable targets of his harsh, oppressive measures. Solving 'the Catalan problem' was a central preoccupation of Franco's dictatorship from the very outset. Leading Catalan political and intellectual luminaries either fled the country, were forced into exile or were executed. Between 1 and 6 May 1939, 266 people were executed by the regime, and it is estimated that between 60,000 and 100,000 Catalan citizens went into exile (McRoberts 2001, 40). The Catalan autonomy statute was outlawed and it became illegal to have ever been a member of Republican parties. This early political and social persecution was accompanied by a process of 'cultural genocide': the

Catalan flag was outlawed, as was public use of the language (public sector employees were faced with dismissal for speaking Catalan), the *Institut d'Estudis Catalans* (IEC) was presumed to have ceased to exist and Catalan completely disappeared from the education system. This was in addition to the extensive propaganda which instructed citizens 'no ladres: habla la lengua del imperio',<sup>8</sup> the censorship of private mail (with correspondence written in Catalan destroyed) and even the obligation for Catalan family names to be translated into Castilian (McRoberts 2001, 41). Such harsh oppression obviously provoked feelings of resistance, which for the most part had to remain clandestine. The 1940s witnessed a gradual decline in nervousness about speaking Catalan, the secret reorganisation of the IEC in 1942, and the slow resurgence of cultural traditions, such as dancing the *sardana*. This last development was due to the official line of the regime as maintaining 'a paternalistic attitude towards folkloric representations of regionalism' (Costa Carreras 2009, 21-22) but did not stretch to manifestations as potentially powerful as language. Franco's overt linguistic policies stressed that Catalan was not a real language, but merely a variety of Spanish, and school textbooks listed the 'major dialects' of Spanish as Catalan, Valencian, Mallorcan and Galician (Hughes 1992, 9). However, as time progressed, subtle relaxation towards the use of Catalan increased. Publishing in Catalan very slowly grew (no books in 1939, six in 1942, thirteen in 1950) (Hughes 1992, 31), but in line with Franco's dialectalisation policies, these did not appear in standard, Fabrian Catalan.<sup>9</sup> The outcome of the Second World War put pressure on Franco to lessen extremist oppression of the people (Costa Carreras 2009, 21), and the *milagro económico* of the 1960s and 70s meant that all eyes were on Spain, and persecution of Spanish citizens by its government could not continue to the same extent (or at least not on such a public level) under the watchful gaze of the international community, which was now investing in Spain's future. Cultural and political developments beneficial to the Catalan language gained momentum in this period. The *Nova Cançó* movement of the 1960s saw an emergence of a new type of popular songs in Catalan by artists such as Lluís Llach and Joan Manuel Serrat. Edicions 62 was established

---

<sup>8</sup>TRANSLATION: Don't bark [like a dog]: speak the language of the empire.

<sup>9</sup>Further consequences of Franco's overt linguistic policies will be discussed in section 2.2.1



in 1962 with an aim to publish paperback versions of the most influential Catalan literary texts. Moreover, while official use of Catalan was still prohibited, 1967 saw the Universitat de Barcelona set up a department of *filologia catalana*, where Catalan was taught as if it were a foreign language, this scenario being applied to selected secondary schools in 1970. However, in spite of all these marginal, gradual concessions accorded through the course of the regime, there can be little doubt that the Franco dictatorship was massively detrimental to the Catalan language, and in 1975, Catalan was in a greatly weakened position with regard to its status and the number of speakers within its traditional territory, as well as the individual speakers' abilities in and attitudes towards the language.

Barcelona-specific developments. Arguably more subtle, social strategies were also employed to weaken the situation of Catalan within Catalonia, especially in Barcelona and its environs. Mass immigration from other parts of Spain (particularly the south) into the industrial centre of Barcelona was heavily encouraged by the regime, the result being a great influx of workers who could not speak Catalan, and had no access to the language. Integration between these *xarnego* migrant labourers and native Catalan-speaking *barcelonins*, whose day-to-day language remained Catalan in the majority of domains, was thus difficult due to the linguistic divide (Conversi 1997, 208). Moreover, the fact that migrants typically filled lower-level positions meant that the two groups were also compartmentalised along class lines. Nonetheless, the massive numbers of monolingual Castilian-speaking migrant workers flocking to the city did not benefit the Catalan language situation. In order to increase its vitality, drastic measures were needed to repair the damaged state of the Catalan language at the end of the dictatorship in 1975.

**From 1975 to the present day.** The post-Franco transition to democracy has transformed Spain from a centralist dictatorship into a devolved parliamentary monarchy. The 1978 Constitution established the fundamental rights of Spanish citizens as members of a modern European democracy, and in its Title VIII, addressed the issue of *la Organización Territorial del Estado*. Power was devolved to seventeen *comunidades*

*autónomas* and special rights and privileges were accorded to areas deemed *nacionalidades históricas*, initially understood to be the Basque Country, Catalonia and Galicia (Conversi 1997, 144). Catalonia’s first autonomy statute was ratified in 1979, its article 3 declaring that ‘la llengua pròpia de Catalunya és el català’.<sup>10</sup> This paved the way for specific linguistic legislation, most notably the *Llei de Normalització Lingüística* (LNL) of 1983 and the *Llei de Política Lingüística* (LPL) of 1998. The full aims and content of these items of language policy will be detailed in depth in section 2.2.1, but for our present purposes, suffice it to say that 1983 was a very important year, since the LNL made provisions for the establishment of Catalan-medium primary and secondary education. Indeed, 1983 can be seen as the watershed for Catalan language policy.<sup>11</sup> Language policy is still of prime importance to the Catalan people and remains to this day at the very centre of the national psyche. On 10 July 2010, Barcelona saw the largest demonstration in Catalan history (estimates range from 1.1 million to 1.5 million attendees). People turned out to protest the ruling of the Constitutional Tribunal which deemed several articles of the 2006 Autonomy Statute unconstitutional; of central concern was the overturning of articles which listed Catalan as the preferred language of administration and education, as well as a countermand stating that it was not a duty of Catalan citizens to know the Catalan language.

Barcelona-specific developments. This most recent protest shows that Barcelona is still the centre of all things Catalan. Furthermore, non-linguistic developments have also shaped Barcelona’s sociocultural landscape in the last thirty years: the 1992 Olympic Games saw mass investment in the city and a marked improvement in infrastructure, while the general growth of Barcelona as a tourist centre (both due to its own cultural attractions and to its proximity to the Costa Brava) translates into a total of a 21.7 million visitors to the city in a twelve month period in 2009-10 (Castellsaguer 2010).

---

<sup>10</sup>TRANSLATION: Catalan is the language of Catalonia.

<sup>11</sup>Indeed, this date will serve as the primary marker in delimiting the two age groups in my forthcoming experiment.

**Current issues surrounding Barcelona’s sociolinguistic make-up.** Barcelona thus has a long history of Catalan-Castilian bilingualism, and current surveys conducted by IDESCAT (the Catalan institute for statistics) reveal the extent to which each language is used by the population of Catalonia.<sup>12</sup> A whole wealth of statistics is available concerning the usage of the official and non-official languages in Catalonia: for example, how often each language is used and with whom, as well as detailing attitudes towards each language, and which language citizens ‘identify’ with. In 2008, this survey data revealed the rates of claimed language competence<sup>13</sup> in each of the two official languages to be as follows:

	Can understand	Can speak	Can read	Can write
Catalan	94.6 %	78.3%	81.7%	61.8%
Castilian	99.9 %	99.7%	97.4%	95.6%

Table 1.1. Catalan and Castilian language competence in the population of the autonomous community of Catalonia (IDESCAT 2008a, IDESCAT 2008b, IDESCAT 2008c).<sup>14</sup>

Knowledge of both Catalan and Castilian is clearly widespread, but the complementary statistics on language use show that people do not necessarily conduct their lives bilingually, and nor do they identify as bilinguals first and foremost. Such information is telling as it allows us to see to what extent this seemingly homogeneous

<sup>12</sup>The statistics examined in this section are for the population of Catalonia as a whole, and not just Barcelona. This data was chosen as it was analysed along a great number of diverse variables (such as ‘language of identification’). Moreover, this section has been a discussion the sociolinguistic background of Catalonia as a whole, as well as Barcelona, and so the inclusion of these statistics is justified.

<sup>13</sup>The term ‘competence’ has been employed since it is widely used in applied linguistics and foreign language teaching to refer to (normally L2) users’ abilities in the four recognised skill areas of reading, speaking, listening and writing. This is also how the term is used in the European Indicator of Language Competence (cf. Commission of the European Communities 2005) and the Common European Framework of Reference for Languages (CEFR). As such, the term does not refer to the Chomskyan generativist notion of *linguistic competence* which addresses speakers’ innate, cognitive faculties concerning language, and is contrasted with *linguistic performance* which deals with how language is used for communicative purposes.

<sup>14</sup>Raw data taken from tables *població segons coneixement del català* and *població segons coneixement del castellà*. Percentages then calculated based on total population figures provided in annexed data, such as *població segons llengua d’identificació*.

bilingual community is actually made up of people who consider themselves members of different linguistic groups, which may have different social and political interests. The 2008 survey data concerning ‘language of identification’<sup>15</sup> is as follows:

Castilian	46.5%
Catalan	37.2%
Both	8.8%
Arabic	2.4 %
Other	5.1%

Table 1.2. Language of identification of the population of the autonomous community of Catalonia (IDESCAT 2008c).

At first glance, this data seems to reveal that there are more citizens of Catalonia who primarily identify as Castilian speakers than Catalan speakers. However, the reality is considerably more complex, and when the above results are classified according to the place of birth of the population, a different image emerges, as can be seen below:

Place of birth	Castilian	Catalan	Both	Arabic	Other
Catalonia	29.2%	58.9%	11.7%	0%	0.1%
Elsewhere in Spain	84.1%	8.7%	6.5%	0%	0.7%
Outside of Spain	52.6%	3.5%	2.3%	14%	24.1%

Table 1.3. Language of identification of the population of the autonomous community of Catalonia, classified according to place of birth (IDESCAT 2008d).

Speaker place of birth is clearly a key factor, with those born within Catalonia more readily identifying as Catalan speakers than those born elsewhere. This could be attributed to the fact that a large number of residents of Catalonia were brought

---

<sup>15</sup>As shall be shown in section 5.1.5, ascribing one language or another to a given participant can be highly complex. In the IDESCAT data, the term ‘language of identification’ is simply taken to mean ‘the language the participant considers his/hers and identifies as his/her own.’

in as a labour force from other parts of Spain during the Franco regime (the aforementioned *xarnego* migrants). It could also be attributed to the fact that currently a great number of immigrants have settled in Catalonia, having come from outside of Spain. The number of citizens born outside of Spain who identify themselves as Castilian speakers points to the high level of immigration from other Hispanophone countries, while the presence of Arabic in the above tables indicates a considerable influx from North Africa and the Middle East. The acculturation of this new immigrant population, both socially and linguistically, is a key concern for the current Catalan administration. The Generalitat de Catalunya (specifically the Secretaria de Política Lingüística in conjunction with the Secretaria per a la Immigració) offers many online resources designed to increase immigrants' competence in Catalan, with such materials being available in many languages, including Arabic, Romanian, Tamazight, Chinese and Urdu, reflecting the diverse origins of the new immigrant population. There is a new body of scholarly work which seeks to address how other linguistic and cultural communities are to be considered as part of the Catalan sociolinguistic situation (cf. Sepa Bonaba 2000 for discussion of Sub-Saharan African immigrants, Ouakrim 2000 for discussion of the Tamazight-speaking community, and Golden 2000 for discussion of Chinese-speaking immigrants, inter alia), and this thesis does not address such issues due to considerations of space. But it should be noted that the acculturation of immigrant communities in Catalonia, whatever their linguistic background, is a key issue in contemporary Catalan sociolinguistics.

Another issue surrounding Catalan and Castilian language use in Catalonia is that of intergenerational language transfer: to what extent is each language being passed from parent to child? Are different generations exhibiting different tendencies, as regards the language used with their parents?<sup>16</sup> The IDESCAT linguistic statistics shed some light on such issues, as seen in the following table:

---

<sup>16</sup>For more details as to how examining the behaviour of different generations allows us to draw conclusions about developments in apparent time, see section 4.1.2.1.

Age group	Catalan	More Catalan than Castilian	Both equally	More Castilian than Catalan	Castilian	Other
15-29	30.1%	2%	2.2%	2.7%	49.6%	13%
30-44	26.5%	1.4%	1.7%	1.5 %	56%	12.2%
45-64	31.8%		1.4%	1.2%	58.9%	5.7%
65+	40%		1.5%		51.8%	2.7%

Table 1.4. Language of use with parents,<sup>17</sup> classified by participant age group. Gaps represent unavailable data (IDESCAT 2008e, IDESCAT 2008f).

At a cursory glance, the above table reveals a number of interesting points worthy of further discussion. Younger participants are more frequently using a language other than Catalan or Castilian, which is in keeping with the aforementioned increase in immigration from non-Hispanophone countries. Moreover, there seems to be remarkably little difference between generations concerning intergenerational transfer in Castilian, though this could be attributed to immigration from Hispanophone countries (i.e. that the younger participants who are speaking to their parents in Castilian are from outside of the Catalan-speaking area). However, younger generations seem to be using Catalan with their parents a lot less than their older counterparts used to. These statistics, while interesting, raise a great many questions, not only about intergenerational language transfer, but cross-generational language use in general. For example, are different generations exhibiting different patterns of language use? While a full examination of this issue is beyond the scope of this thesis, we shall see that the upcoming fieldwork experiment does go some way to addressing how each of the official languages is potentially used by members of different age groups.

Finally, it is clear that centuries of Catalan-Castilian bilingualism have resulted in the widespread knowledge of both languages within the community, but it does not follow that these two languages are of equal status or function. Indeed, the afore-

---

<sup>17</sup>This table consists of average percentages compiled from two IDESCAT tables: *població segons usos lingüístics amb la mare per grups d'edat* and *població segons usos lingüístics amb el pare per grups d'edat*. For this reason, and due to abstentions, the rows in the table do not exactly total 100%.

mentioned Francoist oppression dictated that the only official language be Castilian, meaning that Catalan was relegated to low-function, interpersonal use. This led to a situation of *diglossia*, whereby one language (termed in Ferguson's original 1953 article as the H-code, cf. Ferguson 1953, 34-5) fulfils high-prestige, official roles, while the other language (the L-code) serves lower-prestige functions. Franco's dictatorship firmly established Castilian as the H-code with Catalan as the corresponding L-code. However, the Generalitat has implemented measures that clearly aim to rectify this previous imbalance. The current Catalan autonomy statute still dedicates numerous measures to the 'promotion and diffusion of Catalan' (Parlament de Catalunya 2006, article 50) in light of the language having previously been overlooked, be that in the education system, the media, local administration, or other sectors.<sup>18</sup> While the current sociolinguistic reality of Barcelona viewed as a post-diglossic scenario is undoubtedly an interesting approach for future research, this thesis will focus directly on contact and policy, rather than returning repeatedly to the notion of diglossia. However, the functional imbalance of Castilian and Catalan as enforced by Franco, and resultant democratic legislation aiming to manipulate this societal language dominance configuration, is going to prove of importance, and should always be borne in mind. The roles of each language during the twentieth century, and how that has impacted on policy (and to a lesser degree, contact) will be of importance in the upcoming chapters.

In short, the current state of Catalan-Castilian bilingualism in Barcelona, and Catalonia as a whole, invites a great deal of interesting sociolinguistic discussion. The wealth of IDESCAT linguistic census statistics allow insight into many facets of language use, such as whether community members consider themselves Catalan-speakers, Castilian-speakers or indeed bilinguals. We have seen that many factors can condition this, not least the fact that many residents of Catalonia are of immigrant origin, and the social and linguistic acculturation of these people is a present-day challenge faced by the Generalitat. Geographical origin is not the only important factor, however, and this brief glance at statistics concerning family language use has

---

<sup>18</sup>All of this shall receive greater attention in chapter 2.

shown differences in trends between generations of speakers, and such an examination of older and younger participants may prove revealing in the upcoming fieldwork experiment. Finally, given the historical nature of this overview, the (comparatively) recent Francoist measures greatly limited the functions of Catalan, and created a situation of *diglossia*, wherein Castilian assumed all high-prestige roles. The current situation, if it is to be viewed in light of its context, is in many ways a reaction to this, with legislation aiming for the more equal functional distribution of codes, though more details of the linguistic legal framework are to be given in chapter 2.

\* \* \*

Barcelona has thus found itself firmly at the centre of the Catalan-Castilian bilingual community for the best part of a millennium. It has witnessed the rise of a Catalan-speaking empire followed by its fall and decline. Later cultural resurgence made nineteenth- and twentieth-century Barcelona once again a vibrant artistic centre, but this was soon to be brutally crushed and repressed by the Franco dictatorship of the mid-twentieth century. Since 1975, linguistic legislation has contributed to an increase in status of the Catalan language, while Barcelona's increasing appeal as a tourist destination ensures that all eyes are on the Catalan capital. Issues pertaining to language are still very close to the hearts of *barcelonins*, and during my fieldwork collection phase, over a million Catalans took to the streets of the city's central boulevards of Passeig de Gràcia and Gran Via de les Corts Catalanes in defense of their linguistic rights. It is of little wonder that these people are driven to such action, in light of the turbulent sociolinguistic history of Catalan. Indeed, as supported by the demographic data, we now find ourselves in a new period for the Catalan language, with new challenges to face. Having looked at the past, and having started to address contemporary issues, this thesis will continue to throw light on aspects of the current linguistic situation in Barcelona.



## 1.2 The importance of language policy and language contact in Barcelona

As stated, this thesis chooses to focus on language policy and language contact in present-day Barcelona, in order to paint a detailed picture of the contemporary sociolinguistic situation. We have already examined the historical background and context which has led to the current state of Catalan-Castilian bilingualism in the city, and already sketched the linguistic demographic make-up of Barcelona, as well as some of the resultant issues. It is now time to state why the study of language policy and language contact are of particular relevance to this linguistic community.

Section 1.2.1 will offer a brief overview of existing sociolinguistic studies of Catalonia, highlighting the importance of both policy and contact to this particular community. It will be revealed that the present study, in addressing both policy and contact, fills a noticeable gap in the extant body of literature on Catalan sociolinguistics. Moreover, I intend to show that there are considerable theoretical implications for a study which examines both policy and contact. Section 1.2.2 will discuss how language policy and language contact exemplify key components of two, maximally differentiated phenomena that can bear upon a given sociolinguistic situation. This will be illustrated through a three-dimensional model, to be explained in due course.

Therefore, this section will show that the study of language policy and language contact in Barcelona should prove highly informative. This is not just because these phenomena are present in the linguistic community under examination (to be shown in section 1.2.1), but also because a study which takes both of these factors into account may allow us to ascertain more information about sociolinguistic processes in general, due to theoretical implications (to be shown in section 1.2.2). This would therefore allow any findings to be potentially generalised beyond the linguistic community of Barcelona.

### 1.2.1 Policy and contact: Existing studies of Catalonia.

This section will explain why the chosen foci of this thesis, namely language policy and language contact, are relevant and informative when examining the bilingual Catalan-Castilian linguistic community of Barcelona. Firstly, by drawing on the historical background already given, I will underline the importance of the presence of policy and contact in this community. Then I will provide a brief overview of extant literature on Catalan sociolinguistics, showing how work addresses either policy or contact, but never both satisfactorily. Thus, the foci of this work will allow it to fill an existing gap in research on Catalan sociolinguistics.

The historical overview in section 1.1 tells us that there have been radical developments in policy, both general and linguistic, since 1975. The centralist Franco dictatorship decreed that Castilian alone should be the official language of the entire territory of Spain. However, the 1978 Constitution and subsequent linguistic legislation enforced measures ensuring recognition and status for languages such as Catalan. The last thirty years have thus seen a great shift in language policy. Moreover, section 1.1 reveals that Barcelona (and indeed, Catalonia as a whole) has been characterised by a degree of Catalan-Castilian bilingualism since the *decadència* of the fifteenth to eighteenth centuries, if not before. The extent to which this bilingualism is currently the case is supported by the demographic statistics concerning Catalan and Castilian language competence, to be found on page 35. The fact that the members of this community are highly competent in both languages indicates that there is a degree of contact between the two languages, if we follow the widely-held definition of language contact given by Weinreich in 1953, namely that ‘two or more languages [are] in contact if they are used alternately by the same persons.’<sup>19</sup> We can thus see how the contemporary situation in Barcelona is characterised by both developments in language policy and a degree of language contact.

Literature to date detailing the sociolinguistic situation in Catalonia has reflected the importance of policy and contact, and there is a wealth of material available. However, until recently, the body of literature examining language policy and/or

---

<sup>19</sup>This definition of language contact will be examined in greater depth in section 3.1.

contact-induced phenomena in Catalonia often fell into one of several traps:<sup>20</sup>

1. Studies which attended to contact-induced phenomena did not take the policy situation into account when analysing their data. As such, they concentrated on micro-level speaker practices, not placing emphasis on the macro-to-micro (i.e. societal-level to speaker-level) transfer. These works may take the form of classificatory exercises, aimed at examining the characteristics of language use in the area (cf. Szigetvari 1994, Seib 2001, Barber Casanovas 2007). Alternatively, they may be in-depth studies which make little to no reference to top-down developments in language policy (cf. Hernández García 1994, Turell Julià 1994, Torres 2003, Vila i Moreno 2003), or indeed works which do outline the language planning situation, but simply as an element of the historical background of the case study (cf. Grupo Interalia 1998). My work will aim to examine the consequences of the presence of both policy and contact-induced phenomena, thereby potentially addressing the thus-far neglected macro-micro relationship.
2. Studies which primarily focused on language policy and planning were largely descriptive and again, did not examine the macro-to-micro transfer which would relate policy advances to contact-induced phenomena. Such work would basically chart the progress of Catalan language policy and speculate as to where it might be heading (cf. Sabater 1984, Nymark 2000, Pitarch i Almela 2002, Segura i Ginard 2004 *inter alia*).
3. There is some work which does, to a degree, address both the consequences of policy decisions (through the related concern of overt prestige) and their impact on individuals' language use. Such work therefore arguably addresses the link between macro-level decision making and micro-level speaker phenomena. However, this work does not, or cannot, offer current academics a complete picture of the situation, since it talks about Catalan and Castilian as two immovable entities: speakers either use one language or the other. These studies

---

<sup>20</sup>It should be noted that a full review of all relevant literature will form part of chapters 2 and 3. Some works are outlined and cursorily critiqued here merely to demonstrate the shortcomings of existing work addressing policy and contact, and by extension, the relevance of my thesis.

talk of language *choice*, and examine what factors may condition a particular speaker's decision to use one language or the other. Thus, Catalan is viewed as one system and Castilian another, and any interaction between them, and any impact on the linguistic-internal structure of either system, is overlooked (cf. Pujolar 1997). Such discussion treats Catalan and Castilian as polar opposites, and does not address contact-induced phenomena. So, while there is work which discusses how overt prestige (which is at least in part dictated by what policy decrees to be normative) influences speaker language use, contact is not considered. Therefore, once again, policy is (indirectly) considered, but not contact.

4. Other studies are of somewhat limited use simply by virtue of when they were conducted. Calsamiglia and Tusón's 1980 work examines the speaker-driven phenomenon of code-switching, framing the situation with information about new and prospective developments in language legislation, thereby addressing both policy and contact (although not in equal measure). However, this study is over thirty years old and in need of updating as it was conducted before important *normalització* legislation was implemented. Kathryn Woolard's 1989 study of speaker-level attitudes was conducted when language policy in Catalonia was in a similarly embryonic stage.

In light of the potential shortcomings of the until recently extant body of work as detailed above, it is of little surprise that academics, both independently and collectively, have arrived at the conclusion that the current sociolinguistic situation in Catalonia is in dire need of reanalysis. Moreover, the research community currently finds itself in a prime period to undertake research, given the amount of time that has now elapsed since the first Catalan linguistic legislation was put in place. As will be outlined in chapter 2, the LNL of 1983 allowed for the establishment of Catalan-medium education in Catalonia. As such, we are now blessed with a whole generation of Catalans not only schooled in the Catalan language, but who have now left full-time education and entered the workforce. It is therefore possible to witness the full consequences of schooling conducted entirely through Catalan not only for those

students within the system, but for those young adults who have passed through Catalan-medium education in its entirety and have progressed to employment in a range of fields. Thus, we currently find ourselves in an opportune moment to examine the Catalan linguistic community, not only in terms of the contribution of current research to the fields of Catalan and Hispanic Studies, but also since such investigation has great potential to inform the evaluative study of language policy and planning. Such fortuitous timing has not gone unnoticed by the wider research community. Indeed, Noemí Ubach Codina points out that ‘el 2005 féu vint-i-cinc anys de l’inici de la immersió lingüística a Catalunya i vam creure que era un bon moment per fer-ne un balanç’ (Ubach Codina 2009, 15).<sup>21</sup>

Since the inception of this thesis, a spate of academic output seeking to evaluate Catalan language policy, and specifically language-in-education policy has been released. The twentieth edition of *Treballs de Sociolingüística Catalana* (2010) was specifically dedicated to the topic. In May 2011, the collection *Democratic Policies for Language Revitalisation: the Case of Catalan* was released, comprising a series of articles which broach various topics concerning the recent policy developments felt by the Catalan language in all areas where it is spoken. This encompasses a thorough review of language-in-education policies to date (Vila i Moreno 2011), the role of Catalan in the legal system (Vernet and Pons 2011) and the sphere of business and consumer affairs (Branchadell and Melià 2011).

My work will form part of this ‘new wave’ of scholarly output. Arguably, it is studies such as Calsamiglia and Tusón (1980) and Woolard (1989) that current academics wish to recast, as they take into account both government-level decision making (i.e. policy) and speaker-level phenomena (e.g. contact). Such a holistic approach clearly merits further investigation, given the current situation in which Catalan finds itself, thanks to language policy having now had thirty years to impact upon Catalan society. Recent studies have been published regarding the use of Catalan and Castilian in the classroom and how this may (or may not) reflect children’s actual use of the languages for interpersonal communication. However, these works, while considering

---

<sup>21</sup>TRANSLATION: 2005 marked twenty-five years since the start of linguistic immersion in Catalonia and we thought that it was a good time to evaluate its progress thus far.

individuals' language use as a result of government language legislation, thereby exploring a facet of the macro-to-micro transfer, often fall into one of the traps outlined above, namely that they continue to treat Catalan and Castilian as two immovable entities, without investigating how the one language can impact on the form of the other (cf. Vila i Moreno and Galindo Solé 2009, Querol i Puig 2009). Therefore, although policy is being evaluated, contact-induced phenomena remain largely ignored.

However, language contact has not gone completely unnoticed by the 'new wave' of Catalan language policy studies. Indeed, in their evaluation of Catalan language use in primary education, Vila i Moreno and Galindo Solé recognise that 'el català d'infants i jovent està profundament interferit pel castellà' (Vila i Moreno and Galindo Solé 2009, 22),<sup>22</sup> yet go no further in identifying how the two languages are impacting on one another. Even the most recent analyses of Catalan language policy (cf. Vila i Moreno 2011) make no reference as to what may constitute 'Catalan' or 'Castilian', presumably using these terms to refer to the normative variety of each code, even though this may not be what is actually being used in the contexts under discussion. This highlights the gap in the current trend of academic discourse which seeks to evaluate Catalan language policy almost thirty years after its implementation, i.e. speaker language use is not being examined in any more detail than saying 'speaker x uses Catalan/Castilian in a given context'. There is no debate as to what constitutes 'Catalan' or 'Castilian'. To fill this gap, and in keeping with the surge in interest about the consequences of Catalan language policy, it is of interest to discover whether proposed developments in policy bear upon what speakers consider to constitute 'Catalan' or 'Castilian' utterances. What better way to 'fer un balanç' of Catalan language policies than to see if they are reaching their intended audience, and thus affecting this particular element of speaker language competence?

---

<sup>22</sup>TRANSLATION: The Catalan used by children and young people shows profound interference from Castilian.

### 1.2.2 Policy and contact: Theoretical implications. ‘Top-down phenomena’ and ‘bottom-up phenomena’.

It has thus far been made clear that language policy and language contact are of great relevance when studying the bilingual Catalan-Castilian community of Barcelona. However, their pertinence is not restricted to the examination of this single language situation. This section will highlight the theoretical importance of policy and contact by demonstrating that they exemplify key components of maximally differentiated sociolinguistic phenomena, which will be termed *top-down phenomena* (TDP) and *bottom-up phenomena* (BUP) respectively. These will be fully defined in due course but for the present, suffice it to say that TDP and BUP will consist of background situations (such as policy and contact) that may lead to innovative linguistic behaviour, as well as any resultant instances of language usage.

Firstly, this section will engage in a thorough review of literature concerning sociolinguistic variation phenomena that could influence speech situations. The resultant findings will then be formulated into a three-dimensional model which details the various components of these phenomena, and explains plainly how and why my notions of TDP and BUP are maximally distinct. Then, examples will be given of the background situations that make up TDP and BUP, as well as the ways they can be realised as variation in language form (i.e. lexical and structural content) and language function (i.e. use in different situations in society). Finally, the extent to which the terms TDP and BUP will be of importance to this thesis will be fully stated.

The theoretical implications of a study which addresses policy and contact should be made clear: by demonstrating that maximally differentiated variation phenomena can co-occur<sup>23</sup> in this particular linguistic community, a case is to be made that future work should allow for the discussion of simultaneous linguistic phenomena. The consequences of the presence of both TDP and BUP in one linguistic community may prove enlightening, and could open the way for similar study in other speech situations.

---

<sup>23</sup>And potentially, as we shall see in upcoming chapters, exert competing pressures on one another.

\* \* \*

Before entering into discussion of TDP and BUP, I should issue a disclaimer. The literature about to be examined talks of language *change from above* and *change from below*. However, the present work does not presume to address issues of language change. It will become clear that this thesis merely discusses language policy and language contact as phenomena linked to variation. Variation can be diachronically stable, or can indicate an ongoing development over the course of time. Such linguistic *changes in progress* can be identified if a trend is detected in the proportion of variants of a variable in real or apparent time. These changes in progress may then become fully embedded language changes or not. Empirical investigation can allow us to examine possible trends in usage of linguistic variants. If such trends imply a change in progress, they may potentially lead to fully embedded language changes, but this cannot be determined at such an early stage. As such, these are not issues of language change under discussion, but are rather potential pre-precursors to processes of embedded change. Indeed, in the case under examination, we do not yet know if a change has taken place, and so this thesis cannot examine language change *stricto sensu*.<sup>24</sup>

It was indeed my original intention to explore issues of language change, and specifically, the evolution of the terms ‘change from above’ (CFA) and ‘change from below’ (CFB). However, as time progressed, it became clear that the issues raised in

---

<sup>24</sup>The only way such lack of clarity could be avoided would be to adopt terminology akin to Chomsky’s (1986) notions of I-language and E-language. I-language is a speaker-internal notion, referring to individuals’ competences, while E-language is used to refer to linguistic systems found in societies, often denoted using names like English, Hindi, Catalan or Castilian.

This thesis could have decided to talk about changes in speaker-internal competences, that is to say in I-languages, since the upcoming experiment addresses developments in individual speaker judgements as to what constitutes normative language use (this will be discussed in due course). This would help, since at no point would it be claimed that there had been a change in E-language, that is to say, a change embedded in the systems of Catalan or Castilian.

However, the literature on language change does not often employ the I-language/E-language distinction. Moreover, and arguably more importantly, the adoption of these terms comes with a degree of theoretical baggage. Chomsky problematises the very notion of E-languages, such as Catalan or Castilian. The upcoming experiment depends on the presence of these systems, and indeed, it would be very difficult to address issues of contact, while maintaining that ideas such as ‘normative Catalan’ and ‘normative Castilian’ were of limited credibility.

Therefore, the adoption of these terms simply results in a different set of problems. So issues of ‘change’ have been excluded from discussion in this thesis.



this literature review did not relate to my thesis in terms of ‘change’. This was because it would not be feasible to devise an experiment which would ascertain whether a definitive change had indeed taken place in the language systems used in this linguistic community. Nevertheless, the issues which will be raised in this discussion are still of great theoretical importance, even if my thesis does not discuss language change as did Labov, Romaine and others in this section. It will become evident that, upon examination of the literature, scholars differentiated between the terms ‘change from above’ and ‘change from below’ along one or two of the following three axes (which shall be explained in due course):

- + planned / – planned.
- + metacommentary / – metacommentary.
- + overt prestige driven / – overt prestige driven.

In my work, these three axes will prove highly important, since they will allow for the maximal differentiation of phenomena (such as, for example, policy, contact and their consequences); but the main difference here is that these phenomena are **not** changes. They may arguably bring pressures to bear on linguistic communities. They may potentially feed into larger-scale, embedded speaker practices. As such, they may well be pre-precursors of a fully embedded change in a linguistic system.<sup>25</sup> But all of this is mere speculation. This thesis simply offers a perspective on how language policy and language contact have both impacted on members of the bilingual Catalan-Castilian community of Barcelona. Whether these findings can later be developed into work that tests if any linguistic changes have been embedded in the community is a question for the future.

In short, this literature review will reveal the presence of three axes, which I will use to create a model to categorise sociolinguistic phenomena (including, in this linguistic community, policy and contact). These axes are not being used to discuss

---

<sup>25</sup>If ‘language change’ is interpreted as a change in a linguistic system such as Catalan or Castilian, then the precursors would be individual-level speaker practices, i.e. the proven, attested use of new linguistic items or constructions (as in different from those used previously). Policy or contact, which may condition these practices, are thus the pre-precursors.

changes embedded in a linguistic community, since that would be beyond the scope of the thesis. They will nevertheless be shown to be just as helpful in addressing the phenomena under examination. Therefore, I am not directly following in the footsteps of Labov, Ferguson, Romaine and the others mentioned here, whose work aims to define language changes. I am merely using this literature as a point of departure, since it raises concerns (namely the three axes) that are indeed of great theoretical relevance to this study of policy and contact.

\* \* \*

The literature on the terms ‘change from above’ and ‘change from below’ reveals that the terms have been used in many subtly different ways to represent two different processes of language change. However, it appears that all the definitions examined classify the two processes of change along one or two of three binary axes, mentioned above and repeated below:

- + planned / – planned.
- + metacommentary / – metacommentary.<sup>26</sup>
- + overt prestige driven / – overt prestige driven.

I shall highlight (in boldface) these binary distinctions as they become apparent in various academics’ uses of the terminology under review. While we know that language change is not the focus of the thesis, these axes could still reveal what scholars deemed to be the most important criteria by which to determine that two processes were indeed opposites. Then based on this information, I will be able to arrive at definitions for two maximally differentiated, related phenomena (i.e. TDP and BUP).

Since one of the primary aims of this thesis is to evaluate Catalan language policy, it makes sense to first examine Ferguson’s 1987 interpretation of the notion of *change from above*, given its focus on planning. The literature review will then develop from here:

---

<sup>26</sup>This will receive full explanation in due course, and refers to whether speakers are able to offer explicit metacommentary on a particular phenomenon.

Level 1 (conscious)	CFA	Stereotypes	Speakers can offer meta- commentary, and stylistic variation is shown
level of consciousness			
Level 2 (subconscious)	CFB	Markers	Speakers cannot offer metacommentary, but still show stylistic variation
level of awareness			
Level 3 (unconscious)	CFB	Indicators	Speakers cannot offer metacommentary and show no stylistic variation

Figure 1.2: Labov’s stereotypes, markers and indicators as CFAs / CFBs

Change from above... takes place when an active language ideology, an individual language planner, or a language planning agency initiates an explicit standardising proposal (Ferguson 1987, 304).

Here what defines a change as being ‘from above’ is its level of **planning**. If a language change is explicitly planned as detailed by Ferguson, then it is deemed ‘from above’. However, this term was not invented by Ferguson. CFA and its counterpart CFB were already in widespread use, having been popularised in William Labov’s 1972 *Sociolinguistic Patterns*. His terminology was coined with specific reference to sound change, and was used as follows:

[Changes from below are] below the level of social awareness. The variable shows no pattern of stylistic variation in the speech of those who use it, affecting all items in a given word class. The linguistic variable is an *indicator*... [A c]hange from above [is a correction] towards the model of the highest status group. (Labov 1972, 178-9)

Here, we have two factors that are clearly at play. Firstly, and most importantly, is the notion of stylistic variation and the related concern of speakers being able to offer **metacommentary**. This becomes particularly apparent when we look at the related terminology of *indicators*, *markers* and *stereotypes*. Labov uses his term *stereotype* to refer to variables on which speakers can offer metacommentary, and which show

stylistic variation; such variables, when they are language changes, are CFAs. He uses the term *marker* to refer to a variable on which speakers cannot offer metacommentary, but which still show a degree of stylistic variation; according to Labov's initial definitions, such variables, when they are language changes, are CFBs. Finally, he uses the term *indicator* to refer to variables of which speakers are neither aware nor conscious, which thus do not show stylistic variation, and obviously speakers cannot offer metacommentary thereon; such variables, when they are language changes, are CFBs. All of this information is presented succinctly in figure 1.2, which reveals that the key distinction between CFAs and CFBs is, in fact, the ability of speakers to offer metacommentary, rather than stylistic variation. The most important differentiating factor between CFAs and CFBs must be a quality not shared by stereotypes (classified as CFAs) and markers (classified as CFBs). This cannot be stylistic variation, since both stereotypes and markers are stylistically malleable. The key difference between these types of variable is that speakers can offer metacommentary on stereotypes (CFAs), but not on markers or indicators (CFBs). A key axis of distinction is therefore whether speakers can provide metacommentary on the variable in question, which will be represented as  $+/-$  **metacommentary**.

Secondly, Labov's definitions take into account the notion of changes as **overt prestige driven**. Introduced by Labov and further elucidated by Trudgill and Chambers (1998) and Meyerhoff (2006), *overt prestige* is the quality 'associated with a variant that speakers are aware of and can talk about in terms of standardness, or aesthetic and moral evaluations like being 'nicer' or 'better' (Meyerhoff 2006, 292). This can be contrasted with items high in *covert prestige*, which is 'a norm or target that is oriented to without the speaker even being aware they are orienting to it.' (Meyerhoff 2006, 288). Labov's instances of CFA seek to approximate 'the model of the highest status group', i.e. the set of linguistic items highest in overt prestige.

It is thus clear that when Labov's initial pair of definitions is viewed in conjunction with Ferguson's planning-based view, three important components become apparent : my axes of  $+/-$ **planned**,  $+/-$ **metacommentary** and  $+/-$ **overt prestige driven**.

As with many theoretical terms, CFA and CFB have developed through time, their meanings subtly modified by the numerous academics who have employed them in numerous bodies of work. Such almost imperceptible alterations in meaning must be addressed in order to provide an account of how any related concepts might be most effectively put to use. It will become clear that by only paying attention to one (or occasionally two) of my three axes, academics have misinterpreted Labov's initial definitions, in themselves clear, leading to confusion regarding what these terms actually refer to. I shall now review existing usage of the terms, starting with definitions which only consider one axis, then progressing to two. Then, I will demonstrate how these axes need not necessarily refer to change, but can also categorise related variation phenomena, and propose definitions for my own notions of *top-down phenomena* and *bottom-up phenomena* which take into account all three axes. Finally, I will give examples of how such variation is realised in linguistic communities.

**A one-axis approach.** Perhaps the most egregious misinterpretations are made when only one axis is used to delimit the two processes, as in the following entry in Crystal's *Cambridge Encyclopedia of Language*:

[In Labov's 'fourth floor' experiment] /r/ had been unconsciously recognised as a marker of high prestige, and was beginning to be used in careful lower-class speech. The direction of change was 'from above'... [While in Labov's fieldwork on Martha's Vineyard] islanders came to imitate the way the fishermen were speaking, because [again subconsciously] they admired their traditional character and way of life. In due course, the change spread throughout the island – a change 'from below' (Crystal 1987, 332).

It is clear that, of the three axes, Crystal's distinction only pays attention to +/- **overt prestige driven**. As such, the processes arrived at are not maximally distinct, since (for example) these two changes are identical in their means of implementation, both arising from speaker contact. Moreover, in ignoring the +/- metacommentary dimension, Crystal overlooks Labov's primary criterion for distinguishing between the processes in the first place. Other definitions have also chosen to only focus on one dimension, but with arguably less misleading consequences:

In order to understand the role of social class in language change, it is essential to understand the distinction between changes that take place below the level of consciousness, so-called *changes from below*, and those that take place above the level of consciousness, or *changes from above* (Wolfram and Schilling-Estes 1998, 188).

This definition only pays attention to the  $+/-$  **metacommentary** dimension, as it talks about the ‘level of consciousness’.<sup>27</sup> Wolfram and Schilling-Estes have overlooked Labov’s mention of the ‘highest social group’ thereby ignoring the  $+/-$  overt prestige driven dimension, but as we can see, this interpretation of CFA and CFB is a lot more faithful to Labov’s original ideas than Crystal’s was. This reinforces my claim that the metacommentary axis is the most central to Labov’s definition, since if one divides solely along this axis, a far less distorted interpretation of Labov’s ideas is yielded, as in Wolfram and Schilling-Estes (1998) than if one were to only use his secondary dimension of overt prestige, as in Crystal (1987).

**A two-axis approach.** In other cases, researchers have adhered to Labov’s (in my view) dual-axis definition, as follows:

Sociolinguists have distinguished between ‘change from above’ and ‘change from below’ to refer to the differing points of departure for the diffusion of linguistic innovations through the social hierarchy. Change from above is conscious change originating in more formal styles and in the upper end of the social hierarchy; change from below is below the level of conscious awareness, originating in the lower end of the social hierarchy (Romaine 2003, 103).

While Romaine’s definition has clearly addressed both the  $+/-$  **metacommentary**<sup>28</sup> and  $+/-$  **overt prestige driven** dimension, as a theory, it falls down on other grounds. Although Labov stipulated that CFAs are forms which aim to approximate linguistic variables rich in overt prestige, the social class of the group motivating the change was never explicitly stated, much less offered as the prime differentiating factor between the two changes. While Romaine’s stipulations that CFA and CFB

---

<sup>27</sup>Indeed, as has been argued here with reference to Labov’s work, ‘consciousness’ may be a more accurate term than Labov’s ‘awareness’ as it captures Labov’s own distinction with indicators and markers as CFBs, and stereotypes as CFAs (see figure 1.2).

<sup>28</sup>While metacommentary itself has not been explicitly mentioned, Romaine’s definition leans on the intrinsically related notion of stylistic malleability

must occur above and below the level of conscious awareness respectively, and that CFA has its origins in more formal styles, are accurate reproductions of Labov's use of the terms, her (main) claim that 'below' and 'above' refer to different points in the social hierarchy is something of a problematic departure. Aside from the fact that 'below' and 'above' now seem to be speaker-external rather than speaker-internal concepts, her definition seems to confuse theoretical claims with empirical findings. While CFAs often begin in higher social classes (as these are typically the groups which use forms high in overt prestige) and CFBs often originate in lower social groups (perhaps due to members of the lower classes being more numerous, or at least, more frequently studied), this can only be ascertained from testing Labov's terms in action. If this class element is included in the definitions themselves, some arguably analogous processes are excluded from analysis by Romaine which could be included by Labov. For example, let us think of a process which begins above the level of conscious awareness (a stereotype or marker) and attempts to approximate a form high in overt prestige, but that originates in lower social classes; admittedly, such processes are rare, but when one is establishing a theory, this is immaterial – Romaine's comparatively reductive definition cannot account for this, whereas Labov's can. As such, while it is important for a definition of these two processes to be based on multiple axes (this allows for maximally differentiated processes, as we shall continue to see), it must not be reductive, and thereby unable to account for certain phenomena. A good example of this is Ellis' use of the terms under discussion:

When a group drifts toward a new sound or vocabulary item and they are unconscious of it, this is a change from below [consciousness level]. The introduction of language change is from above when the change is consciously introduced by higher status groups (Ellis 1999, 150).

Ellis' definition references the central +/- **metacommentary** component and discusses the secondary +/- **overt prestige driven** dimension through his mention of 'higher status groups'. His notion of CFA reflects a subtle departure from Labov's ideas, also present in Romaine's definition. CFA is now actively introduced *by* higher status groups, rather than simply constituting a change which aims to replicate or

approximate the language use of these higher status groups. This ties in with Ferguson's development, and my third axis, of  $+/-$  **planned**.

**Towards three axes.** When trying to arrive at maximally distinct types of language change, scholars have thus far chosen to make divisions along the following three axes, whether they are aware of it or not:

- $+/-$  **metacommentary.** This refers to the degree to which speakers are aware and/or conscious of a particular linguistic variable. A variable can be an *indicator*, with speakers possessing no consciousness or awareness, and thus they cannot offer any metacommentary thereon. *Markers* are variables of which speakers are not conscious, but on some level are aware, since they are subject to stylistic manipulation. However, the lack of speaker consciousness means that, once again, they cannot offer any metacommentary on these variables. Finally, *stereotypes* are variables of which speakers are both aware and conscious: they are open to stylistic variation and speakers can offer metacommentary on them. Thus, we have *indicators* and *markers* on one side of the divide, being considered CFBs, and possessing the quality '– metacommentary'. On the other side, we have *stereotypes*, considered CFAs, and being '+ metacommentary'. Labov's initial definition posited the degree of speaker conscious awareness as the principal distinction between his two types of language change.
- $+/-$  **overt prestige driven.** This refers to the fact that some language changes seek to reproduce variables that are high in overt linguistic prestige, while others do not. Occasionally, this has been developed by scholars (cf. Romaine 2003) into the notion of *social class of the group motivating the change* as being of primary importance. My nomenclature here is very deliberate. If a phenomenon is to be deemed –overt prestige driven, this can mean one of two things: either it is driven by covert prestige, or it is not motivated by prestige at all.
- $+/-$  **planned.** If a phenomenon is explicitly planned by a body, agency or ideology, then it is deemed '+ planned.' This was of paramount importance in



Ferguson's (1987) notion of 'change from above'.

These three axes were used in the literature as a means of classifying processes of language change, but they are equally applicable to related phenomena. At no point is it said that the phenomena which are to be categorised by these axes must be language changes that are already embedded in a particular linguistic variety. Indeed, the individual linguistic variables which may be deemed '+ planned' or '- overt prestige driven' merely need to be attested to some degree, so that they can be examined; but at no point do they necessarily need to constitute language changes. These variables could be examples of phenomena that we have already referred to as pre-precursors of language change (see page 49). As such, a variable could be a result of a situation of language contact, for example, and as we shall see, can be classified according to the three axes. The variable does not have to be considered an enacted language change. If it becomes a language change one day, then so be it. But the fact that it has not yet been fully embedded in a community does not stop such a variable from being classified as '-planned', '-overt prestige driven' and '-metacommentary', for example. As previously stated, it became clear that the thesis could not go so far as to answer questions of language change; but there is no strict need for the three axes to be used to classify embedded changes, since other related variables and phenomena can also bear these labels. So let us keep our three-axis definition and use it to its best advantage.

Therefore, I propose that only through a three-dimensional division along all of the above axes can we identify two, *maximally distinct* phenomena which can bear upon sociolinguistic situations. For the purposes of creating a sound theoretical definition, these axes are treated as bifurcating. Obviously, in reality it is not always possible to define a situation as + or - a particular property; but the present concern is the development of a theory that allows us to arrive at the most differentiated phenomena possible, and so the various 'shades of grey' will be temporarily overlooked.<sup>29</sup>

---

<sup>29</sup>Further dimensions to my model were considered, but ultimately rejected in the interests of parsimony, notably +/- responsive (i.e. is the phenomenon a response to external sociolinguistic or cultural factors or pressures, or is it internally motivated?). This was deemed unnecessary as it overlapped with the +/- planning axis. Phenomena which were -planning showed no differentiation

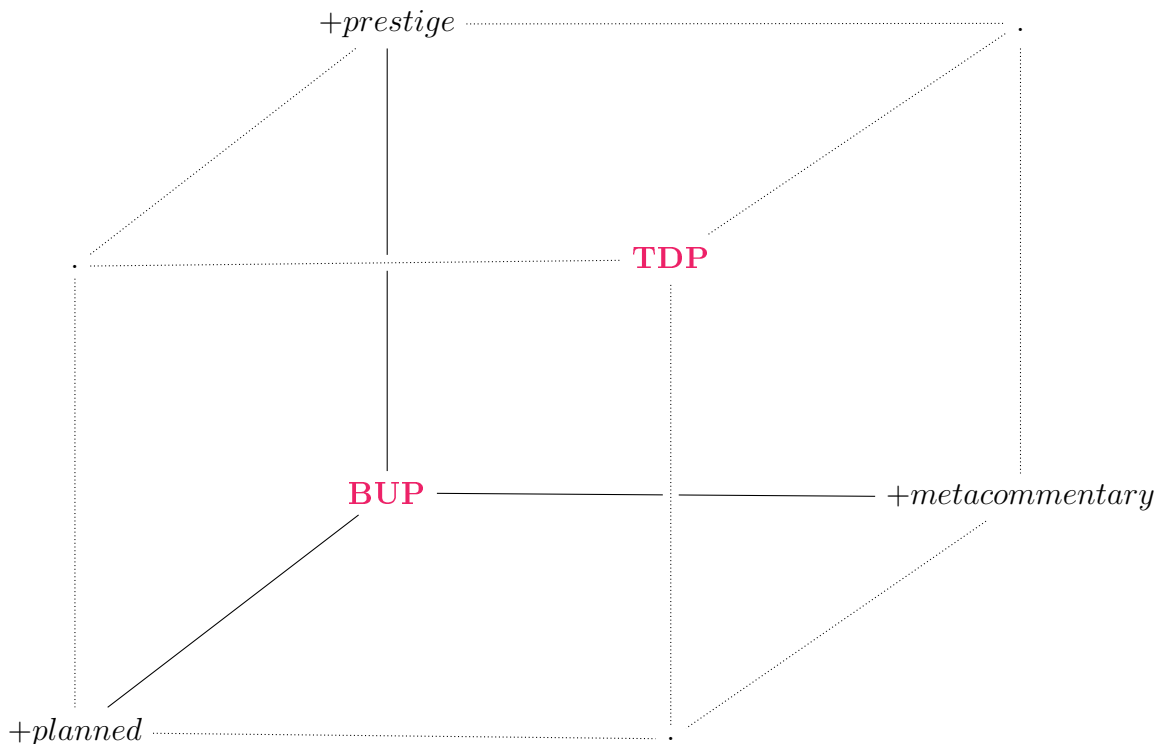


Figure 1.3: Graphical representation of how TDP and BUP need to relate to the dimensions of *metacommentary*, *overt prestige* and *planning*.

To achieve maximal distinctiveness, one phenomenon will score positively on all three axes: this will be the *top-down phenomenon*. The other will score negatively on all three axes: the *bottom-up phenomenon*. This can be represented graphically as in figure 1.3.

At the start of this section, it was stated that these phenomena would encompass situations which give rise to the use of linguistic items of structures. This is the case for both TDP and BUP and should be borne in mind when reading the upcoming definitions. Now that we know the criteria that TDP and BUP must fulfil on a theoretical level, the following definitions are proposed:

---

if they were + responsive or –responsive. As such, the important distinction lies with the +/-planning axis, not the +/-responsive axis. It should also be noted that, contra Romaine, there is no two-way division based on the social class of the motivating group. While TDPs will be deemed to require official support (and as such, usually originate in the upper echelons of society), this is not in direct opposition to BUPs, which can emerge from any sector of society. As such, no bifurcating axis could be formed.

**TOP-DOWN PHENOMENA (TDP)** : This language use is implemented above the level of consciousness, as the result of actions by an officially sanctioned organisation or body, which are driven by conscious decision-making and aim to approximate language usage rich in overt prestige.

**BOTTOM-UP PHENOMENA (BUP)**: This language use is implemented spontaneously and below the level of consciousness.

These definitions outline two parts to TDP and BUP. On one hand, we have the situations (such as policy or contact) which encourage the introduction of certain linguistic variables. On the other hand, we have the variables themselves. These variables could be innovations of a lexical, syntactic or phonological nature, for example. They are the aforementioned pre-precursors to change, i.e. individual tokens that could well embed themselves in a given linguistic variety one day, constituting a change. The phenomena defined above thus consist of both the situation and any potential linguistic variables emanating therefrom.

It must now be ascertained exactly how the above definitions fulfil the theoretical criteria as per figure 1.3. Table 1.5 extracts fragments from the definitions in order to demonstrate how each of the axes is addressed:

Phenomenon	+/- metacommentary	+/- overt prestige driven	+/- planned
TDP	‘above the level of consciousness’ <b>+ metacommentary</b>	‘aim to approximate language use rich in overt prestige’ <b>+ overt prestige</b>	‘result of actions by an officially sanctioned organisation or body’ <b>+ planned</b>
BUP	‘below the level of consciousness’ <b>-metacommentary</b>	no mention of prestige <b>-overt prestige</b>	‘implemented spontaneously’ <b>-planned</b>

Table 1.5. Tabular representation of how my definitions of TDP and BUP relate to the dimensions of *metacommentary*, *overt prestige* and *planning*

TDP and BUP therefore refer to clearly differentiated phenomena. But if this thesis is to progress, these terms must index more than mere theoretical abstractions: they need to come alive in a way such that they can be examined by researchers. We know that these ‘phenomena’ consist of both background situations and concrete linguistic variables or tokens. In order to progress, we have to not only know what the situations are, but also what the concrete linguistic instances could be.

Starting with the background situations, TDP have been defined as driven by ‘a conscious decision-making process... by an officially sanctioned organisation or body’: as such, language policy and planning (to be fully explained in chapter 2) is a logical choice as an example of such a situation. BUP have been defined as arising ‘below the level of consciousness’. Language contact (to be fully explained in chapter 3) should therefore fit the bill as a potential background situation, since Weinreich’s definition (already mentioned on page 42) refers to contact as the use of two or more languages by the same persons. In theory, speaker multilingualism could lead to cross-linguistic interference below the level of consciousness, and so we could posit that language contact is a potential background situation for BUP.<sup>30</sup>

<sup>30</sup>This will also be borne out in the upcoming paragraphs, when we will see that certain linguistic phenomena arising from language contact fit our definitions of BUP

So, now that an idea has been given as to the background situations, attention must be turned to the linguistic instances that form part of the TDP and BUP. It would be helpful to determine whether these can be realised both in form (structural and lexical content) and function (use in society), and if so, how. In theory, there is nothing preventing TDP and BUP being manifested in language form or in language function. Nothing within either definition dictates that one phenomenon refers to form and the other to function, and indeed examples of TDP and BUP in form and/or function can be found. For example, if a specific lexical item is consciously introduced into a given speaker's vocabulary by means of language corpus planning (potentially as a natively-derived term to serve as an alternative to a foreign borrowing), then this is an example of TDP in form. More broadly, given that these phenomena are linked to the examination of trends of linguistic variation, this could be expressed as the introduction of specific variants of linguistic variables. Analogously, results of language status planning (if a given language is introduced into a new domain and there is a concomitant increase in overt prestige) can be analysed as instances of TDP in language function.<sup>31</sup> Less specifically to the Catalan situation, this could be rendered simply as altering the norms of language use. BUP linguistic variables are most often tokens of language form, such as the introduction of lexical borrowings or instances of grammatical convergence which arise as a result of language contact. It can also constitute a function-based phenomenon, although this is generally secondary to developments in form, such as if a particular language were to lose overt prestige (and as such be excluded from use in high-prestige functions) as a result of extensive language mixture. Examples of the different manifestations of TDP and BUP, as well as the different background situations which cause them, can be summarised as follows:

---

<sup>31</sup>For brief definitions of *corpus planning* and *status planning*, see the list of terms and abbreviations beginning on page 16.

	TDP	BUP
Background situation	Language policy/planning	Language contact
Form-based phenomenon	Introduction of specific variants of linguistic variables	Systematic introduction of variant of a given variable through language contact
Function-based phenomenon	Altering norms of language use	Decrease in overt prestige (and thus exclusion from high-prestige functions) of a variety through BUP in language form

Table 1.6. Examples of TDP and BUP

In short, the numerous uses of the terms ‘change from above’ and ‘change from below’ in the literature to date have revealed three common and intertwining axes, which academics have used to differentiate that which is ‘from above’ and that which is ‘from below’. I believe that a definition which takes into account all three dimensions allows for maximally distinct, yet still clearly comparable phenomena: my TDP and BUP. TDP and BUP were then exemplified so as to facilitate the study of them at work in linguistic communities. Examples of all facets of TDP and BUP – the background situations, manifestations in form and in function – are given in table 1.6.

\* \* \*

In conclusion, the existing literature concerning ‘language change from above’ and ‘language change from below’ has revealed there to be three important factors which serve to differentiate between these processes, namely degrees of planning, motivation by overt prestige, and speakers’ abilities to offer metacommentary on the language use in question. Such factors are equally applicable to other phenomena, not just changes, which may impact upon linguistic situations. These phenomena consist of

a background situation, as well as potential manifestations in language form and function; all of these components are classified along the three axes revealed in this section, namely ‘+/- planning’, ‘+/- overt prestige’ and ‘+/- metacommentary’. The most differentiated of these phenomena are obviously those which score positively on all three axes compared with those which score negatively on all three axes, which I have called *top-down phenomena* and *bottom-up phenomena* respectively. Examples of these phenomena are given in table 1.6, and can be seen to result from language policy and language contact. This is exactly what this thesis will be focusing on: policy and contact (and any resultant phenomena) in the Barcelona linguistic community.

The theoretical interest of the work is thus clear. It is already clear that both policy and contact co-exist in the Catalan/Castilian linguistic community in Barcelona. We therefore know that the two background situations of the maximally differentiated phenomena of TDP and BUP co-occur in this one linguistic community. By exploring this further, the present work will shed more light on what pressures these two phenomena exert, both on each other, and on language usage in this community. Moreover, through the demonstration that two maximally differentiated phenomena may (or may not) exert competing linguistic forces, this project opens the way for further study of linguistic communities where multiple sociolinguistic phenomena are at work.

Finally, now that the theoretical merit of a study which explores the consequences of policy and contact has been stated, I must state the extent to which the concepts of TDP and BUP are going to be employed in this thesis. At the start of chapters 2 and 3, which will concern policy and contact respectively, the relevant sections of table 1.6 will be reproduced. This is simply so that what needs to be examined is made clear, in order that later, if required, comment could be passed on the effectiveness of the model. Table 1.6 thus serves to give chapters 2 and 3 structure. Moreover, chapter 6 will provide a discussion of the upcoming experiment findings. These will be primarily concerned with offering conclusions regarding policy and contact, since that is the crux of the thesis. However, one conclusion (conclusion 6) will relate these findings to the model of phenomena presented in this section, in order to speculate

as to its future usefulness. The three-axis model does not therefore directly impact on any upcoming experimental findings. Rather, it provides theoretical justification for the study of policy and contact in a given community, as well offering a means by which the present findings might readily be generalised to other linguistic situations, thus encouraging future research.

\* \* \*

It is now clear why a thesis which focuses on both policy and contact in Barcelona is of interest. Section 1.2.1 has demonstrated that such work fills an overlooked gap in present literature. Section 1.2.2 has revealed that there are interesting theoretical implications to such a work, namely that policy and contact constitute key elements of two, maximally differentiated phenomena, and the study of how these two co-occur might prove enlightening. The next section briefly addresses the practicalities of how such a study is to take place.

### **1.3 Ways in which to study language policy and language contact in Barcelona**

The present work studies the presence of language policy and language contact in Barcelona, and aims to examine any ways in which these could bear upon the linguistic community. While section 1.2 has shown why this is of potential interest, it still remains to be seen how this is to be empirically tested. Since policy and contact have been revealed to form part of maximally differentiated phenomena (TDP and BUP), it would be logical to assume that they exert different, maybe even competing, pressures on the community. Indeed, the previous section has already given examples of how TDP and BUP may be realised in language form and function, and so why should this not be the case in Barcelona? This thesis will attempt to discover exactly what the consequences of policy and contact are in Barcelona, in order to gain a better insight into the details of language use in this linguistic community.



The next two chapters will examine the theory behind language policy and language contact, applying these findings to Barcelona. At the end of chapter 3, based on all of this information, research questions will be devised, which will be answered by a fieldwork experiment. The questions and the experiment must successfully synthesise all of the findings regarding language-in-education policy (chapter 2) and contact-induced language phenomena (chapter 3), and allow for simultaneous experimental testing of policy, contact and their consequences in the Barcelona Catalan-Castilian bilingual community.

A number of related concepts will be used in the creation of the upcoming research questions and experiment. Contact-induced linguistic phenomena will be shown (in chapter 3) to permeate attested language use in Barcelona. Such evidence of language contact is, however, often not regarded as ‘standard’ Catalan or Castilian by the relevant linguistic regulatory bodies. As such, this *non-normative language use* will prove of key importance to the fieldwork. In order to test the extent to which this type of language usage is recognised as non-normative by speakers, and to see if this is at all correlated with the introduction of language policies promoting normative variants, elements of speaker ‘competence’ in the normative variants of each language is to be tested. ‘Competence’ is used here as in L2 literature to refer to skills in speaking, writing, reading and listening, and is what the Catalan government terms *coneixement*. Any developments in *coneixement* might be linked to developments in language policy, which could prove enlightening. But all of this shall receive attention in due course, over the next two chapters. For our present purposes, suffice it to say that a set of research questions will be composed, based on the key theoretical and practical concerns to be raised over the next two chapters. These questions will be answered by a fieldwork experiment which will employ a number of related concepts such as *non-normative language use* (to be fully explained in due course), in order to shed light on the consequences of maximally differentiated phenomena in Barcelona, namely language policy and language contact.

## 1.4 Thesis plan

Chapter 1 has clearly established all the situational and theoretical information needed to proceed with the study. The first section explored the sociolinguistic and historical background of Barcelona, thereby cementing in the reader's mind that this work will constitute a significant contribution to the field of Catalan studies, and thus elevating Barcelona above the status of mere 'case study'. The second section asked why this thesis has chosen to focus on policy and contact, addressing both the existing literature concerning the linguistic situation in Barcelona, and broader theoretical concerns. The chapter then concluded with the aims and objectives of the thesis as a whole, before offering the present chapter-by-chapter outline.

Structurally, chapters 2 and 3 are almost identical to one another. Chapter 2 is devoted to language policy and chapter 3 to language contact. The first section of each of these two chapters addresses theoretical concerns, detailing the most important existing academic work on the topic and explaining how this has led to current perspectives. The second section then places all of this in the Barcelona context, ensuring a balance between sociolinguistic theory and Catalan studies. Therefore, each of the first three chapters contains 'literature review', 'history of Barcelona' and 'theoretical background' components, but rather than being ordered as such, they are arranged thematically. Chapter 1 introduces the reader to the basics, while chapters 2 and 3 talk about each of the types of linguistic phenomenon under discussion, relating everything to the city being studied, Barcelona.

Chapter 3 concludes with a series of research questions that the fieldwork will need to answer in order to effectively analyse the situation at hand. Chapter 4 addresses these questions in order to devise a unique experiment methodology which takes all the relevant issues into account. Chapter 5 will offer a full statistical analysis of my experiment results, and reveal what sampling techniques, tests and methods were used, and why. Chapter 6 will then constitute a discussion of any issues presented by the results, and provide answers to my research questions, as well as detailing how future research may use my thesis as a point of departure for further study both within Catalonia and beyond.

## Chapter 2

# Language Policy in Barcelona

[A] strong reason for language planning [in post-1975 Spain] was to ensure the survival of the languages... Not only had Catalan, Basque and Galician given way to Castilian Spanish in many public and private domains... but the languages themselves had become Castilianised in various ways – through extensive lexical borrowing and grammatical interference... It is clear that language planning in Spain embraces both status and corpus planning. As regards status planning, many of the measures taken have been designed to promote the use of the regional languages in most spheres of public life, with special emphasis given to education. As far as corpus planning is concerned, the policies followed have led to... the codification and elaboration of the languages, so as to equip them adequately for their use as official languages. (Hoffmann 1995, 69-71)

The motivations behind any particular language policy are complex and varied, and any attempt to reduce them to a simple sound bite ultimately glosses over numerous important details. Nevertheless, the above citation offers a succinct and edifying overview of a few of the mainsprings behind Spain's language policy decisions of the last thirty years. Moreover, Hoffmann's summary of language planning motivation in late twentieth century Spain provides examples of both status and corpus planning,<sup>1</sup> placing everything within the Spanish context, and that is fundamentally what this chapter sets out to achieve.

This chapter discusses language policy (and, as we shall see, planning) and poten-

---

<sup>1</sup>These terms will receive greater attention on page 71. Definitions can also be found in the list of terms and abbreviations, beginning on page 16.

tial resultant phenomena. Table 1.6 in chapter 1 exemplifies my notion of top-down phenomena (TDP), including both background situations and manifestations of the phenomena (in language form and function) as they might appear in a linguistic community:

	TDP
Background situation	Language policy / planning
Form-based phenomenon	Introduction of specific variants of linguistic variables
Function-based phenomenon	Altering norms of language use

Table 2.1. Partial reproduction of table 1.6, pertaining to TDP.

The present chapter will use the above schema merely as a guide, in order to ensure that all matters which may prove pertinent to the upcoming experiment are discussed.<sup>2</sup> Following table 2.1, the chapter will cover discussion of language policy and planning, plus a number of ways in which policy decisions can impact on language form and function. Section 2.1 will address all the relevant theoretical history and debate, following the order of background situation (policy/planning) → form-based phenomena → function-based phenomena. Section 2.2 will follow exactly the same order, but focusing on the linguistic community at hand, Barcelona.

Finally, section 2.3 will summarise the most important findings concerning language policy in Barcelona and any consequences thereof in order to achieve this chapter's ultimate goal: the creation of a research question on the topic of language policy in Barcelona. This research question will be tested by the upcoming fieldwork

---

<sup>2</sup>This chapter does not seek to make any claims as to the nature of TDP themselves, or the theory behind this. Discussion of this will be limited to conclusion 6 in chapter 6. This chapter (and indeed chapter 3 simply retains this table since it allows for discussion of a range of relevant topics.

experiment.<sup>3</sup>

## 2.1 Language policy and resultant phenomena (TDP): Theory

This section will allow us an understanding of language policy and potential resultant sociolinguistic phenomena in form and function, as per the schema in table 2.1. Section 2.1.1 provides a historical background of language policy and planning theory. Section 2.1.2 focuses on one subtype of language planning (acquisition planning) and thoroughly explains the theoretical links between language planning and instances of TDP in form and function, before section 2.1.3 offers a brief summary.

### 2.1.1 Language policy and planning

Up to this point, the thesis has been employing the term ‘language policy’ as per the definition given in the list of terms and abbreviations. However, in order to fully understand what is meant by language policy, the related term ‘language planning’ must also be explored in depth. The following section aims to provide a historical overview of the study of language policy and planning, and will ultimately reveal how we reached the current situation of defining policy and planning as two intrinsically linked concepts, which can both be related to the extra-linguistic sociopolitical contexts to which they apply.

**The development of ‘planning’ and ‘policy’.** Over the last half-century, theoretical debate has revealed a distinction between the notions of *language planning* and *language policy*. Different scholars have used these terms to refer to subtly different processes, which has led to many different interpretations of the terminology at hand. As such, I shall use the following set of definitions by Ager as a point of departure, since it neatly distinguishes between planning and policy:

---

<sup>3</sup>Chapter 3 will be in the same vein, but will culminate in a research question concerning language contact in Barcelona.

Language planning has... come to mean the ways in which organised communities, united by religious, ethnic or political ties, consciously attempt to influence the language(s) their members use, the languages used in education, or the ways in which Academies, publishers or journalists make the language change. Language policy is official planning, carried out by those in political authority, and has clear similarities with any other form of public policy (Ager 2001, 5).

Ager's adoption of the idea that policy is 'planning in action' offers a simple framework within which to examine the two processes, their histories and how they interact with one another. I shall now outline the principal theories of language planning, before conducting a similar analysis of the history of the study of language policy, thus making it clear how Ager arrived at the above definition.

**Language planning.** The term 'language planning' was allegedly first used by Uriel Weinreich, as the title of a seminar given at Columbia University in 1957 (Eastman 1983, 114). In the years immediately preceding usage of the term, dialogue began on the clearly analogous topics of *language engineering* and *glottopolitics* (Cooper 1989, 29). Moreover, the notion of planning the form, and by extension the function, of language has existed for centuries, as evidenced by the establishment of language academies such as the *Académie française* in 1635 and the *Real Academia Española* in 1713, and the philological section of the *Institut d'Estudis Catalans* in 1911. Nevertheless, it was not until the late 1950s that discussion of 'language planning' started to enter scholarly debate. Einar Haugen's 1959 article *Planning for a Standard Language in Modern Norway* brought the issues surrounding 'the activity of preparing a normative orthography, grammar, and dictionary for the guidance of writers and speakers in a non-homogeneous... community' (Haugen 1959, 8) to the fore for the first time. In her history of the study of language planning, Carol Eastman maintains that in the 1960s, the main focus was on addressing language problems and 'linguistically complex situations' (Eastman 1983, 115). This was epitomised by such works as Joshua Fishman's *Language Loyalty in the United States* (1966), which explored problematic cases of multilingualism, using nascent developments in language planning study to better deal with the issues at hand. Subsequently, in the late 1960s

and early 1970s, Eastman claims that the field, while still focusing on the potential solutions to language problems, subtly shifted its emphasis towards language change and the study of standardisation (Eastman 1983, 121). My brief history of language planning will chiefly focus on two of the most important developments to emerge from study of the field: Kloss' distinction between *corpus planning* and *status planning*, and Haugen's *standardisation model*. Far from being outdated, these two theoretical constructs are still current in discourse on language planning, and as such merit attention.

A little-known 1969 report by Heinz Kloss regarding the language situation in Québec coined two terms which have come to characterise language planning to this day: *corpus planning* and *status planning*, which were defined thus:

Planning with regard to languages is usually understood to mean that some agency, person, or persons are trying to change the shape or the corpus of a language by proposing or prescribing the introduction of new technical terms, changes in spelling, or the adoption of a new script... These innovations have one thing in common, that they modify the nature of the language itself, changing its corpus as it were. We may thus speak of language *corpus planning*. There exists, however, another dimension of planning where one busies oneself not with the structure and form of language but with its standing alongside other languages or vis-à-vis a national government. Those concerned with this type of language planning... are primarily interested in the status of the language whether it is satisfactory as it is or whether it should be lowered or raised. Here we can speak of language *status planning* (Kloss 1969, 81, my emphasis).

*Corpus planning* is thus concerned with language-internal factors, while *status planning* measures address language-external factors. This two-way division has proved of great use to language planners and scholars alike, and is still widely considered the primary axis along which to categorise instances of language planning, as witnessed in more recent explanatory texts (cf. Bourhis 1984, 3; Tonkin 2005, 120-6 inter alia).<sup>4</sup>

---

<sup>4</sup>Kloss' two-way corpus / status distinction shows clear influence of an earlier notion of his which distinguished between two types of language: *abstand* languages and *ausbau* languages. *Abstand* languages are seen as 'languages by [intrinsic] distance', i.e. an *abstand* language is sufficiently structurally differentiated to be considered a language in its own right. *Ausbau* languages are 'languages by development', which have been 'shaped or reshaped, molded or remolded... [by] deliberate language planning' (Kloss 1967, 29). This distinction is clearly tied in with the internal

However, this dichotomy arguably resulted in ambiguous and problematic cases, notably in the field of language acquisition. Cooper (1989) thus proposes a third axis of study, *acquisition planning*, which is concerned with increasing the number of speakers of a given language, as well as increasing existing speakers' competence in the language. Cooper argues that there is a fundamental difference between planning which 'is directed towards increasing a language's uses' and planning which 'is directed toward increasing the number of users' (Cooper 1989, 33), considering the former as status planning and the latter as acquisition planning. Language policies which aim to promote the use of a particular language on an international level (such as the spreading of English by the *British Council*, or of French by the *Alliance française*) are thus instances of acquisition planning, as they focus on gaining new (chiefly second-language) speakers. Cooper also regards initiatives with 'language maintenance' goals as instances of acquisition planning, on the grounds that said languages will be '[acquired] by the next generation' (Cooper 1989, 159).<sup>5</sup> This new form of planning was further fragmented by Haarmann (1990), yielding *prestige planning*, to refer to planning activities 'aimed at promoting a positive view of a language' (Sallabank 2008, 125). As such, current research often divides language planning into two, three or four subcategories,<sup>6</sup> the most frequent distinction being made between *status planning* and *corpus planning*, with some academics also choosing to identify *acquisition planning* and/or *prestige planning* as separate phenomena. It should be noted however that, when employing as many as four subcategories, 'it is acknowledged that in practice they cannot be implemented without overlap' (Sallabank 2008, 125).

---

/ external dichotomy that Kloss is addressing with corpus and status planning, in that abstand languages are distinct because of their language-internal factors, while ausbau languages are distinct because of language-external concerns. We should take great care not to equate language-internal concerns with 'form' and language-external concerns with 'function'. These issues will become more apparent when examining Haugen's standardisation model.

<sup>5</sup>Such strategies are not aimed at acquiring new speakers, but rather at fostering a high level of competence within the heritage community. An example of this is given as the *kohunga reo* language nests whereby Māori elders provided language immersion experiences for the youth of the community (Cooper 1989, 158).

<sup>6</sup>If not more. Ager (2005, 1) distinguishes yet further between *prestige planning* and *image planning*, stating that 'image planning [is] a stage in identity formation and consolidation, and prestige planning [is] attitudinal change'. These concerns will be addressed when I go on to discuss the motivations behind language planning and policy later in this section.



Another important early development which has greatly influenced subsequent study of language planning is Einar Haugen's *standardisation model*. Haugen's 1966 article *Dialect, language, nation* stressed the importance of language as a nation-building tool by addressing the form / function and language-internal / language-external dichotomies which were current in debate and popularised by Kloss as *ab-stand* versus *ausbau*, or later as *corpus* versus *status*. He proposed the following standardisation model, consisting of four stages which, if followed, allowed for development of a linguistic variety from mere 'vernacular' to a fully-fledged 'standard' (Haugen 1966, 507): <sup>7</sup>

	<i>Form</i>	<i>Function</i>
<i>Society</i>	Selection	Acceptance
<i>Language</i>	Codification	Elaboration

'Selection' refers to the need to agree on a model of language as a basis for any ensuing norm (for example standard Italian as based on Florentine speech practices, standard French as based on Parisian language etc). 'Codification' refers to the '[development of] the form of a language, i.e., its linguistic structure, including phonology, grammar, and lexicon' (Haugen 1966, 504), and is primarily witnessed through the creation of dictionaries and prescriptive grammars. 'Elaboration' is described by Haugen as an attempt to achieve 'maximal variation in function' (as opposed to codification, which aims for 'minimal variation in form') (Haugen 1966, 505); in short, making a language fit for purpose, ensuring it has the resources to be employed in as many situations as possible. 'Acceptance' covers the processes which ensure that the new standardised language is used within the target linguistic community; this can be achieved by providing training and job opportunities for language users, and increasing its official use in government. When we compare the standardisation model with the corpus / status distinction offered by Kloss three years later, we see that Haugen's four way classification of language policy incorporates not only the form / function bi-

---

<sup>7</sup>It should be noted that these four stages need not be regarded as consecutive. It is normal for the selection process to precede the other three, but (for example) acceptance measures can continue long after a language has been codified and even elaborated.

furcating axis, but also the internal / external dimension, by detailing processes such as elaboration, which are fundamentally language-internal (*corpus planning*) but are concerned with the development of the function of the language (as well as its form). Haugen's standardisation model, like Kloss' corpus / status distinction, remains a fundamental tenet of language planning over forty years after its initial conception, and is still extensively used, in its original state, by academics focusing on several speech situations around the world (cf. Deumert 2004, Lodge 2007, Stolt 2009).

In short, language planning has been studied in an attempt to provide solutions to linguistically complex situations. The diverse nature of such linguistically problematic scenarios has resulted in the need to delimit different types of language planning in order to achieve different aims and objectives. Kloss' 1969 distinction between corpus (language-internal) planning and status (language-external) planning is a theoretical concept that is still observed to this day, although admittedly often modified, with new dimensions such as acquisition planning and prestige planning gaining ever more currency. Haugen's 1966 four-stage standardisation model continues to be adhered to by current scholars wishing to examine language situations whereby it is necessary to consider issues of form and function, as well as concerns of whether a process is language-internal or language-external. Language planning is thus the study of the ways in which languages can be consciously influenced and/or changed, and how they do this because of numerous internal or external pressures. It is now necessary to turn our attention to how such planning can be implemented through concrete political action.

**Language policy.** The rise of the theoretical study of language planning is correlated to an increase in the use of terminology to describe related concepts, most notably *language policy*. Ager's view outlined above (which I share) that language policy can be reduced to 'language planning in action' is widely held, although many definitions of the term are subtly different from one another. Carol Eastman's introduction to language planning talks about official language policies as 'based on decisions about what languages are to be designated as national, official and world languages' before

declaring that language policies exist in all societies, whether this is overtly stated or not, and that indeed such ‘implicit’ policies exist in the absence of language planning (Eastman 1983, 6). As such, language planning in an Eastman sense is the study of language-based governmental decision making, which can then go on to inform policy decisions: indeed ‘the language planner has a responsibility to offer advice to policy-makers’, although this advice may often go unheeded (Eastman 1983, 7). She stipulates that language policy needs to be formulated, codified and elaborated before being implemented (Eastman 1983, 7). The formulation stage is the deliberative process behind the policy: deciding what needs to be done and how. The codification stage refers to the preparation of the policy, and ensuring its appropriateness for the linguistic community at hand. The elaboration of policy is a continual process which ensures that the policy is fit for purposes in all spheres of activity. Only when these three processes have been undertaken (or partially undertaken in the case of elaboration) is the policy then implemented (Eastman 1983, 7-8). Obvious parallels to the Haugen standardisation model are not accidental: in adopting a four-stage framework clearly analogous to that proposed by Haugen, Eastman reinforces that language policy and planning are fundamentally interconnected processes.

A slightly different perspective is taken in Schiffman (1996), which adopts Bugarski’s view that ‘language policy refers... to the policy of a society in the area of linguistic communication... Language planning is understood as a set of concrete measures taken within language policy to act on linguistic communication in a community, typically by directing the development of its languages’ (Bugarski 1992, 18 in Schiffman 1996, 3). Here, as in other definitions, language policy refers to the concrete policies implemented by an authorising body; language planning however is seen as a set of measures visible within these policies, as opposed to a framework of principles which informs policy-making decisions. In spite of these subtle differences, whether we adopt Eastman’s, Schiffman’s or Ager’s perspective, we are left with the notion of language policy as a phenomenon which is distinct from, yet intrinsically linked to, language planning. As such, we can now proceed to examine existing typologies of language policy; just as the corpus / status distinction was proposed for

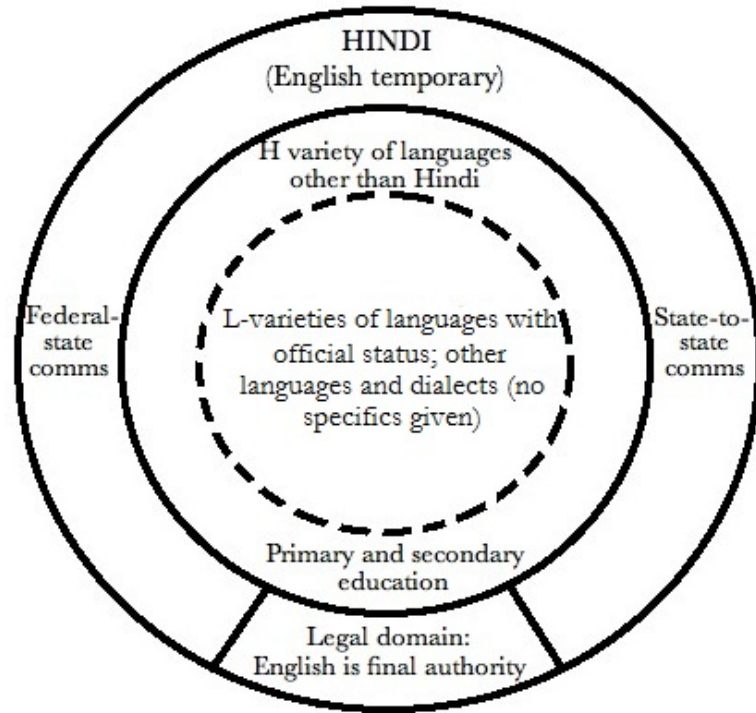


Figure 2.1: Schiffman's taxonomical language policy model applied to India

language planning measures, so too have various types of policy been identified.

Providing a typology of language policy has proved tricky, and I will show that complex solutions to the problem, while enlightening, bring their own problems, and that a simple approach may prove more beneficial. The most comprehensive attempt at a thorough taxonomy of language policies was made in Schiffman (1996). He acknowledges the dearth of extant work before summarising and critiquing previous endeavours by Kloss, Stewart, Ferguson and Fasold, and then presenting his own approach. Schiffman's main criticism lay with the fact that typologies of language policy failed to take into account the sociolinguistic reality of the situations they were attempting to influence, and when an attempt *was* made to account for the language situation, this still fell short due to lack of consideration of the registers and repertoires of the speakers (Schiffman 1996, 30-46). To rectify this, Schiffman proposes a model visualised as up to three concentric circles for each polity, with the innermost circle addressing domains where linguistic varieties low in overt prestige are used, the middle circle examining domains where varieties higher in overt prestige are used, and

the outermost circle tackling specialist registers (Schiffman 1996, 48). Schiffman's example for India is reproduced in figure 2.1. The use of such a model benefits from taking into consideration extra-political factors which may condition domain-specific language use, and also shows awareness of related concepts such as diglossia and societal multilingualism. However, providing so much detailed information within a simple typology of language policy may render comparisons across polities more difficult: similarities between two polities under examination may be obfuscated by differences in the allocation of codes, which may not even be related to language policy. Indeed, this highlights another slight flaw with Schiffman's taxonomy, in that a lot of details of domain-specific language use and societal multilingualism might not be related to the community's language policy, but would still merit inclusion on the model; and if such language practices *are* connected with the language policy of the polity at hand, one could not ascertain whether they are causes or in fact consequences of the policies implemented. Moreover, this approach is much more focused on language-external issues such as the functional allocation of codes (indeed, each concentric circle represents a different code), and so policies that are instances of corpus (language-internal) planning in action might not be so easily addressed by the model. A complex approach therefore throws up problems of its own. However, a simpler solution to the issue of providing a helpful typology of language policy is also found in Schiffman's work. He refers to existing taxonomies which categorise according to, on one hand the *features* of the policies in question, and on the other the *kinds* of policies (Schiffman 1996, 28). The concentric circle model was an attempt to detail the *kinds* of policies, but if we look at existing attempts at detailing policy *features*, a simpler, but still useful, typology is presented. Schiffman reviews Kloss' (1940, 1977) attempts at describing language policies by highlighting a number of features, and then supplements this list with some features of his own. In order to describe language policies, the following features are identified:

- *Promotive versus tolerance.* Originally identified by Kloss, this addresses the fundamental difference between 'promotive' policies which actively foster the use of a given language through the provision of opportunities (such as teacher

training, statutory guarantees, funding etc.) and ‘tolerance’ policies which do not offer said provisions, thus just allowing the language to exist by virtue of not directly impeding its growth. Rather than constituting a bifurcating division, policies lie on a continuum between these two extremes, many being classed as ‘mixed’ (Schiffman 1996, 28-9).

- *Egalitarian versus restricted.* Kloss also distinguishes between ‘egalitarian’ policies, which accord all languages within a given linguistic community equal status, treating all citizens as bilinguals, and ‘restrictive’ policies, which do not (Schiffman 1996, 29).
- *Personal versus territorial language rights.* Kloss discusses the difference between the rights that language legislation bestows on the beneficiaries of the policies. ‘Personal’ rights are held by the citizens themselves, wherever they may go within the area under jurisdiction of the policy, while ‘territorial’ rights are geographically dependent, and the same citizen has different rights according to where s/he is (Schiffman 1996, 29-30).
- *De jure versus de facto.* This distinction was added by Schiffman, and refers to the openly-stated aims and remit of governmental policy (‘de jure’) and the grass-roots practices that result, and may differ greatly from the policy itself (‘de facto’) (Schiffman 1996, 30).<sup>8</sup>

This simple four-point schema allows us great insight into the nature of individual language policies and facilitates comparison across different policies and/or polities. Moreover, it refers explicitly to the policies themselves rather than the polities to which they apply. Although this was a criticism of Schiffman’s, we have seen that including too many details of the sociolinguistic reality at large, while useful, brings its own complications. Other academics have attempted other taxonomies of language policy, and have arrived at similar solutions: Jacques Leclerc’s typology distinguishes

---

<sup>8</sup>Schiffman also added a fifth distinction, between *overt* and *covert* policies, where overt policies make explicit mention of particular codes in a given situation. However this binary opposition buys us relatively little since, even by Schiffman’s own admission, there is a great deal of overlap with *de facto* and *de jure*. As such, overt and covert policies will not be referred to separately.

assimilation policies, non-interventionist policies, sectorial policies, strategic multilingualism policies and several more (Leclerc website). In the interests of parsimony, I have chosen to observe Schiffman/Kloss' typology since it has fewer categories and the different features are set in binary opposition, allowing for easier analysis.

It should be noted that, while I have chosen to follow the definitions of language policy as laid out by Ager, Schiffman and Eastman (inter alia) since they readily allow for a feature-based typology of policies, other academics adhere to different schools of thought. Most notably, Bernard Spolsky's recent work *Language Policy* offers a thoroughly modern conceptualisation of the terminology at hand. Like Schiffman, Spolsky sees the dangers of considering language policy as existing in a vacuum, and as such seeks to highlight the importance of extra-linguistic, situational factors (Spolsky 2004, 41). He stresses the importance that policy applies to all language varieties and to linguistic communities of all sizes (Spolsky 2004, 40). But most radically, Spolsky modifies the existing interpretation of policy as planning in action, instead choosing 'language policy' as the overarching term which is divided into three components:

1. *Language practices* which he defines as 'the sum of the sound, word and grammatical choices that an individual speaker makes, sometimes consciously and sometimes less consciously, that makes up the conventional unmarked pattern of a variety of a language' (Spolsky 2004, 9).
2. *Language beliefs and ideology* which he describes as 'a... community's consensus on what value to apply to each of the language variables or named language varieties that make up its repertoire' (Spolsky 2004, 14).
3. *Explicit policies and plans*, which result from language planning activities (Spolsky 2004, 39). This last point is what linguists to date had referred to as 'language policy', and ties in with our existing conceptualisation of 'policy is planning in action'.

Arguably the future of language policy and planning study lies in this direction, and it is certainly through attempts to reconcile instances of policy with the world

around them that we can truly understand more about the field as a whole. Other approaches (such as Schiffman's concentric circle model) have tried to include extra-linguistic factors and have unwittingly complicated the situation. Spolsky's tripartite division manages to bring in various sociopolitical and ideological factors into the study of the field, while remaining clear and simple. It should be noted however that this is not language policy as we know it, but rather a combination of policy and planning. We should also be aware that Spolsky is not 'reinventing the wheel' here: he simply uses 'language policy' as an umbrella term to cover the links between what we know as policy, planning and any related extra-linguistic factors. I maintain that Spolsky's nomenclature is potentially misleading, as it requires a distinction to be made between 'language policy' (the global term) and 'explicit policies' (what has until now been referred to as 'language policy'). I thus do not adopt his use of the terminology, but fully believe that his approach is a dynamic and innovative, yet simple, way to synthesise the study of policy and planning.

This current trend, which not only seeks to amalgamate the two intertwining threads of policy and planning, but also to incorporate the relevant sociolinguistic and political context is also expressed in Ricento (2006). He concludes that we can all benefit if we think of language issues as 'personal' rather than 'abstract' (Ricento 2006, 21), thus relating language policy issues to the social reality of the linguistic community in question. Moreover, in the same volume, Nancy Hornberger recapitulates a thoroughly integrative framework of policy and planning first proposed in the early 1990s, synthesising the contributions of many of the academics already mentioned (Haugen, Kloss and Cooper *inter alia*). She directly combines policy and planning, listing various policy goals (such as *officialisation, education/school, foreign language literacy* etc.) and categorising them following Cooper's (after Kloss) three-way status / corpus / acquisition planning distinction, as well as Haugen's four-way standardisation model (Hornberger 2006, 29). In promoting such a thoroughly integrative approach, Hornberger is supporting the view that the lines between policy and planning are at best blurry, and that current scholarly output should focus on a more holistic approach, which not only encompasses social realities, but accords



less importance to the distinction between policy and planning. Such a view runs in tandem with Ager's view of policy as 'planning in action'; Hornberger's approach simply highlights that policy and planning constitute two inalienable elements which can be examined using the same framework.

In short, language policy has often (though not exclusively) been used to refer to instances of planning in action, i.e. to concrete items of linguistic legislation. Few attempts have been made to typologise language policy, and those which have follow distinct trends. Some typologies seek to categorise language policy by describing its features, in my opinion the most successful being the classificatory list outlined in Schiffman (1996) drawing on Kloss (1940, 1977), for its simplicity and clarity. Current typologies seek to not only include the parallel fields of policy and planning under one framework (cf. Hornberger 2006), but also stress the point that policy and planning does not occur in a vacuum, and that the sociolinguistic reality of the situation should ideally be taken into account (cf. Spolsky 2004).

**Conclusions on planning, policy and motivation.** We now have a clearer idea of what has been meant by the concepts of language planning and language policy. Traditionally, language planning described the collection of efforts aimed at addressing linguistically problematic situations and issues of language standardisation. Language policy has widely been used in the literature to refer to the concrete items of legislation and linguistic policies. More recent analyses have largely respected this nomenclature (with the notable exception of Spolsky 2004), but have sought to provide typologies which break down any theoretical barriers between the two concepts, viewing them as fundamentally linked (cf. Hornberger 2006) as well as integrating the sociolinguistic reality of the polity under discussion (cf. Schiffman 1996, Spolsky 2004). Having used Ager's interpretation of the terminology as a point of departure, it is clear that such a view falls in line with both contemporary and classic schools of thought on language policy and planning, since it outlines two distinct, yet evidently related and overlapping, fields. Therefore, in light of the information presented in this section, I will adhere to Ager's definition, and thus language planning is to be defined as

the ways in which communities influence language form and function, while language policy is simply official planning (Ager 2001, 5).

Another way in which Ager's approach to the field reflects contemporary trends is his way of addressing the sociolinguistic situation of the polities in question. Ager (2001) outlines various potential motives for the implementation of different linguistic legislation, thereby taking into account the cultural and political background of each case. He distinguishes between five political motivations.<sup>9</sup> *Identity*-driven initiatives are most readily visible through the desire of communities to 'become or to remain nations, and preferably nation-states' (Ager 2001, 13), and indeed Catalonia is cited as a case in point. *Ideology*-based motivation (Ager 2001, 40-55) is centred on a political desire for societal-level developments. *Image* as a motivating factor (Ager 2001, 56-76) refers to the desire to engender or protect a positive perception of a language outside of its linguistic community, as evidenced by the work of the *British Council* for English, the *Alliance française* for French and the *Institut Ramon Llull* for Catalan. *Insecurity*-driven policies (Ager 2001, 77-86) are motivated by a fear of 'the outsider', and are often manifested as oppressive policies against the use of minority group languages (on the status planning dimension) or a zero-tolerance attitude to foreign borrowings (on the corpus planning dimension). Finally, *inequality*-based action (Ager 2001, 87-107) is implemented as an attempt to redress a previously-administered injustice or imbalance, as seen with attempts to rid English of gender-biased terms: post-colonial language planning strategies in Catalonia can also be seen as instances of inequality-driven legislation (to compensate for Francoist oppression), thus showing how one set of policies can have multiple motivations. This simple, five way overlapping taxonomy of motivations for policy and planning helps to provide further insight into instances of linguistic legislation and how they relate to the concrete social realities in which they are found.

In short, it has become clear that planning and policy are two distinct, yet germane concepts, which are relatable to the sociolinguistic reality of the community to which

---

<sup>9</sup>In addition to one speaker-centred motivation, which addresses the integrative and instrumental motivations for language learning.

they apply through examination of their motivations. Ager's initial notion of policy as 'planning in action' has been placed within the historical context of language policy and planning study, and his work has proved even more valuable as he goes into explicit detail regarding the numerous motivations behind language political decisions. It is now necessary to link this cause to concrete top-down linguistic phenomena. The next section focuses on my chosen domain of acquisition planning, and details its link to resultant sociolinguistic phenomena, such as those given in table 2.1 as examples of TDP.

### 2.1.2 Acquisition planning and its relation to TDP

My research has chosen to focus on language-in-education policy in post-Francoist Catalonia, examining the speaker-level consequences of developments in governmental decision-making regarding the medium of instruction (MOI). As will be shown in section 2.2.1.1, the last thirty years have seen a shift from Castilian medium to almost exclusively Catalan medium education. But what does this mean? What do we know about the linguistic consequences of acquisition planning, or of specific instances of this planning in action, i.e. language-in-education policy? This section will examine existing approaches to acquisition planning in order to provide a better framework within which to analyse the Catalan policies under discussion. It will also explain the links between acquisition planning and any resultant phenomena through a simple schema.

**Acquisition Planning.** As already discussed in section 2.1.1, Cooper's (1989, 33) introduction of *acquisition planning* into the existing two-way status / corpus distinction intended to overcome problems with the schema. Tonkin (2005, 127) brings to our attention that education can be used to address issues of both corpus and status: corpus planning subsumes the teaching of concrete linguistic forms, whereas status planning covers issues such as the choice of medium of instruction. As such, a third type of planning was adopted, so as to avoid classifying the same policies under two different planning frameworks. In his definition, Baker (2006, 52) reminds us

that acquisition planning is not solely concerned with educational reform, and that ‘language reproduction in the family’ is equally to be taken into account.

So when this planning is ‘put into action’, it is translated into different types of policies, the most frequently discussed being *language-in-education policy* (cf. Kaplan and Baldauf Jr 1997, Spolsky 2004 inter alia). Spolsky (2004, 46-8) states that such policy initiatives are used to tackle a number of issues, and in his view the most important of these is the decision on which language is to be the MOI in the polity in question. Issues of MOI were of political importance even before the advent of the academic study of language policy and planning. This is attested in an early UNESCO monograph on the fundamentals of education, where it is taken as ‘axiomatic... that the best medium for teaching is the mother tongue of the pupil’ (UNESCO 1953, 6)<sup>10</sup>, although this is accompanied by the caveat that ‘it is not always possible to use the mother tongue in school and, even when possible, some factors may impede or condition its use’ (UNESCO 1953, 11 in Bratt Paulston and Heidemann 2006, 298). The many facets and complicated motivations behind MOI decisions have led to many subtypes of education which meet (or fail to meet) the many and varied sociolinguistic needs of the communities to which they apply. If there is a debate as to the MOI in a given situation, then more than one language is available for selection. We are thus dealing with (at least in some respect) a multilingual situation and as such, Baker’s (2006, 215-6) typology of bilingual education offers a comprehensive selection of the various results of different possible policy decisions related to MOI choice. He identifies three forms of education, each of which draws on the notions of majority and minority language, which should not be viewed as greater or lesser in terms of number of speakers, but rather in terms of power, access to resources and institutional support:

1. Firstly, monolingual forms of education for child speakers of a minority language can either take place solely in the majority language, aiming at linguistic (and often cultural) assimilation, or solely in the minority language, with a view to

---

<sup>10</sup>A stance later problematised by Spolsky who claims that spoken home varieties are rarely, if ever, the same as those taught in schools, and should not be considered as such just because they share the same name (Spolsky 2004, 46).

segregating the minority community and denying them access to the symbolic capital represented by use of the majority language. As such, this form is subdivided into:

- (a) *Submersion programs*, where language minority students are subjected to 100% exposure to the majority language and receive no special consideration (Baker 2006, 216).
  - (b) *Submersion programs with pull-out classes*, where language minority students receive extra ‘compensatory’ classes, still in the majority language (Baker 2006, 219).
  - (c) *Segregationist (apartheid) programs*, where language minority students are denied access to the majority language (Baker 2006, 221).
2. Secondly, ‘weak’ forms of bilingual education make more provisions for language minority students, and do not have complete monolingualism as their aim:
- (a) *Transitional programs* are similar to submersion programs except that the minority language is used by students in the early stages of education, with the aim at the gradual phasing out of the minority language (Baker 2006, 221).
  - (b) *Mainstream programs with foreign language teaching* again are similar to submersion programs, but may include the teaching of the minority language as a foreign language (Baker 2006, 223).
  - (c) *Separatist programs* are rare, and constitute a minority language community detaching itself from the majority language in a statement of autonomy (Baker 2006, 224).
3. Finally, ‘strong’ forms of bilingual education primarily differ with regard to the students which follow them, and aim to secure students’ bilingualism and biliteracy.
- (a) *Dual language programs* contain a balanced mixture of children from different language backgrounds, who follow an entirely bilingual curriculum

(Baker 2006, 242).

- (b) *Immersion programs* are followed by majority language students, who are studying through a language other than their mother tongue (Baker 2006, 242).
- (c) *Heritage language programs* are followed by a mixture of children from different language backgrounds, and are delivered entirely through the (usually hitherto minoritised) language of the community (Baker 2006, 242).
- (d) *Mainstream bilingual education programs* are simply bilingual programs that are implemented in a society which is already bilingual in the languages in question (Baker 2006, 250) (in other words, there is more than one ‘majority’ language being used).

This model is greatly useful when examining issues surrounding MOI choices, as it does more than just classify policies and programs on the sole basis of the language selected as the vehicle for classroom interaction. It also explores the numerous motives behind different policy choices and describes the outcomes that result from the implementation of such decisions. As such, when this taxonomy is applied to the Catalan language policy situation in section 2.2.1, we will be able to infer a great deal of potentially useful information, and more easily compare the Catalan case to other analogous situations worldwide. It should be remembered that, while my later study of language-in-education policy in Catalonia will primarily be concerned with the choice of Catalan as the MOI, the field does not end there. In addition to the selection of codes for domain-specific use, acquisition planning addresses other areas, such as the teaching of foreign languages (also mentioned, somewhat cursorily, in Baker’s taxonomy) (Spolsky 2004, 47), and the correction of non-standard language in (at least *prima facie*) monolingual contexts (Bratt Paulston and Heidemann 2006, 301).

In short, acquisition planning (and the related, applied concept of language-in-education policy) refers to conscious attempts to influence the acquisition of a particular linguistic variety, be that in the home or in school. Of these two branches, more attention is paid to language planners’ endeavours to manipulate language use

in the education system. Indeed, the education sector is often viewed as ‘the transmitter and perpetuator of culture [and]... the site of language planning’ (Kaplan and Baldauf Jr 1997, 123). Within the study of language in education, a great deal of importance is attached to the choice of a particular medium of instruction and the associated benefits, or detriments, that this entails for the languages within the polity under examination.

**Acquisition Planning and resultant phenomena (TDP).** If the most important issues within acquisition planning can be reduced to MOI choice, then we must establish how these decisions lead to top-down phenomena, as exemplified in table 2.1. The example of a TDP in form resulting from language planning is the ‘introduction of specific variants of linguistic variables’, and a relevant example given on page 61 is the ‘introduction of a native alternative to a foreign borrowing through corpus planning’. The link between acquisition planning and corpus planning has been stated above (see page 83), i.e. the acquisition of concrete linguistic forms. All of these pieces fit together neatly as shown in the following schema:

1. A process of corpus planning decides upon the normative status of a given lexical or grammatical item, to be used in the system chosen as MOI.



2. This item is then diffused using the education system, and the linguistic community is thus made aware of the item’s normativity.



3. The potential use of this item, and awareness of its normativity, to the detriment of other, non-native (or at least non-normative) alternatives constitutes an example of TDP in language form.

The prime example of TDP in language form is therefore the widespread acceptance (and potential use) of a lexical or grammatical item deemed by an official body to be normative, and the rejection of all alternative, non-approved forms. The choice

of a given code as MOI can also entail other, closely related TDP in language form: for example, if the language is to be used as a vehicle for education, it needs to be fully elaborated in order to fulfil that role. Therefore any creation of new terminology (or any elaboration process in general) is a language form-based phenomenon that has arisen from conscious corpus planning initiatives, i.e. a TDP.

TDP also occur in language function, and on page 61, we have seen the example of an ‘increase in overt prestige of a variety through status planning’. This can be very transparently linked to acquisition planning initiatives, particularly the choice of MOI. If a language is chosen as the vehicle of education, it gains prestige as a consequence: if this language is used in schooling, the presumption is that it can be used as a marker of academic achievement, that it will be later required for access to the job market, and so on.

### 2.1.3 Summary

The section thus far has given an overview of language policy and planning, as the prime example of conscious decision-making which can provide the background situation to TDP. A great many important developments have occurred over the last fifty years, and many of the early advances are still applicable today, such as Haugen’s 1966 standardisation model and the status / corpus planning distinction introduced by Kloss in 1969, and later modernised by Cooper to include the third strand of acquisition planning. This last notion of acquisition planning proved highly useful when establishing a theoretical link between language policy and planning on one hand, and instances of TDP on the other. The use of a given code as the vehicular language of teaching means that any officially-backed decisions concerning the form of the language are more quickly and readily diffused through the linguistic community, giving TDP in form. Moreover, it was deduced that the decision to make a given language the medium of instruction in the education system could positively impact upon the overt prestige of said language, motivating TDP in function.



## 2.2 Language policy and resultant phenomena (TDP) in Barcelona

Thus far, this chapter has offered an overview of the history and theory of language policy and planning, paying particular attention to the development of acquisition planning. This provided us with a number of theoretical taxonomies and frameworks – such as Haugen’s standardisation model, Cooper’s status / corpus / acquisition distinction (after Kloss) and Baker’s taxonomy of multilingual education systems – within which to analyse concrete instances of language policy and planning. Moreover, these instances of planning and policy, particularly acquisition planning, have been related to potential manifestations in language form and function through my notion of TDP.

It is now necessary to apply all of this information to the community under examination. Table 2.1 offered examples of the background situation as well the concrete linguistic phenomena that make up TDP. Section 2.2.1 will draw on the theoretical background given in 2.1 and offer an in-depth analysis of language policy and planning initiatives in Barcelona, thus addressing the relevant background situation in the linguistic community. Section 2.2.2 will then examine the potential manifestations of phenomena in language form and function of in Barcelona that can be attributed to language policy development (introduction of native alternatives to borrowings, increase in overt prestige etc.).

### 2.2.1 Language policy in Barcelona

This section will apply the theoretical background and framework within which to examine language policy and planning, to better understand the speech situation at hand, i.e. Barcelona.<sup>11</sup> This summary will also underline the particular importance of language-in-education policies in present-day Catalonia, thus reaffirming the im-

---

<sup>11</sup>It should be pointed out that, particularly in the recent age of post-Franco democracy, the linguistic policies followed in Barcelona are those of the autonomous community of Catalonia as a whole. As such, the important linguistic legislation outlined is applicable to the entire autonomous community of Catalonia (and not just the city of Barcelona), unless otherwise stated.

portance of the role of acquisition planning in the recent linguistic history of the community. The thesis' focus on acquisition planning will therefore be shown to be entirely congruent with the specific needs, practices and history of the bilingual community of Barcelona. This section is largely chronological in structure, and builds on the more general sociolinguistic history of the region offered in section 1.1, concerning itself solely with the development of linguistic legislation.

As we have already seen, overt linguistic policies with aims detrimental to the status of the Catalan language have existed for centuries. Within the aforementioned *Decrets de Nova Planta* of 1716 (see page 25), Castilian was introduced as the official language of justice and administration, and Catalan language universities were closed in Barcelona, Tarragona, Lleida, Girona and Vic (Enciclopèdia Catalana, s.v. *Decrets de Nova Planta*). Such early examples of status planning served the dual purpose of elevating the standing of Castilian, while diminishing that of Catalan.

The next important innovations came with the resurgence of the vitality of Catalan in the nineteenth century after the *Renaiçença* and *Modernista* movements. It should be pointed out that, if we return to Haugen's standardisation model, the four stages are not necessarily sequential: this does not only mean that the processes can come in any order, but that stages (and potentially, the whole cycle) can be repeated as the speech situation demands. The reforms of Catalan at the early twentieth century encompassed elements of all four of Haugen's stages of selection, codification, acceptance and elaboration. This was largely due to the work of Pompeu Fabra, which is still recognised in today's democratic Catalonia:

Es sabut que els catalans comptem... amb una llengua codificada – codificació que, gràcies a Déu (més ben dit, gràcies a Pompeu Fabra), va més enllà de la mera forma gràfica dels mots (Segarra 1985, 7).<sup>12</sup>

Fabra's selection work (in the Haugen sense) comprised decisions made as to which variety of Catalan would go towards the fully codified language. One of the prime crite-

---

<sup>12</sup>TRANSLATION: It is known that we Catalans are blessed with... a codified language – codification which, thank God (or more accurately, thanks to Pompeu Fabra), goes beyond the mere spelling of words.

ria for establishing a unified, codified Catalan was that of diasystematicity (Costa Carreras 2009, 43), i.e. that as many diatopic variants of the language be included as is useful and necessary. For example, the decision to mark feminine plural nouns with the suffix *-es*, as in *les cases* is more representative of the pronunciation of Western Catalan (Eastern Catalan not differentiating between unstressed *-a* and *-e*). However, verbal morphology is more indicative of Eastern Central Catalan: compare the standard *canto* (Eastern Central Catalan) with forms such as *cant* (Balearic Catalan) and *cante* (Valencian). The codification efforts of the period are almost exclusively found within Fabra’s work, the main aim of which was to fully codify Catalan through normativising orthography, syntax and morphology and managing linguistic diversity.

Turning to acceptance, it was necessary for Fabra’s norms to be diffused through the community, so they could be potentially assimilated by language users. One of the ways in which this was achieved was through the aforementioned *converses filològiques* (see page 28), a total of 915 short articles which appeared in the Catalan press, and were thereby readily accessible. Each brief text addressed a particular facet of language use, instructing readers on what was normatively permitted. In a tradition which we will see continues to this day, much of this work was dedicated to the eradication of Castilianisms. This could be undertaken on a general level such as in *converses* which set out to identify many different types of Castilianisms, some deemed more ‘obvious’ than others, warning readers that ‘no [volem] acontentar-nos amb una depuració soma de la llengua [catalana], consistent en la sola extirpació dels castellanismes més aparents’ (Fabra 1919a ; cf. also Fabra 1919d).<sup>13</sup> The vast majority of the *converses*, however, deal with very specific instances of language use. There are articles dedicated to particular examples of verbal or nominal morphology where interferences from Castilian may appear. For example, the use of a faux Catalan suffix *-ès* corresponding to the Castilian *-ez*, is to be avoided: normative Catalan requires *-esa*, as in *escassesa* (‘scarcity’, Cast. *escasez*) or *senzillesa* (‘simplicity’, Cast; *sencillez*) (Fabra 1919b). Other articles aim to clarify the complexities of the clitic pronoun system of normative Catalan (cf. Fabra 1923a, Fabra 1923b), while

---

<sup>13</sup>TRANSLATION: We [do not want to] content ourselves with a shallow purification of the [Catalan] language, consisting only of the eradication of the most obvious Castilianisms.

yet more discuss how normative Catalan often requires a different verb tense than might be used in Castilian (cf. Fabra 1920, Fabra 1925a). When added to this, topics as wide-ranging as the avoidance of over-Catalanised forms such as *Manxester* and *Xicago* (Fabra 1925b) and the correct usage of voiced and voiceless phonemes intervocally (Fabra 1919c), we are presented with a great deal of information concerning what does and does not constitute Fabrian normative Catalan. Moreover, these *converses* were short, written in an engaging style, and appeared almost daily in newspapers such as *La Publicitat* for the best part of ten years. Readers were thus exposed to normative Catalan, and were informed as to how to avoid non-normative Castilianisms.<sup>14</sup> Given their potential to distribute information regarding normative language use, it is of little wonder that articles similar to the *converses* have undergone a resurgence in recent years. These descendants of Fabra’s short articles will be addressed on page 106. However, the *converses* were not the only means of diffusing the new normative standard for Catalan. Acceptance of these planning efforts, and an increase in status for Catalan in general, was also achieved through related governmental policies that allowed for greater opportunities to use the language in official scenarios, a fact also recognised in the contemporary literature:

És cert que si avui tenim una ortografia unificada ho devem en paraules de Miracle «al patriotisme, a l’entusiasme, a la visió claríssima i al sentit de responsabilitat de tot un Enric Prat de la Riba», o sigui a una obra de govern, i no pas exclusivament a la personalitat de Pompeu Fabra (Segarra 1985, 9).<sup>15</sup>

Finally, ongoing processes of functional, language-internal elaboration were also addressed in Fabra’s early twentieth-century work, whereby ‘synonymous or equivalent constructions [were] distributed on a stylistic scale between greater and lesser formality’ (Costa Carreras 2009, 53). As such, it is clear that the first three decades

---

<sup>14</sup>Present-day tactics to avoid Castilianisms will be addressed later in this section, while details as to the nature of the interference itself can be found in chapter 3, particularly on page 144ff.

<sup>15</sup>TRANSLATION: If today we have a unified spelling system, we owe it to, in the words of [Catalan historian Josep] Miracle, ‘the patriotism, enthusiasm, clear vision and sense of responsibility of a certain Enric Prat de la Riba’. That is to say, [we owe it] to the work of governmental bodies, and not just to Pompeu Fabra.

of the twentieth century saw a great leap in Catalan language planning of all types. This aided the status of the language and furthered its ability to be officially used in Catalan-speaking regions, including Barcelona.

However, all was to change with the repression which would characterise mid-twentieth century Spain. As outlined in section 1.1, apart from a brief hiatus during the Second Spanish Republic (1931-6), Catalans were subjected to a number of cruel and oppressive policies under two successive military dictatorships, that of Primo de Rivera (1923-30) and Franco (1939-75). Official language policies of this period, whose aim was arguably to eliminate the use of Catalan in line with the trope of *España: una, grande y libre*, adopted two guises: overt oppression and dialectalisation. Overtly oppressive language policies included the removal of Catalan from all official domains through the abolition of the Catalan Autonomy Statute in April 1938, and a ministerial decree in August of the same year which deemed all declarations or contracts made in any language other than Castilian to be invalid (Solé i Sabaté and Villaroya 1994, npn), not to mention the threat of dismissal or imprisonment for any public sector workers caught using Catalan. Other such overt policies prohibited publication in Catalan, with catastrophic consequences for the industry: in 1933, Catalan language publishing comprised 23.1 % of Spanish publications, whereas in 1975 this figure was just 4.5% (Solé i Sabaté and Villaroya 1994, npn). Such oppression was in addition to government banners and slogans placed around Catalonia, such as the following placed in Lleida: *Si eres patriota habla español... si no lo eres, fastídiate y háblalo también*.<sup>16</sup> Dialectalisation policies were also employed, in a ‘divide and rule’ strategy to undo any semblance of unity between Catalan-speaking peoples. An official government textbook from 1939 to be used in schools detailed the language situation in Spain as follows:

- ¿Se habla en España otras lenguas más que la castellana?
- Puede decirse que en España se habla sólo la lengua castellana, pues aparte de ésta tan sólo se habla vascuence que, como la lengua única, sólo se emplea en algunos caseríos vascos y quedó reducido a funciones de dialecto por su pobreza

---

<sup>16</sup>TRANSLATION: If you’re a patriot, speak Spanish... If not, get over it and speak Spanish anyway.

lingüística y filológica.

– ¿Cuáles son los dialectos principales que se hablan en España?

– Los dialectos principales que se hablan en España son cuatro: el catalán, el valenciano, el mallorquín y el gallego (taken from Solé i Sabaté and Villaroya 1994).<sup>17</sup>

This dialectalisation runs deeper than the mere denial of Catalan as a language. It is no accident that such efforts appear after the early-twentieth century boom in Catalan language planning detailed above. As such, at the outset of the Franco era, Catalan was a fully-codified language in a very strong position and in order to destroy usage of the language, this unity had to be undone. This was arguably (at least temporarily) successful on several fronts. Franco's dialectalisation policies meant that Fabrian codified norms were becoming ever more distant from the everyday language, while purists and traditionalists, appreciating the lamentable state in which the language found itself, considered deviation from Fabrian norms as sacrilegious. Thus a cleavage began between (in the words of Salvador Espriu) 'a moribund tongue [with] indigestible words..., almost unintelligible by now to many of us' and the language spoken on the streets (Costa Carreras 2009, 22). Therefore, applying the models outlined in the first part of the chapter, Franco's policies arguably encompassed both status and corpus planning measures: overt oppression seeking to reduce the status of Catalan by limiting its function, and dialectalisation tackling language-internal issues of form. In short, the large-scale repression of Catalan by the Franco regime was grounded in concrete political decisions which sought to undermine and ultimately eliminate the language from a unified Spain.

After the death of Franco in 1975, Spain entered a period known as the *transición*, which saw its development from a centralist fascist dictatorship to a democratic monarchy made up of seventeen *comunidades autónomas*. The policies which enshrined such a change were replete with overt discussion of how the language situ-

---

<sup>17</sup>TRANSLATION: – Are there other languages spoken in Spain aside from Castilian?

– We can say that in Spain, only Castilian is spoken, since otherwise there is just Basque which, as the sole language of communication, is only used in a few Basque hamlets. The (Basque) language itself is limited to a mere dialect, due to its linguistic and philological poverty.

– What are the main dialects spoken in Spain?

– There are four main dialects spoken in Spain: Catalan, Valencian, Mallorcan and Galician.

ation should be addressed throughout the country, beginning with the 1978 Spanish Constitution, the document on which the modern democracy is based. Article three of the preamble accords Castilian the status of official language of the state, detailing that all Spaniards have the duty to know it and that they have the right to use it, before stating that ‘the other Spanish languages’ (*las demás lenguas españolas*) will be co-official in their respective autonomous communities (Gobierno de España 1978). The following year, the Generalitat de Catalunya drafted and approved an autonomy statute which, predictably, placed a great deal of emphasis on the role and status of the Catalan language. Article three accords Catalan alone the title of ‘language of Catalonia’ (*llengua pròpia de Catalunya*), before going on to state that both Catalan and Castilian are official languages of the region. The Generalitat then outlines its mission statement to:

Guarantee the normalised and official use of both languages, to take the necessary measures to assure that they are widely known, and create conditions such that the two languages are totally equal when it comes to the rights and duties of the citizens of Catalonia (Generalitat de Catalunya 1979, my translation).

This last declaration is arguably an attempt to redress what is seen by many as an imbalance resulting from the 1978 constitution whereby all citizens of Spain (and therefore Catalonia) have the ‘duty’ to know Castilian, but merely the ‘right’ to know (languages such as) Catalan. These early items of democratic legislation set the tone for the type of Catalan language policy which has been emblematic of the last thirty years, namely status planning measures that focus on Haugen’s acceptance stage, since they ultimately aim to create greater opportunities for the use of the Catalan language.

Once the new Spanish state was established and the roles of the Generalitat de Catalunya were delimited, and once the system had been in place for a few years, the foundational piece of linguistic legislation regarding the role that the Catalan language was to play was drafted and approved: the *Llei de Normalització Lingüística* (LNL) of 1983. This law openly aims to rectify the damage perpetrated on the Catalan language since it lost its official status with the *Decrets de Nova Planta*, and to

safeguard the use of Catalan in administration, education, the media and numerous other domains. The law serves to build upon article three of the 1979 autonomy statute through the normalisation of the use of Catalan in all fields, and summarises its main objectives in article two of its preamble (Generalitat de Catalunya 1983a):<sup>18</sup>

Atesa la situació lingüística de Catalunya, són, doncs, objectius d'aquesta Llei:

- a. Emparar i fomentar l'ús del català per tots els ciutadans.
- b. Donar efectivitat a l'ús oficial del català.
- c. Normalitzar l'ús del català en tots els mitjans de comunicació social.
- d. Assegurar l'extensió del coneixement del català.

The document is divided into titles concerning officiality of use, teaching and learning of Catalan, language and the media, governmental support, and the use of Aranese.<sup>19</sup> The LNL is a textbook example of status planning in action, and focuses on Haugen's acceptance stage. When reviewing Schiffman's typology of policies, the LNL is largely promotive in that it actively offers opportunities which favour the growth of Catalan, and egalitarian in that it aims for the co-officiality of Catalan and Castilian. It arguably addresses questions of territorial, rather than personal, rights, as it strives for 'la normalització de l'ús de la llengua catalana en tot el territori de Catalunya'<sup>20</sup> (Generalitat de Catalunya 1983a, 'lleis' section), and is de jure in nature. This document was the most influential and instrumental in setting up the current Catalan language immersion education system, and its Title II (which addresses the school system) will be explored in greater depth in section 2.2.1.1.

The LNL was followed in 1998 by the *Llei de Política Lingüística* (LPL), which aimed to consolidate and expand upon the original 1983 document. The new law would take into account the sociolinguistic situation in which Catalonia found itself,

---

<sup>18</sup>TRANSLATION: **a.** To protect and encourage the use of Catalan for all citizens. **b.** To effectively implement the official use of Catalan. **c.** To normalise the use of Catalan in all media of social communication. **d.** To ensure the extension of Catalan language competence.

<sup>19</sup>Aranese is a variety of Gascon spoken in the Vall d'Aran in north-western Catalonia. Its use is protected by the 1979 autonomy statute and subsequent legislation.

<sup>20</sup>TRANSLATION: The normalisation of Catalan language use in the entire territory of Catalonia.



thanks to the ‘transcendental changes’ brought about by the LNL and related legislation (Generalitat de Catalunya 1998b). Much like the LNL, the LPL was divided into individual chapters concerned with language officiality, education, media and governmental support, but was remodeled to include separate sections regarding the forms of personal and place names, and socioeconomic activity. Three characteristics of the LPL highlight a change in tone from previous language legislation. Firstly, a greater attention to detail (both with regard to content shared with previous documents, and to new material) is evident in the 1998 law. This can be seen in a more extensive discussion of administrative language use, such as the stipulation that Catalan be employed for internal government communication (chapter 1, article 9), and in an elaborated discussion of language and the media (chapter 4, articles 26 and 27). Secondly, the LPL places greater explicit emphasis on the co-officiality and proposed equal status of Catalan and Castilian, which can be seen in the discussion of linguistic rights. Article 4.1 of the preamble of the 1998 LPL states how everyone in Catalonia has the right to know the two languages and to express themselves in whichever they choose (‘a Catalunya, tothom té dret a... conèixer les dues llengües oficials [i d’]expressar-se en qualsevol de les dues llengües’), while the corresponding article 2.1 of the 1983 LNL simply outlines citizens’ rights to know and use Catalan (‘el català és la llengua oficial de Catalunya... tots els ciutadans tenen el dret de conèixer-lo i d’expressar-s’hi’). Thirdly, if we examine the policies using the aforementioned Haugen and Kloss frameworks, the LPL moves beyond status planning and includes some corpus measures which fit into Haugen’s elaboration stage. This can be witnessed in article 29 of chapter 4 of the LPL, which calls for the translation into Catalan of new multimedia and I.T.-based products which use technologically innovative terminology, and by extension, the standardisation of such terms in Catalan. Again, the specifics of this law as regards education policy will receive greater attention in section 2.2.1.1.

Aside from these four major pieces of democratic legislation (namely, the 1978 constitution, the 1979 autonomy statute, the 1983 LNL and the 1998 LPL), a number of smaller decrees have been put in place over the last thirty years to more accurately

determine the roles to be played by Catalan and Castilian. Such legislation addresses the use and function of language in a multitude of fields, ranging from administration of several kinds (state level, local level, judicial and military), to signposting, private enterprise and personal names as they appear on the civil register. These laws often focus on micro-level details of language use, such as the requirement for official printed state documentation to be bilingual (Generalitat de Catalunya 1999b, article 8), the choice of either official language for minute-taking in government meetings (Generalitat de Catalunya 1983c, article 86) or the need for local police officers to provide certificates demonstrating Catalan language proficiency (Generalitat de Catalunya 1991, article 24). These decrees are clear examples of status planning, as they address the relative standing of Catalan and Castilian, and their use in society, and can be seen to correspond to Haugen's acceptance stage, given that they create opportunities for the wider application of the language in society. However, governmental initiatives have not ended there. It was decreed that the Institut d'Estudis Catalans (IEC) should be 'la institució acadèmica encarregada legalment d'establir la normativa de la llengua per a tota la comunitat lingüística catalana' (Institut d'Estudis Catalans 2009, xi).<sup>21</sup> As such, the IEC (in particular its *secció filològica*) has been greatly active both in ensuring acceptance and undertaking elaboration duties, through the creation of grammars, dictionaries, articles and services through which speakers can answer any questions related to normative Catalan (such as Optimot, to be explained on page 104). Moreover, in 1985, the Generalitat de Catalunya, along with the IEC, created TERMCAT, whose remit comprises linguistic elaboration:

El TERMCAT té com a missió garantir el desenvolupament i la integració de la terminologia catalana en els sectors especialitzats i en la societat en general, mitjançant la creació contínua d'eines i de recursos innovadors i de qualitat. <sup>22</sup>  
(Termcat website).

---

<sup>21</sup>TRANSLATION: The academic institution legally charged with establishing the [Catalan] language norm for the entire Catalan linguistic community.

<sup>22</sup>TRANSLATION: TERMCAT's mission is to guarantee the development and integration of Catalan terminology in specialist sectors, as well as the society in general, through the continual creation of innovative, high-quality tools and resources.

The work of bodies such as the IEC and centres like TERMCAT will be explored in depth on page 103, where the precise nature of manifestations of TDP in language form in Barcelona will be discussed. However, for the time being, we can still see that, in addition to the 1983 LNL and the 1998 LPL, the Generalitat de Catalunya laid down numerous measures to guarantee both the elevated status and elaborated corpus of Catalan.

The last five years have seen yet further modernisation of language policy, not least through the approval of a new Catalan autonomy statute in 2006. However, new developments were seen by many Spanish political forces (above all those of the right-wing *Partido Popular*) as having gone too far, and as such were deemed unconstitutional and subject to reinterpretations announced by the Constitutional Tribunal on 28 June 2010. Objections were made to the continued reference to Catalonia as a ‘nation’ and to the linguistic rights and duties of its people. Since the aforementioned 1978 Constitution, it has been all Spanish citizens’ *duty* to know Castilian, but within Catalonia, people merely have the *right* to use Catalan. Autonomy statutes and linguistic policies have since attempted to redress what is seen as an imbalance, with the 1998 LPL making this an explicit mission statement (section 3 of the preamble). The 2006 text sought to resolve the issue once and for all in Catalan’s favour as follows:

El català és la llengua oficial de Catalunya. També ho és el castellà, que és la llengua oficial de l’Estat espanyol. Totes les persones tenen el dret d’utilitzar les dues llengües oficials *i els ciutadans de Catalunya tenen el dret i el deure de conèixer-les* <sup>23</sup> (Parlament de Catalunya 2006, article 6.2, my emphasis).

Such a bold declaration that all Catalan citizens had the *duty* to be bilingual in Catalan and Castilian was seen as diametrically opposed to the 1978 Constitution. The subtle shift in focus from territorial rights to personal rights (see page 78) should also be noted: whereas the LPL stipulated ‘a Catalunya, tothom té dret a...’ (in Catalonia, everyone has the right to...) (Generalitat de Catalunya 1998b, article 4), the new

---

<sup>23</sup>TRANSLATION: Catalan is the official language of Catalonia, as is Castilian, the official language of the Spanish state. All people have the right to use both official languages and *the citizens of Catalonia have the right and the duty to know them.*

statute talks about the personal rights of the ‘ciutadans de Catalunya’ (the citizens of Catalonia) (Parlament de Catalunya 2006, article 6.2). Perhaps unsurprisingly, the Tribunal deemed the article unconstitutional, since the linguistic duties of Spanish citizens were already laid out (see page 95). In protest, over a million people took to the streets of Barcelona on 10 July 2010 in the largest demonstration in Catalan history. As such, it is evident that language policy in Catalonia is still evolving, and has the potential to galvanise the populace into impassioned collective action.

### **2.2.1.1 Present-day language-in-education policies in Barcelona**

Many of the developments in Catalan language policy that have been thus far described have directly impacted on the education system of the region. The teaching of Catalan had been prohibited since the Reial Cèdula of 23 June 1768 (Generalitat de Catalunya 2001a). Between 1939 and 1975, the Francoist dictatorship implemented an exclusively monolingual policy for children in compulsory full-time education, and the attitude towards minority languages is succinctly expressed in the quote from the school textbook cited above, i.e. that they are merely non-standard dialects of Castilian Spanish, and as such have no place in the classroom. This monolingual Castilian stance was about to go through an almost complete U-turn, and by the school year 1999/2000, 94% of primary education and 85% of secondary education would be either Catalan-medium or predominantly Catalan-medium (Vila i Moreno 2008, 34-5).

Article 15 of the 1979 Autonomy Statute transfers full responsibility for education to the Generalitat. When this is viewed in conjunction with the document’s opening gambit of declaring Catalan the ‘llengua pròpia’ of Catalonia, it is clear that future legislation was to be concerned with the spread of Catalan through the school system. Title II of the 1983 LNL (dedicated to education) opens by declaring Catalan the ‘llengua pròpia’ of the education system, just as it is of the region as a whole. However, article 14.2 goes on to state that in the early years, children have the right to receive education in whichever of the two official languages is habitual for them. Indeed, article 14 stresses that both languages must be taught (as subjects) through-

out the system, and that students must leave with a high level of competence in both Catalan and Castilian. Nevertheless, the section ends by reminding us that the ‘normal language’ of the education system is to be Catalan, whether in the classroom or outside. The 1998 LPL explicitly restates the preferential status accorded to Catalan in the classroom (article 21), and goes into more detail regarding the responsibilities of universities, who are also to do all they can to promote and protect the use of Catalan in their internal and external business. While the LNL and the LPL provided the basic tenets of what a largely Catalan-medium system was to achieve, other laws and decrees detailed how Catalan was actually to be used on a day-to-day basis, most notably the DOGC 362 of 1983, which was implemented through several measures passed in late 1983 and early 1984. Article 6 of this legislation explains what is to be understood by the term ‘llengua pròpia’, i.e. that the generalised and progressive use of Catalan be spread to all education levels. Articles 7 and 8 address the use of the language within the EGB (*Educació General Bàsica*, i.e. until the age of fourteen), covering how the language is to be gradually introduced for those whose habitual language is Castilian, and what specific subjects are to be given in Catalan: for example, children in the early years of schooling will be taught either social sciences or natural sciences through the medium of Catalan, while those later on will receive Catalan-medium education for both. Article 9 discusses how Catalan is to be used for those in compulsory education over the age of fourteen, stating that Catalan must be used for two of the four principal subject areas of natural sciences, art, history and mathematics. The Spanish education system as a whole proceeded to undergo a series of reforms, such as the LOGSE (*Ley Orgánica General del Sistema Educativo*) of 1990 which did away with the EGB and introduced a primary/secondary distinction, as well as the bachillerato as a post-secondary, pre-university two-year qualification. The Catalan decrees and statutes have been updated in accordance with state-level reforms, and as recently as July 2009, article eleven of the 12/2009 Law continues to stress that Catalan is the default language of teaching in the school system. The Catalan-medium education system that is in place today in Barcelona was therefore set up as a result of the 1983 LNL, and subsequent decrees (primarily the DOGC 362

of 1983) which secured the details of how exactly it was to be put into practice.

When we examine recent Catalan language-in-education policies, it is clear that they do not fit neatly into one of the subcategories listed in Baker's typology, given in section 2.1.2. The decision that allows Castilian-dominant children to receive Castilian-medium education in the early years falls in line with what Baker terms 'transitional programs', but it should be noted that the aim is not the phasing out of Castilian, since all legislation decrees that students should leave the system with a high level of competence in both languages. The Catalan case constitutes an example of Baker's 'strong' forms of bilingual education, and while in 1983, the schemes proposed by the Generalitat would have fitted largely within the description of 'heritage language program' (due to the hitherto minoritised status of Catalan), nowadays the system is more akin to a simple 'mainstream bilingual program'. Confusingly, the term 'language immersion program' (Cat *immersió lingüística*) is often employed with reference to the Catalan situation, though this usage does not correspond to Baker's typology (where this refers to education through a non-majority language).

This thesis' focus on language-in-education policy is therefore not only grounded in the common-sense notion that developments in the schooling system have significant bearing on the intergenerational transmission of a language, but also in the fact that Catalonia (and by extension, Barcelona) has recently undergone a radical overhaul regarding the medium of instruction used in school. A focus on language-in-education policy in Barcelona should therefore yield interesting results, based on these two overlapping factors.

### **2.2.2 TDP in Barcelona: Details of phenomena in language form and function**

Table 2.1 outlines background situations, as well as the potential manifestations in language form and function that together exemplify TDP. Having now looked at the TDP background situation in Barcelona, i.e. language policy and planning, it makes sense to examine the phenomena in both form and function that have arisen as a result of recent language planning strategies. Just as in section 2.1, we examined

the background and then the phenomena on a theoretical level, so too here have we already addressed the background in section 2.2.1, before now moving on to discuss the potential manifestations in the context of Barcelona.

**TDP in language form in Barcelona.** On page 61, the specific example of a type of TDP in language form was ‘the introduction of a native alternative to a foreign borrowing through corpus planning’. In many ways, the result of such linguistic divergence between the native ‘host’ language (in this case, Catalan) and the non-native ‘graft’ language (here, Castilian) is the polar opposite of the instances of contact-induced language convergence, to be discussed in section 3.2. However, it is important to remember that the two processes are not strictly parallel, and that indeed the choice of examples of divergence for TDP and convergence for BUP is largely determined by the facts of the linguistic community at hand. It just so happens that Catalan corpus planning has been largely concerned with increasing the linguistic distance between Catalan and Castilian, and that most of the contact-induced language phenomena attested in Barcelona are of convergence between the languages. Obviously, there is an intrinsic link here: if convergence between Catalan and Castilian were not occurring, there would be little need for corpus planning measures to increase (or at least maintain) a degree of dissimilarity between the codes. The most important thing is that the linguistic phenomenon arises from conscious decision-making, in this case language policy (and more specifically when looking at language form-based phenomena, corpus planning). As such, the choice of policy-motivated language divergence (and indeed, of language planning as a background phenomenon of TDP in general) is merely an example that is particularly pertinent to the Catalan situation, given the emphasis on divergence between Catalan and Castilian.<sup>24</sup>

In the last thirty years, Catalan has undergone several such types of TDP in form, which can for the most part be classified as ‘attempted *rectification* of Castilianisms’ and/or ‘neologisms coined for elaboration purposes’.

---

<sup>24</sup>Although policy-motivated divergence has been chosen here, language situations can also be characterised by policy-motivated language convergence. Indeed, this is even witnessed in the Catalan speech situation, where the work of Fabra and the IEC often sought to converge the different diatopic varieties of Catalan to form one, normative variety of the language.

Some work arguably comes under both of these headings, in that through the full elaboration of Catalan, and through ensuring that speakers are made aware of ‘native’ Catalan items or constructions, there should no need to resort to Castilianisms. In such work, the IEC and its *secció filològica* is a key player, ensuring the normative Catalan standard remains as free of Castilian influence, and as fully elaborated, as possible. The IEC’s current output covers a wide range of fields. Dictionaries are produced, not just of non-specialist vocabulary, but also tailored to the specific needs of experts in fields as diverse as geology (*Diccionari de geologia*), the law (*Diccionari jurídic català*), forestry (*Vocabulari forestal*) and biology (*Vocabulari de biologia de la reproducció*). An IEC *Gramàtica de la llengua catalana* appears on their website, in addition to the service Optimot (in conjunction with the aforementioned TERM-CAT), which allows Catalan language users to pose a number of questions concerning correct language usage. Vocabulary, grammar and translation queries can be solved by Optimot using an advanced search engine, and should the online system not satisfactorily answer a given question, users have access to a personalised help service, whereby they can consult an expert for help with their exact issue. The IEC also offers advice on how Catalan texts should transcribe languages such as Hebrew and Russian which use other alphabets, as well as providing information on how Catalan family names should be correctly spelled (as a countermeasure to the aforementioned Francoist policy that all surnames should be written using Castilian orthography). A useful tool compiled by the IEC is the *Nomenclàtor oficial de toponímia major de Catalunya*, a reference work which includes over 52,000 names of geographical features in Catalonia, covering everything from streams and mountain ranges to population centres and transport links. This is deemed important since ‘els noms de lloc... [són] un patrimoni cultural col·lectiu que tenim l’obligació de preservar i transmetre a les generacions futures, i sobretot de procurar que no siguin anorreades per noms forasters... que reflecteixin una voluntat de substitució cultural que s’ha produït... durant els règims polítics especialment hostils a la nostra llengua que hem hagut de patir’ (Institut d’Estudis Catalans 2009, xi).<sup>25</sup> So as we see here, the elaboration

---

<sup>25</sup>TRANSLATION: Place names... [are] a form of collective linguistic heritage which we have the obligation to preserve and pass on to future generations. Above all, we must ensure that these [place



and maintenance of Catalan, even when considering topics such as toponymy, are considered by the IEC and the Catalan government as a direct means of regaining control after former (chiefly Francoist) oppression.

Other output is perhaps more easily classified as either ‘rectification of Castilianisms’ or ‘creation of neologisms’. Focusing on the first of these categories, another means by which the IEC is ensuring that Catalan stay free of Castilian influence (and of particular interest to this work) is through the creation of *propostes per a un estàndard oral de la llengua catalana*. This work is in two parts, covering phonetics and morphology, and its role can be summarised as follows:

La present proposta s’insereix en una obra de continuació de la tasca codificadora de la llengua catalana iniciada i duta a terme per l’Institut d’Estudis Catalans. Aquesta codificació mancava d’una ortoèpia; el fascicle sobre fonètica és ja una contribució en el sentit d’esmenar aquesta mancança... Aquesta proposta es refereix – i es limita – a l’estàndard *oral*, que és el que afecta directament els mitjans auditius i audiovisuals, l’ensenyament en general i qualsevol altre ús públic de la llengua... *Aquesta proposta no s’ha d’interpretar, doncs, com una modificació de la normativa gramatical i ortogràfica en vigor* <sup>26</sup> (Institut d’Estudis Catalans 1999b, 9-10, emphasis in original).

This work is therefore an explicit attempt to regulate oral language, since a great deal of existing codification work (by Fabra and others), by targeting aspects such as spelling, was aimed at normativising written Catalan. Phonetic proposals concern the normative pronunciation of vowels in the Catalan system, final consonant devoicing, and when some consonants can be omitted altogether (the frequent pronunciation of the verb *prendre* as /'pɛn.drə/, for example, is permitted). Morphological proposals concern the normative use of clitics and personal pronouns, as well as an extensive

---

names] are not wiped out by foreign names... which would reflect a desire for cultural oppression such as... during those political regimes which were especially hostile to our language, which we have had to suffer.

<sup>26</sup>TRANSLATION: The current proposal forms part of the ongoing work to codify the Catalan language, which was started and carried out by the Institut d’Estudis Catalans. This codification was lacking a pronunciation dimension; the instalment concerning phonology aims to partly fill this gap. This proposal addresses the *oral* standard (and nothing more), which directly affects television and audio media, the education system in general and any other public use of the language... *This proposal should not therefore be interpreted as a modification of the existing grammatical and orthographical standard.*

list of approved verb conjugations.<sup>27</sup> The work, unlike Fabra’s *converses* discussed on page 91, makes very little reference to concrete Castilianisms, instead promoting the Catalan variant that is to be normatively favoured. Given the flexibility and wide use of oral language, the *propostes* take into consideration both diatopic and diaphasic factors, which are termed the *àmbit* and the *registre* of language use, respectively (Institut d’Estudis Catalans 1999a, 12). *Àmbit* is divided into *àmbit restringit* for instances of language use acceptable in only limited geographical areas, and *àmbit general* for usage that is less regionally marked. *Registre*, by dealing with both informal and formal scenarios, claims to ‘respond to a social and psychological reality’ (Institut d’Estudis Catalans 1999a, 12), since normative Catalan is to be used in all scenarios, not just those where the speaker is forced to be particularly erudite. This latter notion demonstrates a curious interplay with my earlier TDP axes of ‘+/- metacommentary’ and ‘+/- overt prestige driven’. The *propostes* note that, when discussing an oral standard, drawing distinctions between registers may appear ‘surprising’, the presumption being that scenarios with lower degrees of formality would not be governed by this oral standard Catalan. However, the theory outlined in chapter 1 perfectly accounts for this. My axes do not refer to the registers in which such usage takes place; whether they are formal or informal is of no consequence. The usage introduced by this oral standard is still ‘+ overt prestige driven’ as it adheres to the norm of a government-sanctioned body, and it is still ‘+ metacommentary’ as speakers are *consciously* encouraged to use it. Other theories might struggle if they were to hinge upon related notions such as *register* or *style shifting*, since even the *propostes* acknowledge that lower registers or speech styles might not be usually characterised by such an oral standard. However, this oral standard still scores positively on all three of my axes, since it was never specified in which registers ‘+ metacommentary’ and ‘+ overt prestige driven’ phenomena are to occur.

Another means by which Castilianisms are ‘rectified’ is through something of a revival of Fabra’s tradition of *converses filològiques*, discussed on page 91. In recent years, leading Catalan linguists have published many short articles in the Catalan

---

<sup>27</sup>This classification of normative and non-normative use will prove of use when designing the upcoming experiment, and shall be revisited in chapter 4.

press, explicitly warning against Castilianisms in several facets of language usage. Notably, leading member of the IEC Albert Jané wrote a column entitled *El llenguatge* for the newspaper *Avui*, publishing nearly 2,750 articles between 1976 and 1985, addressing such diverse topics as lexical barbarisms (cf. Jané 1976b, Jané 1980), use of punctuation (cf. Jané 1976c), onomatopoeias (cf. Jané 1976a) and toponymy (cf. Jané 1976d). Similarly, Joan Solà published several hundred short articles in his column *Parlem-ne* which appeared in *Avui*'s cultural supplement from 1991 until just before his death in 2009. These articles, much like Fabra's *converses* make explicit reference to Castilianisms and how to avoid them, and have the advantage that they are accessible to a wide readership; indeed, *Avui*'s current circulation, now that it has merged with the other major newspaper *El Punt* is estimated at around 54,000 copies, in addition to approximately 600,000 monthly internet hits (Avui 2009).

In addition to this, there are other, non-IEC *diccionaris de dubtes i barbarismes* on the market which aim to help speakers avoid the use of non-normative Catalan, be that through the use of contact elements from Castilian or otherwise. The leading Catalan publishing house Edicions 62 released a new such dictionary in 2008 (Paloma i Sanllehí and Rico i Busquets 2008) which advises readers on questions of Catalan grammar, word choice and spelling (the *dubtes*) and how to avoid non-native lexical items (the *barbarismes*). Grammatical advice includes an explanation of the function of the Catalan pronoun *res* and the quantifier *gens*, which in Castilian can often both be translated by *nada* ('nothing'); the dictionary entry attempts to avoid the overgeneralisation of *res* and to maintain the grammatical distinction between the two items. Issues of word choice include the prohibition of so-called colloquial forms such as *poguer* or *sapiguer* (in favour of the normatively-approved *poder*, 'to be able to' and *saber*, 'to know') or 'dialectalisms' such as *vai* or *vaic* (in favour of *vaig* or *vàreig*, AUX.1SG). Orthographical advice is offered for words such as *cartolina* ('cardboard'), where the unstressed vowel can lead to the non-normative *cartulina*, which is the normative Castilian spelling. Finally, half the dictionary is given over to *barbarismes* which have come into Catalan largely as contact elements (see section 3.2.2) and how to avoid them, such as the non-normative *minifalda* (normative *minifaldilla*,

‘miniskirt’) and *gaviota* (normative *gavina*, ‘seagull’). Such ‘rectification’ of non-normative language use as published in such dictionaries will prove of vital importance to the design of my experiment, to appear in chapter 4.

Finally, turning our attention to the creation (more than the promotion) of neologisms, the aforementioned TERMCAT terminology centre is a leading actor in the elaboration of Catalan. The aims and objectives of the centre are supported by Generalitat statutes and decrees, the most recent being the Decret 108/2006, article 5.3 of which lists TERMCAT’s responsibilities as including the ‘planning and coordination of research into Catalan terminology’, ‘offering of terminological support to the Generalitat’ and ‘placing their information and resources at the disposal of society in general’ (Generalitat de Catalunya 2006, my translations). TERMCAT’s elaboration work covers specialist fields from agriculture to zoology, encompassing themes as diverse as gastronomy, transport, leisure and the internet. They offer several, fully elaborated online specialist Catalan dictionaries, and have approved or created thousands of vocabulary items, including *sac de cops* (notably closer to the English *punch bag* than the Castilian *saco de arena*), *cabraboc* (a word of Catalan creation, where Castilian borrows the English *freemartin*) and *cabells d’àngel* (closer to the French *cheveux d’anges* than either the Spanish *cabelleras* or the English *tinsel icicles*). Understandably, a lot of the terminology approved by TERMCAT is etymologically similar to Castilian, due to the languages’ inherent closeness, but this does not detract from the centre’s aim to ensure that the two languages are equally elaborated.

**TDP in language function in Barcelona.** On page 61, we see the Barcelona-specific example of TDP in language function as ‘an increase in overt prestige of a variety through status planning’. As we have seen in section 2.2.1, the last thirty years have seen great advances in status planning for the Catalan language. The overarching aim of this Generalitat-approved linguistic legislation was to improve the status of Catalan and to redress the balance upset through Franco’s attempts at cultural genocide. The increase in overt prestige triggered by this legislation can be seen

in a number of ways, such as the requirement to demonstrate Catalan competence in order to gain job opportunities, the ever-increasing use of Catalan as a vehicle of cultural expression and the birth of Catalan as a language of international communication. These new instrumental values of Catalan constitute a top-down phenomenon in language function.

A great deal of legislation has been passed which requires job applicants to demonstrate a high level of competence in Catalan in order to be considered, above all for public sector roles. Teachers (Generalitat de Catalunya 2009, article 119.5), local government officers (Generalitat de Catalunya 1990, Generalitat de Catalunya 1994) and police officers (Generalitat de Catalunya 1991) must all provide accreditation of their level of Catalan, while civil servants from other regions (i.e. outside Catalonia) who work in close collaboration with Catalan government bodies are also required to possess enough Catalan to conduct business effectively (Generalitat de Catalunya 2003b). However, such legal demands also extend to the private sector, including areas which might not be the expected beneficiaries of language status planning: the Generalitat's Decret 348/2001 regulates self-tanning centres and explicitly states that all safety warnings must be written in, at the very least, Catalan alone (Generalitat de Catalunya 2001b), while 1999's DOGC 3046 targets dry cleaners, specifying that their price lists must also be written in Catalan (Generalitat de Catalunya 1999a). These requirements have led to a concomitant need to regulate accreditation of Catalan language competence. This is therefore also governed by Generalitat decrees (Generalitat de Catalunya 2001a).

The overt linguistic prestige of Catalan has not only been raised by a general need to use the language in order to get ahead in life, as detailed above. Linguistic legislation passed by the Generalitat has also allowed the language to blossom as a vehicle for cultural expression. The 1998 LPL dedicated an entire chapter to media and cultural industries (*els mitjans de comunicació i les indústries culturals*) wherein it was stated that the 'normal language' of any television, radio or written media projects operated by the Generalitat would be Catalan, and that everything would be done to promote and foster the use of Catalan in the theatrical and musical domains. As

ever, the general mission statement of the LPL was supported by specific laws which dictated the details of how this was to be achieved. Decret 269/1998 addresses radio programming so that at least 50% of radio programming was to be in Catalan and that at least 25% of songs would be in Catalan or Aranese (Generalitat de Catalunya 1998a). Decret 15/2003 states that at least 50% of Generalitat programming should consist of content originally produced in either Catalan or Castilian, and that at least 50% of this ring-fenced output should be in Catalan (Generalitat de Catalunya 2003a). These decisions are then enforced based on the Llei 8/1996 which details the fines to be given to television corporations for not complying with the above legislation: not complying with linguistic quotas is considered a serious infraction and incurs a penalty of between 2,000,001 and 15,000,000 pesetas (approximately 12,000 to 90,000 euros, Generalitat de Catalunya 1996). Recently, language status planning promoting the cultural value of Catalan has taken a new development with the passing of the 2010 Catalan Cinema Law which aims to regulate the production, distribution and international promotion of Catalan films through allocating funds to the cinematographic industry, as well as determining sanctions to be imposed on companies which do not comply (Generalitat de Catalunya 2010). In short, the Generalitat's language policy has assured the introduction of Catalan into cultural and artistic spheres, in turn contributing to an elevated level of overt prestige.

Finally, status planning has raised the profile of Catalan through attempting to make it a language of international communication. The 1998 LPL stated the Generalitat's new mission to encouraging agreements with other Catalan-speaking regions in Spain, Andorra, France and Italy to contribute to the protection and widespread use of the language (Generalitat de Catalunya 1998b, segona disposició adicional), as well as promoting the use of the language in academic and research institutions across the European Union (Generalitat de Catalunya 1998b, tercera disposició adicional). This legislation has come to life, as evidenced by the work of the *Institut Ramon Llull*, who operate under the auspices of the *Generalitat de Catalunya* and the *Govern de les Illes Balears*. Their aim is to promote Catalan language and culture on a global scale, and is achieved in a multitude of ways. Catalan language teachers are sent to 171

universities in 37 countries, and funding is granted to international research projects, such as the present work. This elevated profile of Catalan has been made possible thanks to status planning undertaken by governmental bodies, and thus constitutes an instance of TDP in language function.

### 2.2.3 Summary

This section has followed the schema laid out in table 2.1 very closely, applying the background situation to, and different manifestations of TDP to the situation in Barcelona. The history of language policy in Barcelona was given, with particular attention paid to recent developments in education policy, and concerns regarding the choice of Catalan as the medium of instruction. The consequent TDP in form and function were then exemplified. The form of Catalan is currently subject to a degree of structural and lexical manipulation mediated by organisations such as TERMCAT and publications such as the *diccionaris de dubtes i barbarismes*. The function of Catalan is also affected, due in large part to recent linguistic legislation (which makes sense, as we have seen that language policy is a potential background to manifestations of TDP): the increase in overt prestige of Catalan is evidenced through several scenarios within the society whereby demonstration of a high level of competence in Catalan is a pre-requisite for social, economic and cultural advancement.

## 2.3 Language policy in Barcelona: Overview and research question

This chapter has explored the role played by language policy, both on a theoretical level (section 2.1) and in direct relation to the linguistic community under examination, Barcelona (section 2.2). Language policy was given as an example of a background situation which could result in numerous top-down linguistic phenomena, these being classified ‘+ metacommentary’, ‘+ overt prestige driven’ and ‘+ planned’, as per the model given in chapter 1. These were stated in full in table 2.1 at the start

of the chapter.

The background situation of language policy in Barcelona, as well as specific examples of these resultant phenomena were given in section 2.2. We have seen that there has been a great deal of new linguistic legislation over the last thirty years, the aims of which have been made explicit. It has been stated that the broad linguistic goals of the *Llei de Normalització Lingüística* are to ‘protect and encourage the use of Catalan for all citizens... to effectively implement the official use of Catalan... [and] to ensure the extension of Catalan language competence’ (Generalitat de Catalunya 1983a). Moreover, we have seen that a great deal of concrete linguistic phenomena have arisen as a result of language policy, such as the elaboration and diffusion of a normative standard Catalan, as exemplified by the *propostes per a un estàndard oral, converses filològiques* or the *diccionaris de dubtes i barbarismes*, detailed in sections 2.2.1 and 2.2.2. But, put simply, to what extent is all of this actually happening?

Having outlined the aims of Catalan linguistic policies, it would be enlightening if they were critically tested. Is the linguistic community under investigation characterised by new oral standards proposed by organisms such as the Institut d’Estudis Catalans? Is Catalan ‘protected and encouraged’? Is its competence extensive? As for any resulting phenomena, we know that these normative standards are in existence, but are people aware of them?<sup>28</sup> Does normative language use characterise the linguistic community? Some of these questions may be quite broad-brush, and a doctoral thesis cannot hope to answer them all. But this work seeks to paint an accurate portrait of the sociolinguistic reality in Barcelona through the examination of two key factors: policy and contact. It was already stated in chapter 1 that researchers are currently seeking to *fer un balanç* of the Catalan language policy situation. Moreover, I have already outlined the merits of looking at language-in-education policy, in section 2.1.2 on a theoretical level and section 2.2.1.1 on a Barcelona-specific level. Therefore, maybe a project which offers an evaluation of Catalan language-in-education policy could make a valuable contribution to the field.

We are already aware of some key policy goals, expressed in both linguistic leg-

---

<sup>28</sup>The notion of ‘awareness of non-normative language use’ will be of key importance to the creation of the experiment, and will be explored in depth in section 3.3.



isolation, and in supporting projects such as the *propostes per a un estàndard oral*: Catalan is to be protected, standardised, its knowledge extended. So, I aim to devise an experiment that can offer some conclusions regarding the fulfilment of these goals, thereby painting a clearer portrait of the linguistic community under examination. So, as regards language policy, I simply intend to ask:

**POLICY RESEARCH QUESTION**

To what extent is current Catalan language-in-education policy meeting its aims?

## Chapter 3

# Language Contact in Barcelona

Els catalanoparlants utilitzem cada cop més expressions traduïdes o copiades literalment del castellà. Sovint el que parlem és més a un calc literal del castellà que no pas un idioma amb entitat pròpia. Però per què succeeix això? Doncs senzillament perquè quan parlem català agafem com a referència el castellà i oblidem que el català té les seves pròpies expressions i regles. Ens falta una referència en català i l'agafem del castellà. El que acabem parlant no és català, sino castellà amb accent català o un dialecte del castellà (El català com cal website).<sup>1</sup>

The viewpoint expressed in the above citation, while clearly partisan in nature, outlines an arguably emblematic feature of Catalan as spoken in present-day Catalonia, namely linguistic interference from Castilian. In addition to highlighting the existence of contact-induced language usage, the above fragment also aims to provide reasons for such phenomena, albeit in a rather repetitive, sweeping manner. While this thesis may not aim to provide to reasons why speakers use forms influenced by contact, greater familiarity with the situation of language contact between (and as we shall see, consequent convergence of) Catalan and Castilian, will certainly allow us to paint a clearer picture of the current speech situation in Barcelona. Indeed, the emotive

---

<sup>1</sup>TRANSLATION: We Catalan speakers use more and more expressions which are literally translated or copied from Castilian. Often, what we speak is more like a literal calque of Castilian than a self-sufficient language. But why does this happen? Well it's simply because when we speak, we take Castilian as a frame of reference, and we forget that Catalan has its own expressions and rules. We lack a frame of reference in Catalan and take it from Castilian. What we end up speaking isn't Catalan but Castilian with a Catalan accent, or even a dialect of Castilian.

language of the above citation should serve as an indicator that issues arising from the study of language contact in Barcelona are able to generate heated debate among certain Catalan speakers, and are thus of key importance to both scholars and language planners who wish to understand the community. The following chapter will thus provide valuable insight for those engaging in research into Catalan sociolinguistics.

As in chapter 2, the schema found in table 1.6 is to be used to structure the present chapter:

	BUP
Background situation	Language contact
Form-based phenomenon	Systematic introduction of variant of a given variable through language contact
Function-based phenomenon	Decrease in overt prestige (and thus exclusion from high-prestige functions) of a variety through BUP in language form

Table 3.1. Partial reproduction of table 1.6, pertaining to BUP.

This chapter will allow the reader greater familiarity with language contact and any resultant phenomena on both theoretical and Barcelona-specific levels, and will ultimately culminate in a contact-based research question to be answered by the upcoming experiment. Following the same structure as chapter 2, table 3.1 will be used as a guide. All elements of BUP will be addressed, starting with the background situation of language contact, followed by any manifestations in language form and function. Firstly, section 3.1 will undertake this on a theoretical level, with section 3.1.1 addressing the theoretical developments in the field of language contact, before section 3.1.2 looks at any resultant phenomena in language form and function, paying particular attention to what will be defined as *contact-induced linguistic phe-*

*nomena* (CILP). Then section 3.1.3 will explore language contact in urban settings worldwide, before section 3.1.4 summarises the section's theoretical findings. Section 3.2 will follow the same order, but focusing on Barcelona: section 3.2.1 will look at language contact in Barcelona, before section 3.2.2 addresses resultant phenomena in language form and function. Section 3.2.3 summarises the Barcelona language contact information presented.

Finally, section 3.3 will extract the key developments made in this chapter in order to formulate a contact-based research question. This question will be answered by the fieldwork experiment, and alongside the policy research question presented at the end of chapter 2, will allow this thesis to paint a clear picture of the sociolinguistic reality in Barcelona, encompassing both language policy and language contact.

## **3.1 Language contact and resultant phenomena (BUP):**

### **Theory**

This section will allow us an understanding of language contact and potential resultant sociolinguistic phenomena in form and function, as per the schema in table 3.1. Section 3.1.1 will address the history of the concept of language contact over the last sixty years, highlighting important aspects which have remained constant as well as interesting developments. Section 3.1.2 then relates language contact, which was given as an example of a background situation of BUP, to concrete linguistic phenomena in both language form and function. Section 3.1.3 applies this to several urban scenarios in order to underline how the study of language contact in cities can be revealing (after all, that is what this thesis will do in the case of Barcelona), before section 3.1.4 summarises the findings of the section.

#### **3.1.1 Language contact**

Language contact has been given as a defining characteristic of the linguistic situation in Barcelona, but has not yet been explored on a theoretical level in this thesis.

This section will review and synthesise the key literature on the subject in order to provide a comprehensive understanding of language contact and of its development as a concept. The history of the discipline of *contact linguistics* will be charted, in order to examine how this notion has developed and how scholars have addressed language contact and any resultant linguistic phenomena. The first major work on the subject was Weinreich's 1953 *Languages in Contact*. The literature review will therefore start with this piece of work, before examining how subsequent academics in the 1960s, 70s and 80s contributed to research in the field. Finally, I will discuss contemporary views on language contact, offering a number of perspectives from the 1990s and the 21st century, tying all of this in with the present thesis.

**Weinreich: *Languages in Contact*.** Largely considered the magnum opus of contact linguistics, Weinreich's 1953 *Languages in Contact* is the first comprehensive book-length study of the field, and defines language contact as follows:

Two or more languages will be said to be in contact if they are used alternately by the same persons. The language-using individuals are thus the locus of the contact (Weinreich 1953, 1).

Language contact is thus simply a situation in which two languages are present. It is not a process or an activity, but merely the co-existence of two languages. Within this work, Weinreich clearly distinguishes between *mechanisms of interference*, *psychological elements*, *sociocultural elements* and *other elements*. As such, my analysis will take all of these into consideration. The work serves the dual purpose of on the one hand outlining the concept of language interference arising from contact and providing a typology thereof; and on the other highlighting various psychological and sociocultural factors important in situations of language contact. The emphasis placed on such psychological and sociocultural considerations is indicative of the author's view of the bilingual speaker as the locus of language contact,<sup>2</sup> while the

---

<sup>2</sup>Note that the speaker is the locus of *contact* and not of *change*, which occurs at a societal level. This falls in line with the thesis' view of language change as occurring when an innovation is eventually embedded within a society.

typology of language interference accords importance to language as a system, and not just language users.

The concept of *language interference* was defined by Weinreich simply as ‘instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of language contact will be referred to as interference phenomena’ (Weinreich 1953, 1). He then discusses *mechanisms* and *structural causes* of such interference (Weinreich 1953, 7), identifying the presence of *phonic interference* (Weinreich 1953, 14), *grammatical interference* (Weinreich 1953, 29) and *lexical interference* (Weinreich 1953, 47) and providing several cross-linguistic examples of each. In my view, this aspect of his work, while undeniably useful, is chiefly a classificatory exercise and as such does not constitute the most interesting of his theoretical advances. These advances focus on the bilingual language user and provide both speaker-internal (psychological) and speaker-external (societal) background to language contact. Weinreich’s definition of interference as divergence from normative language use will prove of great importance in the creation of my experiment, and will be addressed in further detail in due course.

Various psychological factors form a crucial part of Weinreich’s analysis, with the bilingual individual as the primary starting point. He claims ‘it is clear that even socio-cultural factors regulate interference through the mediation of individual speakers’ (Weinreich 1953, 71), thereby placing psychology in a position of utmost importance. A large part of Weinreich’s psychological analysis hinges on the dominance of a given language in a bilingual speaker, and how this can conceivably affect language contact processes (an implication being that a dominant language will exert more influence on a speaker’s utterances). Criteria for ascertaining the *relative status* of languages within the bilingual user include relative proficiency, mode of use (i.e. reading, writing, speech etc.), order of learning and age, usefulness in communication, emotional involvement, function in social advance and literary-cultural value. Based on these factors, he proposes a *dominance configuration* to determine relative importance of codes in a specific bilingual’s brain (Weinreich 1953, 74-80).<sup>3</sup>

---

<sup>3</sup>This notion of language dominance will also prove important when analysing the results of the fieldwork experiment, and can be seen in section 5.1.5.

Weinreich also appreciated the importance of sociocultural factors, placing them in conjunction with the aforementioned psychological aspects. The *social functions* of given languages are deemed important, as it is claimed that languages are more susceptible to interference when they are used in certain domains (Weinreich 1953, 87).<sup>4</sup> Certain components of Weinreich's analysis in fact bridge the gap between the sociocultural and the psychological. The *symbolic importance of a standardised language* (Weinreich 1953, 99) is a prime example, as it gives rise to issues of prestige. Although prestige is often seen as a sociocultural issue, it is inextricably linked to language attitudes, which are arguably psychological in nature. Language standardisation has been discussed at great length in chapter 2, being closely linked to notions of language policy. Concerns of policy and contact are not therefore always independent of one another, which will become clear in our overall discussion in chapter 6.

In short, Weinreich's seminal text laid the theoretical foundations of contact linguistics. His notion of the bilingual individual as the locus of language contact influenced the direction of his work, and that of subsequent scholars. It led to a heavy emphasis on psychological factors as well as sociocultural concerns, since one must be aware of not only the speaker's mental processes but of his/her social environment.

**Subsequent developments: 1953-88.** Having discussed *Languages in Contact*, it is now logical to provide an overview of how these views and concepts were developed. The following section aims to address relevant advances from Weinreich (1953) to Thomason and Kaufman (1988), although naturally, the following collection of theoretical developments is far from exhaustive, for considerations of space.

While I have argued that Weinreich's discussion of *mechanisms of interference* was comparatively pedestrian and classificatory, subsequent developments have proved enlightening, as academics try to unlock the predictive power that the knowledge of interference mechanisms may bring. In short, if mechanisms and constraints could

---

<sup>4</sup>The nine *domains of language use*, as defined by Schmidt-Rohr (1932) are as follows: the family, the playground, the school, the church, literature, the press, the army, the courts, and administration (Weinreich 1953, 87).

be ascribed to instances of language interference, then surely it could be ascertained where and when these instances might occur. Moravcsik (1978) provides potential constraints on forms of borrowing, and argues their ‘explanatory value’ while claiming this all potentially allows one to explain instances of language contact. These concepts were more recently developed into *hierarchies of borrowability* (to be discussed on page 122). Similarly, Poplack (1980) posits constraints on code-switching, i.e. ‘the systematic (though usually outwardly spontaneous) alternation of two or more languages’ (Mahootian 2006, 1). Attempts to mechanistically predict contact-induced linguistic phenomena (CILP) have been proposed by Thomason and Kaufman (1988) and Thomason (2001), and have been largely dependent on psychological and sociocultural variables, among other factors. Markedness (a psychological construct), intensity of contact (a sociocultural element), typological distance between language varieties, the presence or absence of imperfect language learning, and match in pragmatic function are all considered important determinants in predicting the occurrence and nature of instances of CILP (Thomason and Kaufman 1988; Thomason 2001, 66-7). Therefore, while the mechanistic content of Weinreich’s work was largely classificatory, later academics adapted and created new mechanisms in an attempt to predict instances of language interference, most frequently borrowing, code-switching and CILP.

As stated, Weinreich regarded the bilingual individual as the chief locus of language contact, thus according importance to *psychological* theories of bilingualism, and such concepts were both developed and challenged. The emphasis placed on the study of the language user was frequently maintained in later studies (Haugen and Weinreich 1958, Moravcsik 1978). However, certain academics started to question whether this was a feasible approach, notably Caudmont (1982), who asks whether language contact always implies individual bilingualism, claiming that:

Le postulat de Weinreich n’est pas acceptable, car il prend le terme de bilinguisme dans un sens trop large et ne tient pas compte des contacts accidentels qui n’impliquent souvent ni la connaissance, ni l’utilisation de l’autre langue



(Caudmont 1982, 379-80).<sup>5</sup>

I agree that while it is valid that language contact does not always imply individual bilingualism, it is a little cavalier to dismiss Weinreich's notions out of hand. In taking the bilingual as the point of departure for his study, Weinreich is merely claiming to examine the phenomena that result from individual bilingualism. That the concept of language contact has developed after the publication of his work is incidental, and makes his observations no less valid. Moreover, this shows that Weinreich was successful in stimulating the study of contact linguistics to a point where the field grew to encompass phenomena beyond his initial subject matter. Indeed, most viewed Weinreich's work as a springboard, although some (e.g. Caudmont 1982) saw new advances as contradictory to the arguments presented in *Languages in Contact*.

Weinreich's work also highlighted sociocultural concerns within contact linguistics: once again, existing terms were developed and new concepts introduced. The sociocultural dimension as a whole remained significant in many later works, most notably Thomason and Kaufman (1988, 4) which 'remind[s] the linguistic profession that languages are a product of, and a vehicle for, communication among people.' A move away from the focus on individual bilingualism to different multilingual or multidialectal situations allowed completely new sociocultural backgrounds to be the subject of contact linguistics, such as in Trudgill's seminal 1986 work *Dialects in Contact*, which widened the focus to include mutually intelligible dialects. The boom in the study of code-switching scenarios is another example. Within this field, widespread debate has ensued regarding the effectiveness of microsociological versus macrosociological approaches to the analysis of code-switching data.<sup>6</sup> While code-switching undoubtedly constitutes a fascinating area of study within language maintenance, and is arguably emblematic of Catalan speech practices, it shall receive only limited attention here, since this thesis has chosen to focus on the related, yet distinct phenomenon of CILP.

---

<sup>5</sup>TRANSLATION: Weinreich's position is untenable, since his interpretation of bilingualism is too broad, and he doesn't consider random incidents of contact which imply neither knowledge nor use of the other language.

<sup>6</sup>This dichotomy was addressed in Hawkey (2008). See Myers-Scotton (1993) for a leading macrosociological approach (the markedness model) and Wei (1998) for a microsociological stance (conversation analysis, as pioneered by Auer).

The exact differences between these two phenomena will become clear in my upcoming typology of consequences of language contact (fig. 3.1, page 131). However, such a topic warrants its own full-length doctoral thesis, and has indeed been the subject of many to date.

In short, mechanisms of language contact were explored in greater detail in an attempt to predict instances of interference phenomena. The psychological bias of Weinreich's work also received much attention, scholars further investigating the emphasis on the study of the bilingual individual as point of departure, as well as searching for new ways to measure language dominance. Sociocultural concepts were elaborated, and research into related sociological areas such as code-switching grew exponentially. *Languages in Contact* thus stimulated sufficient interest to launch a dynamic field of study, as exemplified by the above developments.

**Recent advances: 1988 - present.** Contact linguistics continues to be a dynamic area of study, and recent developments have followed the example of previous research in pushing the boundaries of the field and attempting to accurately explain and account for phenomena of language contact. In recent years, book-length works by Thomason (2001) and Winford (2003) have provided an outline of the entire field of contact linguistics, thus bringing the topic into the mainstream of academic research. As before, existing concepts have been developed and new theories proposed.

The aforementioned desire to predict interference phenomena has continued to evolve in recent work. Previous scholars tried to impose constraints on borrowing and other interference phenomena (cf. Moravcsik 1978). Contemporary thinking suggests that such constraints are futile, and it is in fact possible to borrow any lexical or structural feature from one language into another (Thomason 2001, 63). However, while all features are in theory borrowable, it is more probable that some features or lexical items will be transferred than other. Thus scholars established *hierarchies of borrowability* or *borrowing scales* (Thomason 2001, 71; Winford 2003, 30; Siemund and Kintana 2008, 15). Such scales express the fact that lexical items are more easily diffused than structural features, and content words are the most eas-

ily borrowed. These hierarchies outline contact situations and indicate the form and extent of borrowing likely to be found. However, one must remember that all items can be diffused from one language to another, and ‘any borrowing scale is a matter of probabilities, not possibilities’ (Thomason 2001, 71). Additionally, the continued study of imperfect language learning (Thomason 2001, 66) and a renewed interest in the degree of speaker bilingualism (Siemund and Kintana 2008, 4) aim to offer insight into the nature of instances of language interference. Moreover, although attempts at predicting CILP based on markedness and typological distance (as key factors within borrowing scales) are still being developed, they have been supplemented with research regarding speaker attitudes, which concludes that CILP is in essence unpredictable (Thomason 2001, 77-85). In short, the development of existing study has largely centred on the wish to better understand and thus predict occurrences of language interference.

Nelde (1997) distinguishes *language conflict* from language contact, thus creating the new, complementary discipline of *conflict linguistics*. Language conflict ‘arises from the confrontation of differing standards, values, and attitude structures’ (Nelde 1997, 292). Academics have largely chosen to continue the pursuit of contact (not conflict), perhaps due to the similarities between the two concepts: ‘Nelde’s law’ claims ‘there can be no language contact without language conflict’ (Nelde 1997, 292) and therefore one may question to what extent it is useful to study conflict linguistics as a separate model. Winford (2003, 33-7) draws attention to the notion of equality between social groups in contact situations. He makes it clear that, when examining interference phenomena, one must take into account the relative statuses of the linguistic communities at hand, as traditionally dominant/subordinate configurations (which can arise for reasons of race, conquest etc.) may have different outcomes to situations in which the language groups are social and historical equals. It is therefore clear that contact linguistics continues to attract a great deal of academic attention, and many academics are concerned with developing a mechanism by which instances of language interference can be predicted or at least comprehensively accounted for. Notions continue to arise and offer new perspectives on how contact linguistics is

viewed.

How are scholars currently defining language contact, and how does this fit in with the outlook offered in the present thesis? Many have chosen to move away from Weinreich's initial emphasis on the speaker as locus of language contact, adopting a more macrosociological approach, as in Crystal (1997):

[Language contact is] a term used in sociolinguistics to refer to a situation of geographical continuity or close social proximity (and thus of mutual influence) between languages and dialects (Crystal 1997, 87).<sup>7</sup>

Modern definitions of language contact, and resultant situations of bi- or multilingualism tend to adopt a more holistic stance, encompassing both individual and societal factors, as follows:

Co-occurrence of two or more languages either in the individual (bilinguality) or in society (bilingualism) (Hamers and Blanc 2000, 370).

A situation in which more than one language exists in a given area or... community which may lead to speakers of one language deliberately or subconsciously introducing into their own language features of the other language (Llamas et al. 2007, 219).

Hamers and Blanc's definition is entirely congruent with my belief that language contact is merely a background situation whereby two languages are present. They remind us that there is a relationship between this occurring on the individual level and on the societal level, though these two are sufficiently distinct as to require two different terms. Other current scholarly output reflects and supports this view of language contact, and by extension societal bi- and multilingualism, as a situation and not a process, as in Dewaele et al. (2003, 1), where bilingualism is simply described as

---

<sup>7</sup>This entry also makes cursory reference to the bilingual language user, but not as the initial *locus* of language contact. Instead, the emphasis is placed on the languages (and not their speakers).

‘the presence of two languages.’ The first part of Llamas et al.’s definition only refers to contact on a societal level, not concerning itself with the speaker until discussion of the processes which may result from the situation of contact. It should also be noted that this description of the phenomena which arise from contact, which will receive full attention in the next section, does not fully match my notion of BUP. Speakers being able to ‘deliberately or subconsciously’ introduce contact elements directly contradicts my idealised notion that speakers cannot offer metacommentary on this type of language use (see figure 1.3 for the complete model).<sup>8</sup> As such, the above definitions of language contact and related notions (such as Hamers and Blanc’s *bilinguality* and *bilingualism*) are entirely congruent with the outlook of the thesis. Societal-level perspectives are helpful since I am treating language contact as a background situation, analogous to language policy and planning; such background situations can relate to entire linguistic communities. Individual speaker-level perspectives are also required since the upcoming experiment looks at individual language use, and not the definitive embedding of contact-induced language forms within an entire society (since this would arguably constitute language change, which is beyond the remit of the thesis).

In short, this section has charted the development of the use of the term *language contact* – and of the field of *contact linguistics* – so that we are familiar with the theoretical context before discussing the background situation of contact in Barcelona. Weinreich’s initial idea of the speaker as the locus of contact, alongside the importance of sociocultural factors, can still be seen in contemporary definitions of contact (and related phenomena), which embrace both individual speaker-level and wider societal-level phenomena.

---

<sup>8</sup>Despite the occasional emphasis on speaker-level situations, this thesis has chosen to use the term ‘language contact’ over the occasionally preferred ‘speaker contact’ (p.c. Christopher Pountain). The latter term does not necessarily imply multilingualism since in theory, two monolinguals can provoke speaker contact.

### 3.1.2 Language contact and its relation to BUP

Language contact has been given as an example of a background situation which can result in concrete linguistic manifestations of BUP. This section will explicitly state the theoretical relationship between language contact and instances of BUP in language form and language function. This relationship (particularly concerning form-based phenomena) will be further illustrated by an innovative typology of the consequences of language contact (figure 3.1, page 131).

**Language contact and BUP in language form.** Many existing works have attempted to break down language contact into identifiable phenomena. It should be remembered that, based on the existing literature, I view language contact as a situation which is the catalyst for linguistic phenomena: the phenomena are not a subset of language contact. As we shall see, my typology of the consequences of language contact is subtly different as it highlights the fact that individual phenomena are the result of a background situation of language contact (as per my definition) and as such are not types of language contact, but rather processes emanating therefrom. This subsection will define phenomena deemed as instances of BUP (namely CILP), and reject the inclusion of other, related practices (such as code-switching).

Before examining the relevant literature, it should be pointed out that some authors in the field (e.g. Thomason 2001) use the term contact-induced language *change* (CILC), thus it is reproduced here where appropriate. These authors discuss the ability of language contact to provoke full linguistic changes embedded in a system, and therefore talk of CILC. This thesis does not address fully embedded linguistic changes, but rather phenomena that could arguably one day lead to change (see page 49). Thus, I talk of contact-induced linguistic *phenomena* (CILP), unless quoting other authors who do indeed discuss language change. CILP is fundamentally very similar to CILC: they are both brought about by language contact, but CILP does not necessarily constitute an embedded linguistic change, and encompasses related contact-induced precursors to potential changes. Since the difference between CILC and CILP is not in the nature of the phenomena, but rather in their degree of embed-

ding in a particular community, they would occupy the same place on a typology. For this reason, I will sometimes refer to CILC/CILP when I want to index both certain authors' existing discussions (CILC) and the phenomena under examination in this thesis (CILP).<sup>9</sup>

Recent studies of the field have proposed useful typologies of language contact, distinguishing several independent but undeniably related outcomes. Thomason (2001) posits a three-way split into CILC, *extreme language mixture* and *language death*. *Extreme language mixture* refers to the creation of pidgins, creoles and mixed languages, while *language death* encompasses language loss, attrition and grammatical replacement (Thomason 2001, 60). Winford (2003) suggests a similar three-way separation into *language maintenance*, which comprises borrowing, structural convergence (in other words, Thomason's CILC) and code-switching; *language creation*, akin to Thomason's extreme language mixture; and *language shift*, comparable to Thomason's language death. While both typologies closely resemble one another, I feel that Winford's 2003 work is slightly more useful for our current purposes. I consider *language maintenance* more appropriate as a top-level division of the consequences of language contact, since it is broader than mere CILC/CILP, encompassing borrowing, code-switching and potentially other phenomena. Similarly, I consider *language shift* less limiting than *language death*, and *language creation* less restrictive than *extreme language mixture*. Moreover, I consider the term *extreme language mixture* ambiguous, since it could feasibly refer to cases of structural convergence or CILC/CILP as opposed to its intended meaning of creolisation (indeed, it is highly likely that the author of the opening citation of this chapter would class contemporary Catalan, as characterised by convergence with Castilian, as an instance of *extreme language mixture*). These views will be expressed in my forthcoming typology of language contact.<sup>10</sup>

Thus far, language contact has been explored and its resultant phenomena identified,

---

<sup>9</sup>For full definitions of CILC and CILP, see the list of terms and abbreviations starting on page 16.

<sup>10</sup>For full explanation of all terms not directly related to my thesis, see Thomason (2001) and Winford (2003).

including CILC/CILP. However, CILP is yet to be fully defined, and its relationship to BUP in language form has still not been explained. Thus far, I have taken CILP to refer to linguistic phenomena that have arisen through the situation of contact; but this is potentially a bit vague, as it would cover everything mentioned in Winford and Thomason's typologies, from language shift to language creation. It is clear therefore that the literature often uses the term CILC to designate something a little more specific, more akin to Weinreich's notion of 'language interference'. CILC has thus been employed as something of a grab-bag term, defining a number of related linguistic interference processes that have arisen from multilingualism within a linguistic community. Such processes are highlighted by Silva-Corvalán when discussing 'phenomena characteristic of bilingualism and multilingualism' as referring to practices such as simplification, overgeneralisation, transfer, convergence and analysis (Silva-Corvalán 1996, 1). *Simplification* and *overgeneralisation* refer to similar processes, namely the extension of usage of a particular form, where a competing construction or item is available; however, simplification implies the decrease in usage of the competing form and overgeneralisation does not (Silva-Corvalán 1996, 2). *Transfer* is a potentially vague concept and is consistent with Weinreich's interference, i.e. 'the incorporation of language features from one language into another' (Silva-Corvalán 1996, 4), and can lead to *convergence*, namely greater structural similarity between the two codes (Silva-Corvalán 1996, 4-5). Finally, *analysis* 'is the process which underlies either the preferential use or the creation of periphrastic constructions as opposed to synthetic ones' (Silva-Corvalán 1996, 5).<sup>11</sup> These processes are often used in the literature in the context of language change, but they are equally applicable to this thesis. Processes of simplification, overgeneralisation etc. can appear in speech patterns without necessarily having been fully enshrined in the linguistic community as a change. As mentioned before, the nature of these phenomena would be the same, the only difference between CILC and CILP being the degree to which said phenomena have become embedded in the linguistic community. The phenomena themselves

---

<sup>11</sup>It should be noted that, despite the fact that processes of *convergence* are frequently found as instances of CILC/CILP which lead to BUP, this does not imply that processes of *divergence* (whereby two codes become structurally more dissimilar) is impossible in these circumstances. Indeed, one should avoid the facile and erroneous distinction of convergence=BUP versus divergence=TDP.



are the same, whether they have reached a point of full embedding or not. It is reasonable to assume that these phenomena of simplification, overgeneralisation, transfer, convergence and analysis will characterise the speech of a bilingual community such as Catalonia.

The debate about whether to regard code-switching as germane to these other speaker-driven processes is understandably complex. Processes of code-switching have been set aside from my discussion of CILC/CILP, since the study of this practice has arguably developed into a field in its own right, and the search for rules governing spontaneous conversational code-switching has resulted in the creation of various models to account for language choice in given circumstances. This falls outside of the scope of this thesis, as spontaneous mixture of codes in a given utterance is fundamentally different to the systematic incorporation of a particular lexical, phonological or syntactic feature into a code other than the language of its origin (as observed in *convergence*, *transfer*, *analysis* etc). While both processes are witnessed in the Catalan context (cf. Calsamiglia and Tusón 1984 for an account of Catalan-Castilian code-switching in Barcelona), I maintain that studies of code-switching address different, albeit related issues; as such, code-switching is not to be included within my definition of CILC/CILP (and thus remains outside the scope of this study). Gysels neatly summarises the phenomena to be included, stating ‘lexical borrowing, phonological and syntactic interference and other contact phenomena may be considered as essential features of (a contact language)’ (Gysels 1992, 45). As such, a collection of speaker-driven processes including simplification, overgeneralisation, analysis, transfer, and convergence (and theoretically divergence) make up CILC (if the processes are fully embedded and enshrined as a change) or CILP (if they are not necessarily fully embedded).

On page 61 and in table 3.1, we can interpret the example of BUP in language form as the ‘systematic introduction of a (potentially foreign) variant through language contact’, which is perfectly in line with my definition of CILP. Processes of systematic transfer and convergence are the clearest reflections of my definition, but under certain circumstances, my example of BUP in language form could be inter-

preted as simplification, overgeneralisation or analysis.<sup>12</sup> Code-switching has been excluded from my example of BUP, as I have stipulated that the introduction be systematic. By this I mean that the borrowing must be somewhat generalised, and used to the detriment of the native item with equivalent meaning in most cases, as opposed to being unpredictable one-off instances of language use. With such borrowings, speakers are largely unaware that they are mixing different languages, thus cementing their position as ‘-metacommentary’. Code-switching is a lot less predictable and systematic: while some items are more or less likely to be switched cross-linguistically (see page 122 for mention of hierarchies of borrowability), by its very nature it addresses the spontaneous alternation between two codes. Note that code-switching is still arguably a BUP in language form. Its spontaneity is congruent with its position within figure 1.3 as unplanned. The case for code-switching being ‘-metacommentary’ may be somewhat less clear cut. With code-switched lexical items, the native item is still very much present to be used by the speaker. Thus, he/she could choose to use the native item if the social situation dictated, and therefore the code-switched item is ‘+metacommentary’, and less in line with the criteria of ‘- on all axes’ held by BUP (although this interpretation of code-switching slightly problematises its spontaneity). With CILP, the native alternative to the non-native item is less readily available. In any case, this thesis has chosen to focus on CILP as an example of BUP in language form, and not code-switching. This is not because code-switching is any less interesting or valid, but because it opens up a whole new field of questions and analyses that, for considerations of space and relevance, were not deemed necessary for this work.

All of the above information can be neatly synthesised into the typology of the consequences of language contact given in figure 3.1. The different types of process which can result from language contact are laid out largely as per Winford’s (2003) classifications, and CILP (which equates to my example of BUP in language form) has been positioned accordingly.

---

<sup>12</sup>Divergence is a little more problematic. In a situation of language contact, if a linguistic form in one code becomes markedly different from the form found in the other code, this implies at least a degree of consciousness, thus making its status as ‘-metacommentary’ and ‘-planned’ (see model 1.3) a little more doubtful.

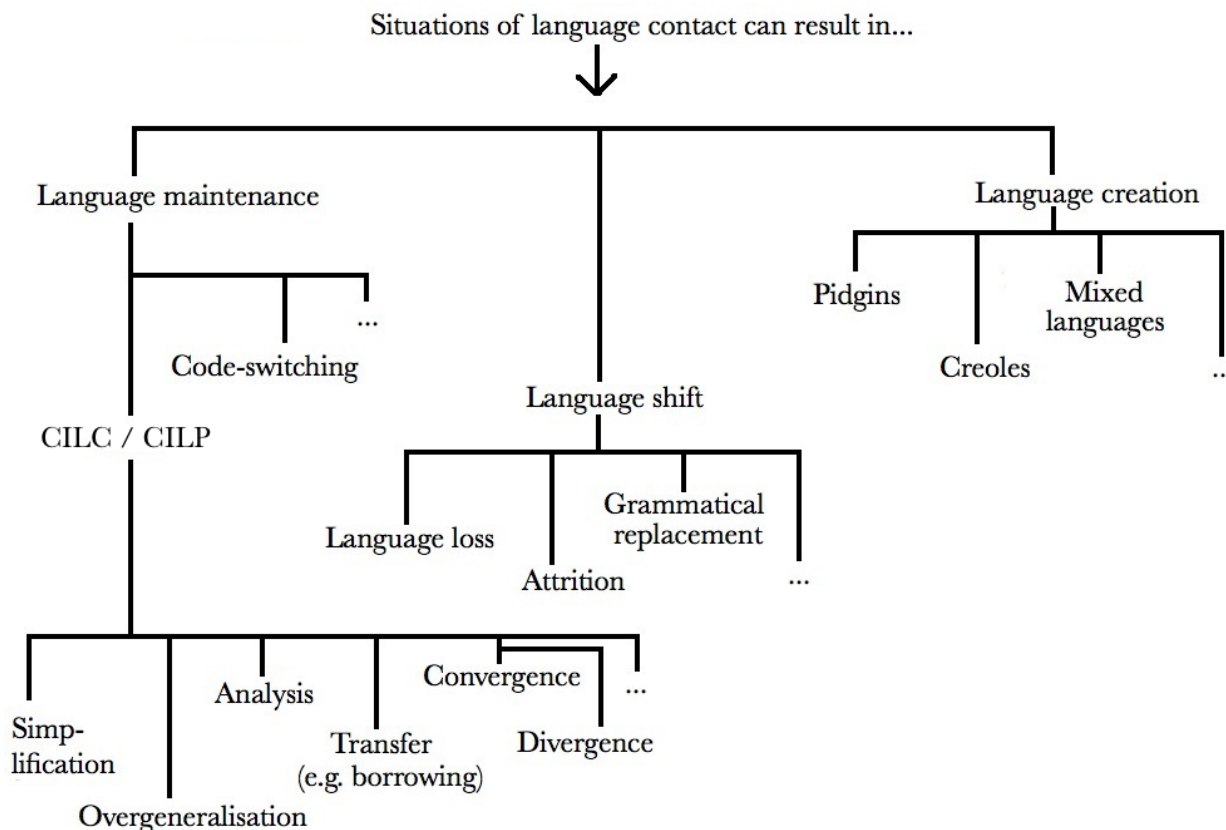


Figure 3.1: Non-exhaustive typology of the consequences of language contact

**Language contact and BUP in language function.** Although examples of BUP might be more evident when looking at language form, there are several ways in which BUP can interact with language function. The example given in table 3.1 was of a ‘decrease in overt prestige through BUP in language form’. The lay opinion that contact varieties, be they characterised by heavy lexical or structural borrowing or by extensive code-switching, are ‘lazy’ or ‘degraded’ is widespread. Moreover, such a decrease in overt prestige is implicit in the very aims of language policy as detailed in section 2.1.2. The fact that schools make pains to teach in the normative variety of a given language, that specialist publications are dedicated to ‘correct’ language use, and that government funds are allocated to bodies which ensure the autonomy and future of languages through the creation of native neologisms, shows how deviation from this normative use is not to be tolerated. Moreover, our literature review of language contact has revealed that such divergence from norms, labelled by Weinreich in

1953 as ‘language interference’, has been a central concern of contact linguistics since its inception. If these practices did not incur such stigma and negative value judgements, arguably not so much attention would have been accorded to their ‘correction’ (see section 2.2.2 for examples of this in the Barcelona linguistic community). This decrease in overt prestige thereby leads to an exclusion from high-prestige functions, giving our function-based phenomenon.

However it is not always necessarily the case that BUP brings about this particular development in function. Certainly in multilingual Western societies this is often the case, but other such configurations are conceivable and should not be ruled out. For example, let us imagine a linguistically complex community which exhibits two languages in a diglossic relationship : an H code fulfilling high prestige functions, and an L code fulfilling low prestige functions. In such a case, one can imagine that a contact variety of the L code characterised by interference from the H code could acquire prestige relative to the non-contact L variety due to its association with the H language. If a speaker were to use an L variety which demonstrated their competence in the H variety, overt linguistic prestige might be accorded to such language use. These situations are not widely attested, but the example of Singapore is given in section 3.1.3.

As such, my example of BUP in function was selected because it is frequently attested. This is not to say that it is the only possible example of a BUP in language function, but as with all entries in table 3.1 merely a useful paradigm through which to test the model.

### **3.1.3 Testing the theory: BUP in the city**

In order to test the link between language contact and processes of BUP, it will be helpful to see what conclusions have been reached by existing studies of language contact. Language and dialect contact situations come in many shapes and sizes, from Trudgill’s (1986) rural Norfolk communities to the urban New York scenarios studied by Poplack (1980). As this thesis focuses on a situation of urban language contact, I will now examine analogous situations worldwide in order to ascertain how

they fit in with the model already proposed. Firstly, I will address the interaction of cities with the notion of language contact, before then addressing their interaction with concrete bottom-up phenomena in form or function, as already identified (so chiefly CILP). Finally, different examples of urban environments around the world will show how the relationship between BUP and the city can be manifested in many different ways. All of this will support my choice of Barcelona as an interesting and enlightening example of a city characterised by language contact and resultant linguistic phenomena in form and function.

**Cities and language contact.** As stated, cities have been utilised as backdrops for a great many language contact studies. Mackey (2005) explains this as follows:

Cities have long been the chief locus of language contact, since they are in essence restricted areas dependent on face-to-face interaction. The population of a city is continually being renewed so that at any given time, within a plurilingual city, there may be representatives for (various) phases of transitional bilingualism (Mackey 2005, 1304).

Such intensity is an obvious advantage of urban study. It would be a logical assumption that if one is studying contact phenomena, one would wish to examine the areas with most contact, i.e. most people, in order to observe the greatest amount of innovation. Another related aspect is identified by Romaine:

Overall, then, we can expect the sociolinguistic consequences of urbanization to be quite complex because urbanization tends to promote diversity as well as uniformity. Urban environments are often sites of contact between languages as well as dialects (Romaine 2000, 66).

By drawing attention to the process of urbanisation, Romaine highlights more interesting aspects of urban contact linguistics, namely the fact that city-building brings various social groups within a relatively confined space. This diversity can lead to potentially interesting linguistic practices, and provide valuable academic insight into

the phenomena at hand. Nelde's highly related notion of linguistic conflict (see page 123) also benefits from urban study since:

It is precisely in modern urban society that conflicts result essentially from the normative sanctions of the more powerful group, usually the majority, which demands linguistic adaptation to the detriment of language contact, and thus preprograms conflict with those speakers who are unwilling to adapt (Nelde 1997, 291).

This situation is certainly applicable to Barcelona, in light of the Francoist sanctions placed on Catalan language usage during the dictatorship (see sections 1.1 and 2.2.1), and moreover can be traced back to the intensity of contact. Arguably, the power differential described by Nelde is also relevant in rural societies, but is more readily observed in urban situations, as a result of the two reasons outlined by Mackey and Romaine. Firstly, the contact is more intense due to the greater concentration of people and as such, conflict is perceived more easily; and secondly urban centres allow many different social and ethnic groups to unite and thus power distinctions will occur between different sectors of society. In short, cities can yield fruitful results due to the nature of urban contact phenomena. However, this is not to say that rural studies are without their merits, as proven by Trudgill (1986): this merely serves as explanation for the choice of an urban centre such as Barcelona in the search for illuminating findings regarding Catalan-Castilian language contact.

**Cities and BUP.** Having demonstrated the links between cities and language contact, it is now necessary to analyse the connection between urban environments and the study of BUP, and particularly my example of BUP, CILP. This can then be applied to various linguistic communities, including Barcelona. Trudgill and Chambers' seminal work *Dialectology* popularised the concept of NORMs. These are non-mobile, old, rural males and are considered bastions of linguistic traditionalism in that they largely do not use innovative forms (Trudgill and Chambers 1998, 29-30).<sup>13</sup> There-

---

<sup>13</sup>Such subjects had often been the focus of studies concerning diatopic variation before this work. However, Trudgill 1986 was the most famous work to formalise and consolidate the concept of NORMs as subjects of linguistic study.

fore, if rural areas are sites of linguistic stagnation, urban settings should play host to linguistic innovation. Indeed, Trudgill and Chambers support this notion by stating ‘(NORMs) should be rural because urban communities involve too much mobility and flux’ (Trudgill and Chambers 1998, 29). The idea that such mobility and migration would encourage language contact (and resultant phenomena) is still a fundamental motivation behind continued study of bi- and multilingual situations (Dewaele et al. 2003, 1).

Therefore, cities encourage innovative linguistic behaviour (as argued by Trudgill and Chambers) as well as language contact (as argued by Mackey, Romaine and Nelde): as such, urban environments are surely a logical choice for the study of CILP. This viewpoint is reinforced in a classic, pre-Weinreich text focusing on evolutionary linguistics. Even before the emergence of contact linguistics as a field, Jespersen (1946) recognises two opposing tendencies in language evolution, one favouring language diversification and the other favouring unification. He claims that philologists had only thus far been considering diversifying factors, at the expense of the study of unification. In Jespersen’s view, there were a great many unifying forces, among them urbanisation (Chambers 2002, 249-50). These unifying forces in language evolution can be interpreted as two languages influencing one another in their development, i.e. BUP. One can thus rely on texts from various linguistic disciplines (evolutionary linguistics and contact linguistics) to support the link between BUP and urban scenarios.

**Examples of BUP in the city.** Having outlined the benefits of research into urban contact linguistics, the following section will demonstrate various manifestations of language contact and BUP in large cities. Please note that it should not be thought that urban language contact situations only fit into one or another category (this includes the cities mentioned below). Moreover, this is not an exhaustive typology and other possible outcomes may arise in cities characterised by language contact. However, these case studies provide evidence that urban contact linguistics, and more specifically CILP (or indeed in some cases, the closely related CILC), can take many forms and result in greatly divergent scenarios.

Use of a language variety can be favoured for reasons of *capital*. In certain contact situations, one particular code is chosen as it bestows potential socioeconomic advancement on its speakers. Morrison and Lui (2000) adopt Bourdieu's (1976) notion of *capital* in order to analyse the usage of English in Hong Kong (characterised by Chinese-English language contact). The social prestige of English as a worldwide language (its linguistic capital) and the possession of a certain cultural background knowledge in order to effectively employ the use of English (cultural capital) combine and convert into access to economic capital. Thus, the use of a particular language can be favoured, such as English in Hong Kong, as a tool of socioeconomic advancement. This idea can be extended to assess the relative socioeconomic value of linguistic varieties affected by CILP, as in Fasold's (1984) analysis of the situation in Singapore. English, Chinese, Tamil and Malay are all spoken in Singapore and since these different languages are imbued with varying amounts of linguistic, cultural and socioeconomic capital, there will be consequences for related contact varieties. For example, in Singapore, Standard English is the most prestigious code. A variety of English influenced by Malay thus loses capital when compared to Standard English. Concomitantly, a Chinese variety influenced by English would offer more socioeconomic privileges than Standard Chinese, due to this English input.<sup>14</sup> Therefore, CILP does not necessarily decrease the capital of a language variety: in some cases, it depends on what language this external influence comes from. In short, certain urban contact situations are characterised by the use of one language over another for reasons of socioeconomic advancement. This does not necessarily imply instability of the variety with less capital, but merely a restriction to certain domains of usage. Moreover, this process not only affects standard languages, but also varieties characterised by CILP.

Processes of BUP can also lend a linguistic variety a certain urban-ness. In certain urban language contact scenarios, varieties characterised by CILP can be viewed favourably. While this thesis does not attempt to address the issue of language attitudes and identities, it is nonetheless interesting to note that in Dakar, Senegal, a contact variety of Wolof and French known as Urban Wolof has come to denote a

---

<sup>14</sup>For more information, see Fasold's (1984, 50) discussion of *linear polyglossia*.



favourable urban identity (Dreyfus and Juillard 2001, 682; McLaughlin 2001, 159). This *dakarois* identity is used in contrast with those not from the city, and implies that non-users of Urban Wolof form an out-group, differentiated as less sophisticated and ‘edgy’. This is thus another example of the potentially instrumental value of certain language varieties characterised by CILP, although this situation refers to a social instrumentalism. While in Hong Kong and Singapore, contact vernaculars can connote socioeconomic advancement, in Dakar they promote a sense of solidarity and co-identity. In short, urban contact varieties can impinge upon both socioeconomic and ethnolinguistic concerns.

Certain cities provide interesting examples of how socioeconomic and ethnolinguistic issues can interact with one another, yielding various results. The history of Montreal shows the original socioeconomic dominance of English language groups followed by a rise in prominence of Francophone sectors of society (Levine 1990, 177). The situation remains complex because not only do clear ethnolinguistic (Anglo versus Franco) divisions persist, but also the status of English or French as instrumental in terms of economic advancement is far from clear. While the 1970s and 1980s saw the exponential growth of Francophone businesses (Levine 1990, 185), as well as the introduction of powerful linguistic legislation favouring French (notably Bill 101 in 1977), the position of English as a world language contributes a certain prestige (Levine 1990, 226). Urban language contact in Montreal thus raises many questions, largely due to ambiguity surrounding the solidarity and prestige values of English and French.

Arguably the case studies highlighted so far have concerned linguistic phenomena that result from language contact and can impact upon a community, that is to say CILP, our example of BUP. In some cases, however, we are arguably dealing with the closely related phenomenon of CILC, given that these practices can sometimes become fully embedded in a linguistic community after generations. In Brussels, French and Flemish are seen as indicative of their respective heartlands (i.e. Wallonia and Flanders). French is therefore the language of socioeconomic dominance as Wallonia was historically more prosperous than Flanders (Treffers-Daller 2002, 51); so while

there may have been an instrumental motive for Flemings to learn French, the reverse was not the case. However, this is not reflected in contact varieties as in Singapore. Ethnolinguistic identity can be seen to beat socioeconomic mobility when examining the phonology of each variety: Brussels French is characterised by Flemish traits, while Brussels Flemish does not contain French traits (Treffers-Daller 2002, 62), the reverse of what one might expect. In short, urban language contact situations need not function like Hong Kong, Singapore or Dakar, which have arguably less ambiguous manifestations of the ethnolinguistic-socioeconomic dichotomy. Each situation is different and CILC/CILP can often provide valuable understanding of how the various languages interact, as in Brussels.

In conclusion, the study of BUP in multilingual cities may prove useful, due not only to the theoretical links outlined above, but also to the manifold phenomena that often arise as a result of urban language contact. Both classic linguistic works (Jespersen 1946; Trudgill and Chambers 1998) and modern texts (Romaine 2000; Mackey 2005) provide examples of links between urbanisation and the study of language contact and BUP. Case studies from around the world also highlight the role played by CILP/CILC in urban situations in increasing the socioeconomic clout of a particular variety (as in Singapore contact varieties), in conveying a sense of ethnolinguistic solidarity (as in Dakar), and in reflecting the simultaneous presence of both factors (as in Brussels). The study of the urban environment may thus yield equally helpful results when examining BUP in Catalonia.

### **3.1.4 Summary**

Section 3.1.1 addressed the theory and development of the concept of language contact, the example BUP background situation. The original ideas laid out in Weinreich (1953) have largely remained constant, namely that the bilingual individual is the locus of contact, but that societal factors also bear upon the situation. Section 3.1.2 established the link between the background situation of language contact and the concrete manifestations of BUP in language form and function. The chief example

given of BUP in language form arising from contact situations is CILP, which covers a number of linguistic phenomena such as simplification, overgeneralisation, analysis, transfer, convergence and divergence, and is fully defined on pages 127-129. The link between the background situation of language contact and resultant processes such as CILP is visually represented in figure 3.1, page 131. Finally, section 3.1.3 applies this notion of CILP to urban scenarios around the world, showing how many revealing and diverse linguistic situations can arise from a backdrop of language contact. In short, this section has explored the BUP background situation of language contact, going on to look at resultant linguistic phenomena. Now that this has been undertaken on a theoretical level, we will turn our attention to the linguistic community of Barcelona, looking at language contact and any resultant manifestations of BUP in form and function.

## **3.2 Language contact and resultant phenomena (BUP) in Barcelona**

The example background situation of BUP given in table 3.1, language contact, has been fully described and theoretical links have been made between this and resultant linguistic manifestations in form and function. Moreover, the choice of an urban setting for a study of language contact has been justified. It is now necessary to apply all of our existing knowledge to the analysis of the linguistic community at hand, Barcelona. Section 3.2 follows a structure which will now be familiar from the last two chapters. Section 3.2.1 addresses our background situation, language contact, in Barcelona. Section 3.2.2 then examines manifestations of this phenomenon in language form and function in Barcelona. Finally, the findings will be summarised in section 3.2.3.

### 3.2.1 Language contact in Barcelona

As outlined in section 1.1, language usage in Barcelona has been characterised for centuries by contact between Castilian and Catalan. As such, a great many sociolinguistic studies make reference to some form of contact phenomena (be it code-switching, transfer etc.) since it is nigh on impossible to examine an area of language contact without doing so. Indeed, it is illogical to analyse the language practices of a community without paying attention to its history, since ‘the sociolinguistics of society is about the social importance of languages to groups of people’ (Fasold 1984, 1) and thus one needs background knowledge of said community in order to identify the significance of their linguistic habits. Therefore, any sociolinguistic study of Catalonia must take into account centuries of language contact and should therefore at least mention contact phenomena (see Calsamiglia and Tusón 1984, 119; Sabater 1984, 32 and Bastardas i Boada 1996 for examples). Many early works unwittingly adopt Nelde’s concept of language conflict in light of the oppression suffered by the Catalan language, and the metaphor of fighting for one’s language was prevalent. Often however, such emotive terminology merely referred to the dynamism of the situation, as in Tovar (1968) where we are reminded that ‘naturalmente las lenguas no luchan ni compiten [porque] no son entes vivos’ (Tovar 1968, 7).<sup>15</sup> Other early authors include Badia i Margarit (1964), Ninyoles (1969) and Coromines (1954), whose works covering Catalan-Castilian conflict and contact were published during a period when the mere use of Catalan was outlawed. Widespread academic study was thus inhibited.

The 1970s and 1980s saw a growth in production, not only due to greater linguistic freedoms in Catalonia, but also to the development of the field of contact linguistics (see section 3.1.1). In 1984, the *International Journal of the Sociology of Language* dedicated an entire issue to Catalan sociolinguistics, thereby addressing many language contact phenomena, including an in-depth analysis of Barcelona language practices (Calsamiglia and Tusón 1984). This continued to be a productive

---

<sup>15</sup>TRANSLATION: Of course languages do not fight or compete [because] they are not living beings.

field of study, and was extended to encompass the examination of language attitudes, as in Woolard's 1989 matched-guise experiment in Barcelona.

Subsequently, work started to address the issue of the language choices made by different groups. Pujolar (1997) systematically assesses the qualities attributable to each language (*les veus del català i del castellà*) before outlining reasons for the selection of a particular code (*tries de llengua*), referring specifically to two Barcelona youth groups, or *penyes*. Similarly, Vila i Moreno (2003) examines specific linguistic domains and language choice therein. This approach is indicative of certain Catalan language contact studies: it is commonplace to address Catalan and Castilian as two immovable entities. Such works detail how the circumstances under which either Catalan or Castilian may be used. While greatly useful, such work gives little indication as to the nature of CILP, as it does not address what is meant by 'Castilian' or 'Catalan', merely using the terms as epithets to cover widely divergent contact vernaculars. Language varieties are thus pigeonholed as one or the other, according to their position on a Castilian-Catalan spectrum. While such a binary classification may at first sight seem reductive, it can prove useful when quantifying CILP. Despite the existence of innumerable contact varieties (indeed each idiolect is arguably characterised by different levels of language interference), Castilian and Catalan are two codified, recognisably distinct language varieties, supported by different official bodies and sharing only a limited degree of mutual intelligibility. So, while such a binary division can sometimes be reductive, it will be initially employed in the next section as a device to allow us to examine existing literature.<sup>16</sup>

The most recent output concerning language contact in Catalonia takes existing work in interesting new directions. Galindo Solé (2004) discusses code-switching, as many had before, the fundamental principles of the study being cited directly from Weinreich's *Languages in Contact*. However, this work differs from previous studies in its focus on the discursive functions of switching between Catalan and

---

<sup>16</sup>See page 43 for discussion of how this existing approach can sometimes be reductive. In the upcoming methodology it will be employed to a degree, insofar as it is necessary to label a given instance of language as either 'Catalan' or 'Castilian' in order to test speakers' judgements as to how accurate these labels are. Thus, this binary division is used as an initial means to gain insight into how the labels are actually used.

Castilian, drawing on notions such as voicing. Nevertheless, this work still (arguably through necessity) takes up the idea of Catalan and Castilian as two immovable entities, overlooking a great deal of potentially interesting, more systematic interference phenomena. Other work addresses contemporary sociolinguistic concerns in a Catalan-Castilian context, such as the increasingly globalised presence of English. De Rosselló i Peralta and Boix Fuster (2004) looks at attitudes towards the use of Catalan, Castilian and English in classes at the Universitat de Barcelona. A number of factors are discussed and explored, including the pressure of international students, as well as requirements that students have knowledge of English to communicate their work internationally; all of this combines to ensure that often the MOI is anything but Catalan. The presence and contact of multiple languages including English is problematic for policy decisions intending to advance Catalan. As such, current studies which examine the contact between Catalan and Castilian in given domains encompass many wide ranging concerns, including even (what we described in chapter 1 as) the maximally distinct phenomena of language policy.

Even when examining this limited cross-section of studies of language contact in Barcelona, it is clear that a number of different concerns have been raised. Early studies focused more on the conflict between dominant Castilian and oppressed Catalan, making heavy use of combative metaphors. Later, this developed into studies of code-switching and attitudes, but often suffered from treating the two languages as immovable entities (however useful a device this may prove to be, even to this thesis). Most recently, contact work has factored in concerns from outside the field, such as how contact and policy decisions may be related. Therefore, this thesis, in aiming to paint a complete sociolinguistic picture of Barcelona by addressing both policy and contact, arguably has its parallels with other contemporary studies of this linguistic community.

### 3.2.2 BUP in Barcelona: Details of phenomena in language form and function

Having addressed language contact in Barcelona, it is now necessary to turn our attention to manifestations of bottom-up phenomena in both form and function, as exemplified in table 3.1. This section will firstly examine phenomena in language form, then those in language function in the Catalan-Castilian linguistic community of Barcelona.

When looking at language form, this thesis is limiting itself to the discussion of linguistic convergence, that is to say, of Castilianisms in Catalan and Catalanisms in Castilian. Table 3.1 exemplifies BUP in language form as the systematic introduction of a (potentially foreign) variant through language contact. Section 3.1.2 determines that phenomena classified as CILP were prime candidates for this manifestation in language form. CILP has been defined as a collection of related processes and phenomena, namely simplification, overgeneralisation, analysis, transfer, convergence and divergence, all of which are defined on pages 127-129 and appear in figure 3.1. Linguistic convergence has been specifically selected for study over these other phenomena, given that Castilianisms in Catalan, as well as Catalanisms in Castilian are widely attested in this linguistic community, and have been for centuries, as we shall see in this section. The other phenomena which make up CILP were rejected for various reasons. Instances of simplification and overgeneralisation per se have been avoided, since such an experiment would have to look at the frequencies of usage of tokens, in order to determine whether they could be classified in such a way. Such a methodology may not ultimately prove as appropriate to the Barcelona linguistic community as one which tests convergence, which we already know characterises the speech situation. Analysis would lead to the examination of a rather limited selection of tokens, while some instances of transfer (borrowing) are potentially fully integrated and normatively accepted, and thus indistinguishable from non-borrowed items, save for their etymology. Finally, instances of contact-induced divergence which are ‘– metacommentary’, and thus are maximally distinct from our policy-based phenomena, are not characteristic of the Barcelona speech situation. Divergence between Catalan and

Castilian is generally as a result of policy, not contact, and so would not be applicable here. Moreover, as already indicated, the related phenomenon of code-switching was not chosen since this raises questions which would employ a completely different set of theoretical concerns. Such a study, while interesting, has not been undertaken for reasons of space. In short, examination of BUP in language form has been limited to discussion of instances of contact-induced linguistic convergence, i.e. Castilianisms in Catalan and Catalanisms in Castilian. This is because all the other phenomena which constitute CILP (or the related concern of code-switching) are either of less direct relevance to the linguistic community, or respond to completely different issues.

**BUP in language form in Barcelona: Castilianisms in Catalan.** It is inevitable that Catalan would have received considerable influence from Castilian, given the latter's historical position as language of power. Many Castilian lexical and grammatical items are fully integrated into the Catalan language, and the *Institut d'Estudis Catalans* maintains strict rules regarding the inclusion of certain Castilianisms in their monolingual Catalan dictionary. The presence of such elements is undeniable, and as such the *Gran Enciclopedia Catalana* provides the following definition and brief history of *castellanismes*:

[El terme *castellanisme* es refereix a un] element lingüístic del castellà usat en una altra llengua. Generalment és del tipus lexical. El català començà a rebre castellanismes cap al s XIV, però és al s XV que n'augmenta el nombre. Durant el Segle d'Or castellà penetraren mots cortesans després abandonats, i de tota altra mena d'esferes. De fet, malgrat aquesta invasió, els termes genuïns no desaparegueren completament. Mentre en algunes contrades restava el mot genuí per a designar un objecte concret, en d'altres, on aquest no existia, s'introduïa el castellanisme corresponent. El calc sintàctic, que s'havia limitat a la prosa de certs escriptors, ha començat a arrelar en el parlar del poble (*Enciclopèdia Catalana*, s.v. *castellanisme*).<sup>17</sup>

---

<sup>17</sup>TRANSLATION: [The term *Castilianism* refers to a] linguistic element of Castilian origin used in another language. It is generally lexical in nature. Catalan began to receive Castilianisms towards the fourteenth century, but it is in the fifteenth century that they increase in number. During the Spanish Golden Age, words entered the language, be it from the Spanish royal court or any number of other fields, which were later abandoned. However, despite this invasion, the original terms did not disappear completely. While in some regions the original word remained in use to designate a particular object, in others the corresponding Castilianism was introduced. Syntactic calquing,



Many early modern treatises of Catalan grammar, including the *meisterwerks* of Pompeu Fabra, give examples of such traits. Before contact linguistics was even established as a discipline, Fabra (1924) outlines examples of phonic interference (*el[s]altres* instead of *el[z]altres*), lexical calquing and transfer (*relloc* from Castilian *repuesto*, instead of the normative *recanvi*), morphological issues (inclusion of the morpheme *-ero* in Catalan words as in *bombero*), and syntactic characteristics (functional shift of the marker *baix* to occupy the position of *sota*, following the model of Castilian *bajo*).<sup>18</sup> Lamuela and Murgades (1984) point out that Fabra attributes the above contact phenomena to socio-political factors, namely the position of Castilian as dominant language (Boix i Fuster and Vila i Moreno 1998, 228).<sup>19</sup>

In short, the presence of features of Castilian origin in Catalan speech has long been acknowledged and as such, Castilianisms have been the focus of much academic and non-academic literature. Pericay and Toutain's 1986 *Verinosa llengua* advocated the standardisation and diffusion of a form of Catalan characterised by more Castilianisms as this represents 'el català que ara es parla' (the Catalan spoken nowadays). Such varieties were later derided and dubbed *català light* or *catanyol* in works such as Pazos i Noguera's 1990 *L'amenaça del català light* (the threat of *català light*) and Alexandre's 2006 *TV3 a traïció* (the treachery of TV3).<sup>20</sup> Recent contributions have discussed the role played by register in such language use, focusing on the heavily Castilianised character of Barcelona Catalan slang. Pujolar (1997, 200) states 'la major part del vocabulari de l'argot que s'utilitza a Barcelona prové del castellà. L'argot català ha deixat d'existir.'<sup>21</sup> Barber Casanovas (2007, 23) supports this, arguing 'no és gens estrany que l'argot català sigui majoritàriament manllevat de l'espanyol.'<sup>22</sup>

---

which was previously limited to the prose of certain writers, has begun to take root in general speech.

<sup>18</sup>This specific example will go on to form part of my experiment, as seen on page 187

<sup>19</sup>Admittedly, Fabra did not employ the above Weinreichian division, but the various contact phenomena were nonetheless highlighted.

<sup>20</sup>TV3 is the primary Catalan-language television network.

<sup>21</sup>TRANSLATION: Most of the vocabulary of Barcelona slang comes from Castilian. Catalan slang has ceased to exist.

<sup>22</sup>TRANSLATION: It is not at all strange that Catalan slang should be mostly borrowed from Spanish.

Concrete instances of Catalan-Castilian linguistic convergence and the manifestations of CILP are provided in Boix i Fuster and Vila i Moreno (1998, 257), and Echenique Elizondo and Sánchez Méndez (2005, 287, 296). The pre-pilot questionnaire (see section 4.1.2.2) also aims to identify the nature of Castilianisms (and indeed Catalanisms as outlined in the next paragraph). The precise form of Castilianisms will not be discussed here since this chapter merely addresses theoretical concerns; they shall instead receive attention in chapter 4, which examines methodological issues pertinent to the content of my fieldwork.

**BUP in language form in Barcelona: Catalanisms in Castilian.** Castilianisms have thus been attested for centuries, and have constituted the focus of both academic and non-academic attention since at least the period of Catalan language standardisation in the early twentieth century. However, scholars have recently broadened their approach to include contact phenomena in the opposite direction, i.e. Catalanisms in the Castilian of Catalonia. Seib (2001) suggests possible motives for this shift in focus as follows:

En esta zona de intenso contacto lingüístico, no obstante, por razones políticas e históricas, los estudios hasta hace pocos años se centraban únicamente en los efectos que ejercía el español sobre el catalán; *la influencia del catalán sobre el español no se focalizaba* (Seib 2001, 4, my emphasis).<sup>23</sup>

Such ‘political and historical’ reasons likely refer to the political developments addressed by many authors (Szigetvari 1994, ii; Grupo Interalia 1998, 4; Seib 2001, 4-5; Galindo Solé 2003, 1; Echenique Elizondo and Sánchez Méndez 2005, 298-301 inter alia). Catalan linguistic normalisation efforts of the 1980s contributed to a shift in language prestige and thus potentially a shift in the focus of study. These recent political and social developments are of great importance in the current Catalan linguistic situation (see section 2.2.1). Such relationships between processes of TDP and

---

<sup>23</sup>TRANSLATION: However, in this area of intense linguistic contact, for political and historical reasons, until recently studies solely focused on the effects that Spanish had on Catalan; *Catalan’s effect on Spanish was not addressed*.

BUP will be explored in due course, notably in the discussion in chapter 6. For our current purposes, it should be remembered that BUP do not occur in a vacuum, and that there is always a socio-political context.

The late twentieth and early twenty-first centuries have thus seen a surge in the production and publication of studies of Catalanisms, including Sinner's (2002) doctoral thesis *El castellano de Cataluña*, which provides a near-exhaustive typology of specific Catalanisms and details circumstances of their usage (cf. also Seib 2001 for similar classificatory work). Sinner's personal website ([www.carstensinner.de](http://www.carstensinner.de)) has since provided a forum for other scholars to publish papers addressing Catalanisms, and as such, one can now consult lexicons of Catalanisms (Szigetvari 1994; Barber Casanovas 2007), as well as studies regarding syntactic calquing (Grupo Interalia 1998) and the literary use of Catalanisms (Barrera i Aguilera 2000). Moreover, the Grupo Interalia was founded in 1994 within Barcelona's Universitat Pompeu Fabra, in order to investigate Catalan-Castilian language interference, i.e. both Catalanisms and Castilianisms. In short, the growth in the study of Catalanisms is patent.

Turning to the usage of Catalanisms, Colón (1989, 82) primarily views lexical transfer as a function of geolinguistic variables, citing Western Catalan's geographical proximity to Castilian-speaking areas as reason for their use of the normative Castilian *primo carnal*, while Eastern Catalan favours the calque *primo hermano* (cf. Catalan *cosí germà*). Although more recent studies make little reference to the geographical axis of enquiry, they do not aim to predict or explain the usage of Catalanisms in terms of sociolinguistic variables (gender, register, age etc). The Grupo Interalia (1998) describes the transfer of the two-way Catalan deixis system into Castilian (which is normatively split three ways), yet do not specify such transfer as a function of given sociolinguistic variables. Similarly, Galindo Solé (2003) examines the impact of Catalanisms on the system of Barcelona Castilian as a whole, rather than pinpointing the contexts in which such interference phenomena occur. In other words, a greater emphasis is placed on listing examples of Catalanisms than on determining when and why they may arise. Once again, consult chapter 4 for specific examples of Catalanisms, as well as Castilianisms.

In short, in recent years there has been a shift in focus regarding the study of CILP in Catalonia. The discipline now includes not only the analysis of Castilianisms, but also of Catalanisms. Such a shift in emphasis may be attributable to a number of factors, and many academics cite recent political developments such as 1983's *Llei de Normalització Lingüística* as a potential cause. See section 2.2.1 for a full discussion of linguistic legislation in Catalonia.

When looking at BUP in language form, it is clear that Barcelona has a long history of language contact, which has largely manifested itself as Castilianisms in spoken Catalan. This phenomenon has been attested since the fourteenth century and has been maintained due to steady and constant contact between Catalan and Castilian. However, recent years have seen an increase in the study of Catalanisms in spoken Castilian. This raises several questions. Did people not consider these phenomena worthy of attention before? While it may seem counterintuitive to suggest that, until this point such phenomena had been monodirectional, how can one be sure that Catalanisms were a widespread phenomenon prior to this period (despite receiving little to no academic attention)? Has there been a development in the nature of Catalan-Castilian language interference in recent years? Indeed, have socio-political concerns played a role? All of these questions merit further exploration.

**BUP in language function in Barcelona.** These are not prevalent with regard to Catalan, in light of the major TDP in language function that have greatly improved the position and overt prestige of Catalan. As such, they do not characterise the speech situation in Barcelona as much as BUP in language form, and consequently may not be deemed of as much relevance to the study. Varieties characterised by contact elements are still stigmatised, but this stigma has not spread to the language as a whole. Due to the extensive linguistic legislation which promotes the overt prestige of normative Catalan (see section 2.2.1), I maintain that normative Catalan as a system does not suffer from a decrease in overt prestige because contact varieties of it may be used in certain situations. Rather it is the contact varieties themselves

that are stigmatised: a variety of Catalan which contains heavy lexical or structural interference from Castilian will be accorded low overt prestige and deemed unsuitable for certain situations, but the standard language retains its high level of prestige. It will be interesting to ask if the situation is the same for non-normative Castilian in the region. All being equal, non-normative Castilian should be lower in overt prestige than the standard language. However, this may not be the case, and will receive more attention in my upcoming experiment.

### 3.2.3 Summary

This section has applied the theoretical findings of section 3.1 to the community under examination. As in chapter 2 with regards to language policy, this section first addresses the background situation, language contact (section 3.2.1) before tackling the manifestations of BUP in both language form and function (section 3.2.2). Studies of language contact in Barcelona have focused on concrete processes such as code-switching and borrowing but almost all have maintained the distinction between Catalan and Castilian as two immovable entities. This distinction, while somewhat artificial since each idiolect arguably falls on a continuum between the two codes, will allow us to clearly examine the bidirectional CILP between the two languages, by letting us differentiate between Castilian-influenced items in Catalan on one hand, and Catalan-influenced items in Castilian on the other. Contact-influenced linguistic convergence can affect both languages, although Catalanisms in Castilian have only recently received extensive attention. Consequences of language contact in language function also throw up questions of the relevance of bi-directionality. As with language form, both codes can of course influence one another, but again, the impact on the Castilian of the region is comparatively overlooked.

Evidently there has been a development of late, and people are now waking up to the idea that linguistic convergence can work both ways in this community. The developments in language policy outlined in chapter 2 have led to a radical shift in the linguistic landscape of Catalonia, and so people are having to adjust their perspective. Thus, policy and contact may be bearing upon one another. In choosing

to address both of these concerns, the thesis will be able to comment upon any potential competing pressures exerted by policy and contact (such observations will be fully explored in the discussion in chapter 6).

### 3.3 Language contact in Barcelona: Overview and research question

This chapter has discussed at length the presence of situations of language contact as well as resultant phenomena in both language form and function. Section 3.1 examined this on a theoretical level, employing examples from linguistic communities worldwide, while section 3.2 focused specifically on the Catalan-Castilian community of Barcelona. As a result, we not only can confirm that this speech situation is characterised by the presence of extensive contact between Catalan and Castilian, but we also can give details as to the specific nature of some of the contact phenomena present in Barcelona. For reasons detailed in section 3.2.2, the thesis has chosen to focus its attention on instances of contact-induced linguistic convergence between Catalan and Castilian. Castilianisms in Catalan and Castilianisms in Castilian were subsequently explained at length, and are comprehensively detailed in studies such as Boix i Fuster and Vila i Moreno (1998) and Echenique Elizondo and Sánchez Méndez (2005) (for Castilianisms), and Szigetvari (1994) and Sinner (2002) (for Catalanisms). Moreover, many examples of Castilianisms in Catalan can be found in official Catalan language policy initiatives such as the *propostes per a un estàndard oral* and the many *diccionaris de dubtes i barbarismes* found on the market (for greater discussion of these, see chapter 2).

It is now required to devise a research question which allows us to examine contact phenomena in the Barcelona linguistic community, and to justify the choice of such a question. Our contact research question should allow us insight into the role played by instances of contact-induced linguistic phenomena in the speech situation, and should be empirically testable by the upcoming experiment. Moreover, in order to provide a

more complete picture of the linguistic community, our contact question should relate to the policy question given at the end of chapter 2. It could thus be made clear the degree to which the co-existing pressures of language policy and language contact are intertwined.<sup>24</sup>

In light of the theoretical and practical concerns raised thus far, I have chosen to examine the degree to which speakers are aware of instances of linguistic convergence arising from contact, when such convergence does not form part of the normative version of either language. That is to say, I am looking at speaker *awareness of non-normative language use*, specifically non-normative Catalanisms in Castilian and Castilianisms in Catalan. There are many reasons to look at awareness of contact-induced linguistic phenomena, not least because this may allow the research community a first impression as to how such linguistic instances may be perceived. In order for future studies to look at contemporary instances of language contact in Barcelona in more detail, we must first know if speakers are even aware that they exist, and if so, to what degree. Moreover, by focusing on awareness of language use and integrating some of the findings from both chapters 2 and 3, some interesting potential issues arise (particularly when bearing in mind the axis of ‘+/- metacommentary’, as it appears in the model in chapter 1):

1. Language policy finding: Normative Catalan has been chosen as the medium of instruction. One of the education system’s aims is to ensure that students leave the system with a high level of Catalan language competence. Therefore, attempts are made to ‘rectify’ the use of Castilianisms.
2. Language contact finding: Given the ethos of the new Catalan-medium school system, Castilianisms in Catalan are being highlighted and discouraged. This means that explicit metacommentary is being offered on certain instances of non-normative contact-induced linguistic convergence.
3. Problem: Castilianisms in Catalan are instances of BUP and therefore occur

---

<sup>24</sup>If the phenomena resulting from policy (TDP) and contact (BUP) are shown to bear upon one another, there may be interesting consequences for the model given in chapter 1, allowing as it does the simultaneous representation of such maximally distinct phenomena in one single language situation: any such consequences could be addressed in the discussion in chapter 6.

below the level of conscious awareness. This is why they are ‘– metacommentary’. But people are commenting on them. Which means they are now *above* the level of conscious awareness. So surely they are now ‘+ metacommentary’. How can this be explained?

Given that the instances of language contact we are examining constitute non-normative use, they are targeted by the policy measures which dictate that normative Catalan should be used in classrooms. The direct conflict between policy and non-normativity – specifically here, contact-induced linguistic convergence – is clear. If people are made aware that they exhibit non-normative language usage, and are also made aware of the fact that this practice is low in overt prestige, they will arguably try to eliminate said usage from their speech. At least that is the idea, and can be summarised as follows:

Awareness that you		Awareness that this		Desire to eliminate
exhibit a linguistic	+	trait is low in	=	this trait from
trait		overt prestige		your repertoire

In short, this is how policy is tackling contact. But to what extent is this happening? And how can this be tested? In order to test speakers’ desire to eliminate contact elements from their repertoire, a study of language attitudes would be required, and falls outside the remit of the thesis. However, testing speaker awareness of a particular phenomenon is easier. As shown in figure 1.2 in chapter 1, Labovian *stereotypes* occur above the levels of consciousness and awareness. A defining feature of such stereotypes is that speakers can offer metacommentary about them when provoked. Therefore, if speakers outwardly identify the occurrence of a particular linguistic phenomenon, it logically follows that they are both conscious and aware of it.

In conclusion, this study of language contact has chosen to examine speakers’ awareness of non-normative instances of contact-induced linguistic convergence. This is because firstly, it allows us a basic insight into how contact-induced language is viewed



in this community. Secondly, there may be interesting consequences for the model of sociolinguistic processes given in chapter 1, since there may prove to be some flexibility as regards the ‘+/- metacommentary’ axis. Thirdly, such a question would tie in with our existing question on language policy, since the language policy decision to promote normative linguistic varieties is contingent on awareness of non-normative language use.

Finally, in order to test any link with policy initiatives, we would need to examine if awareness of non-normative language use was different for people depending on whether they had benefited from the language-in-education reforms which promote normative Catalan that were introduced in the early 1980s. That is to say, is it possible to witness cross-generational differences?<sup>25</sup>

Bearing all of this in mind, our contact research question is as follows:

#### **CONTACT RESEARCH QUESTION**

To what extent are Catalan-Castilian bilinguals aware of non-normative instances of contact-induced linguistic convergence? Has there been any development in this awareness in recent years?

---

<sup>25</sup>Related methodological concerns will be addressed in section 4.1.2.1.

# Chapter 4

## Experiment Methodology

The previous three chapters have addressed the sociohistorical and theoretical concerns of the study and culminated in a series of general research questions which are now to be tested. This chapter addresses those research questions and formulates an innovative and unique experiment methodology which will ultimately allow for the successful and insightful examination of the Catalan-Castilian bilingual community of Barcelona. Section 4.1 highlights the specific nature of the phenomena to be discussed and concludes with a number of criteria that the experiment must fulfil. Based on this, a pilot study was conducted, and will be discussed in section 4.2. Finally, these findings are developed into a workable fieldwork methodology, presented in section 4.3.

### 4.1 Methodological Context

This section will outline the methodological background which will inform the creation of the fieldwork experiment. Firstly, the research questions posed in chapters 2 and 3 will be addressed, and the specific components of language policy and language contact in Barcelona that require attention will be identified. Then, once the most relevant aspects have been outlined, a number of experimental concerns will be addressed, which tackle issues pertaining to policy and contact. Then, when all this information has been presented, we will have the necessary tools to formulate the

pilot study (in section 4.2), and finally the final fieldwork (in section 4.3).

### 4.1.1 Addressing the research questions

The previous three chapters have provided us with a series of research questions to be tackled by the thesis, re-stated below:

#### POLICY RESEARCH QUESTION

To what extent is current Catalan language-in-education policy meeting its aims?

#### CONTACT RESEARCH QUESTION

To what extent are Catalan-Castilian bilinguals aware of non-normative instances of contact-induced linguistic convergence? Has there been any development in this awareness in recent years?

These research questions seek to examine specific facets of language policy and language contact in Barcelona. These have been outlined in the previous chapters. As regards language-in-education policy, the measures to promote Catalan from forbidden language to chief vehicular medium of instruction (see section 2.2.1.1) have resulted in a great shift in classroom language use. This ties in with Catalan governmental policy aims, which state that children should have a high level of competence in both Catalan and Castilian on leaving the system (see section 2.2.1). Schooling therefore aims to increase Catalan language competence, and thus is the primary vehicle for the ‘attempted *rectification* of Castilianisms’, given as a key concern in section 2.2.2. Concerning language contact, we have seen that non-normative linguistic convergence between Catalan and Castilian (that is to say, the presence of Castilianisms in Catalan and Catalanisms in Castilian) has been chosen as a focus of study. Moreover, the research question seeks to address speaker awareness of such non-normative linguistic convergence, for the reasons outlined in section 3.3, namely that this is a means of examining the direct conflict between policy and non-normativity, thus allowing us to potentially generalise any findings to the discussion of policy, and that this may have interesting consequences for the model of language change presented in chapter

1. These facets can be summarised as follows:

Language policy in	Language contact in
<b>Barcelona:</b> see section 2.2.2.	<b>Barcelona:</b> see section 3.2.2.
<ul style="list-style-type: none"><li>• Establishment of Catalan as medium of instruction.</li><li>• Attempted <i>rectification</i> of Castilianisms.</li></ul>	<ul style="list-style-type: none"><li>• Speaker awareness of Castilianisms (in Catalan).</li><li>• Speaker awareness of Catalanisms (in Castilian).</li></ul>

Now that the relevant concerns have been determined, it is necessary to formulate an experiment that integrates and analyses all of them. The next section will discuss the theoretical and practical issues underpinning the design of the upcoming experiment, first focusing on how to effectively include the language policy points under discussion, before turning its attention to language contact.

#### 4.1.2 Designing the experiment

So firstly, the experiment will be analysing developments in the Catalan education system, i.e. the choice of a normative Catalan which is free of Castilianisms as chief medium of instruction. How can we go about evaluating the education system? These are our language policy concerns.

Secondly, speaker awareness of said Castilianisms must receive attention, along with their less-examined counterpart, Catalanisms (in non-normative Castilian). The extent to which such contact elements are employed and potentially later embedded will be determined in order to gain a firmer grasp on how they function within the codes of the linguistic community. But what is the exact form of non-normative contact-induced linguistic convergence in the community? These are our language contact concerns.

In order to design an effective experiment which addresses these two issues, practical problems need to be solved with regard to experiments testing policy and contact.

Only then, with full knowledge of all the facts, can a successful fieldwork experiment be undertaken.

#### 4.1.2.1 Language policy concerns : Apparent time studies

The primary language policy concern is testing Catalan-medium education and how it has facilitated the ‘rectification’ of Castilianisms, thus ensuring a high level of student competence in normative Catalan. As has already been discussed in section 2.2.1.1, the choice of Catalan as medium of instruction was largely implemented through 1983’s *Llei de Normalització Lingüística* (LNL). For this new system to be evaluated, it must be contrasted with the old one. Therefore, the education system before 1983 needs to be contrasted with the system after 1983. Short of going back in time, how can this be achieved? This section addresses this concern and provides us with the theoretical tools to undertake such an experiment.

In order to compare pre-*normalització* and post-*normalització* behaviour, I have chosen to examine younger and older generations. The technique of comparing the linguistic practices of different generations as a means of drawing further conclusions is known as an ‘apparent time’ study. Such studies are frequently favoured by sociolinguists since the alternative (‘real time’ or ‘longitudinal’ studies) requires the analysis of the same (or a directly comparable) sample of people in actual time, often unfeasible when the relevant time period can be of decades or more. Therefore, any data gathered to support hypotheses regarding policy effectiveness since the LNL of 1983 must be drawn from an apparent time fieldwork investigation.

Apparent time studies are dependent on the ‘apparent time hypothesis’, which states that ‘the linguistic usage of a certain age group will remain essentially the same for the people in that group as they grow older’ (Chambers 2002, 212). The hypothesis has been tested and fundamentally proven through comparison with real time findings (cf. Trudgill 1986 on Norfolk; Cedergren 1987 on Panama).<sup>1</sup> Based on such empirical observation, the apparent time hypothesis predicts that elements learned by participants during the education process – such as, for example, the ability

---

<sup>1</sup>This is despite some notable caveats, such as forms of *age grading*, whereby use of innovative forms is distributed on a U-shaped curve, with greater frequency in early and later life.

to identify non-normative language use – will be representative of the school system, and thus the linguistic community, of the era when the subject was a schoolchild. The apparent time hypothesis is therefore appropriate for my work, and will be employed for considerations of efficacy, resources and time.

#### 4.1.2.2 Language contact concerns : The pre-pilot questionnaire

It has already been established that this thesis will primarily test speaker awareness of instances of non-normative language use, particularly convergence between Catalan and Castilian. However, in order to empirically test this, we need to be familiar with the nature of such interference elements. In order to gain insight into what the speech of *barcelonins* actually looks like, a pre-pilot questionnaire was distributed to twenty native Catalan-Castilian bilinguals. The results are presented here, complete with supporting quotes from the questionnaires. After examining this preliminary data, it will be clearer which specific instances of non-normative language use should be included in the upcoming fieldwork experiment.

The questionnaire itself can be found in Appendix A. The questions were designed to encourage broad answers from participants as I merely wished to identify general traits of language use seen to be characteristic of Barcelona youth speech, which I could then refine and develop. The questions took two different criteria into account. Firstly, the *generation of speaker*. In line with the apparent time methodology, I chose to divide the linguistic community into, on one hand those under the age of thirty-five (i.e. those largely educated through the post-*normalització* policy of *immersió lingüística*) and, on the other hand, the remainder. Secondly, *the language under consideration*. I aim to study all language varieties in Barcelona and how they are used, and therefore I need to gather information on both Barcelona Castilian and Barcelona Catalan. The final question was asked so that respondents would provide me with as much metalinguistic commentary as possible. I did not want respondents to fail to mention phenomena they consider irrelevant, in case such information should prove beneficial to my research. The information obtained could therefore be categorised as follows:

		Generation of speaker	
		Younger	Older
Language under consideration	Catalan	Information about younger Catalan speech practices	Information about older Catalan speech practices
		Castilian	Information about younger Castilian speech practices

Table 4.1. Proposed findings of pre-pilot questionnaire

The questionnaire was chiefly distributed by email, thus eliminating the need for travel and any concomitant expenses. However, with three respondents, face-to-face interviews were conducted. Respondents were obtained through personal contacts of members of the Centre for Catalan Studies at Queen Mary, University of London. Aside from considerations of practicality (ease of acquiring participants), this had the secondary benefit of an increased metalinguistic awareness of the Catalan situation on the part of the respondents. Given their links to the Centre, the majority are engaged in some form of Catalan language work, be that in a university department of *filologia catalana* in Catalonia or as a Catalan language teacher in the UK. In light of the above considerations, respondents were not controlled for sociolinguistic variables such as age, gender, social class or place of origin. Given the preliminary nature of this questionnaire (i.e. that it does not constitute my actual fieldwork), any restrictions placed on respondents may prove counterproductive. At this early stage, it would be unwise to only contact members of the intended target age groups, since questionnaire respondents cannot be used in later experiments due to their familiarity with the project. In short, I did not want to exhaust all contacts at this preliminary stage as I am reserving potentially suitable candidates for participation in the forthcoming fieldwork. Basic respondent details are as follows:

	Barcelona	Catalonia (elsewhere)	Comunitat Valenciana	Balearic Islands	Sardinia (l'Alguer)
Male (N=11)	3	6	0	1	1
Female (N=9)	8	0	1	0	0
<b>Total (N=20)</b>	<b>11</b>	<b>6</b>	<b>1</b>	<b>1</b>	<b>1</b>

Table 4.2. Questionnaire respondent details, categorised by sex and place of origin / residence. <sup>2</sup>

The totality of linguistic features outlined in the responses is as follows. More common responses (those mentioned by more than one participant) will then be explored in greater detail.

Linguistic feature	Nature of feature	Cited in _ responses
Voicing / devoicing of S	Phonological	3
Mismatch of Catalan and Castilian vowel systems	Phonological	3
Affrication of fricatives in Catalan	Phonetic / Phonological	2
Velarisation of L	Phonetic	2
Pronoun difficulties in Catalan	Syntactic	8
Syntactic calquing	Syntactic	6
<i>Apostrofació oral</i> and related issues	Syntactic	1
Lexical calquing	Lexical	10
'Poverty' of youth vocabulary	Lexical	5
Emergence of neologisms	Lexical	4
Loss of old vocabulary	Lexical	2

<sup>2</sup>Several participants stated within their responses that they had spent many years away from their place of birth and as such, sole reference to their place of origin would have been misrepresentative of their speech habits. In these cases, they have been classified according to their place of residence within the Catalan speaking world.



Youth ignorance of different registers in Catalan	Pragmatic	2
Youth codeswitching	Sociolinguistic	1
Youth code crossing	Sociolinguistic	1
Youth using fewer idioms in Catalan	Sociolinguistic	1
Youth using more ‘correct’ Catalan	Sociolinguistic	1
Elders using more ‘correct’ Catalan	Sociolinguistic	1

Table 4.3. Linguistic features specific to Barcelona, according to questionnaire respondents.

*Phonetic and phonological features.*

- *Voicing / devoicing of S.* Intervocalic -s- in standard Peninsular Castilian is pronounced as the voiceless apical alveolar fricative sibilant /s̺/. In standard Barcelona Catalan the same grapheme is rendered as the voiced apical alveolar fricative sibilant /z̺/. Thus, *casa* (house) is pronounced [ˈka.s̺a] in Castilian and [ˈka.z̺ə] in Catalan. This interference is already attested in the literature, (cf. Echenique Elizondo and Sánchez Méndez 2005, 300) with the Castilian of Catalonia being characterised by the presence of this intervocalic voiced sibilant, while the Catalan of Castilian-dominant speakers shows the reverse tendency. This is a particularly innovative phenomenon since use of /z̺/ in Castilian arguably constitutes the adoption of a new phoneme, as this sound does not form part of the standard Castilian phonemic inventory (though it can occur allophonically in some dialects before a voiced consonant). Questionnaire respondents also referred to this feature:

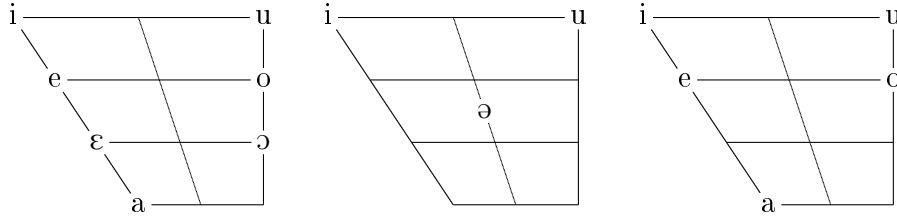


Figure 4.1: Vowel charts showing (from left to right): the seven-vowel system of Catalan in stressed position, the three-vowel system of Catalan in unstressed position, and the five-vowel system of Castilian in stressed or unstressed position

Les persones catalanoparlants habituals pateixen *alguns* problemes d'interferència lingüística del castellà... sens dubte, el fenomen [fonològic] principal és la confusió de s sorda i sonora <sup>3</sup> (respondent 1e).<sup>4</sup>

- *Mismatch of Catalan and Castilian vowel systems.* The Catalan vowel system is more extensive than that of Castilian, consisting of eight vowels (seven in stressed position, three in unstressed) as opposed to the five found in Castilian (stressed or unstressed). These are fully detailed in figure 4.1. Bilingual speakers' appropriation of the vowel system of one language into the other is characteristic of language contact varieties found in Catalonia (Echenique Elizondo and Sánchez Méndez 2005, 299), and is highlighted by respondents:

Tinc un parell d'amigues [d'origen castellanoparlant espanyol]... Les dues parlen el català amb greus problemes fonètics. No distingeixen... les vocals obertes de les tancades (respondent 5a).<sup>5</sup>

- *Affrication of fricatives in Catalan.* Word-initial x- in Catalan is normatively pronounced as the voiceless palato-alveolar fricative sibilant /ʃ/. Thus *xerrar*

<sup>3</sup>TRANSLATION: People who frequently use Castilian evidence *some slight* (original emphasis) linguistic interference issues which come from Castilian... without doubt the main [phonological] phenomenon is the confusion of voiceless and voiced s.

<sup>4</sup>In order to preserve anonymity, pre-pilot questionnaire respondents have been indexed using a system of letters and numerals, which indicate to the researcher both their identity and how they were introduced to the study.

<sup>5</sup>TRANSLATION: I have a couple of friends [of Spanish, Castilian-speaking origin]... They both have serious phonetic problems when speaking Catalan. They don't distinguish... between open vowels (i.e. /ε/ and /ɔ/) and closed ones (i.e. /e/ and /o/).

(Eng *to chat*) is rendered as [ʃə.'ra] and *xocolata* (Eng *chocolate*) as [ʃu.ku.'la.tə]. However, potentially under the influence of the Castilian cognates *charlar* and *chocolate*,<sup>6</sup> this initial sound is often affricated to /tʃ/ in Barcelona Catalan, giving [tʃə.'ra] and [tʃu.ku.'la.tə]. This is attested in the questionnaire:

Pel que fa a la fonètica, les generacions més joves afriquen *xocolata* (respondent 1xy.1).<sup>7</sup>

- *Velarisation of L*. This feature is due to the presence of an allophone in the Catalan phonetic inventory that is absent from Castilian. Catalan contains an alveolar lateral approximant [l] as well as a velarised alveolar lateral approximant [ɭ]. The latter is an allophonic variant of the phoneme /l/ and appears in syllable codas. However, Castilian only contains the non-velarised variant, and thus *sol* (sun) is pronounced as [sol] in Castilian and [sɔɭ] in Barcelona Catalan. As such, the presence of this velarised allophone in Castilian is often recorded in the speech of Catalan-Castilian bilinguals (Echenique Elizondo and Sánchez Méndez 2005, 299):

[Un] dels trets més fàcils de detectar en persones grans que normalment parlen en català quan parlen en castellà [és la] *ela velar* (respondent 1xy.1).<sup>8</sup>

### *Syntactic features.*

- *Pronoun difficulties in Catalan*. Catalan contains a more extensive series of weak clitic pronouns than Castilian, including the pronominal adverbs *hi* and *en*. *Hi* can function as a place adverb (Cat *hi sóc* = Eng *I am there*) or as a pronoun which replaces a complement preceded by any preposition other than

<sup>6</sup>Although such an observation is highly speculative, and further evidence is required before such a claim could be made.

<sup>7</sup>TRANSLATION: With regard to phonetics, younger generations affricate [the initial sounds in words such as] *chocolate*

<sup>8</sup>TRANSLATION: [One] of the easiest traits to pick out (when listening to) older people, who normally speak Catalan, when they speak Castilian [is the] velar L.

*de* (Cat *si jo renunciava a tenir-ho, hi renunciaries també?* = Eng *if I gave it up, would you give it up too?*) (Enciclopèdia Catalana, s.v. *hi*). Similarly, *en* can function as a partitive (Cat *en vols?* = Eng *do you want some of it?*) or as a pronoun representing a complement preceded by *de* (Cat *abans tractava d'aquesta pel·lícula, però ara no en tracta* = Eng *he used to deal with that film, but now he doesn't (deal with it)*) (Enciclopèdia Catalana, s.v. *en*). As such, these items are cognate with the French *y* and *en* respectively, but have no counterparts in Castilian. This asymmetry leads to an elimination of these weak pronouns in Catalan, described in the questionnaire responses:

Els locutors de TV3 i altres TV locals en català també estàn tenint més i més problemes amb els pronoms febles... i el pitjor exemple de tot això és que tots estem patint aquest 'virus'... Vaig aprendre els pronoms febles en adquirir la llengua de petita i a l'escola ens vam treballar per acurar-ne més l'ús... però ara m'adono sovint que el deixo de dir i m'autocorregixo... Fins i tot un amic meu que és filòleg i a més té el nivell D de català... quan parlem i m'adono que es deixa un pronom, li dic 'VIRUS' (respondent 5a).<sup>9</sup>

- *Syntactic calquing* is often seen in situations of language contact (cf. Heine and Kuteva 2005), whereby one language adopts grammatical structures present in another with which it is in close contact: these adopted syntactic structures often co-exist with the native structures, and can even supplant them, leading to CILC. Such syntactic calquing may prove difficult to identify in languages as typologically similar as Catalan and Castilian, but is nonetheless recognisable. Indeed, the above 'pronoun difficulties' are an example of calquing a Castilian syntactic structure (i.e. one without the pronouns *hi* and *en*) into an otherwise Catalan utterance. Thus, my description of features as instances of 'syntactic calquing' refers to examples whereby syntactic characteristics of one language

---

<sup>9</sup>TRANSLATION: Presenters on TV3 and other local Catalan-language channels are having more and more problems with the weak pronouns... and the worst thing about all this is that we're all suffering from this 'virus'... I learned the weak pronouns when I acquired my language as a little girl, and at school they made us work at refining our use of them... but now I realise that I'm forgetting them and I correct myself... Even a friend of mine who's a language specialist and has a level D in Catalan ... when we speak and I realise he's dropped a pronoun, I say 'VIRUS'.

are used in a sentence which (at least primarily) contains lexical items from the other language. This does not cover instances of ‘pronoun difficulties’, which are being treated as separate syntactic interference phenomena, due to their widespread nature. Again, this has been attested in my questionnaire responses:

En pocs anys, he notat com algun amic meu [catalanoparlant de tota la vida], ha passat de dir ‘On ets? Quedem a dos quarts de nou?’ a dir ‘On es-tàs? Quedem a les vuit i mitja?’. La forma correcta en català és la primera. En la segona, les paraules són catalanes, però l’estructura, l’esquelet de la llengua, és castellana (respondent 1m).<sup>10</sup>

### *Lexical features.*

- The terms *lexical borrowing* and *lexical calquing* refer to closely related interference phenomena. Borrowing involves the inclusion of foreign lexical material with at most phonological / orthographic adaptation. A calque does not include any extraneous form, but imitates with native morphemes a structure from another language. Lexical calquing is an arguably basic practice in situations of language contact and it has been claimed that extensive lexical borrowing is a prerequisite for similar syntactic calquing, described above (Winford 2003, 53-4). As such, a great deal of language interference phenomena takes the form of lexical borrowing. In the Catalan context, alien (usually Castilian) lexical items are described as *barbarismes* (barbarisms) and the *Diccionari de dubtes i barbarismes* mentioned in section 2.2.2 states that ‘un barbarisme és una forma lingüística d’origen estranger no assimilada a la llengua pròpia i normativament rebutjada.’<sup>11</sup> Several participants specified that a great deal of interjections in

---

<sup>10</sup>TRANSLATION: In the space of a few short years, I have noticed how a certain friend of mine [a lifelong Catalan speaker], has gone from saying ‘Where are you? Shall we meet at half eight?’ [using Catalan structure] to saying ‘Where are you? Shall we meet at half eight?’ [using Castilian structure]. The correct version in Catalan is the first one. In the second, the words are Catalan but the structure, the skeleton of the language, is Castilian.

<sup>11</sup>TRANSLATION: A barbarism is a linguistic item of foreign origin that has not been assimilated into the native language and is [thus] excluded from normative use.

Catalan actually consisted of such barbarisms, for example *bueno* (instead of the normative Catalan *bé*), *pues* (Cat *llavors*) or *vale* (Cat *val*). It should be noted that half of all participants mentioned lexical interference in some form as characteristic of Barcelona language use, and thus this feature arguably merits further exploration in the upcoming fieldwork:

Observo un grau major de castellanització tant en el lèxic i la sintaxi com en la fonètica (respondent 1b).<sup>12</sup>

[Entre els joves habitualment castellanoparlants, es produeix] el fenomen de la verbalització a la catalana de paraules castellanes... per exemple, *hígat* per *fetge* (respondent 1e).<sup>13</sup>

- *‘Poverty’ of youth vocabulary.* This term describes a belief, held by older speakers, that younger members of the linguistic community use a somehow ‘less rich’ form of language. This does not necessarily mean that older people are merely referring to an absence of high-register vocabulary among youth speech; indeed, the lack of a youth Catalan slang has been cited in the questionnaires, which is potentially interesting as this is occasionally seen as a diagnostic of the early stages of language death (p.c. Christopher Pountain):

Hi ha una mena d’argot castellà que ha anat passant al català – per exemple, ja fa anys que vaig començar a sentir ‘esto raya’, quan alguna cosa és avorrida, que els adolescents catalanoparlants ja tradueixen com ‘això ratlla’. Evidentment, l’argot juvenil no és res de nou, però el que és curiós ara és la inexistència d’un argot català, que recorre per tant als calcs del castellà (respondent 1r).<sup>14</sup>

---

<sup>12</sup>TRANSLATION: I observe a great degree of Castilianisation in lexis, syntax and phonetics

<sup>13</sup>TRANSLATION: [Among the habitually Castilian-speaking youth, we’re starting to see] the production of Castilian words in a Catalan way... such as *hígat* (from Cast *higado* = Eng *liver*) instead of *fetge* (the normative Catalan word for *liver*).

<sup>14</sup>TRANSLATION: There is a type of Castilian slang which has gradually moved into Catalan – for example, I started hearing ‘that’s a drag’ (Castilian) years ago, to describe something boring, which Catalan teenagers were translating as ‘that’s a drag’ (Catalan). Of course youth slang is nothing new, but what I think is strange is the lack of a slang of Catalan origin. [The language therefore] has to resort to calques from Castilian.

It should be noted however that such a belief system is potentially difficult to quantify: how is one to judge the ‘richness’ of vocabulary of a particular age group? How can one determine how many lexical items are used by a given speaker, let alone an entire age group?

- *Emergence of neologisms and loss of old vocabulary.* These two interconnected terms are relatively transparent and refer to the common-sense notion that as new concepts are created, new words are required for them. Similarly, as old concepts diminish in usage, so does the related word. It is therefore logical to assume that youth speech might be more replete with neologisms, as they are often perceived as more technologically aware; moreover, they will have had less call to use terminology which refers to outdated concepts.

El jovent... fa un ús més gran dels neologismes (respondent 1s).<sup>15</sup>

*Pragmatic features.*

- *Youth ignorance of different registers in Catalan.* This is linked to the above ‘poverty of youth vocabulary’, but concerns pragmatic, and not lexical issues. It is based on the same ideas about youth language (that it is somehow ‘less rich’) and is similarly difficult to quantify in experimental terms.

Hi ha una tendència, en la qual m’incloc, que descuida tot tipus de formalismes (respondent 1w).<sup>16</sup>

Having conducted this pre-pilot questionnaire, the nature of BUCs in the community are considerably clearer. The practices highlighted are by and large instances of non-normative language use and, as expected, many derive from language contact between Catalan and Castilian. Interference between languages – both lexical and

---

<sup>15</sup>TRANSLATION: The youth... make greater use of neologisms.

<sup>16</sup>TRANSLATION: There is a tendency (among young people), and I include myself in this, that neglects formalities of any kind.

syntactic – is prevalent, as are difficulties with Catalan weak pronouns and general cross-linguistic phonological interference. In order for the upcoming experiment to be particularly representative of the linguistic community under examination, all of these issues are to receive attention.

### **4.1.2.3 Summary**

Taking into account all of the considerations brought up in this section, it must be made clear what the upcoming experiment needs to test and how it should go about doing so. Firstly, it must test the effectiveness of the education system, using the apparent time method as detailed in section 4.1.2.1. And secondly, it must test speaker awareness of instances of non-normative language use as discovered in the pre-pilot questionnaire in section 4.1.2.2, namely lexical and syntactic interference, and cross-linguistic phonetic and phonological interference. Based on this information, a pilot study was devised and conducted, and will now be examined in full.

## **4.2 Pilot study**

The previous section has determined that, in order to answer our contact research question, the experiment must test speaker awareness of non-normative linguistic convergence between Catalan and Castilian. In order to then comment on language policy, it will be based upon an apparent time methodology to test two age groups characterised by different language-in-education policy decisions. Thus, based on the findings from the pre-pilot questionnaire and the decision to employ an apparent-time methodology, a pilot study was formulated. The upcoming methodology is unique and was devised by the researcher. Therefore, it required some fine-tuning and refinements so that the final fieldwork experiment would be as effective as possible. In this section, we shall briefly discuss the pilot study, addressing the procedure of the experiment (section 4.2.1), then the specific stimuli used (section 4.2.2), before providing some details concerning the four participants (section 4.2.3). After this, the results of the pilot study will be briefly outlined (section 4.2.4) and then discussed (section 4.2.5).



Based on this discussion, the final experiment will be formulated (section 4.3)

### 4.2.1 Procedure

The pre-pilot questionnaire highlighted the importance of lexical and syntactic interference, phonetic and phonological interference, and further issues regarding more general language usage such as register and range of vocabulary. The pilot study wished to tackle as many of these issues as possible, and it was deemed that the best way for this to be achieved was to divide the experiment into two parts:

Part One. The first part was the reading aloud of two passages. Participants were given two passages and asked to read them aloud. One passage was written in Catalan and contained Castilianisms, while the other was written in Castilian and contained Catalanisms. Participants were required to read both passages aloud and to comment on anything they considered out of place. In doing so, they would potentially identify the instances of non-normative linguistic convergence contained therein. This identification of non-normativity is how speaker awareness is to be tested: if speakers are able to offer metacommentary on such instances, then they are (to use the Labovian terminology highlighted in chapter 1) *stereotypes*, and thus appear above the level of conscious awareness. So, participants can be judged to be aware of a particular instance of non-normative linguistic convergence if they offer commentary on it. Each text contained nine tokens of non-normative linguistic convergence, although participants were not informed of this. Additionally, there were six tokens in each text designed to elicit phonetic / phonological phenomena outlined in the pre-pilot questionnaire. For this reason, the texts had to be read aloud.

Part Two. The second part was a brief discussion based on a visual stimulus. In addition to the lexical, syntactic, phonetic and phonological phenomena addressed by part one, the pilot study wished to elicit some conversational data, in order to address the concerns of register and vocabulary mentioned above. A short (i.e. no longer than five minute) conversation was deemed to be the best means to provide such information. Further details of this stimulus and why it was chosen are given in section 4.2.2.

An obvious methodological concern was the choice of a language of administration for the experiment. The choice to deliver the experiment in Catalan or Castilian may feasibly bias the responses of the participant in the first part of the exercise. Moreover, the discussion part can only be conducted in one language or the other with a given participant, and not both: two languages would require two different stimuli images, and would result in a great deal of conscious language switching between Catalan and Castilian, which may arguably render the experiment too artificial. Such concerns cannot be sidestepped, but can be controlled for. Therefore, any introductory preamble to the experiment took place in the language of the first reading passage of part one; the participant was then shown the second reading passage and could offer metalinguistic commentary in whichever language they chose (though it was thought likely that they would instinctively switch to the language of the stimulus). In fifty percent of cases, the first stimulus passage in part one was the Catalan one, and in fifty percent Castilian, so as not to bias the results. The second stimulus passage was then presented, and the brief discussion in part two continued in the language of this stimulus passage (so as to avoid unnecessary conscious switching between Catalan and Castilian). In short, the experiment was administered as follows:

- CONDITION A (50% of cases):
  - Preamble: Castilian
  - Part one, passage one: Castilian
  - Part one, passage two: Catalan
  - Part two image discussion: Catalan
  
- CONDITION B (50% of cases):
  - Preamble: Catalan
  - Part one, passage one: Catalan
  - Part one, passage two: Castilian
  - Part two image discussion: Castilian

The above procedure aims to tackle the key concerns raised in the pre-pilot questionnaire. Part one seeks to provide information concerning lexical and syntactic borrowing, as well as phonetic and phonological interference. More information on how this is exactly achieved will be given when the stimuli themselves are examined (section 4.2.2). Part two then seeks to provide more conversational data, in order to potentially provide some tentative information regarding the points concerning vocabulary and register also mentioned in the pre-pilot questionnaire. How all of this ties in with the apparent time hypothesis will be addressed when the pilot participants are discussed (section 4.2.3). We will now proceed to a discussion of the pilot study stimuli.

## 4.2.2 Stimuli

In this section, the content of the reading passages and the conversation stimulus image will receive attention.

### 4.2.2.1 Pilot study part one: Reading passages

As stated above, the two reading passages each contained nine instances of non-normative linguistic convergence, as well as six tokens intended to elicit the phonetic / phonological phenomena outlined in the pre-pilot questionnaire. The reading passages are given in full in appendix B. This section will now discuss each of the thirty tokens (fifteen in the Catalan text, fifteen in the Castilian text) in turn.<sup>17</sup>

#### Catalan pilot reading passage: Tokens of non-normative use

1. *Baix*. This example of syntactic calquing appears in line 4 of the pilot dialogue in the phrase ‘baix la televisió’. It was used again in the ultimate fieldwork, and as such, receives full attention on page 187.

---

<sup>17</sup>Where a token that appears in the pilot study is eventually also used in the final fieldwork experiment, only a very limited description is provided here, and a reference to its full description is provided.

2. *Fa molt fred*. This potential instance of syntactic calquing appears in line 5 of the pilot dialogue. It has been argued that this is a syntactic calque based on the Castilian *hace mucho frío* (Eng *it is very cold*), since normative Catalan makes use of the preposition *de* in this circumstance (cf. El català com cal website), giving *fa molt de fred*. However, subsequent research revealed that the Institut d'Estudis Catalans dictionary (s.v. *molt*) permits both usages, and so this was ultimately excluded from the fieldwork experiment.
3. *Tiburó*. This instance of lexical borrowing occurs in line 8 of the pilot dialogue. Normative Catalan prescribes *tauró* (Eng *shark*). In this instance, the Castilian item *tiburón* has been adapted to *tiburó*, following the model of words like *informació* (Catalan) compared to *información* (Castilian) (Paloma i Sanllehí and Rico i Busquets 2008, s.v. *tiburó*, *tiburón*). The low frequency of this word in general usage led to it standing out as unnatural, and so this was ultimately excluded from the fieldwork experiment.
4. *Tonteria*. This instance of lexical borrowing occurs in line 13 of the pilot dialogue. It was retained for the final fieldwork experiment and receives full attention on page 185.
5. *Tinc 17*. This instance of syntactic interference (to use the terms in the pre-pilot questionnaire, of 'pronoun difficulties') occurs in line 14 of the pilot dialogue, and is an example of the omission of the partitive pronoun *en*, required in normative Catalan. This was used again in the ultimate fieldwork, and receives full attention on page 191.
6. *No he pensat*. The instance of syntactic interference appears in line 15 of the pilot dialogue, and once again deals with the omission of a Catalan pronoun, this time the locative *hi*. This was used again in the final fieldwork, and is fully explained on page 187.
7. *Acera*. This instance of lexical borrowing appears in line 20 of the pilot dialogue, and was retained for the final fieldwork. As such, it receives full attention on

page 184.

8. *Quebrant-se*. This instance of lexical borrowing also appears in line 20 of the pilot dialogue, and was also retained for the final fieldwork. As such, it receives full attention on page 184.
9. *Menys mal*. This instance of lexical calquing appears in line 22 of the pilot dialogue, and was also retained for the final fieldwork. As such, it receives full attention on page 186.

#### Catalan pilot reading passage: Phonetic and phonological environments

10. *Xerrem*. This example occurs in line 1 of the pilot dialogue, and tests the degree to which initial fricatives are pronounced as affricates. Upon reading this word aloud, subjects' pronunciation of the initial sound was evaluated. This initial sound can either be pronounced as the normative Catalan fricative, giving [ʃə.'rem] or as the arguably Castilianised affricate, giving [tʃə.'rem].
11. *Casa*. This example occurs in line 2 of the pilot dialogue, and tests the degree to which intervocalic -s- is voiced (as in normative Catalan) or voiceless (as in an arguably more Castilianised realisation, based on the normative Castilian pronunciation). The normative Catalan pronunciation is thus ['ka.zə], while the Castilianised variant would be closer to ['ka.sə].
12. *Xocolata*. This example occurs in line 3 of the dialogue, and as in token 10, tests the degree to which the initial sound is affricated. A normative Catalan pronunciation would be [ʃu.ku.'la.tə], whereas a more Castilianised version would be [tʃu.ku.'la.tə]
13. *Per això no fa calor*. The reading aloud of this phrase (in line 6 of the pilot dialogue) in Catalan allows us to test the mismatch between Catalan and Castilian stressed vowel systems. As already stated, the seven-vowel stressed vowel system of Catalan differentiates between open-mid and closed-mid vowels. This phrase, pronounced normatively as [pər ə'ʃɔ no fa kə'lo] contains both

the open-mid back vowel /ɔ/ and the closed-mid back vowel /o/. A pronunciation uninfluenced by Castilian would maintain the difference between these two sounds, whereas a more Castilianised pronunciation would not.

14. *Anglès és*. This appears in line 9 of the pilot dialogue. As in token 13, this examines speakers' ability to maintain a difference between open-mid and closed-mid vowels. This time, the front vowels /ɛ/ and /e/ are tested. If the two different vowel qualities are evident in the pronunciation of [əŋ.'glɛz es], then the pronunciation is arguably less influenced by Castilian (where such a contrast does not exist).
15. *Maduresa*. This appears in line 16 of the pilot dialogue. As in token 11, this tests the voicing of intervocalic -s-. A normative Catalan pronunciation would be [mə.ðu.'rɛ.zə], whereas a Castilianised variant would be closer to [mə.ðu.'rɛ.ʝə].

#### Castilian pilot reading passage: Tokens of non-normative use

1. *Tanto me da*. This example of potential calquing appears in line 2 of the pilot dialogue. The Castilian expression *me da igual* (Eng *I don't care*) is rendered in Catalan as *tant me fa*, which has arguably been calqued into Castilian as *tanto me da* (Seib 2001, 108). This construction, while not exclusive to Catalan-speaking regions, has been attributed to Catalan language contact, which would indeed explain its frequency in areas of Catalan-Castilian bilingualism. However, given its dubious status as an item having arisen from contact, it was not included in the final fieldwork study.
2. *Consciente que*. This example of syntactic calquing from Catalan into Castilian is found in line 4 of the pilot dialogue. It was ultimately used in the final experiment, and is examined in full on page 197.
3. *El Iñaki Gabilondo*. This example of syntactic calquing appears in line 5 of the pilot dialogue, and consists of using the definite article before a personal name. This was used in the final experiment, and is examined in full on page 194.

4. *Mediana*. This example of calquing appears in line 5 of the pilot dialogue. It was used in the final experiment and is thus examined in full on page 193.
5. *Habían*. This instance of potential syntactic calquing appears in line 6 of the pilot dialogue. Castilian *haber* and Catalan *haver-hi* function as impersonal verbs, expressing the notion of ‘there is/are was/were’ etc. In Castilian, this verb does not agree in number with the relevant noun, thus giving *había una chica* (Eng *there was one girl*) and *había dos chicas* (Eng *there were two girls*), and not \**habían dos chicas* (showing plural agreement). Catalan *haver-hi* shows a tendency for agreement, thus giving *hi havia una noia* (Eng *there was a girl*) and *hi havien dues noies* (Eng *there were two girls*, showing plural agreement). This example thus uses Castilian plural agreement of impersonal *haber* following this usage found in Catalan. While erroneous use of *habían* is not exclusive to Catalan-speaking areas, its relatively high frequency has been attributed to language interference from Catalan (Blas Arroyo 1991, 276-7). However, due to its widespread use outside Catalan-speaking areas, this item was not eventually included in the final fieldwork.
6. *Hagan fuera*. This calque occurs in line 9 of the pilot dialogue. It was used in the final fieldwork and is explained on page 192.
7. *Diferente*. This occurs in line 11 of the pilot dialogue and was employed in the final fieldwork. As such, it receives full attention on page 191.
8. *Piquen a la puerta*. This example of lexical borrowing appears in line 16 of the pilot dialogue and was used in the final fieldwork. It is fully explained on page 193.
9. *La sabes muy larga, tú*. This calque is found in line 18 of the pilot dialogue, and was eventually used in the final experiment. It thus receives full attention on page 7.

Castilian pilot reading passage: Phonetic and phonological environments

10. *Café*. The reading aloud of this token aims to test any mismatch between the Catalan and Castilian stressed vowel systems. The Catalan lexical item *cafè* requires an open-mid front vowel [kə.'fɛ] as opposed to the Castilian *café*, which contains a closed-mid front vowel [ka.'fe]. Thus, a Catalanised pronunciation may evidence a more open vowel in this position.
11. *Económica*. As in token 10, this token aims to test mismatch between the two vowel systems. The Catalan lexical item *econòmica* requires an open-mid back vowel in stressed position [ə.ku.'nə.mi.kə] as opposed to the Castilian *económica*, which is pronounced with a closed-mid back vowel [e.ko.'no.mi.ka]. As above, a Catalanised articulation may evidence a more open vowel in this position.
12. *Fantasía*. This instance potentially constitutes a reversal of the process to be observed in instances 11 and 15 of the Catalan pilot dialogue. The voiced apical alveolar fricative sibilant /z̞/ is not found in varieties of Castilian unaffected by Catalan (except allophonically before a voiced consonant). Thus a Catalanised pronunciation may contain this sound, as in [fan.ta.'z̞i.a], as opposed to the normative Castilian [fan.ta.'zi.a].
13. *Cosa*. As in instance 12, a Catalanised pronunciation may be ['ko.z̞a] as opposed to the standard ['ko.ɣa].
14. *Atrasado*. Again, as in instance 12, giving the arguably Catalanised [at.ra.'z̞a.ðo] instead of [at.ra.'ɣa.ðo]
15. *Alfombra*. This token also tests the extent to which participants show interference between the vowel systems, but this time in unstressed position. The three-phoneme system of vowels in unstressed position in Catalan contains the schwa, which is absent from the Castilian five-vowel system. Thus, an unstressed /a/ sound in Catalan would be rendered as /ə/. As such, the standard Castilian pronunciation of [al.'fom.bra] may be articulated as [əl.'fom.brə] by a Catalan-Castilian bilingual, using one or both schwas.



#### **4.2.2.2 Pilot study part two: Conversation stimulus image**

The stimulus image of a kitchen was chosen, and is given in appendix B. The conversation was initiated with me asking the participant to name as many items in the image as possible and to talk thereon. A vocabulary list of twenty-seven items within the image was drawn up, containing the normative terms in Catalan and Castilian. This list serves to evaluate the participants' utterances, and is a quick means of determining whether participants are consistent in the choice of language for different items. A home scene was chosen so as to provide participants with images of familiar, easily nameable objects.

#### **4.2.3 Subjects**

In keeping with the apparent-time methodology outlined in section 4.1.2.1, participants were found from two age groups. The younger group members were aged between 25 and 35 at the time of experiment. This age range was chosen since the oldest members would have been born in 1974 or 1975, and therefore underwent the majority of their education after reforms were introduced in 1983. The lower age limit was set at 25 so that participants would have long since left high school, and thus were not too recently exposed to the kind of explicit promotion of normative Catalan language use that is being tested. The older group members were aged between 55 and 65 at the time of the experiment. This age range was chosen merely to ensure that the two groups were directly comparable. All older participants would have been completely educated during Franco's regime, the medium of instruction therefore being Castilian.

Pilot study subjects were obtained through personal contacts, and consisted of two members of each age group. Participants were provided with 'personal details' forms (see appendix D) detailing their age, schooling, family background, language use and level of education. The younger group consists of a 30-year-old female (henceforth, participant P1.1) and a 25-year-old male (participant P1.2). The older group consists of a 59-year-old male (participant P2.1) and a 64-year-old male (participant P2.2).

All participants currently live in Barcelona and all had undergone both primary and secondary education in the Barcelona metropolitan area.

#### 4.2.4 Results

The following results will not be subject to a statistical analysis, since they come from only four participants. They are however given here for the sake of completeness, and will trigger some brief discussion in section 4.2.5, from where the final methodology can be formulated.

##### Part one: Reading aloud two passages

Participant code	Catalan non-norm. tokens ( /9 ) <sup>18</sup>	Catalan phono ( /6 ) <sup>19</sup>	Catalan total ( /15 )	Castilian non-norm. tokens ( /9 )	Castilian phono ( /6 )	Castilian total ( /15 )
P1.1 (younger)	7	3	10	5	5	10
P1.2 (younger)	7	5	12	6	6	12
P2.1 (older)	4	0	4	4	6	10
P2.2 (older)	4	4	8	3	5	8

Table 4.4. Pilot study results

Part two: Conversation based on image stimulus. Each participant described the stimulus image (P1.1 and P2.1 in Castilian, P1.2 and P2.2 in Catalan), mentioning at least ten items on the vocabulary list. However, three of the participants used items only from the appropriate language list, without mixing. Only one participant (P1.1) used the terms *fogones* (Eng *hob*, from Catalan *fogons*, instead of the Castilian *hornilla*) and *pica* (Eng *kitchen sink*), and the latter was quickly corrected to the standard *fregadero*. Additionally, participants occasionally misappropriated the Catalan term *rentaplats* (Castilian *friegaplatos*, Eng *dishwasher*) to refer to the kitchen sink (Cata-

<sup>18</sup>Each mark is given for identification of the token of non-normative use.

<sup>19</sup>Each mark is given for adherence to the normative pronunciation of the language in question.

lan *pica* / Castilian *fregadero*). While the use of the term *rentaplats* may result from a generalisation into Catalan of the lexical similarity between the Castilian items *friegaplatos* and *fregadero*, more data would be needed to elevate such observations beyond mere speculation. Participants' descriptions lasted on average approximately three minutes, and as such, did not provide enough data so as to be amenable to satisfactory analysis. Moreover, since participants were not conversing with a native Catalan/Castilian speaker, issues of speaker accommodation may interfere with habitual speech practices. As such, the conversations recorded have not been subjected to any empirical analysis due to these fundamental flaws in the nature of the data obtained.

#### 4.2.5 Discussion

The purpose of the pilot study was to produce a final fieldwork methodology that would accurately and concisely test the research questions at hand. Any elements within the pilot study which were judged to be extraneous or unsuccessful would be removed, thus resulting in a refined fieldwork experiment. As already stated, this methodology is unique and devised by the researcher, and so a degree of modification was required in order for the final fieldwork experiment to be as effective as possible. This discussion will determine what elements were to form part of the final fieldwork experiment, and what were to be discarded.

The results given in section 4.2.4 cannot be subjected to a full statistical analysis, since they come from only four participants. However, at first glance, younger and older participants are performing differently, particularly when asked to identify tokens of non-normative linguistic convergence. This potential difference between age groups' performance when asked to identify tokens of non-normative language use (not including the phonetic / phonological environments) may prove interesting. This can therefore remain in the ultimate experiment.

One criticism levied at the pilot study (by participants and colleagues) is that the dialogues are stilted and unnatural. Therefore, a complete rewrite of the two dialogues was necessary before they could be used in the ultimate experiment. Less-frequently

used tokens such as *tiburó* and *alfombra*, as well as dubious cases like *habían* and *tanto me da* do not appear in the final experiment stimuli. Moreover, since the final experiment is limited to just lexical and syntactic tokens, care was taken to ensure the same amount of each type of interference in the two texts. As will be shown in the final methodology, each stimulus text eventually contained ten lexical tokens and five syntactic tokens. The fact that both final stimuli texts contained the same number of tokens of each type of interference, as well as the fact that low frequency items were removed from both texts, ensured that the two were more readily comparable.

The items which sought to elicit phonetic and phonological convergence phenomena from participants also do not appear in the final experiment, for numerous reasons. On a practical level, it is hard to perceive small phonetic differences (such as vowel quality or the presence of intervocalic voicing) without specialist recording equipment. This is not only problematic for reasons of availability, but also because the experiment then would become considerably more formal and difficult to conduct. Moreover, these phonetic / phonological elicitation are the reason that the texts needed to be read aloud. This often proved complicated when participants were asked to identify non-normative language use, since when reading aloud, pilot subjects did not focus on the content of the dialogues, instead just reading automatically. Thus, when asked to error-spot, the task became stilted, with participants having to stop and start (particularly if they found an obvious error during the first read-through). Therefore, due to these numerous practical concerns, the fieldwork experiment does not contain the phonetic / phonological elicitation, and as such, participants were not required to read the texts aloud.

Finally, part two of the pilot (the brief conversation based on a stimulus image) did not feature in the ultimate experiment. This data would have required a completely different analysis, and in light of the promising nature of part one where participants were asked to identify non-normative linguistic convergence, the reason for the presence of the conversational part of the experiment was unclear. It was deemed best to focus on the most promising element of the pilot study, discarding any peripheral elements (such as the conversation), so as to provide one clear analysis for later dis-

cussion. Moreover, as already mentioned, the nature of the data obtained may not be reliable, since the researcher is not a native speaker of Catalan or Castilian, and as such, issues of accommodation may have interfered with habitual speech practices. Conversational data is undoubtedly interesting, and can form the basis of future work in this field, but its relevance to the research questions and to the other parts of the experiment is questionable. Therefore, this part was eliminated from the final study.

In short, the primary finding to take forward from the pilot study is that the ultimate experiment should consist solely of the identification of non-normative linguistic convergence. The stimuli dialogues should be reformulated, so as to appear more natural and native-like, with low-frequency vocabulary or constructions removed. Everything else is to be removed: the elicitation of phonetic / phonological phenomena (and thus the need for the dialogues to be read aloud) and the conversation based on a stimulus image. Based on these conclusions, the next section will outline the final fieldwork methodology in full.

## 4.3 Fieldwork experiment methodology

On the basis of the pilot study just outlined, a unique and innovative fieldwork methodology was subsequently formulated. This section explains the experiment by outlining the procedure (section 4.3.1), discussing the stimuli (section 4.3.2) and offering details regarding the subjects (section 4.3.3).

### 4.3.1 Procedure

The experiment must test participants' awareness of non-normative language use in both Catalan and Castilian. This will be achieved using the following unique and innovative technique. Each participant will be required to read two prepared texts which mimic colloquial spoken language: one in Catalan, one in Castilian. Within each text will have been placed fifteen instances of non-normative language use.<sup>20</sup>

---

<sup>20</sup>While this methodology is completely original and of my own creation, similar studies have been conducted which test speakers' awareness of non-normative language use as a means of gaining

The participant will be asked to read through both texts and identify anything they consider to be ‘incorrect’. As stated on page 169, speakers can be deemed to be aware of a given instance of non-normativity if they offer commentary on it. Thus, asking participants to offer commentary on instances of non-normative linguistic convergence is how this experiment tests awareness of such tokens. The exact procedure is as follows:

1. The participant will be introduced to the experiment in either Catalan or Castilian. This choice of language will bear no relation to the specific linguistic history of the participant.<sup>21</sup> This introduction will take a few minutes and will consist of me explaining the upcoming procedure, and ensuring that the participant has no doubts or queries about the experiment.
2. The participant will be handed the first text and asked to read it through and identify any ‘errors’, or items that appear strange and worthy of comment. Reading of the texts can be either out loud or silent, as the participant chooses. If the introduction phase took place in Catalan, then this first text will be in Catalan. If the introduction was in Castilian, this will be in Castilian. Both texts can be found in their entirety in Appendix C.
3. The participant will be handed the second text and asked to read it through and identify any ‘errors’. If the participant has already read the Catalan text, then this will now be the Castilian text, and vice versa.
4. The participant will then be asked to re-read the first text and identify anything they may have initially missed. They will be told that they can point out instances they may have already identified, as it is better to mention something twice than to not mention it at all for fear of repetition.

---

knowledge of the wider linguistic practices of a given community (cf. Paquot 1988, Maurais 2008 regarding Québec). Indeed, the main objective of the commission responsible for Maurais’ 2008 study was to ‘propose a way in which to tackle [issues surrounding] the quality and knowledge of the [French] language which would permit [the Office québécoise de la langue française] to reliably monitor the evolution [of questions about the linguistic situation in Québec]’ (Maurais 2008, npn, my translation).

<sup>21</sup>Moreover, care will be ensured that there is an approximate balance between the final size of the groups interviewed in each language.

5. The participant will then be asked to re-read the second text and identify anything they may have initially missed.
6. The participant will be asked to complete the personal details questionnaire, in whichever language they choose. This can be found in Appendix D.
7. The participant will be asked to sign two copies of the consent form. One will then be given to the participant for their personal records, and the other will be kept on file. This can be found in Appendix E.
8. The participant will be asked for any contacts who might be willing and able to complete the experiment.

The use of texts in both languages provides a more complete picture of the situation in Barcelona than use of just one, Catalan text. This will allow for a more thorough discussion, and for tentative conclusions to be drawn regarding the linguistic practices of both older and younger participants.

### **4.3.2 Stimuli**

The two texts were written so as to mimic spoken language, since it is in this register that contact elements most frequently occur. In formal and written registers, the use of non-normative language is arguably more salient due to its informal (and thus inappropriate) nature. As such, there were advantages and disadvantages to the decision to use written, yet informal, texts. On the one hand, the fact that the stimuli are written may make contact elements appear more salient than desirable, as participants would not be used to seeing these elements on paper. On the other hand, if the texts were spoken, practical problems arise, such as having to stop and start the tape if the participant hears something, or if they are not sure. On balance, written texts ensure a much smoother running of the experiment, and the potential increased salience of the instances of non-normative language use is hopefully negated (or at least mitigated) by the informal, colloquial nature of the dialogues. I wrote the two dialogues which were subsequently proof-read by two native Catalan speakers (for the

Catalan text) and three native Castilian monolinguals (for the Castilian text). The texts can be seen in full, along with English translations, in Appendix C.

Concerning the nature of the non-normative language use to be used in the experiment, the pre-pilot questionnaire indicated that the most salient features of bi-directional Catalan-Castilian language interference were lexical, syntactic and phonetic / phonological. However, in light of discussion based on the pilot study, this experiment has chosen not to examine phonetic and phonological traits (see section 4.2.5). The experiment thus includes only lexical and syntactic examples of non-normative language use. As stated, each text contained fifteen instances of non-normative language use. For ease of comparison, each text included ten tokens of lexical interference and five tokens of syntactic interference.

#### 4.3.2.1 The Catalan text

##### Lexical tokens

1. *Despedida*. This is a direct lexical borrowing from Castilian (Cast *fiesta de despedida*, Eng *leaving do*). The normative Catalan items are *fiesta de comiat* and *fiesta de acomiadament* (Paloma i Sanllehí and Rico i Busquets 2008, s.v. *despedida*; Corporació Catalana de Mitjans Audiovisuals 2005, s.v. *despedir-se*; Bon català website, s.v. *despedir-se*). The contact item *fiesta de despedida* is found in line 4 of the fieldwork text.
2. *Acera*. This is another example of a lexical borrowing from Castilian (Cast *acera*, Eng *pavement*). The normative Catalan item is *vorera* or *voravia* (Paloma i Sanllehí and Rico i Busquets 2008, s.v. *acera*, *cera*; Bon català website, s.v. *acera*). This appears in line 8 of the text.
3. *Quebrant-se*. This is another direct borrowing. The Castilian item *quebrar* (Eng *to break*) is rendered normatively in Catalan as *trencar* (Paloma i Sanllehí and Rico i Busquets 2008, s.v. *quebrar*; Bon català website, s.v. *quebrar*), and can be found in line 8 of the stimulus.



4. *Jerès*. This item represents an instance of orthographic interference from Castilian, which in turn results in what is normatively concerned an incorrect pronunciation. The Castilian item *jerez* (Eng *sherry*) is rendered normatively in Catalan using a word-initial *x*-, giving *xerès*. This means that the normative Catalan item is pronounced as [ʃə.'rɛs] or, in some dialects arguably already characterised phonologically by contact features, as [tʃə.'rɛs]. However, the non-normative use of word-initial *j*-, would make the first sound voiced, giving [ʒə.'rɛs] or [dʒə.'rɛs] (which is still not the same word-initial sound as Castilian, which is realised as /x/). The pronunciation of *xerès* with a voiced word-initial phoneme, has therefore potentially come through orthographic interference from Castilian (Paloma i Sanllehí and Rico i Busquets 2008, s.v. *jerés*). The item can be seen in line 10 of the dialogue.
  
5. *Ordenador*. The Castilian item *ordenador* (Eng *computer*) is translated as *ordinador* in normative Catalan. The direct transfer of Castilian *ordenador* into Catalan to convey the meaning of *computer* would only be minimally different phonetically ([ur.ðə.nə.'ðo] instead of the normative [ur.ði.nə.'ðo]), but could result in semantic ambiguity. This is because *ordenador* already exists in Catalan, meaning *that which orders / ordains* (which is also rendered in Castilian as *ordenador*, making the Castilian item polysemous). This lexical transfer has been attested (Paloma i Sanllehí and Rico i Busquets 2008, s.v. *ordenador*; El català com cal website, s.v. *ordenador*) and is shown in line 11 of the text.
  
6. *Guinyant*. The Castilian item *guiñar el ojo* (Eng *to wink*) is rendered as *fer l'ullet* or *picar l'ullet* in normative Catalan. The contact element *guinyar l'ull* is discouraged (Paloma i Sanllehí and Rico i Busquets 2008, s.v. *guinyar l'ull*) and can be seen in line 13 of the text (in its gerundive form *guinyant*).
  
7. *Tonteries*. The Castilian item *tontería* (Eng *stupid thing*) is developed from the adjective *tonto* (Eng *stupid*), which is not present in normative Catalan. Thus, *tonteria* is not admissible in normative Catalan. The *Diccionari de dubtes i barbarismes* (s.v. *tonteria*) recommends alternatives including *ximpleria* and

*poca-soltada*. Note also that the example has removed the accent from the Castilian original, in line with Catalan accent rules, and formed the plural using the Catalan *-es* ending, rather than the Castilian *-as*. This can be found in line 14 of the stimulus.

8. *Menys mal*. This calque is widely reported, and its use is often corrected by people promoting normative Catalan. The Castilian *menos mal* (Eng *thank goodness*) is normatively rendered as *encara sort* or *encara bo* in Catalan (El català com cal website, s.v. *menys mal*; Bon català website, s.v. *menys mal*) and appears in line 15 of the dialogue.
9. *M'he liat*. The Castilian verb *liarse* (Eng *to get confused*) is translated into normative Catalan using the verb *embolicar-se*, the synonymous Castilian expression *hacerse un lío* appearing as *fer-se un embolic* (Paloma i Sanllehí and Rico i Busquets 2008, s.v. *lío*). The contact Catalan elements *liar-se* and *fer-se un lío* are normatively discouraged, and the former can be seen (in the present perfect tensed form *m'he liat*) in line 24 of the text.
10. *Millons*. The Castilian *millón* (Eng *million*) is *milió* in normative Catalan. This minor orthographical difference actually represents two different phonemes. The Castilian item is pronounced with the palatal lateral /ʎ/, making the word [mi.'ʎon], while the Catalan word uses the sequence of alveolar lateral followed by high front vowel /li/, giving [mi.'ljo]. The use of a palatal lateral element in this Catalan word is thus non-normative and consequently discouraged (Paloma i Sanllehí and Rico i Busquets 2008, s.v. *milló*). The item appears in line 26 of the text, in the plural form *millons*. Through following Catalan pluralisation rules, the item is distinct from both normative Catalan *milions* and normative Castilian *millones*.

## Syntactic tokens

11. *No he pensat*. Eight respondents of the pre-pilot questionnaire cited ‘pronoun difficulties in Catalan’ as a prime example of Catalan-Castilian language interference. Responses chiefly made reference to the so-called ‘weak pronouns’ (*pronoms febles*), notably the adverbials *en* and *hi*. For a full discussion of this type of language interference, see page 163. In this specific example, found in line 7 of the stimulus text, normative Catalan requires the locative pronoun *hi*, since the complement of the verb *pensar* is introduced by the preposition *a*. The locative pronoun is absent from normative Castilian, and so it is also omitted from the contact Catalan utterance as glossed below:

Normative Catalan

no \*(hi) he pensat  
NEG pro.LOC AUX.1SG thought

Normative Castilian

no he pensado  
NEG AUX.1SG thought

Contact Catalan

no he pensat  
NEG AUX.1SG thought

‘I wasn’t thinking’

12. *Baix*. Normative Catalan dictates that *baix* can only serve as an adjective, adverb or noun (Eng *low, under*), and not as a preposition. The corresponding preposition (Eng *under(neath), beneath*) is *sota*. However, non-normative Catalan uses *baix* as a preposition. This shift in function has been noted since studies by Pompeu Fabra and attributed to syntactic interference from Castilian, modelled on the adjective / adverb *bajo* and the adverb *abajo* (Boix i Fuster and Vila i Moreno 1998, 227). An example of this non-normative usage can be found in line 10 of the dialogue. The glosses below represent this, detailing prepositional usage in the examples labelled (a) and adverbial usage in those labelled (b):

Normative Catalan

- a.     sota/\*baix la     taula  
           underneath the.F table
- b.     volar baix  
           fly.INF low

Normative Castilian

- a.     bajo            la     mesa  
           underneath the.F table
- b.     volar bajo  
           fly.INF low

Contact Catalan

- a.     baix            la     taula  
           underneath the.F table

‘Under(neath) the table’ (a) / ‘To fly low’ (b)

13. *Metge*. This lexico-syntactic calque constitutes an example of hypercorrection and consists of the semantic spread of the Catalan item *metge* across syntactic categories. The Castilian item *médico* fulfils both a nominal (Eng *doctor*) and adjectival function (Eng *medical*). In normative Catalan, the noun is *metge*, and the adjective is *mèdic*. However, this distinction is sometimes lost in contact Catalan. There is arguably a widespread belief that Catalan masculine adjectives are merely the same as their Castilian cognates without the final *-o*, modelled on such examples as Catalan *solter* and Castilian *soltero* (Eng *single*) or Catalan *fresc* and Castilian *fresco* (Eng *fresh*). Perhaps wishing to maintain the distinctiveness of Catalan, in this example speakers are aware of the Catalan noun *metge* as a translation of Castilian *médico*. They are however seemingly unaware that this translation only holds for the nominal meaning of *médico*, and not the adjectival one. Upon hearing the normatively approved adjective *mèdic*, the speaker in question presumes that this is an erroneous application of the rule of omitting word-final Castilian *-o* to obtain a Catalan adjective. Knowing that the Catalan item *metge* exists, the speaker applies it as if its meaning mapped exactly onto that of Castilian *médico*. Thus, in contact Catalan the term *metge* is generalised across both nominal and adjectival functions. The irony of this situation is that, in wishing to avoid what is believed to be

a Castilianism (i.e. that Catalan *mèdic* is merely Castilian *médico* without the final *-o*), the speaker is committing another Castilianism (i.e. the functional generalisation of *metge* to serve as both noun and adjective, based on the Castilian *médico*). This calque has been attested (cf. Paloma i Sanllehí and Rico i Busquets 2008, s.v. *mèdic*) and is found in line 17 of the stimulus text. The following glosses represent this, with adjectival usage shown in the examples labelled (a), and nominal usage in the examples labelled (b).

Normative Catalan

Normative Castilian

a. el certificat mèdic/\*metge  
the certificate medical/doctor

a. el certificado médico  
the certificate medical

b. el metge/\*mèdic és jove  
the doctor/medical be.3SG young  
Contact Catalan

b. el médico es joven  
the doctor be.3SG young

a. el certificat metge  
the certificate medical

‘The medical certificate’ (a) / ‘The doctor is young’ (b)

14. *Lis.* The Catalan system of third person direct and indirect object pronouns is different from that of Castilian. The two systems are shown below in table 4.5.

Language	Direct object (m)		Indirect object (m)	
	Singular	Plural	Singular	Plural
Catalan	EL	ELS	LI	ELS (HI)
Castilian	LO	LOS	LE	LES

Table 4.5. Third-person direct and indirect object pronouns in normative Catalan and normative Castilian

In third-person singular pronouns, there is a clear differentiation between direct and indirect object markers in both languages. However, in third-person plu-

ral marking, normative Catalan does not necessarily distinguish between direct and indirect object pronouns (while the locative *hi* can be placed after the pronoun *els* to express the notion of indirectness, this is not obligatory). Castilian however makes a compulsory distinction between the direct object marker *los* and the indirect object marker *les*. Contact varieties of Catalan have developed a third-person indirect object marker which is distinct from the direct object marker (as in Castilian), and which is modelled on the distinction made between direct and indirect objects in third-person singular Catalan pronouns. This pronoun is *lis*. As the Catalan third person direct object pronoun *el* has the indirect object counterpart of *li*, so too in contact Catalan does *els* correspond to *lis*. Not only does *lis* reflect the singular *li*, but the non-normative four-way paradigm of *el, els, li, lis* closely mirrors the normative Castilian *lo, los, le, les*. This usage is already attested (Corporació Catalana de Mitjans Audiovisuals 2005, s.v. *pronomns febles*), and can be witnessed in line 19 of the text. The supporting glosses for each language are as follows:

#### Normative Catalan

què els (hi) va dir  
 what pro.IND-OBJ pro.LOC AUX.3SG say.INF

#### Normative Castilian

qué les/\*los dijo  
 what pro.IND-OBJ/PRO.DIR-OBJ say.3SG

#### Contact Catalan

què lis va dir  
 what pro.IND-OBJ AUX.3SG say.INF

‘What did he say to them?’

15. *Té 35*. As with instance 11 (page 187), this is an instance of erasure of a weak adverbial pronoun. This time, the partitive *en* is required in normative Catalan. Since there is no equivalent partitive in Castilian, it is also omitted from the contact Catalan utterance, glossed below. For further information regarding the omission of weak adverbials from Catalan, see page 163. This token is found in line 22 of the stimulus.

Normative Catalan

quants anys té?  
 how-many years have.3SG  
 \*(en) té 35  
 pro.PART have.3SG 35

Normative Castilian

cuantos años tiene?  
 how-many years have.3SG  
 tiene 35  
 have.3SG 35

Contact Catalan

quants anys té?	té	35
how-many years have.3SG	have.3SG	35

‘How old is he? He’s 35’

#### 4.3.2.2 The Castilian text

1. *Diferente*. Castilian contains the near-synonymous items *diferente* and *distinto* (Eng *different*), which are both translated into Catalan as *diferent* (the item *distint* having largely fallen out of use, resulting in semantic convergence). Catalan-Castilian bilinguals therefore favour use of *diferente* in Castilian, even in cases where normatively *distinto* is required (Seib 2001, 96), as in line 5 of the experiment text.
2. *Plego*. The Catalan verb *plegar* (Eng *to finish work*) requires a complement when translated into Castilian: *acabar el trabajo*. There is therefore no one-word cognate in Castilian, and it is a stereotypical feature of contact Castilian to employ a lexical borrowing (Szigetvari 1994, 43-4; Seib 2001, 21-2), as in line 8 of the text.

3. *Tú misma*. The use of *¡tú mismo/a!* as an interjection (Eng *it's up to you*), is originally a Catalanism in Castilian speech, calqued on the Catalan *tu mateix/-a* (Szigetvari 1994, 39). Indeed, such interjectional usage was still not normatively recognised in Castilian at the time of this doctoral research (Real Academia Española 2001, s.v. *mismo*), with Castilian alternatives such as *jalla tú!* or *jcomo quieras!* preferred. However, this expression is becoming widespread enough that Castilian speakers from outside Catalan-speaking regions have more or less fully appropriated it in colloquial registers. Nevertheless, due to its greater degree of normative acceptability in Catalan than in Castilian, it is still a potentially interesting case worthy of inclusion, and appears in line 10 of the experiment text.
  
4. *Hicieron fuera*. This is a direct calque of a standard Catalan expression which is not found in normative Castilian. The Catalan *fer fora algú d'un lloc* (Eng *to throw somebody out of somewhere*) is translated into normative Castilian as  *echar a alguien de un lugar*. *Hacer fuera* is thus a calque of *fer fora* (Seib 2001, 96), and appears in lines 11-12 of the experiment text.
  
5. *Orejas*. Normative Castilian makes a clear distinction between *oreja* and *oído*, to refer to the outer and inner ear respectively. The distinction in Catalan between the cognates *orella* and *oïda* is a lot less clearly demarcated, with *orella* often used to cover both terms (Szigetvari 1994, 38). Therefore, contact Castilian often uses the term *oreja*, where normative Castilian would require *oído*, due to the semantic blurring present in Catalan, and this is exemplified in line 13 of the dialogue.
  
6. *Cuello*. In a near-identical case to *oreja* (above), normative Castilian makes a clear distinction between *cuello* and *garganta* to refer to the neck (outer body part) and throat (inner body part) respectively. Again, the Catalan distinction between *coll* (neck) and *gola* (throat) is less clearly defined, and *coll* is sometimes used in both cases (Szigetvari 1994, 12-3). This is shown in line 13 of the text.



7. *La sabes larga*. The Catalan idiomatic expression *saber-la (molt) llarga* (Eng *to have thoroughly prepared a story or alibi*) cannot be succinctly translated into Castilian and as such, Catalan-Castilian bilinguals are encouraged to perform a direct lexical calque, giving *saberla (muy) larga* (Seib 2001, 117). This appears in line 14 of the experiment text.
  
8. *Será*. The English verb *to be* is represented by the Castilian verbs *ser* and *estar*, and the Catalan verbs *(és)ser* and *estar*. However, the usage of Castilian *ser* does not exactly parallel Catalan *(és)ser*. Nor is *estar* always employed in the same contexts in both languages. One such semantic mismatch concerns location not qualified by an explicit time period. Normative Castilian requires the use of *estar* whereas normative Catalan uses *(és)ser* (Seib 2001, 98-100; Wheeler et al. 1999, 523). As such, the use of *ser* in Castilian in this context is deemed normatively incorrect, and is represented in line 16 of the stimulus.
  
9. *Mediana*. This example of lexical calquing is based on the Catalan item *mitjana* (Eng *average*), the corresponding normative Castilian term being *media* (Seib 2001, 103), and appears in line 25 of the text.
  
10. *Picando*. The Catalan expression *picar a la porta* (Eng *to knock at the door*) is habitually translated into Castilian as *llamar a la puerta*. While the Castilian phrase *picar a la puerta* first appeared in a *suplemento* to the RAE dictionary of 1970 (Real Academia Española 2001, s.v. *picar*), it is not widely considered as an instance of normative language use (Szigetvari 1994, 43). Indeed, the comparatively high frequency of this term in Catalan-Castilian bilingual areas potentially indicates an instance of language interference, since Catalan uses the verb *picar* in this context more often than Castilian. A simple Google search gives 29,000 results for *picar a la porta*, alongside 162,000 hits for the alternative Catalan construction *trucar a la porta*. However, in Castilian, while *picar a la puerta* gives 108,000 results, the common alternative *llamar a la puerta* gives over eight million hits. Proportionally, it would seem, the use of *picar* in this context is somewhat Catalan. This interesting case merits further investigation

and appears in line 27 of the text.

### Syntactic tokens

11. *La Meritxell*. Catalan requires the presence of a direct article before proper nouns referring to people when not in the vocative case. While this is also found in certain rural varieties of Castilian, it is not normatively prescribed as in Catalan (Seib 2001, 88). This is exemplified in line 2 of the dialogue, and the contact Castilian text (as well as translations into both normative varieties for comparison) are glossed as follows:

Normative Castilian

soy (\*la) Meritxell  
be.1SG the.F Meritxell

Normative Catalan

sóc \*(la) Meritxell  
be.1SG the.F Meritxell

Contact Castilian

soy la Meritxell  
be.1SG the.F Meritxell

‘It’s Meritxell’

12. *¿Que vamos...?* In Catalan, the unstressed conjunction *que* is often used to introduce questions. Such usage is optional, and not normatively stigmatised. However, in Castilian this question opener is not recognised, and is thus seen as feature of contact Catalan-Castilian speech (Seib 2001, 87). This can be seen in line 14 of the dialogue, and the supporting glosses are as follows:

Normative Castilian

(\*que) vamos a tomar un café?  
that go.1PL to take.INF one.M coffee

### Normative Catalan

(que) anem a prendre un cafè?  
that go.1PL to take.INF one.M coffee

### Contact Castilian

que vamos a tomar un café?  
that go.1PL to take.INF one.M coffee

‘Shall we go for a coffee?’

13. *Sobre*. This lexico-syntactic calque is an example of how certain prepositions can acquire new functions in contact varieties. In normative Castilian, *sobre* is used to represent the idea of something being *on top of* another, explicitly stated object. Thus, turning our attention to the normative Castilian examples below, the use of *sobre* in (a) is grammatical. However, when the second item (i.e. the one which the first item is on top of) is unstated, *sobre* is not appropriate, and *encima* must be used. Therefore (b) is ungrammatical and (c) is acceptable. However, in Catalan, *sobre* fulfils both of these functions (i.e. those of Castilian *sobre*). The contact Castilian example, seen in the gloss below and in line 18 of the stimulus, therefore uses the item *sobre* in the place of Castilian *encima*, following the model of normative Catalan. Moreover, such shift in the use of Castilian *sobre* has already been attested (c.f. Seib 2001, 105-6):

### Normative Castilian

- a. con una estatua de la Sagrada Familia sobre la mesa  
with one.F statue of the.F Sagrada Familia on-top the.F table
- b. \*con una estatua de la Sagrada Familia sobre  
with one.F statue of the.F Sagrada Familia on-top
- c. con una estatua de la Sagrada Familia encima  
with one.F statue of the.F Sagrada Familia on-top

Normative Catalan

amb una estàtua de la Sagrada Família sobre (la taula)  
with one.F statue of the.F Sagrada Família on-top the.F table

Contact Castilian

con una estatua de la Sagrada Familia sobre  
with one.F statue of the.F Sagrada Familia on-top

‘With a statue of the Sagrada Familia on top’

14. *Sin*. In Castilian, the preposition *sin* (Eng *without*) requires a complement, and cannot stand alone phrase-finally. The Catalan cognate *sense* does not obligatorily take a complement where one is implicit. Therefore grammatical interference would transfer the property of Catalan *sense* to occur phrase-finally to Castilian *sin*. This use is attested (Szigetvari 1994, 53) and can be witnessed in line 22 of the fieldwork text.

Normative Castilian

me ha dejado sin \*(nada)  
pro.1SG.ACC AUX.3SG left without nothing

Normative Catalan

m' ha deixat sense (res)  
pro.1SG.ACC AUX.3SG left without nothing

Contact Castilian

me ha dejado sin  
pro.1SG.ACC AUX.3SG left without

‘He left me with nothing’

15. *Consciente que*. In Castilian, the phenomenon of *dequeísmo* is widespread, where the preposition *de* is systematically placed before the conjunction *que* even where it is not required. Similarly, *queísmo* occurs when *de* is systematically omitted before *que*, even where inappropriate. These practices have been the subject of much scholarly attention for many years, and are attested in numerous forms of Castilian (cf. Serrano 1998 for the Canary Islands, Gómez Torrego 1991 for a wider view of Spain, Arjona 1979 for Mexico and Bentivoglio 1981 for Venezuela). However, in normative Catalan, the conjunction *que* can never be preceded by a preposition, and so the combination *de que* is not found. So, under influence from normative Catalan, contact Catalan-Castilian speech is argued to not only exhibit a lack of *dequeísmo* when compared to other varieties of Castilian, but a higher rate of *queísmo* (Seib 2001, 84-5). This appears in line 23 of the stimulus and the supporting glosses are shown below. Note that, due to widespread *dequeísmo*, ‘eres consciente que’ is widely heard in Castilian, but is not normatively permitted, hence the asterisk in the first example:

Normative Castilian

eres    consciente \*(de) que...  
 be.2SG aware    of    that...

Normative Catalan

ets    conscient (\*de) que...  
 be.2SG aware    of    that...

Contact Castilian

eres    consciente que...  
 be.2SG aware    that...

‘You are aware that...’

### 4.3.3 Subjects

Between April and August 2010, the experiment was administered to a total of fifty Catalan-Castilian bilingual participants. Full details are provided in section 5.1.1. Participants were found from two age groups. The younger group members were aged between 25 and 35 at the time of experiment. This age range was chosen since

the oldest members would have been born in 1974 or 1975, and therefore underwent the majority of their education after reforms were introduced in 1983. The lower age limit was set at 25 so that participants would have long since left high school, and thus were not too recently exposed to the kind of explicit promotion of normative Catalan language use that is being tested. The older group members were aged between 55 and 65 at the time of the experiment. This age range was chosen merely to ensure that the two groups were directly comparable. All participants would have been completely educated during Franco's regime, the medium of instruction therefore being Castilian. In total, thirty-two participants were from the younger group and eighteen from the older group. Chapter 5 divides the sample along other variables, providing statistics for participant sex, language dominance and language of interview, *inter alia*.

This method of finding participants obviously has its advantages and disadvantages. On the positive side, the lack of formality in recruiting participants meant that the interviews themselves were usually somewhat informal. Since the texts being evaluated were written renditions of casual speech, any undue formality in the execution of the task may have affected participant results. Furthermore, on a practical level, the fact that the two age groups were thirty years apart meant that participants could suggest their parents or children as future candidates for the experiment. Though this brings the disadvantage that the samples were arguably not entirely independent, since some members of one age group were friends or relatives of members of the other group. However, factoring in all these methodological concerns, this was a practical means for the researcher, as an outsider with very few contacts in Barcelona, to recruit over fifty participants from a cross-section of the Barcelona linguistic community. While the disadvantages of this method should not be overlooked, the positive aspects outweigh them to the degree that this data collection style was appropriate to the task.

Within the experiment, all work was completely anonymised, with each participant being accorded a code which indicated to the researcher their role within the study, as follows:

Age group	Sex	Language of interview	Group member
1 = Older	1 = Female	1 = Castilian	a = First
2 = Younger	2 = Male	2 = Catalan	b = Second c = Third etc.

Table 4.6 Coding system for experiment participants

Therefore, participant 1-2-1-c would be older (1), male (2), interviewed in Castilian (1), and the third such candidate (c). Participants were procured through personal contacts. I found the first participants who then provided the next layer of subjects and so on, until the whole network of fifty participants developed as shown in appendix F.

\* \* \*

This chapter has addressed the research questions laid out in chapters 2 and 3, and identified not only what concretely needs to be tested (based on findings from the first three chapters), but highlighted any particular practical concerns (aided by a pilot study). Then, using this information, a unique fieldwork methodology was formulated. The experiment consists of an exercise which determines subjects' abilities to identify non-normative language use in both Catalan and Castilian, in order to determine whether there are any differences between participants who were schooled through the medium of Catalan and those who underwent Castilian-medium education. This methodology was tested using a pilot study and refined to give the final experiment that was administered to fifty Catalan-Castilian bilingual participants. The next chapter will offer an in-depth statistical analysis of the experiment findings, which will then be discussed in chapter 6 in order to ultimately provide some answers to the initial research questions.

# Chapter 5

## Experiment Results

Between April and August 2010, the fieldwork experiment was administered to fifty Catalan-Castilian bilingual participants, who in this chapter have been subdivided along a number of axes, namely the variables *age group*, *sex*, *participant language dominance*, *language of interview*, *level of education* and *social class*. This chapter presents the bare results of my experiment, first giving overall participant figures before exploring the above variables in order to ascertain their relevance to the study. Then, a types and tokens based analysis will be conducted before an examination of any additional participant information as provided on the Personal Details Form. Finally, all results will be succinctly summarised in order to proceed to the discussion of my findings in chapter 6.

### 5.1 Exercise Results

This section will present the raw results of my fieldwork experiment, as categorised using the variables listed above. Firstly, section 5.1.1 will provide the overall numbers of participants, as well as the numbers in each variable subgroup, and section 5.1.2 will provide the global results for all participants, not stratified by any variable. Sections 5.1.3 to 5.1.8 will then examine the results dissected along each of the variable axes in turn, while section 5.1.9 will detail any multiple interactions between more than one variable. Finally, section 5.1.10 will synthesise and summarise these findings.



The above variables were chosen for analysis in order to understand the data as fully as possible, and to determine whether the results were at all interesting and relevant to the research questions concerning policy and contact. *Age group* is the most fundamental axis, since the sample is primarily divided between younger and older participants. If these two age groups perform differently, then (invoking the apparent time hypothesis), we will be able to go on to discuss cross-generational differences which could allow judgement to be passed on the effectiveness of language policy. This has been touched upon in chapter 4, and will form the crux of the discussion in chapter 6. We are expecting to see significant differences when examining participant age group. However, to be certain that participant age group is indeed the deciding factor, a series of other variables need to be discussed, so that their role can be determined. Ideally, they will be largely discarded, leaving only participant age group as significant. This chapter will reveal whether this is in fact the case.

The reasons for choosing each of the other individual variables, along with examples of past studies that have shown these variables to be important, will be briefly discussed at the start of each upcoming subsection on the variable in question. Many have a long history of sociolinguistic study and have been shown in the past to bear upon linguistic performance. The most obvious of these is *participant sex*: given that the sample is fairly evenly divided between men and women, it would be negligent not to examine if these groups were in fact performing differently. Similarly, *social class* has long proved important in studies of language use, and the examination of this variable also gives rise to the *level of education* axis, particularly relevant given that this study aims to examine consequences for language-in-education policy. The variable of *participant language dominance* is of specific importance to bilingual scenarios. This must be taken into account, given that the experiment is testing the ability to identify normativity in two given languages, and therefore a participant's relative competence in each of these languages may well tie in with the ultimate experiment findings. The other variable, *language of interview*, is more methodological in nature – each interview, by necessity, was conducted in either Catalan or Castilian – but must still be ruled out if we want to ultimately prove the importance of another

variable (in this case, *participant age group*). As stated, each of these variables will be discussed in turn at the start of each relevant subsection.

### 5.1.1 Participant details: Overall frequencies

Overall participant frequencies, as categorised by the aforementioned variables, are as follows:

#### Total participants

N=**50**

#### Participant age group

Younger (25-35 years old): N=**32** (64% of total)

Older (55-65 years old): N=**18** (36% of total)

#### Participant sex

Male: N=**24** (48% of total)

Female: N=**26** (52% of total)

#### Participant language dominance

Catalan: N=**37** (74% of total)

Castilian: N=**13** (26% of total)

#### Language of interview

Catalan: N=**28** (56% of total)

Castilian: N=**22** (44% of total)

#### Participant level of education

Left school at sixteen years old or earlier: N=**5** (10% of total)

Left school at eighteen years old: N=**11** (22% of total)

Gained a further vocational qualification: N=**3** (6% of total)

Completed *estudis de 1r cicle*: N=7 (14% of total)

Completed *estudis de 2n cicle*: N=21 (42% of total)<sup>1</sup>

Gained a doctorate: N=3 (6% of total)<sup>2</sup>

#### Participant social class

Upper middle class: N=18 (36% of total)

Lower middle class: N=17 (34% of total)

Working class: N=12 (24% of total)

Lower class: N=3 (6% of total)<sup>3</sup>

As outlined in chapter 4, participants were selected according to their age group in a form of stratified sampling. A balance between male and female participants was also attempted, as well as between the number of interviews conducted in each language, although these issues were not of primary importance. Perhaps unsurprisingly, no attempt was made to control for participant language dominance, as there was no way of determining this until after the interview (when an informed decision could be made about the participant in question). With all variables involving a two-way subdivision (i.e. *age group* as subdivided into *older* and *younger*), the samples are reasonably balanced, ranging from 48%/52% (sex) to 74%/26% (participant language dominance). Moreover, cross tabulations were performed on all such variables and, with one minor exception, all two-way distinctions fell within an 85/15 split.<sup>4</sup> As such it was not deemed necessary to correct for bias in any of these cases. Many of these issues will be revisited when each variable is discussed in due course.

---

<sup>1</sup>The Catalan terminology has been employed here since the Spanish and British systems do not correspond exactly and thus labels such as ‘undergraduate’ and ‘masters’ would be misleading. *Estudis de 1r cicle* indicates three years spent at university level, while *estudis de 2n cicle* represents four years. These are not equivalent to undergraduate and masters respectively, since *estudis de 1r cicle* does not equate to official accreditation in its own right. It should also be noted however that since this experiment, this terminology has been replaced by *grau* (equivalent to *estudis de 2n cicle*), with *màster* referring to a further postgraduate qualification.

<sup>2</sup>‘Doctorate’ is deemed the next level (as opposed to masters), as it is equivalent to *estudis de 3r cicle*.

<sup>3</sup>See section 5.1.8 for more information on how these labels were ascribed to participants.

<sup>4</sup>In the cross tabulation for *language of interview* against *language dominance*, it was revealed that 85.7% of participants interviewed in Catalan were Catalan-dominant.

### 5.1.2 Overall test scores

Prior to dividing my sample along the different variable axes, it will be useful to examine the overall test scores. These scores represent the mean number of instances of non-normative language use correctly identified by participants across the board, irrespective of age, sex, language dominance etc. Therefore, the maximum possible score would be fifteen for each exercise (since only fifteen instances of non-normative language were placed within each text). The results are as follows:

Lang. of exercise	N	Mean	Std. Deviation	Std. Error Mean
Catalan	50	9.84	3.303	.467
Castilian	50	7.56	2.305	.326

Equality of variances	t	df	Sig. (2-tailed)
Assumed	4.002	98	.000
Not assumed	4.002	87.569	<b>.000**<sup>5</sup></b>

Table 5.1 Overall test results. Scores of all participants and independent t-test results

The upper section of the table provides us with the mean score on each exercise, along with standard deviations and standard error means. The lower section gives the result of a t-test for *independent* samples, which compares the means of each group. Such a test shows us the overall differences between the pair of means without presuming that each exercise was taken by the same sample of participants. Independent t-tests are therefore useful for the purposes of this experiment as they provide results as to the practices of a wider group, without focusing on the specific performance of participants which, while undoubtedly useful, answers different questions, and is thus addressed with a different statistical test.<sup>6</sup>

The independent t-test results show that the difference between the means is

---

<sup>5</sup>statistically significant at the .05 level ; <sup>6</sup>statistically significant at the .01 level

<sup>6</sup>When the sample is subsequently analysed along the variable axes, dependent sampling techniques will also be employed.

statistically significant. A further test (Levene's test for equality of variances) was performed, and in this case, equality of variances cannot be assumed. The difference in means is thus significant at the  $p < .01$  level (achieving a score of .000), showing that there is a less than 1% probability that the higher score for identifying non-normative language use in Catalan occurred by chance.

It is important to mention at this point that my work makes no specific predictions regarding overall results, only those differentiated by age group. So the overall results are only of interest insofar as they provide an insight on how participants performed at the task in general. However, they do not confirm or deny any of my research questions.

It has been shown that, on average, participants identified more instances of non-normative Catalan than of non-normative Castilian, and our independent t-test demonstrates that this is significant at the  $p < .01$  level. However, this alone does not reveal much information about where exactly these differences primarily occur. Are all participants performing in the same way? If so, what trends are to be found? It is now necessary to divide my sample along the variables already outlined (*age group*, *sex* etc.) to determine what factors are most relevant in evaluating how effective participants are at identifying instances of non-normative language use in Catalan and Castilian, and thus answer my research questions.

### **5.1.3 Age group (younger / older)**

As stated, the sample is first and foremost divided into two different age groups: a younger group aged between 25 and 35, who received Catalan-medium education, and an older group, aged between 55 and 65, educated in Castilian. Given that this experiment employs an apparent-time methodology to compare these two age groups so as to potentially later draw conclusions about language-in-education policy decisions (thus answering the policy research question posed at the end of chapter 2), the variable of *participant age group* shall be examined first.

Firstly, and most importantly, I intend to compare the results of older and younger participants to ascertain if one group identifies more instances of non-normative lan-

guage use in Catalan than the other. Secondly, I intend to perform the same comparison, but examining the results of the Castilian experiment in order to offer further insight into the situation. Since statistics are employed to reject hypotheses rather than accept them (i.e. they are used to state what *isn't* happening, from which further conclusions can be drawn), two null hypotheses must be formulated which can then be dismissed (or not), based on the results. Null hypotheses should therefore cover all outcomes other than the one that the test aims to confirm; they can then potentially be rejected, confirming the only remaining possible outcome:

Null hypothesis 1 (to reject): *There will be no difference between the performances of each age group when asked to identify instances of non-normative Catalan.*

Null hypothesis 2 (to reject): *There will be no difference between the performances of each age group when asked to identify instances of non-normative Castilian.*

Independent t-tests allow us to compare the mean scores of each age group in each test (Catalan and Castilian). The results were as follows:

Lang. of exercise	Age Group	N	Mean	Std. Dev.	Std. Error Mean
Catalan	Younger	32	10.75	2.436	.431
	Older	18	8.22	4.037	.952
Castilian	Younger	32	7.72	2.218	.392
	Older	18	7.28	2.492	.587

Lang. of exercise	Equality of variances	t	df	Sig. (2-tailed)
Catalan	Assumed	2.768	48	.008
	Not assumed	2.420	24.122	<b>.023*</b>
Castilian	Assumed	.645	48	<b>.522</b>
	Not assumed	.624	32.029	.537

Table 5.2a. Test results divided by *language of exercise*, then *age group*. Scores and independent t-test results

As we can see, the difference between the means in the Catalan test is statistically significant at the  $p < .05$  level. Therefore, it can be stated that there is less than a 5% probability that the higher results obtained by the younger group came about by chance. We can therefore successfully reject null hypothesis 1, and say that **the younger age group identifies more instances of non-normative Catalan than the older group.**

A different tendency is witnessed between the means in the Castilian test. The results obtained by younger participants, while slightly higher, are not significantly different from those obtained by older participants. It is therefore not possible to reject null hypothesis 2, and as such, **age group is not a significant factor in determining participants' ability to identify non-normative Castilian.**

The data examined in table 5.2a compares the performance of two different groups of participants in the same test: for example, older versus younger participants in the Catalan exercise. However, if we rearrange the data, as in table 5.2b below, we will be able to compare the same participants under two different conditions: for example, younger participants performance in Catalan versus Castilian. This will allow us to answer different questions and pursue further tests, and gives rise to the following null hypotheses:

Null hypothesis 3 (to reject): *There will be no difference in younger participants' performance when asked to identify non-normative language use in Catalan and in Castilian.*

Null hypothesis 4 (to reject): *There will be no difference in older participants' performance when asked to identify non-normative language use in Catalan and in Castilian.*

Therefore, if the data in table 5.2a is rearranged, the following contrasts appear:

Age Group	Lang. of exercise	N	Mean	Std. Dev.	Std. Error Mean
Younger	Catalan	32	10.75	2.436	.431
	Castilian	32	7.72	2.218	.392
Older	Catalan	18	8.22	4.037	.952
	Castilian	18	7.28	2.492	.587

Age Group	Equality of variances	t	df	Sig. (2-tailed)
Younger	Assumed	5.205	62	<b>.000**</b>
	Not assumed	5.205	61.460	.000
Older	Assumed	.844	34	.404
	Not assumed	.844	28.314	<b>.405</b>

Table 5.2b. Test results divided by *age group*, then *language of exercise*. Scores and independent t-test results

The difference between the mean scores of younger participants in each exercise is statistically significant at the  $p < .01$  level, and so we can reject null hypothesis 3 and state that **the younger age group identifies more instances of non-normative Catalan than non-normative Castilian.**

However, there is no significant difference between the results of the older participants in each test. We thus fail to reject null hypothesis 4 and must determine that **older participants do not identify significantly more instances of non-normative language use in one language than the other.**

In addition to the independent sampling techniques already used, this data can be tested using *dependent sampling* methods. Each individual t-test in table 5.2b compares the *same* sample under two conditions (namely, responding to the exercise in Catalan and responding to the exercise in Castilian). A General Linear Model Repeated Measures dependent sampling test allows the comparison of these two different samples, taking into account that *within* each sample, the same members are being tested, and *between* each sample the participants are different. Such a com-



parison allows us to examine the trends of each group. This test could be performed on the data as presented in table 5.2b (as the same participants are being compared under different conditions), but not table 5.2a (where different participants are being compared under the same conditions). This is another benefit of rearranging the data: new tests can be run on the re-organised output. In short, these new tests are asking if, within the sample, participants are responding in a predictable way in one test compared to the other. And are these trends different *between* groups, so that when one compares younger participants' responses in Catalan to their responses in Castilian, a particular trend is revealed? A Repeated Measures test allows us to see whether this trend is different from that which is revealed when comparing older participants' responses in the two languages.

Test	Type III Sum Sq.	df	Mean square	F	Sig.
Sphericity assumed	25.083	1	25.083	7.872	<b>.007**</b>
Greenhouse-Geisser	25.083	1	25.083	7.872	.007
Huynh-Feldt	25.083	1	25.083	7.872	.007
Lower-bound	25.083	1	25.083	7.872	.007

Table 5.3. Interaction of *language of exercise* and *age group*. Repeated Measures test

A significant difference is witnessed in how the two groups of participants are trending. On one hand, younger people are showing a certain behavioural pattern when one compares how each subject performs on one test and then the other; while on the other hand, older people are not exhibiting the same trends. When this is viewed alongside the mean group results found in the independent sampling tests, it can be deduced that these trends are that **younger people reliably perform better in Catalan than in Castilian, while older people show no consistency in how they perform in one language relative to the other.**

The results obtained through independent sampling have determined the following regarding participant age group:

- Younger participants identify more instances of non-normative Catalan than the older group.
- There is no difference between the age groups with regard to identifying non-normative Castilian.
- Younger participants identify more instances of non-normative Catalan than non-normative Castilian.
- Older participants perform similarly when asked to identify non-normative Catalan as when asked to identify non-normative Castilian.

Results obtained through dependent sampling have backed up the final two bullet points in confirming the difference in trends between younger and older participants. However, while dividing response statistics along the axis of age group clearly yields interesting results, it is also necessary to evaluate the impact of other factors, to ascertain what roles they may play and how they may bear upon the findings. If they are seen to play no role, we can determine that age, and by extension MOI choice,<sup>7</sup> is the most important factor in the analysis of these results.

#### 5.1.4 Sex (male / female)

Variationist sociolinguistic studies and their predecessors have recognised the important role played by sex for more than a century. Although the term NORMs was formalised and popularised by Trudgill and Chambers (cf. Trudgill and Chambers 1998, 29),<sup>8</sup> the idea of examining rural, male participants as bastions of linguistic conservatism is witnessed in very early works, such as Gilliéron's *Atlas Linguistique de la France*, published between 1902 and 1910. From these earliest of works, details such as participant sex needed to be considered when devising experiment methodologies. Participant sex has been of importance from pre-Labovian works (c.f. Fischer 1958), up to more recent developments, such as Batterham's sex-differentiated findings regarding New Zealand vowel raising (Batterham 2000). It is therefore necessary

---

<sup>7</sup>See the summary on page 168 for clarification of this link.

<sup>8</sup>For more on the concept of NORMs, see page 134.

to divide my respondents into male and female subgroups in order to investigate whether different sexes are providing different results. Again, the sample will be tested using a combination of independent and dependent sampling techniques where appropriate and, as always, null hypotheses must be formulated, which our statistical tests may allow us to reject:

Null hypothesis 5 (to reject): *There is no difference between sexes' performance when asked to identify instances of non-normative Catalan.*

Null hypothesis 6 (to reject): *There is no difference between sexes' performance when asked to identify non-normative Castilian.*

Lang. of exercise	Age group	N	Mean	Std. Dev.	Std. Error Mean
Catalan	Male	24	10.17	2.929	.598
	Female	26	9.54	3.647	.715
Castilian	Male	24	7.63	2.203	.450
	Female	26	7.50	2.437	.478

Lang. of exercise	Equality of variances	t	df	Sig. (2-tailed)
Catalan	Assumed	.668	48	.507
	Not assumed	.674	47.133	<b>.504</b>
Castilian	Assumed	.190	48	<b>.850</b>
	Not assumed	.190	47.982	.850

Table 5.4a. Test results divided by *language of exercise*, then *sex*. Scores and independent t-test results

As we can see for both exercises, the difference between male and female responses is not significant. It is therefore not possible to reject either null hypothesis 5 or null hypothesis 6, and thus we must determine that **one sex does not perform better than the other in either exercise**. As before, it is possible to rearrange table 5.4a

in order to examine other comparisons and test the following null hypotheses:

Null hypothesis 7 (to reject): *Men (irrespective of age) do not identify more instances of non-normative language use in one language than in the other.*

Null hypothesis 8 (to reject): *Women (irrespective of age) do not identify more instances of non-normative language use in one language than in the other.*

Sex	Lang. of exercise	N	Mean	Std. Dev.	Std. Error Mean
Male	Catalan	24	10.17	2.929	.598
	Castilian	24	7.63	2.203	.450
Female	Catalan	26	9.54	3.647	.715
	Castilian	26	7.50	2.437	.478

Sex	Equality of variances	t	df	Sig. (2-tailed)
Male	Assumed	3.397	46	.001
	Not assumed	3.397	42.713	<b>.001**</b>
Female	Assumed	2.370	50	<b>.022*</b>
	Not assumed	2.370	43.619	.022

Table 5.4b. Test results divided by *sex*, then *language of exercise*. Scores and independent t-test results

As both of the above comparisons reveal statistically significant differences, we can successfully reject null hypotheses 7 and 8 and state that **both men and women identify more instances of non-normative language use in Catalan than in Castilian**. However, this does not reveal any information about results based on sex for two reasons. Firstly, this comparison is not *contrasting* the two different sexes, merely stating how the participants performed in each language. Secondly, this result is to be expected since overall results show that all participants identified more instances of non-normative language use in Catalan than in Castilian (see table

5.1), and so in fact, that both sexes behave this way shows that the male/female split reveals *nothing* different to that which was shown in the overall results.

As before, this data can be examined using dependent sampling methods, since the same participants are being compared under two different conditions:

Test	Type III Sum Sq.	df	Mean square	F	Sig.
Sphericity assumed	1.580	1	1.580	.430	<b>.515</b>
Greenhouse-Geisser	1.580	1	1.580	.430	.515
Huynh-Feldt	1.580	1	1.580	.430	.515
Lower-bound	1.580	1	1.580	.430	.515

Table 5.5. Interaction of *language of exercise* and *sex*. Repeated Measures test

The lack of statistical significance in the above comparison shows that there is no difference in the trends of the two groups. In other words, men and women are showing the same tendencies across the board. Both groups score significantly higher on Catalan, but other than that, very little can be said. Men do not exhibit certain patterns not shown by women, and vice versa.

The results obtained through independent sampling show the following regarding participant sex:

- There is no statistical difference between the performance of the two sexes.
- There is no difference between the performance of each individual sex and the overall performance of the sample.

Results obtained through dependent sampling back up the second bullet point and show that there is no significant difference between the trends of males and of females. As such, it can be determined that sex of participant plays no discernible role in the outcome of this experiment.

### 5.1.5 Language dominance (Catalan / Castilian)

In a bilingual community, it is safe to assume that individuals are not exposed to the two languages in exactly equal proportions. Indeed, Grosjean (2001, 10) points out that ‘the majority (of bilingual speakers) acquire their languages at various times during their lives and are rarely equally fluent in them; many speak one language less well than the other.’ The notion of individual speaker language dominance has been of key importance in the field of contact linguistics, since Weinreich’s notion of *dominance configurations* (see page 118). The Personal Details Form administered to participants contained a number of questions (listed below) pertaining to language acquisition, competence and performance:

- What language do you use with your family?
- What language do you use with friends?
- How old were you when you learned Catalan/Castilian?
- In what setting (home, school etc.) did you learn Catalan/Castilian?
- How would you evaluate your competence at speaking/reading/writing/ understanding Catalan/Castilian?

Based on the answers to these questions, each participant was ascribed a dominant language. In the majority of cases, the answers indicated a clear choice. For example, a participant might say they use one language primarily in social interactions, that they learned that language first, and that they learned it primarily at home. These answers therefore indicated the dominance of that one language, and participants occasionally even rated their competence of that same language higher than the other. However, sometimes the answers to these questions did not so clearly indicate language dominance one way or the other. In these cases, the answers to certain questions were given precedence, namely those answers pertaining to home life. For example, if a participant uses primarily Catalan with their family, if they started to learn Catalan at an earlier age than Castilian, if they learned Catalan at

home but Castilian in school, they were deemed to be Catalan-dominant. Often the most telling answer would be the ages and settings in which participants learned each language. When asked, participants would usually state that they used one language above all with their family, and that this was learned first. This language was thus deemed to be dominant. As a result of this somewhat qualitative approach to determining language dominance there were no ‘balanced bilingual’ participants, since in even the most difficult of cases, one of the answers to questions about language use in the home would indicate a dominance one way or the other.

Participants were divided into two subgroups, Catalan-dominant and Castilian-dominant, and the data was tested using both independent and dependent sampling techniques. The following null hypotheses were formulated:

Null hypothesis 9 (to reject): *There will be no difference between the performances of each language dominance group when asked to identify instances of non-normative Catalan.*

Null hypothesis 10 (to reject): *There will be no difference between the performances of each language dominance group when asked to identify instances of non-normative Castilian.*

Lang. of exercise	Dominance	N	Mean	Std. Dev.	Std. Error Mean
Catalan	Catalan	37	10.57	2.724	.448
	Castilian	13	7.77	4.003	1.110
Castilian	Catalan	37	7.57	2.128	.350
	Castilian	13	7.54	2.847	.789

Lang. of exercise	Equality of variances	t	df	Sig. (2-tailed)
Catalan	Assumed	2.806	48	<b>.007**</b>
	Not assumed	2.337	16.079	.033
Castilian	Assumed	.039	48	<b>.969</b>
	Not assumed	.034	16.959	.974

Table 5.6a. Test results divided by *language of exercise*, then *participant language dominance*. Scores and independent t-test results

In the Catalan exercise, the difference in performance of the two language dominance groups has proven significant at the  $p < .01$  level; we can thus reject null hypothesis 9 and state that **Catalan-dominant participants identify more instances of non-normative Catalan than Castilian-dominant participants**. In the Castilian exercise, there is no significant difference between groups; we thus fail to reject null hypothesis 10 and can deduce that **Castilian-dominant participants do not identify more instances of non-normative Castilian than Catalan-dominant participants**. As in tables 5.2a and 5.4a, the data in table 5.6a can be rearranged in order to test following null hypotheses:

Null hypothesis 11 (to reject): *There will be no difference in Catalan-dominant participants' performance when asked to identify non-normative language use in Catalan and in Castilian.*

Null hypothesis 12 (to reject) *There will be no difference in Castilian-dominant participants' performance when asked to identify non-normative language use in Catalan and in Castilian.*



Dominance	Lang. of exercise	N	Mean	Std. Dev.	Std. Error Mean
Catalan	Catalan	37	10.57	2.724	.448
	Castilian	37	7.57	2.128	.350
Castilian	Catalan	13	7.77	4.003	1.110
	Castilian	13	7.54	2.847	.789

Dominance	Lang. of exercise	t	df	Sig. (2-tailed)
Catalan	Assumed	5.279	72	<b>.000**</b>
	Not assumed	5.279	68.024	.000
Castilian	Assumed	.169	24	.867
	Not assumed	.169	21.664	.867

Table 5.6b. Test results divided by *participant language dominance*, then *language of exercise*. Scores and independent t-test results

The significant difference between the means for Catalan-dominant participants' performance in the two tests allows us to reject null hypothesis 11 and state that **Catalan-dominant participants identify more instances of non-normative language use in Catalan than in Castilian**. While this is of interest, it merely shows that the Catalan-dominant group are exhibiting the same tendencies as the overall sample, i.e. performing better in Catalan. The lack of significant difference between the means for Castilian-dominant participants' performance prevents us from rejecting null hypothesis 12, letting us deduce relatively little. A rejection of null hypothesis 12 would have implied that Castilian-dominant participants acted differently to the overall sample in identifying more Castilian than Catalan; this does not prove to be the case, but the roughly equivalent means, in conjunction with the failure to reject null hypothesis 12 allows us to infer that **Castilian-dominant participants do not consistently identify more instances of non-normative language use in one language than another**.

Once again, this data can be examined using dependent sampling techniques, yielding the following results:

Test	Type III Sum Sq.	df	Mean square	F	Sig.
Sphericity assumed	36.886	1	36.886	12.543	<b>.001**</b>
Greenhouse-Geisser	36.886	1	36.886	12.543	.001
Huynh-Feldt	36.886	1	36.886	12.543	.001
Lower-bound	36.886	1	36.886	12.543	.001

Table 5.7. Interaction of *language of exercise* and *participant language dominance*. Repeated Measures test

The above significant difference between language dominance groups indicates that Catalan-dominant and Castilian-dominant people are exhibiting different trends, when one compares the performance of each participant under the two conditions (namely, the exercise in Catalan and Castilian). This supports the findings in table 5.6b whereby Castilian-dominant participants are shown to be behaving differently to the overall sample by not consistently performing better on the Catalan test. This could therefore account for the significantly different patterns between the groups: in the Catalan-dominant group, the sample is consistently identifying more instances of non-normative Catalan than non-normative Castilian, while in the Castilian-dominant group, participants are behaving less predictably.

Independent sampling methods allow us to draw the following conclusions regarding participant language dominance:

- Catalan-dominant participants identify more instances of non-normative Catalan than Castilian-dominant participants.
- Castilian-dominant participants are at no such advantage with regard to non-normative Castilian, and perform similarly to Catalan-dominant participants.
- Catalan-dominant participants perform analogously to the overall sample, and

identify more instances of non-normative use in Catalan than Castilian.

- Castilian-dominant participants do not exhibit this trend and perform similarly on the two tasks.

Dependent sampling techniques support the final two bullet points in pinpointing a significant difference in trends between the two language dominance groups, confirming the importance of participant language dominance as a factor. The consequences of this will be addressed from section 5.1.10 onwards.

### **5.1.6 Language of interview (Catalan / Castilian)**

As outlined in the methodology in chapter 4, the overall exercise was administered in either Catalan or Castilian. For reasons of consistency, respondents interviewed in Catalan were asked to undertake the Catalan exercise before attempting the Castilian text, and those interviewed in Castilian were asked to start with the Castilian exercise. This division was drawn in order to ascertain whether the language of interview affected participants responses in any way, and if so, how. To this end, the following null hypothesis have been formulated:

Null hypothesis 13 (to reject): *There will be no difference between the performances of participants interviewed in Catalan and those interviewed in Castilian when asked to identify instances of non-normative Catalan.*

Null hypothesis 14 (to reject): *There will be no difference between the performances of participants interviewed in Catalan and those interviewed in Castilian when asked to identify instances of non-normative Castilian.*

Lang. of exercise	Interviewed in	N	Mean	Std. Dev.	Std. Error Mean
Catalan	Catalan	28	10.04	3.121	.590
	Castilian	22	9.59	3.581	.764
Castilian	Catalan	28	7.36	2.147	.406
	Castilian	22	7.82	2.519	.537

Lang. of exercise	Equality of variances	t	df	Sig. (2-tailed)
Catalan	Assumed	.469	48	<b>.641</b>
	Not assumed	.461	41.927	.647
Castilian	Assumed	-.698	48	<b>.488</b>
	Not assumed	-.685	41.334	.497

Table 5.8a. Test results divided by *language of exercise*, then *language of interview*. Scores and independent t-test results

The lack of significant differences in the above table leads to a failure to reject null hypotheses 13 and 14. Therefore, **it cannot be deduced that the language of interview has any discernible effect on the results of each individual exercise**. As before, the data will now be rearranged to give table 5.8b, the results of which will respond to the following null hypotheses:

Null hypothesis 15 (to reject): *There will be no difference in the performance of participants interviewed in Catalan when asked to identify non-normative language use in Catalan and in Castilian.*

Null hypothesis 16 (to reject): *There will be no difference in the performance of participants interviewed in Castilian when asked to identify non-normative language use in Catalan and in Castilian.*

Interviewed in	Lang. of exercise	N	Mean	Std. Dev	Std. Error Mean
Catalan	Catalan	28	10.04	3.121	.590
	Castilian	28	7.36	2.147	.406
Castilian	Catalan	22	9.59	3.581	.764
	Castilian	22	7.82	2.519	.537

Interviewed in	Equality of variances	t	df	Sig. (2-tailed)
Catalan	Assumed	3.742	54	<b>.000**</b>
	Not assumed	3.742	47.877	.000
Castilian	Assumed	1.899	42	<b>.064</b>
	Not assumed	1.899	37.695	.065

Table 5.8b. Test results divided by *language of interview*, then *language of exercise*. Scores and independent t-test results

The significant difference between the mean responses of participants interviewed in Catalan means that we can reject null hypothesis 15, thereby stating that **participants interviewed in Catalan identify more instances of non-normative Catalan than non-normative Castilian**. However, as witnessed in section 5.1.4, this does not offer any new insight; it merely demonstrates that participants interviewed in Catalan are behaving in the same way as the overall sample. The lack of significant difference between the mean findings from the participants interviewed in Castilian means that we cannot convincingly reject null hypothesis 16, and therefore must state that **participants interviewed in Castilian do not identify more instances of non-normative language use in one particular language**. At first glance, this may appear very interesting, as it implies that being interviewed in Castilian causes participants to diverge from the behaviour patterns of the overall sample. However, these findings should be treated with extreme caution: if one looks at the statistics, the relevant significance value is .064, and only just falls outside the necessary  $<.05$  boundary. Therefore, while we cannot say with any certainty that

participants interviewed in Castilian identify more instances of non-normative language use in Catalan than in Castilian (which is what would be expected if language of interview plays no role), there is still a 93.6% chance that this is in fact the case. Strictly speaking however, it is not possible to reject null hypothesis 16 and claim that the results of this subgroup can be predicted. As has been observed with the other factors, the results of table 5.8b can be analysed using dependent sampling techniques:

Test	Type III Sum Sq.	df	Mean square	F	Sig.
Sphericity assumed	5.055	1	5.055	1.403	<b>.242</b>
Greenhouse-Geisser	5.055	1	5.055	1.403	.242
Huynh-Feldt	5.055	1	5.055	1.403	.242
Lower-bound	5.055	1	5.055	1.403	.242

Table 5.9. Interaction of *language of exercise* and *language of interview*. Repeated Measures test.

The lack of significance after the above Repeated Measures test shows that the two groups are trending similarly, when one compares each participant's responses under the two conditions. This supports the findings obtained through independent sampling which indicate that the language of interview is not of importance (indeed, the results are similar to those seen in section 5.1.4 concerning participant sex). Moreover, the fact that the two groups are seen as exhibiting the same tendencies supports our view that, despite not being able to reject null hypothesis 16, participants interviewed in Castilian are behaving in the same way as those interviewed in Catalan, i.e. that they systematically identify more instances of non-normative Catalan than non-normative Castilian, analogously to the overall sample.

Results obtained through independent sampling reveal the following concerning language of interview:

- Language of interview has no effect on the scores obtained in either exercise when comparing the two language of interview groups against one another.

- Participants interviewed in Catalan identify more instances of non-normative Catalan than non-normative Castilian, in line with the results of the overall sample.
- These results do not prove that participants interviewed in Castilian perform in the same way as those interviewed in Catalan, but this should be treated with extreme caution.

Repeated Measures test results indicate that language of interview has no impact, and that participants in the two subgroups trend similarly, revealing that the above issue concerning participants interviewed in Castilian may not prove consequential. In short, language of interview has little to no bearing on the results, participants interviewed in one language performing similarly to those interviewed in the other.

### 5.1.7 Level of education

All of the factors evaluated thus far have resulted in a two-way split of the sample. However, certain variables do not give rise to such clear-cut distinctions and require the application of different dependent and independent sampling techniques. Participants were divided into six educational levels for further analysis:

- Participants who left education at sixteen years of age or earlier.
- Participants who left education at eighteen years of age.
- Participants who gained a further vocational qualification.
- Participants who completed *estudis de 1r cicle*.
- Participants who completed *estudis de 2n cicle*.
- Participants who gained a doctorate.

As with other factors, the aim is to ascertain the role played by level of education in participants' performance, and so relevant null hypotheses must be formulated:

Null hypothesis 17: *There will be no difference between the results of different 'level of education' groups when asked to identify instances of non-normative language use in Catalan.*

Null hypothesis 18: *There will be no difference between the results of different 'level of education' groups when asked to identify instances of non-normative language use in Castilian.*

However, since six different mean scores are being compared for each exercise (rather than just two), a t-test for independent samples is not possible, since repeated t-tests pose an increased risk of resulting in a Type I statistical error (i.e. erroneously rejecting the null hypothesis). Several different independent sampling techniques have therefore been used in conjunction when analysing factors that divide the participants into more than two subgroups. Firstly, a *one-way analysis of variance* (ANOVA) allows us to view the results and determine whether any significant differences appear between groups:



Lang. of exercise	Level of education	N	Mean	Std. Dev.	Std. Error
Catalan	</= 16 y.o.	5	4.80	3.834	1.715
	18 y.o.	11	8.45	2.734	.824
	Vocational	3	8.67	1.528	.882
	1r cicle	7	10.43	3.047	1.152
	2n cicle	21	11.33	2.309	.504
	Doctorate	3	12.67	2.082	1.202
	Total	50	9.84	3.303	.467
Castilian	</= 16 y.o.	5	5.20	2.280	1.020
	18 y.o.	11	6.91	2.468	.744
	Vocational	3	6.00	1.732	1.000
	1r cicle	7	7.86	1.952	.738
	2n cicle	21	8.43	1.912	.417
	Doctorate	3	8.67	2.887	1.667
	Total	50	7.56	2.305	.326

Lang. exercise	Group relation	Sum sq.	df	Mean sq.	F	Sig.
Catalan	Between groups	225.478	5	45.096	6.416	<b>.000**</b>
	Within groups	309.242	44	7.028		
	Total	534.720	49			
Castilian	Between groups	59.944	5	11.989	2.633	<b>.036*</b>
	Within groups	200.376	44	4.554		
	Total	260.320	49			

Table 5.10. Test results divided by *language of exercise*, then *level of education*. Scores and independent ANOVA results

The above results show us that there is a statistically significant difference between the results of the different level of education groups in each language. We can thus reject null hypotheses 17 and 18 and state that **participants with a higher**

**level of education identify more instances of non-normative language use in both languages.** This is supported by a second independent sampling method, the *Pearson correlation*. A post-hoc test was run on the above ANOVA data, which identified a correlation between the total instances of non-normative use identified and the level of education group. This correlation was significant at the  $\alpha=.01$  level for both the Catalan and Castilian exercises.<sup>9</sup> However, such findings are still somewhat vague. Where exactly can these significant differences be found? A third kind of independent sampling test, Duncan's homogeneous subsets, allows us to determine exactly where statistical significance is to be found:

Lang. of exercise	Level of education	N	Subset 1 <sup>10</sup>	Subset 2	Subset 3
Catalan	</= 16 y.o.	5	4.80		
	18 y.o.	11		8.45	
	Vocational	3		8.67	
	1r cicle	7		10.43	10.43
	2n cicle	21		11.33	11.33
	Doctorate	3			12.67
	(Sig.)		(1.000)	(.116)	(.205)
Castilian	</= 16 y.o.	5	5.20		
	Vocational	3	6.00	6.00	
	18 y.o.	11	6.91	6.91	
	1r cicle	7	7.86	7.86	
	2n cicle	21		8.43	
	Doctorate	3		8.67	
	(Sig.)		(.071)	(.077)	

Table 5.11. Test results divided by *language of exercise*, then *level of education*. Duncan's homogeneous subsets

<sup>9</sup> $\alpha$  is also a standard mathematical shorthand to refer to significance values, and will be used interchangeably with 'p' throughout the work.  $\alpha=.01$  is equivalent to  $p<.01$ .

<sup>10</sup>The three subsets represent differences that are significant at the  $\alpha=.05$  level

The above table indicates where significant differences between groups lie. If two values do not appear in the same subset at any point, the difference between them can be assumed to be significant. Therefore, the following significant differences are observed:

- In the Catalan exercise: between participants who left school at sixteen and everybody else.
- In the Catalan exercise: between participants who obtained a doctorate and all non-university attendees (i.e. those who left school at sixteen, eighteen or upon gaining a vocational qualification).
- In the Castilian exercise: between participants who left school at sixteen and those with *2n cicle* or doctorate-level qualifications.

However, due to the low number of participants in each level of education group, these fine-grained differences should be treated with some caution, particularly in the Castilian exercise, where only the subset of participants who left school at sixteen is singled out as significant.

While different independent sampling techniques need to be employed, the Repeated Measures dependent sampling method is still viable when examining participant level of education:

Test	Type III Sum Sq.	df	Mean square	F	Sig.
Sphericity-assumed	29.981	5	5.996	1.782	<b>.136</b>
Greenhouse-Geisser	29.981	5	5.996	1.782	.136
Huynh-Feldt	29.981	5	5.996	1.782	.136
Lower-bound	29.981	5	5.996	1.782	.136

Table 5.12. Interaction of *language of exercise* and *level of education*. Repeated Measures test

The lack of a statistically significant interaction between language of exercise and level of education indicates that the various subgroups are in fact showing similar trends despite obtaining different results. The following can now be said regarding participant level of education:

- There is a significant positive correlation between level of education and number of instances of non-normative language use identified. Participants that were in education for a longer period of time identify more instances of non-normative language use across the board.
- There is no significant difference in trends between the different level of education groups. Therefore, despite a correlation between level of education and number of instances of non-normative language use identified, one cannot predict participants' performance in one exercise relative to the other by simply examining their level of education.

In short, while level of education is important in determining the overall amount of non-normative language use identified by each participant, it is of limited relevance when discussing the difference between respondents' performance in Catalan and Castilian, seeing as all subgroups are exhibiting the same trends (see table 5.12).

### **5.1.8 Social class**

Participant social class is a much less tangible variable than those studied above and is subject to a degree of interpretation on the part of the researcher. In a study of linguistic variation in Philadelphia, Labov (2001, 60-1; 180-2) examined the benefits of different techniques of measuring social class, chiefly asking whether it functions as a sum of individual components (such as level of education, occupation and residence value) or if simply examining one of these factors is sufficient. The overall outcome was that, in most circumstances, a model which takes various constituent parts into account provides a more reliable account than one which merely looks at one factor. As such, I have devised a model to ascribe a given social class to each

participant based on a number of individual sub-variables, namely *participant occupation*, *participant parent occupations*, *participant level of education* and *participant area of residence*. Each participant was awarded a *social class score* out of eighty which factored in each of the above components as follows:

Component	Points value
Participant's occupation	20
Father's occupation	10
Mother's occupation	10
Area of residence	20
Level of education	20
Total	80

Table 5.13. Composition of social class score

Participant's occupation / Participant parent occupations. The UK Office for National Statistics devised its own socio-economic classification (NS-SEC) system in order to determine employment relations and conditions of occupations (Office for National Statistics website), which groups different types of employment into various 'levels', ranging from L1 (employers in large establishments) to L14 (never worked and long-term unemployed), with L15 (full-time students), L16 (occupations not stated or inadequately described) and L17 (not classifiable for other reasons) accounting for any outlying cases. Each participant's occupation was ascribed one of the seventeen levels laid out in the NS-SEC and this was converted into a score out of twenty. The same was done for the occupation of each parent, but in this case the scores were out of ten (as per the above schema). Scores were awarded as follows:

NS-SEC Value	Participant score	Parent score
L1 - L2	20	10
L3	18	9
L4 - L6	14	7
L7 - L9	10	5
L10 - L13	4	2
L14	2	1
L15 - L17	10	5

Table 5.14. Participant scores / Participant parent scores and corresponding NS-SEC values

The NS-SEC was chosen due to its highly stratified nature which allowed for a more fine-grained division of employment types. Its use invites the criticism that it departs from a British perspective of class, which may not be entirely accurate when applied to a Spanish context. Indeed, certain professions are viewed differently in the United Kingdom and in Spain in terms of prestige (for example, local government officers are more highly regarded in Spain). As far as possible, this was corrected when ascribing individual scores to particular professions. In fields which comprise a range of potential positions of differing levels of prestige (such as engineering or technical professions), roles have been interpreted as of intermediate status unless otherwise specified.

Participant level of education. A full breakdown of participant results along the axis of level of education is given in section 5.1.7. As a component of social class, each level was ascribed a score out of twenty as follows:

Level of education	Participant score
Gained a doctorate	20
Completed <i>estudis de 2n cicle</i>	16
Completed <i>estudis de 1r cicle</i>	12
Gained a further vocational qualification	10
Left school at 18	8
Left school at 16 or earlier	4

Table 5.15. Participant scores and corresponding levels of education

Participant area of residence. The *comarca* of Barcelonès is divided into fourteen administrative *districtes*. These were ranked based on the mean household per capita income, obtained from two sources. The figures for the ten *districtes* under the jurisdiction of the Ajuntament de Barcelona were gathered (Ajuntament de Barcelona website) and then integrated with those for the four outlying areas of Badalona, Santa Coloma, Sant Adrià and L'Hospitalet (IDESCAT website), giving the following data, complete with points awarded for each area:

Area of residence	Ajuntament de BCN, RFD Index (2008)	IDESCAT RFDB Index (2004)	Participant score
Sarrià-Sant Gervasi	159.25	175.812	20
Les Corts	152.06	167.874	19
Eixample	114.831	126.831	17
Gràcia	103.28	114.021	15
Sant Martí	91.97	101.535	13
Other (in Catalonia)	not present	100	12
Horta-Guinardó	88.536	99.744	11
Sants-Montjuïc	81.5	89.976	9
Sant Andreu	76.443	84.393	7
Badalona	not present	83.6	6
L'Hospitalet	not present	80.9	5
Ciutat Vella	72.775	80.34	4
Sant Adrià	not present	79.3	3
Santa Coloma	not present	75.8	2
Nou Barris	66.062	72.932	1

Table 5.16 Participant scores and corresponding area of residence, with mean household per capita income figures

Participants' scores were totalled and they were subsequently divided into different social classes, using Thompson and Hickey's (2005) five-way class distinction:



Thompson & Hickey label	Participant score
Upper class	70-80
Upper-middle class	55-69
Lower-middle class	40-54
Working class	30-39
Lower class	0-29

Table 5.17. Participant total scores and corresponding social class

As outlined in section 5.1.1, this distinction divides the sample into eighteen upper-middle class, seventeen lower-middle class, twelve working class and three lower class respondents. Given the small size of the lower class group, and the arbitrary (and therefore intrinsically flexible) nature of the group boundaries, I have decided to combine the working and lower class groups, resulting in the following, three-way distinction:

- Upper-middle class participants. N=**18**.
- Lower-middle class participants. N=**17**.
- Working/lower class participants. N=**15**.

Using these social class categories, dependent and independent sampling techniques were applied to the results in a bid to test the following null hypotheses:

Null hypothesis 19: *There will be no difference between the results of different social class groups when asked to identify instances of non-normative language use in Catalan.*

Null hypothesis 20: *There will be no difference between the results of different social class groups when asked to identify instances of non-normative language use in Castilian.*

As with level of education, the presence of three subcategories means that t-tests are inappropriate and thus the same independent sampling techniques will be employed as in section 5.1.7, namely a one-way ANOVA, Pearson correlation and Duncan's homogeneous subsets.

Lang. of exercise	Social class	N	Mean	Std. Dev.	Std. Error
Catalan	Upper-middle	18	11.56	2.479	.584
	Lower-middle	17	10.24	2.728	.662
	Working/lower	15	7.33	3.395	.877
	Total	50	9.84	3.303	.467
Castilian	Upper-middle	18	8.50	1.689	.398
	Lower-middle	17	7.65	2.370	.575
	Working/lower	15	6.33	2.440	.630
	Total	50	7.56	2.305	.326

Lang. exercise	Group relation	Sum sq.	df	Mean sq.	F	Sig.
Catalan	Between groups	149.883	2	74.942	9.153	<b>.000**</b>
	Within groups	384.837	47	8.188		
	Total	524.720	49			
Castilian	Between groups	38.604	2	19.302	4.092	<b>.023*</b>
	Within groups	221.716	47	8.188		
	Total	260.320	49			

Table 5.18. Test results divided by *language of exercise*, then *social class*. Scores and independent ANOVA results

The significant differences in table 5.18 allow us to reject null hypotheses 19 and 20; it can thus be determined that the differences in mean responses between social classes is significant for both exercises and as such **participants of a higher social class identify more instances of non-normative language use in both languages**. This is supported by a Pearson correlation which proved significant at

the  $\alpha=.01$  level for both the Catalan and Castilian exercises. As before, Duncan's homogeneous subsets pinpoint the exact location of significant differences:

Lang. of exercise	Social class	N	Subset 1	Subset 2
Catalan	Upper-middle	18		11.56
	Lower-middle	17		10.24
	Working/lower	15	7.33	
Castilian	Upper-middle	18		8.50
	Lower-middle	17	7.65	7.65
	Working/lower	15	6.33	

Table 5.19. Test results divided by *language of exercise*, then *social class*. Duncan's homogeneous subsets

It is clear that, for the Catalan exercise, the working/lower class group performs differently from the other two class groups, whose results do not significantly differ from each other. In the Castilian exercise, the only significant difference lies between the performance of the working/lower class group and the upper-middle class group.

As before, the Repeated Measures dependent sampling method is viable when testing participant social class:

Test	Type III Sum Sq.	df	Mean square	F	Sig.
Sphericity-assumed	18.509	2	9.254	2.726	<b>.076</b>
Greenhouse-Geisser	18.509	2	9.254	2.726	.076
Huynh-Feldt	18.509	2	9.254	2.726	.076
Lower-bound	18.509	2	9.254	2.726	.076

Table 5.20. Interaction of *language of exercise* and *social class*. Repeated Measures test

The lack of statistical significance indicates, as with level of education, that the

groups are showing similar trends despite obtaining different results. However, this finding merits further investigation, since there is still a relatively high (though not statistically significant) probability that there are indeed differences in group trends. In order to determine if any significant differences are to be found between participant trends, the following two strategies will be adopted:

- Social class divisions will be overlooked, the focus thereby being placed on the scores obtained by participants.
- Tests will be run on the various component subvariables of social class (participant occupation etc.).

Social class as a scalar variable. For all of the above tests, the social class score obtained allowed a participant to be assigned to a particular class group. However, if the group boundaries (which are arbitrary in nature) are disregarded, we are left with each social class score as a scalar variable. When dependent sampling techniques are applied to the now-scalar social class findings (in order to determine participant trends), the following is observed:

Test	Type III Sum Sq.	df	Mean square	F	Sig.
Sphericity-assumed	103.707	27	3.841	1.137	<b>.383</b>
Greenhouse-Geisser	103.707	27	3.841	1.137	.383
Huynh-Feldt	103.707	27	3.841	1.137	.383
Lower-bound	103.707	27	3.841	1.137	.383

Table 5.21. Interaction of *language of exercise* and *social class as a scalar variable*. Repeated Measures test

It is clear from these results that there is no significant difference between participant trends when viewing social class as a scalar variable. It is therefore possible that the original significance value (which, while high, was still not significant) arose from the imposition of arbitrary social class categories, rather than from any actual

difference in participant habits when the sample was divided along the social class axis.

Testing of social class subvariables. One must also investigate the possibility that one component of the social class variable is responsible for the .076 significance value found when comparing trends of participants in different social class groups. Repeated Measures tests applied to each subvariable yielded the following results:

Subvariable	Type III Sum Sq.	df	Mean sq.	F	Sig.
Participant occupation	31.822	3	10.607	3.337	<b>.027*</b>
Father's occupation	48.061	5	9.612	3.254	<b>.014*</b>
Mother's occupation	5.777	3	1.926	.514	<b>.675</b>
Parents' occupations	36.203	8	4.525	1.308	<b>.267</b>
Area of residence	22.779	10	2.278	.572	<b>.826</b>
Level of education	29.981	5	5.996	1.782	<b>.136</b>

Table 5.22. Interaction of *language of exercise* with *different social class subvariables*. Repeated Measures test. Sphericity-assumed results only

One can observe that the only subvariables which are demonstrating significant differences are *participant occupation* and *participant's father's occupation*, and that even the latter ceases to be significant when combined with *participant's mother's occupation* to give an overall value for *participant's parents' occupations*.<sup>11</sup> The results show that rather it may merely be the subvariable of *participant occupation* that proves significant when examining differences in trends, rather than social class variable as a whole.

<sup>11</sup>It seems counterintuitive to presume that father's occupation has an impact if mother's occupation does not. I have therefore treated this as an anomaly that may have arisen from the fact that the older group's members' mothers' occupations were very limited in range due to the changing roles of women through the twentieth century (leading to a great many participants' mothers' profession being listed as 'housewife'). This smaller range of occupations is seen in the df column, which shows that all participants' mothers' occupations were reduced to four different categories, whereas fathers' occupations were able to be subdivided into six groups. This matter will receive no more attention.

When examining the results, the following can be said about participant social class:

- There is a significant correlation between social class and number of instances of non-normative language use identified. Participants of a higher social class identify more instances of non-normative language use across the board than those of a lower social class.
- No significant difference in trends can be established between different social classes. Therefore, as with level of education, participants' performance cannot be predicted by their social class group membership.

In short, social class behaves analogously to level of education. It is important in determining the overall amount of non-normative language use identified, but all social class groups follow the overall tendency to perform at a higher level in the Catalan exercise than in the Castilian task.

### **5.1.9 Multiple interactions**

As discussed in section 5.1.1, cross tabulations were conducted on the sample when divided into all possible pairings of the six variables, and in only one case was there an imbalance of a ratio greater to 85:15. This does not however rule out potential interactions between the six different variables in their many permutations. It is helpful to be aware of these multiple interactions, since they allow for a more accurate interpretation of the findings by making the researcher aware of any potentially misleading results. If statistically significant results are obtained by dividing a sample along a particular variable axis, checking for multiple interactions allows the researcher to see if said variable is the actual reason for the finding, or if other mitigating factors are at play (such as the interference of another variable). In order to ascertain the relationship between variables, all pairs of variables underwent Pearson correlations (as shown in table 5.23), and any pairings which produced significant correlations are

discussed in more detail below:

	Age group	Sex	Lang. dominance	Lang. of interview	Level of education
Sex	×				
Lang. dominance	×	×			
Lang. of interview	×	×	✓		
Level of education	✓	×	×	×	
Social class	✓	×	×	×	✓

Table 5.23. Statistically significant Pearson correlations between pairs of variables

Age group / level of education. This Pearson correlation is significant at the  $p < .01$  level. Cross tabulations reveal that this is due to older participants with a lower level of education than younger participants. This was unavoidable in a large number of cases due to increased university access in recent years meaning that more younger people attend higher education. It should however be remembered that dependent sampling reveals that, despite more educated people obtaining higher results across the board, this does not affect their performance in one language as related to the other (i.e. all level of education groups show similar trends). Age groups however did show a significant difference in their trending patterns (see sections 5.1.7 and 5.1.3 respectively), thus proving itself to be the more informative variable along which to divide and examine the results.

Age group / social class. This correlation is significant at the  $p < .01$  level, and again refers to older participants belonging to lower social classes. This is hardly

surprising when we consider that level of education (already discussed) is a component part of my definition of social class. Moreover, when one examines the cross tabulation for age group and social class, it is feasible to ascribe any results to age group as opposed to social class: only 4 of the 15 working/lower class group were of the younger age group. In other words, while 26.7% of the working/lower class sample belong to the younger age group, this only equates to 8% of all young participants; and conversely, 73.3% of the working/lower class group are older, which equates to 22% of all older participants. Therefore, results which show that the working/lower class are patterning in a certain way may be attributable to the make-up of said group, i.e. predominantly older participants.

Language dominance / language of interview. This correlation is significant at the  $p < .05$  level, and most likely refers to the 85:15 imbalance mentioned in section 5.1.1. Unfortunately this could not be controlled for as participant language dominance was only determined after each interview. However, this interaction should not give great cause for concern, since language of interview has been proven insignificant in all tests (see section 5.1.6).

Level of education / social class. This correlation is significant at the  $p < .01$  level, but is unsurprising since level of education forms part of my definition of social class.

These multiple interactions have revealed one important finding. Potentially, a lot of results observed when analysing the sample along the level of education or social class variable axes might actually be attributable to participant age group. Findings that show differences between level of education or social class groups may be attributable to an imbalance in the age of participants. While this is a factor to bear in mind, it should not diminish the significance of findings for age group, since these emerged as significant across the board (unlike the analyses for level of education and social class).



### 5.1.10 Summary

Most importantly for the research questions posed in chapters 2 and 3, *participant age group* proved of vital importance: younger respondents consistently identify more instances of non-normative Catalan than older participants, while there is no difference between the performance of the two age groups when asked to pinpoint non-normative Castilian usage. *Participant language dominance* was also revealed to be significant: Catalan-dominant participants identified more instances of non-normative Catalan than Castilian-dominant participants, whereas Castilian-dominant subjects were not at any such advantage in the Castilian exercise, with both language dominance groups performing similarly when asked to identify non-normative Castilian usage. There is a positive correlation between *participant level of education* and overall response totals, insofar as those participants who remained in education for a longer period of time identified more instances of non-normative language use in both exercises than those who left the school system earlier in life. However, participants of all levels of education exhibited the same tendencies (namely, the identification of more instances of non-normative language use in Catalan than in Castilian). This also proved to be the case when examining *participant social class*, i.e. those of a higher social class obtained the highest overall scores, but all class groups performed similarly when comparing the results of the two exercises. Finally, the variables of *participant sex* and *language of interview* were dismissed outright, as divisions along these axes yielded no statistically significant results.

In short, as predicted, the variable which yielded the most significant, pertinent and enlightening results was **participant age group**.

## 5.2 Types and tokens

Each exercise contained fifteen tokens of non-normative language use, which could be subdivided into two types: *lexical* and *syntactic*. A micro-analysis of participant practices comparing the rates of identification of different types and tokens will offer

more insight as to which instances of non-normative language use are perceived as more egregious by native speakers. This may allow us some more fine-grained insight into the nature of contact elements in this linguistic community, which is of particular relevance given the thesis' focus on language contact between Catalan and Castilian. As such, this section will first examine the different *types* of non-normative language use, then the different individual *tokens*. Finally, a brief analysis will be conducted of any 'misidentified' tokens, that is to say tokens that participants erroneously deemed non-normative, to see if this provides any further insight.

### 5.2.1 Type-by-type analysis

In order to investigate the possibility that the type of instance of non-normative language use (i.e. lexical or syntactic) affects its salience to native speakers, it was necessary to conduct independent and dependent sampling tests on the different types. Thus firstly, the overall mean results for a given language were compared to the scaled mean results for the lexical tokens; then, the overall scores were compared to the scaled scores for the syntactic tokens; finally, the scaled scores for the lexical and syntactic tokens were compared against each other. These tests were performed for each of the six variables explored in sections 5.1.3 to 5.1.8. The results showed few differences between lexical and syntactic tokens.

Age group. Independent t-tests comparing the various types of instances of non-normative language use in each language reveal little of significance. In both languages, both age groups' performance remains constant across all types, with no significant differences being found in any of the three comparisons (i.e. overall/lexis, overall/syntax, lexis/syntax). Moreover, if we solely examine either lexis or syntax when comparing younger and older participants' performance in each exercise, the results reflect those in section 5.1.3, with a significant difference between age groups' performance in Catalan lexis and Catalan syntax, but not in Castilian lexis or Castilian syntax. Dependent Repeated Measures tests were also performed on the results for each type in order to determine if there is still an interaction between language of exercise and age group (as witnessed in table 5.3). When comparing each group

while examining Catalan lexis and then Castilian lexis, similar results are found to those in the Repeated Measures test for the overall sample, namely that there is a significant difference in trends between the two age groups. When only examining Catalan syntax and Castilian syntax, the result does not come out as statistically significant at the  $\alpha=.05$  level (sig. = .071). This however is likely due to the fact that there are only five syntactic tokens in each exercise (as opposed to fifteen tokens in each text overall), leading to a low mean square value which makes it more difficult to ascertain statistical significance due to insufficient tokens per participant. The result therefore presumably does not indicate a cross-generational development in participant trends when only looking at syntactic tokens. In short, it can be presumed that, when dividing the results by age group, there is no difference between participant identification of lexical and syntactic tokens.

Sex. T-tests conducted on the different types in each exercise reveals no significant difference between the two sexes' performance, analogous to the overall results found in tables 5.4a and 5.4b. Similarly, dependent sampling tests performed on each of the types revealed no significant difference between the trends of the two sexes in one language relative to the other, as in the overall results in table 5.5. It can therefore be presumed that the type of instance of non-normative language use is not exerting any influence.

Language dominance. No significant difference was found when comparing the different types with the overall findings for either language, with results reflecting those found in tables 5.6a and 5.6b. However, a potentially interesting finding arose when comparing the two different language dominance groups' performance against each other in the two languages. Table 5.6a shows that Catalan-dominant participants perform better in the Catalan task than Castilian-dominant participants. The type-by-type analysis reveals that this is only statistically significant when examining syntax:

Type	Dominance	N	Mean	Std. Dev.	Std. Error Mean
Catalan lexis	Catalan	37	6.81	2.093	.344
	Castilian	13	5.85	2.544	.706
Catalan syntax	Catalan	37	3.76	1.164	.191
	Castilian	13	1.92	1.656	.459

Type	Equality of variances	t	df	Sig. (2-tailed)
Catalan lexis	Assumed	1.351	48	<b>.183</b>
	Not assumed	1.229	18.043	.235
Catalan syntax	Assumed	4.358	48	<b>.000**</b>
	Not assumed	3.684	16.365	.002

Table 5.24. Catalan exercise test results divided by *type*, then *participant language dominance*. Scores and independent t-test results

However, this finding could be spurious, as it is based on only five syntactic tokens per participant, and requires more thorough investigation before a fully informed conclusion can be reached. Dependent sampling techniques reveal a significant difference in trends between language dominance groups for both lexis and syntax in the two languages. This reveals nothing new, as this is also the case for the overall results (see table 5.7). Again, the type breakdown has revealed little about the results, except for the possible finding that Catalan-dominant participants might only be more adept at identifying non-normative use in Catalan than Castilian-dominant participants, if said use is syntactic in nature; but more research is needed to confirm such a supposition.

Language of interview. A type-by-type analysis continues to prove that there is no significant difference between the two language of interview groups. This is borne out by both dependent and independent sampling methods.

Level of education. As demonstrated in section 5.1.7, independent sampling techniques show that participants with a higher level of education identify more instances

of non-normative language use in both exercises. This is seen when examining Catalan lexis, Catalan syntax and Castilian lexis. An ANOVA which solely looks at participants' identification of Castilian syntax does not however yield statistically significant results (sig. = .231), implying that it is not the case that those with a higher level of education identify more instances of non-normative Castilian syntax. This could be due to the relative lack of tokens (only five per exercise). Nevertheless, this finding is potentially interesting, but would require further investigation before any definitive conclusions could be drawn. Dependent sampling techniques performed on the different types support the findings in section 5.1.7, namely that there is no significant difference between the performance of each level of education group.

Social class. Interestingly, an independently sampled type-by-type analysis which takes into account the variable of social class yields similar results to the above level of education findings. ANOVAs examining Catalan lexis, Catalan syntax and Castilian lexis reflected the overall findings in section 5.1.8, while Castilian syntax results did not prove significant (sig. = .104). As before, this could be due to a lack of tokens in the exercise, but nevertheless merits further attention. Dependent sampling techniques performed on the various types also support the findings in section 5.1.8, indicating that there is no significant difference between the performance of each social class group in either exercise.

In short, a type-by-type analysis has revealed relatively little about the salience of different types of non-normative language use. It may be the case that instances of non-normative Castilian syntax are the least readily identified by native speakers, although if this were so, one might expect the findings to become apparent when examining all variables, not just level of education and social class. Therefore, there is a chance that Castilian syntax is less salient, but more testing is needed before any definitive conclusions can be drawn.

### 5.2.2 Token-by-token analysis

As stated in chapter 4, each exercise contained fifteen tokens of non-normative use. A token-by-token analysis of the data will provide considerable fine-grained detail, and reveal which tokens are more salient to native speakers (i.e. which are most frequently identified). Moreover, the rates of identification of each token may reveal what instances of non-normative use are considered as *stereotypes* or *markers* by each generation, as well as shedding light on which tokens can be considered as *indicators*. A high rate of identification for a given token would provide for a plausible argument for its categorisation as a Labovian *stereotype*, since it carries a high degree of conscious awareness on the part of speakers. A mid-range level of identification may imply that a token is a *marker*, while a low level of awareness of the non-normative status of a given token would correspond to an *indicator*.<sup>12</sup> Before examining the data in terms of the six variables, the following table provides the total rates of identification of each token, ranked by frequency:

---

<sup>12</sup>For more information on stereotypes, markers and indicators, see Labov (1972) or page 51 of the present work.

Catalan token (and number) <sup>13</sup>	Identified (/50)	Castilian token (and number)	Identified (/50)
(3) <i>quebrant-se</i>	44	(2) <i>plego</i>	49
(1) <i>despedida</i>	42	(13) <i>sobre</i>	47
(2) <i>acera</i>	42	(10) <i>picando</i>	42
(12) <i>baix</i>	40	(4) <i>hicieron fuera</i>	39
(11) <i>no he pensat</i>	39	(9) <i>mediana</i>	39
(4) <i>jerès</i>	37	(7) <i>la sabes larga</i>	37
(9) <i>m'he liat</i>	37	(11) <i>la Meritxell</i>	35
(14) <i>lis</i>	36	(14) <i>sin</i>	22
(6) <i>guinyant</i>	33	(12) <i>¿que vamos...?</i>	20
(5) <i>ordenador</i>	28	(8) <i>será</i>	18
(10) <i>millons</i>	27	(6) <i>cuello</i>	12
(13) <i>metge</i>	26	(5) <i>orejas</i>	10
(15) <i>té 35</i>	23	(15) <i>consciente que</i>	5
(7) <i>tonteries</i>	20	(3) <i>tú misma</i>	2
(8) <i>menys mal</i>	18	(1) <i>diferente</i>	1

Table 5.25. Frequency of identification for each token

At first glance, the tables show that all tokens were identified at least once, and that the most frequently identified token was spotted forty-nine times. All Catalan tokens show a considerable degree of awareness among both age groups, and so the status of any particular one as a Labovian indicator is doubtful. The majority could arguably be considered markers, with token 3 (*quebrant-se una cama*) as a potential candidate for the status of stereotype. The Castilian exercise contained the most- and least-frequently identified tokens, thereby demonstrating a more erratic pattern of performance. Two obvious potential Castilian-language stereotypes are tokens 2

<sup>13</sup>These are the numbers assigned to each token, as outlined in chapter 4. They will be used in the upcoming discussion and tables.

(*yo plego a las siete*) and 13 (*que tiene una estatua de la Sagrada Familia sobre*). The Castilian text also contains potential Labovian indicators in tokens 1 (*muy diferente de mi vida de hace dos años*), 3 (*tú misma*) and 15 (*eres consciente que*), due to low levels of identification by participants.

For the present purposes, descriptive statistics suffice to provide information as to the rate of identification of each token, and so a cross tabulation was performed on every token against each of the six variables in turn. The results were then converted into percentages of the subgroups in question (e.g. older / younger) and plotted on a clustered bar chart. Since age group is the variable which is most relevant when considering my research questions, only those charts have been included. However, results found when conducting analyses on other variables will also receive limited attention.

Age group. In almost all tokens within the Catalan exercise, the pattern observed in section 5.1.3 is adhered to, namely that older people are performing at a slightly lower rate than younger people (see figure 5.1). The only slight anomalies occur with token 6 (*guinyant l'ull*) and token 13 (*el certificat metge*). *Guinyant l'ull* was identified by proportionally more older participants than younger participants, while *el certificat metge* offers a greatly exaggerated reflection of the overall tendencies, with 68.8% of younger participants identifying the token, as opposed to only 22.2% of older participants. Moreover, an increased awareness of token 14 (*què lis va dir?*) on the part of the younger generation could signify its potential passage from marker to stereotype. In the Castilian exercise, much more erratic patterning is witnessed (see figure 5.2). While all these results balance each other out to give equivalent levels of identification for each age group (see section 5.1.3), rather than seeing similar performance across the board, it is clear that some tokens are identified by almost all participants, some by one age group more readily than the other, and others are hardly identified at all. Particularly interesting are tokens 7 (*la sabes muy larga, tú*) and 14 (*me ha dejado sin*). *La sabes muy larga, tú* was identified primarily by younger participants, while *me ha dejado sin* was more salient to older participants. In short, the token-by-token analysis, when cross tabulated with the participant age



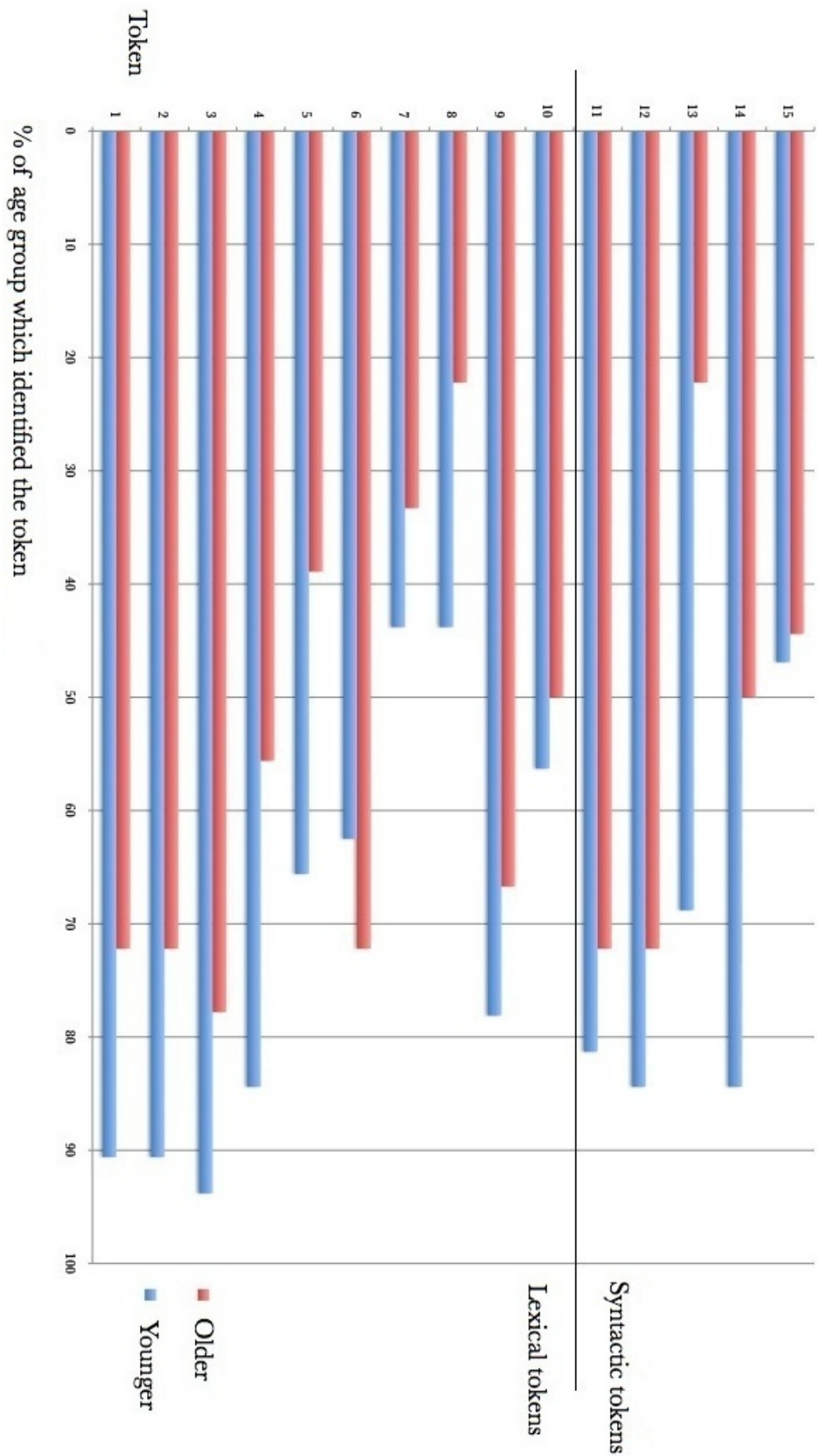


Figure 5.1: Token-by-token analysis : Catalan exercise divided by *age group*

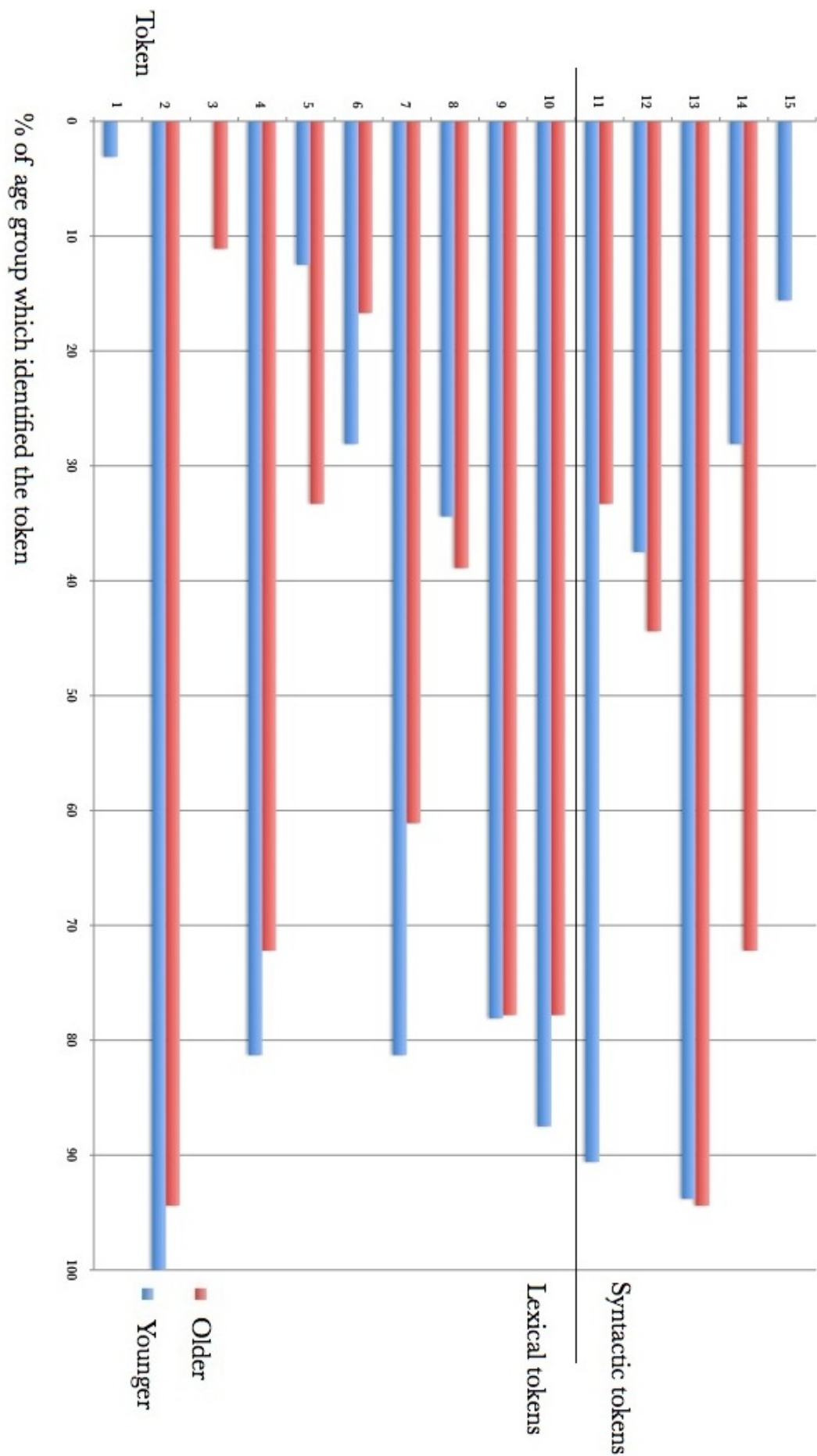


Figure 5.2: Token-by-token analysis : Castilian exercise divided by *age group*

group variable shows that in the Catalan text, tokens were identified more frequently by younger participants across the board, with no serious deviation from the patterns alluded to in section 5.1.3. In the Castilian exercise, the lack of significant difference between the two age groups is brought about by erratic performance, with different age groups identifying different tokens more frequently.

Sex. The results presented in section 5.1.4 indicated that sex of participant played no significant predictive role in the experiment, and the token analysis supports this finding. In the Catalan exercise, men and women perform at a similar level for each token, with only slight anomalies found in tokens 4 (*jerès*) and 13 (*el certificat metge*) which were identified by notably more male than female participants. In the Castilian exercise, again both sexes performed similarly, with the slight anomalies of tokens 5 (*orejas*) and 6 (*cuello*), which were identified by more women than men. There is however no statistical evidence to support a hypothesis that certain tokens are more salient to one sex than another, and as such it can be determined that sex still plays no role in the identification of non-normative language use when examining the data token by token.

Language dominance. As predicted by the results in section 5.1.5, in the Catalan exercise, Catalan-dominant participants more frequently identify each token. This is the case in all instances except token 6 (*guinyant*), providing an interesting anomaly when viewed in conjunction with the token-by-token analysis for age group. In the Castilian exercise, both language dominance groups perform at a similar level for all tokens. The only anomalies are to be found in tokens 8 (*será*, identified by more Castilian-dominant participants) and 11 (*la Meritxell*, identified by more Catalan-dominant participants), which cancel each other out. The token-by-token analysis therefore supports the findings in section 5.1.5.

Language of interview. In the Catalan exercise, the two language of interview groups are performing at similar levels across all tokens. In the Castilian exercise, similar trends are observed, with the slight exceptions that tokens 4 (*nos hicieron fuera*), 5 (*orejas*) and 6 (*cuello*) were identified more frequently by those interviewed in Castilian. However, the overall similar performance of the two groups supports the

findings in section 5.1.6 that language of interview has no bearing upon results.

Level of education. The multitude of educational levels specified in the data make the reading of a standard token-by-token analysis somewhat difficult, particularly in light of the small number of participants in four of the groups. If one disregards these results and focuses on the two most sizeable groups (*left school at eighteen years of age* with eleven participants and *completed estudios de 2n cicle* with twenty-one participants), one can observe the same trends as exposed in section 5.1.7, namely that participants with a higher level of education identify more tokens in both exercises. In the Catalan exercise this is the case for all tokens, while the only notable exception in the Castilian exercise is token 5 (*orejas*). As such, very few tokens imply a trend different to that revealed by the dependent sampling techniques in section 5.1.7.

Social class. In the Catalan exercise, most tokens follow the trends laid out in section 5.1.8, namely that those of a higher social class identify more instances of non-normative use overall. Slight exceptions are to be found with tokens 3 (*quebrant-se una cama*, where all social classes performed similarly), 7 (*tonteries*, where lower-middle classes identified the token most frequently) and 9 (*m'he liat*, where all social classes performed similarly). As ever, the Castilian exercise displayed a slightly more erratic pattern of performance. Disregarding the three outlying tokens (1, 3 and 15), five tokens displayed variations on the established social class trends. Tokens 4 (*nos hicieron fuera*), 12 (*¿que vamos...?*), 13 (*que tiene una estatua de la Sagrada Familia sobre*) and 14 (*me ha dejado sin*) showed all classes performing similarly, while token 5 (*orejas*) strikingly showed an exact reversal in the trend, with working/lower class participants identifying the token most frequently, followed by lower-middle-class, then upper-middle class participants. In short, when examining almost all individual Catalan tokens, the patterns were a microcosm of the overall findings identified in section 5.1.8, with those of a higher social class identifying each token at a higher rate. However, in the Castilian exercise, individual tokens revealed a less predictable pattern, that nevertheless balanced out to give the findings observed in the overall analysis.

In summary, the token-by-token results offer useful insight into participants' performance. The overall results have shown which individual tokens are more and less salient to the participants, while the results broken down by the different variables provide details as to the practices of a particular age group, sex, language dominance group etc. The token analysis cross tabulated with the six variables not only confirms the overall trends revealed in section 5.1, but identified the following important nuances to participant performance:

- The Catalan exercise was performed consistently across all tokens in a way which directly reflects the results in section 5.1. For example, younger participants identified more instances of non-normative language use than older participants. This is reflected by younger participants more frequently identifying each token in a consistent way across the whole exercise.
- The Castilian exercise was not performed consistently or predictably across all tokens. For example, younger and older participants showed no significant difference in their identification of instances of non-normative language use. This was not reflected by a fifty-fifty split between the groups within each token, but rather by a series of erratic results which balanced out when viewing the mean totals for the exercises. So, certain tokens would be predominantly identified by younger participants, and others by older participants, to a degree that eventually resulted in similar average results for each group, yielding no significant difference.

This important distinction not only serves to confirm the findings in section 5.1, but also to provide more fine-grained detail with regard to how participants perform when asked to identify non-normative language use in Catalan and Castilian.

### **5.2.3 Misidentified tokens**

Another point worthy of analysis is that, in being asked to identify non-normative language use, participants occasionally misidentified certain elements as non-normative,

whereas in fact this was not necessarily the case. Indeed, all fifty participants misidentified at least one token within the two exercises as non-normative. Here, we shall very briefly examine the most common ‘false negatives’, and enter into a short discussion on each one.<sup>14</sup> The two most common misidentified tokens in the Catalan and Castilian exercises were selected from the data, and are as follows:

Catalan exercise: ...*deu tenir com uns 45 anys*... This use of *com* with the meaning of ‘approximately’ (line 25), much as in Castilian *como* and English *like*, was viewed by participants as a foreign intrusion into Catalan. Fourteen participants identified this token, ten of whom were from the younger age group. The non-normative status of this item is questionable, since it is not highlighted as non-standard in the works consulted. However, this could well constitute a foreign intrusion from languages such as English, where *like* is witnessed with great frequency as a discourse marker.<sup>15</sup> People may therefore be showing awareness of potential non-normativity coming not only from Castilian, but also from the increasingly present English language. However, a great deal more research on the topic would be required before any claims could be made, and such arguments remain highly speculative.

Catalan exercise: ... *i això?* Twelve participants,<sup>16</sup> only two of whom came from the older group, considered the question *i això?* (line 3) to be a calque of the Castilian *¿y eso?* (Eng *how come?*). Again, the non-normativity of this item is dubious. It is interesting to note that only one participant suggested a more ‘Catalan’ alternative, in the form of *i doncs?* We can see that, in both of the examples from the Catalan text, the vast majority of participants evidencing this form of hypercorrection are younger. This could be tied in with the Catalan school system, and the conclusions

---

<sup>14</sup>This is undertaken here, and not in the upcoming discussion chapter, since this does not directly contribute to the arguments which answer the research questions. This work however could form the basis of a more detailed analysis of related projects in the future.

<sup>15</sup>For work on English *like*, cf. for example Tagliamonte 2005.

<sup>16</sup>Note the low rate of misidentification for both of the Catalan exercise tokens. With just fourteen and twelve participants (mis)identifying them, they were highlighted less frequently than even the least-regularly spotted genuine non-normative Catalan token (*menys mal*, identified eighteen times, see table 5.25).

that will be drawn in chapter 6, but this would have to be substantially developed and tested before any further claims could be made. Once again, any observations concerning misidentification of tokens must remain speculative.

Castilian exercise: ...*mañana me viene muy bien...* Twenty-two participants deemed this construction worthy of comment (line 7, Eng *tomorrow's fine by me*). Twenty participants (fourteen younger, six older) preferred the verb *ir* (Eng *to go*) to the verb *venir* (Eng *to come*), thus favouring *mañana me va muy bien*. Two other participants (one younger, one older) disliked the choice of verb tense, suggesting the future tense, giving *mañana me vendrá muy bien*.

The favouring of *ir* is likely influenced by Catalan, which shows a strong preference for *em va bé* (using *anar*, Eng *to go*) in this context. A simple Google search of Catalan-language documents gives 910,000 results for *em va bé*, while *em vé bé* (using *venir*, Eng *to come*) gives only 1,920 results. Castilian also shows a preference for the verb ‘to go’, but it is nowhere near as marked as in Catalan : 10,100,000 Castilian-language hits for *me va bien*, 2,400,000 for *me viene bien*. It is therefore possible that participants, given the near-unacceptability of the Catalan *venir* in this context, generalised these findings to Castilian, where no such non-normativity exists.

Castilian exercise: ...*escúchame, reina, abrígate bien...* Seventeen participants (fourteen younger, three older) commented on the use of *reina* (Eng *darling*, literally *queen*). This is likely to be a question of formality. Despite the texts aiming to recreate spoken language, certain informal constructions still stood out to participants as such. These items are thus not out of place in the text, but there may have been a confusion of informality of register and unacceptability, and this needs to be borne in mind with future experiments of this type. It should also be remembered that for methodological concerns, these texts, although informal, were written down: such informality may be more jarring on the page than in speech. However, our results are not affected, since participants were not told how many items to look for in each text. Therefore, the signalling of informal items did not *replace* the potential identification

of a non-normative token, but was merely conducted alongside the identification of the non-normative tokens key to the experiment.

These examples of misidentified tokens are interesting for future development but will receive no further attention. They may have appeared due to a number of reasons, be they linguistic (cross-linguistic interference, hypercorrection) or methodological (potential oversights on the part of the researcher and colleagues creating the experiment, the written medium of the test). In order for them to be fully included in a discussion of this linguistic situation, they would need to be submitted to more rigorous testing. While potentially of great interest to future projects, this falls outside of the remit of this thesis.

### **5.3 Language competence: Participant self-report**

Each participant was asked to evaluate their own language competence in both Catalan and Castilian. Participants were required to rate their own competence in both Catalan and Castilian on a scale from one (perfect) to five (no ability whatsoever) in each of the four language skill areas: understanding, speaking, reading and writing. As well as retaining all scores for each of the skill areas, an average score for each language was calculated for each participant. The primary aim of this was twofold: firstly, to ascertain whether the participants were indeed bilinguals, and secondly, to help determine their language dominance (see section 5.1.5). However, these self-report findings have wider applications as they potentially provide interesting insight into speakers' views on each language; moreover, these results can be viewed in conjunction with the experiment results outlined in section 5.1, thus revealing the relationship between how participants' feel they perform in each language, and the scores they actually achieve.

The following self-report language competence findings are complementary to my experiment results, and may provide interesting insight into the nature of speaker perceptions of language contact in the linguistic community. As such, they will be



referenced periodically in the upcoming discussion in chapter 6.

Independent and dependent sampling techniques<sup>17</sup> conducted on the overall results (i.e. those which look at all fifty participants as one group without dividing along any of the variable axes) reveal no significant differences between participants' perceptions of their abilities in each of the languages, both when looking at the overall average scores and when looking at each of the skill areas individually. However, subtle differences in participants' self-report become more apparent when analysing the data in terms of the six variables.

Age group. Independent sampling techniques reveal that the average self-report score was significantly different (sig. = .012) between younger and older participants' as regards their perceptions of their competence in Catalan. Younger people evaluated themselves as more competent in Catalan than older people did. When examining Castilian competence, no such significant difference is found. Moreover, there is no significant difference when comparing younger participants' perceptions of their Catalan and Castilian competence, nor when comparing older participants' perceptions of their Catalan and Castilian competence<sup>17</sup>. Dependent sampling tests imply a difference in trends between the two groups when looking at self-report (sig. = .004) which, when viewed alongside the independent sampling results, may refer to the fact that younger participants' consistently rate their competence higher in Catalan, and older people higher in Castilian. This is borne out by the means for each group. When the average score is broken down into its four component language skills, it is revealed that reading and writing show up significant differences in practices, but not understanding and speaking. With regard to reading, younger people evaluate their competence at reading Catalan at a higher level than older people do (sig. = .049), and this leads to the dependent sampling methods showing a difference in trends between the groups (sig. = .041). Re-arranging the results in order to compare younger participants and older participants (as opposed to results in Catalan and Castilian), when asked to evaluate their writing competence, younger people rate

---

<sup>17</sup>For the variables which produce two subgroups (i.e. *age group*, *sex*, *language dominance* and *language of interview*), t-tests were chosen for independent sampling, with Pearson correlations used for the remaining variables (*level of education* and *social class*). For all variables, Repeated Measures tests were the chosen dependent sampling technique.

themselves significantly higher than older people do in both languages (sig. =.003 for Catalan, sig. = .006 for Castilian). This is not however reflected in the dependent sampling test, which reports no significant difference in trends between the two age groups. Having looked at the self-report findings, it would now be helpful to ascertain how they relate to the different age groups' actual performance in the exercise. Younger participants have a higher perception of their overall competence in Catalan than older people, and this is indeed mirrored in the results, where young people do indeed perform with more accuracy in the Catalan exercise than older people. Similarly, the lack of perceived difference in Castilian competence is reflected with a lack of significant difference in performance in the Castilian exercise. Moreover, the difference in trends between the two groups is borne out when looking at the exercise results (compare with section 5.1.3).

Sex. Neither dependent nor independent sampling techniques reveal any significant differences between or within the two sexes when looking at the overall figures or any of the individual language skills. This reflects the results found in section 5.1.4, where sex of participant was not deemed significant on any level.

Language dominance. Independent sampling techniques reveal no significant differences between or within language dominance groups when examining average results. The only difference reported when looking at the four language skill areas was that Catalan-dominant participants evaluate their spoken Catalan as significantly stronger (sig. = .030) than their spoken Castilian. Dependent sampling techniques however reveal significantly different trends between language dominance groups when looking at average self-report results (sig. = .008), as well as when looking at the skill areas of speaking (sig. =.043) and reading (sig. =.013), with writing just falling short of the  $p < .05$  threshold (sig. =.052). These results do not seemingly reflect the findings witnessed in section 5.1.5, whereby Catalan-dominant participants performed consistently better in the Catalan exercise. The only indication of any perceived difference is found within Catalan-dominant participants' views of their own spoken command of the two languages, none being found in any between-group perceptions. This therefore cannot reflect the results, where the difference is found between the

two language dominance groups (not within them). Moreover, the dependent sampling reveals a significant difference in trends, but the independent results shed no light onto what said trends may be. More information is therefore needed before one can interpret these findings.

Language of interview. Neither dependent nor independent sampling techniques reveal any significant differences between or within the two language of interview groups when looking at the overall figures or any of the individual language skills. This reflects the results found in section 5.1.6, where the language of interview was not deemed significant on any level.

Level of education. Independent sampling techniques reveal a positive correlation (significant at the  $p < .01$  level) between average self-reported competence in Catalan and level of education, participants with a higher level of education evaluating their overall competence in Catalan stronger than those with a lower level of education. An analysis of each of the Catalan language skills shows that similar correlations are found when looking at understanding (significant at the  $p < .05$  level), reading and writing (both significant at the  $p < .01$  level), though not speaking. In Castilian, there is no significant correlation between level of education and language competence when looking at average scores, as well as the individual skill areas of understanding, speaking and reading. However, there is a positive correlation (significant at the  $p < .01$  level) when evaluating the writing of Castilian, with participants of a higher educational level rating their competence more strongly than those of a more limited educational background. Dependent sampling techniques reveal a difference in trends of the various level of education groups, when viewing average scores, as well as the reading and writing skill areas. When viewed in conjunction with the mean results for each group, such difference in trends could refer to the fact that participants of a lower educational level perceive more of a discrepancy between their competences in Catalan and Castilian (reporting a higher competence in Catalan), whereas those of a higher educational level do not perceive such a difference between the two languages. This does not therefore reflect the experiment results found in section 5.1.7, which revealed that all level of education groups performed in a similar manner, i.e. more

accurately in Catalan than Castilian.

Social class. Independent sampling interestingly reveals a negative correlation (significant at the  $p < .01$  level) between average competence in Catalan and social class, with the working/lower class groups claiming higher competence in Catalan overall. Similar correlations are found in the skill areas of reading and writing (both significant at the  $p < .01$  level), although the group who claim the highest level of written Catalan competence is the lower-middle class, as opposed to the working/lower class group. As with level of education, the average results reveal no correlations between social class and reported Castilian competence, the same also being observed with the skill areas of understanding, speaking and reading. Again, self-report on writing competence in Castilian revealed a positive correlation with social class (significant at the  $p < .01$  level), with the upper-middle class group claiming highest competence. Dependent sampling techniques show no significant difference in trends between the groups, both when examining average scores and the individual skill areas. This somewhat reflects the experiment results in section 5.1.8, insofar as there is no significant difference between the trends of each group, although the interesting details about each social class group's perception of its competence (the working/lower class claiming higher competence overall in Catalan, while the upper-middle class claim greater competence in written Castilian) may potentially merit further investigation.

In summary, there are certain parallels between participants' self-reported language competence and their performance in the exercises when the results are dissected along the different variable axes. Younger and older participants' divergent perceptions of their own Catalan language competence fall in line with the results of the experiment, as does the perceived lack of difference in Castilian competence. Moreover, variables such as sex and language of interview, which proved to have no bearing on experiment results, also reveal no significant trends or differences when examining self-reported language competence. Other than this, the self-report competence findings show where participants' perceptions differ from their actual performance, such as the negative correlation for social class groups regarding Catalan competence (contrary

to the findings of section 5.1.8). These findings will receive further attention in chapter 6.

## 5.4 Overview

The experiment results presented in section 5.1 have revealed some interesting details about language practices in Barcelona. When the mean experiment scores for the two exercises were calculated looking at all fifty participants, significantly more instances of non-normative language use were identified in Catalan than Castilian. A variable analysis performed on the data allowed for greater insight, demonstrating that *participant age group* was a greatly influential factor in the results, which will allow for interesting upcoming discussion of the research questions posed in chapters 2 and 3. The two key experimental results pertaining to the *age* variable were as follows:

1. When looking at Catalan, younger participants identified significantly more instances of non-normative linguistic convergence in Catalan than older people (see table 5.2a).
2. When comparing Catalan and Castilian, younger people identified more non-normative language use in Catalan than in Castilian, while older people show no significant difference in performance between the two languages (see table 5.2b).

*Participant language dominance* also proved significant insofar as Catalan-dominant participants consistently perform better in Catalan, but Castilian-dominant participants are at no such advantage. However, the interaction between this variable and *age group* was tested in section 5.1.9 and was not seen as significant. Therefore, the predictive power of the *age group* variable is not diminished. The remaining variables (*sex*, *language of interview*, *level of education* and *social class*) were of limited or no significance.

A types and tokens analysis was performed in section 5.2 to provide more fine-grained detail regarding the experiment results by indicating which types and/or

tokens may be more or less salient to participants. The type analysis revealed that neither particular interference type (lexical or syntactic) was more salient than the other; while the token analysis showed that the Catalan exercise was performed in a relatively consistent manner, but the Castilian exercise was somewhat erratic.

Finally, in section 5.3, participants' self-reported competence in each of the languages revealed some interesting parallels (when examining the results along the axis of *age group*) and divergences (social class observations) when compared to the actual experiment results. So the importance of the *age group* variable is reinforced by the self-report findings (which mirror the experiment results), and these findings will receive further attention in our upcoming discussion of language contact in section 6.1.

In this chapter, the experiment results were analysed in two different ways to two different ends. The variable analysis results in section 5.1 will allow us to answer the research questions, since the differences revealed along the axis of *participant age group* will form the basis of the upcoming discussion in chapter 6. The types and tokens analysis in section 5.2 offers a more micro-level perspective about the nature of individual instances of non-normative language use. The self-report figures in section 5.3 provide interesting complementary information regarding participants' perceptions and metalinguistic awareness, and by extension how this tallies with their actual performance in the exercises.

The key experimental finding was that the *participant age group* variable was the most important, showing us that younger participants demonstrate a higher rate of identification<sup>18</sup> of non-normative Catalan. This will now be applied to the research questions in order to draw conclusions to our research questions of language policy and language contact in Barcelona.

---

<sup>18</sup>And by extension, awareness. For a full explanation of this link between identification and awareness, see page 152.

# Chapter 6

## Discussion

The fieldwork experiment aims to provide answers to the following research questions on language policy and language contact, posed at the ends of chapters 2 and 3 respectively:

### POLICY RESEARCH QUESTION

To what extent is current Catalan language-in-education policy meeting its aims?

### CONTACT RESEARCH QUESTION

To what extent are Catalan-Castilian bilinguals aware of non-normative instances of contact-induced linguistic convergence? Has there been any development in this awareness in recent years?

This chapter will use the findings of the experiment given in chapter 5 to answer these two questions. So that the arguments presented are clear, the language contact question will be answered before the language policy question. Section 6.1 will thus interpret the experiment results to provide answers to the questions concerning Catalan-Castilian bilinguals' awareness of contact-induced linguistic convergence. Then, section 6.2 analyses the results and existing conclusions in order to evaluate the extent to which Catalan language-in-education policy is meeting its aims. Since the fieldwork experiment addressed non-normativity in both Catalan and Castilian in equal measure, this discussion will do the same. Some of the upcoming conclu-

sions and discussion will centre around Catalan in this community, and other around Castilian. Section 6.3 will then examine any related theoretical concerns, including consequences for the model of linguistic processes given in chapter 1, as well as discussing any potential directions for future research. Section 6.4 will summarise all of these findings, providing some final conclusions to the thesis.

## 6.1 Answering the language contact research question

### CONTACT RESEARCH QUESTION

To what extent are Catalan-Castilian bilinguals aware of non-normative instances of contact-induced linguistic convergence? Has there been any development in this awareness in recent years?

The fieldwork experiment undertaken is perfectly suited to answering this question. The tokens of non-normative language use which occur in each text are instances of contact-induced linguistic convergence, be they syntactic or lexical, Castilianisms in Catalan or Catalanisms in Castilian. All of the instances chosen exemplify contact-induced linguistic convergence between Castilian and Catalan, and all are non-normative, insofar as they do not form part of the standard language of the text in which they appeared. Moreover, the experiment tested speakers' conscious awareness of such tokens, since they were asked to identify any language use they considered non-normative, thus offering metacommentary on such instances.

The results in chapter 5 showed that *participant age group* was the most significant variable to examine, and this will be reflected in the upcoming discussion. The performance of the two age groups will be analysed, and contrasted where appropriate, so as to best offer insight into participants' abilities to identify non-normative instances of linguistic convergence. The experiment comprised two texts, one in each language, in order to examine contact phenomena present in both Catalan and Castilian. Therefore, this section will examine both languages, first offering a conclusion



regarding the presence of contact-induced language use in Catalan, then moving on to discuss language contact in Castilian. The discussion in this section will, as far as possible, be limited to the consequences of the findings for the language contact question posed above. Any further arguments which address language policy and related concerns will receive attention in subsequent sections.

\* \* \*

The primary finding of the fieldwork concerned normative Catalan, and was that younger participants identified significantly more instances of non-normative language use in Catalan than the older participants. This is revealed in table 5.2a (see page 207) Therefore, invoking the apparent time hypothesis (see section 4.1.2.1), we can postulate this difference in performance between the two generations is indicative of a development in awareness of non-normativity in Catalan that took place at some point between the formative periods of each generation. This will be of renewed importance when examining the consequences for language policy in section 6.2. For now, it is sufficient to remember the fact that younger participants have benefited from Catalan-medium education, unlike their older counterparts. This idea will be developed in subsequent conclusions.

The first conclusion merely takes the primary fieldwork finding into account, amalgamating it with the facts presented in the above paragraph, and is as follows:

**Conclusion 1: CONTACT IN CATALAN**

**Younger experiment participants, who have received Catalan-medium education, are more adept at identifying non-normative instances of linguistic convergence in Catalan than older participants, who have not received Catalan-medium education.**

These findings are first and foremost applicable to the participants in the fieldwork experiment, hence the explicit reference to study participants in the conclusion. While the sample size is more than adequate for a sociolinguistic investigation of this nature,

more research would be helpful to confirm the present findings and support their generalisation to the whole Catalan-Castilian linguistic community of Barcelona.

This primary experimental finding and subsequent conclusion also shows interesting parallels with the self-reported participant language competence<sup>1</sup> data given in section 5.3. Younger participants' higher performance in Catalan (when compared to older participants) mirrors each group's self-perceived views of Catalan language competence. Younger participants' self-assessment of Catalan language competence is significantly higher ( $\alpha=0.5$ ) than older participants' self-assessment in the four skill areas of reading, writing, speaking and understanding (see page 257). Therefore, younger participants believe themselves to possess a refined knowledge of normative Catalan, and the results of this experiment bear this out, at least for this very specific linguistic task. Similarly, older members of the community did not rate their Catalan language competence so highly, and indeed this is also reflected in the task, since they did not perform so well.<sup>2</sup> However, as we shall see later in this section, there is not always such parallelism between the experimental findings and participant self-report data.

Conclusion 1 is of key importance since it argues that younger people are more adept at identifying instances of non-normative linguistic convergence when they appear in an otherwise normative Catalan text than older people. When viewed in conjunction with the apparent time hypothesis, it seems therefore that younger participants have somehow been made aware of certain instances of non-normative language use, contrary to their older counterparts who are unaware of the non-normative status of these items. This has more far-reaching consequences, when we consider

---

<sup>1</sup>As already mentioned in chapter 1, the term 'competence' is employed in a non-Chomskyan sense here, and refers to language skills, primarily categorised in the four areas of speaking, reading, understanding and writing.

<sup>2</sup>Moreover, while interesting, it should be noted that the self-report findings do not reflect how participants feel they would perform when compared to the other group. That is to say, younger participants did not say they had a higher level of Catalan competence than their older counterparts, but rather that they rated their competence highly in general. Similarly, older participants did not rate themselves as weaker than younger members in Catalan, just weak in general. Had participants been directly asked to give judgements not only on their own competence, but on what they feel other participants' (of their own age group and of the contrasting age group) language competence to be, the parallels with experimental findings may yield further results. This is a potential future direction in which this project could be taken.

what might have caused such a surge in awareness; the role of the education system cannot be underestimated and will be explored further in section 6.2. Moreover, the fact that younger participants can offer metacommentary on certain linguistic items, whereas older people cannot, has interesting implications for the model of sociolinguistic phenomena presented in chapter 1 (particularly when examining the ‘+/- metacommentary’ axis); this will receive further attention in section 6.3.

To directly answer the contact research question using conclusion 1, younger Catalan-Castilian bilingual participants showed a greater awareness of a subset of non-normative instances of contact-induced linguistic convergence— namely Castilianisms in Catalan— than older participants. The statistically significant difference between the performance of the two age groups implies that awareness of non-normative Castilianisms in Catalan may have increased in recent years, based on the apparent time hypothesis.

\* \* \*

Having discussed participants’ awareness of normative Catalan, our results can also provide answers to our contact research question regarding normative Castilian. Conclusion 1 shows that younger participants have a higher rate of identification of non-normative language use appearing in Catalan than older people, leading on from table 5.2a (see page 207). Table 5.2a also showed us that there was no significant difference between the two age groups’ performance when asked to identify non-normative Castilian. When the data in table 5.2a was re-arranged according to language of exercise, table 5.2b (see page 208) revealed that younger participants’ high rate of Catalan performance was significantly higher than their Castilian performance. Table 5.2b also shows us that there was no significant difference between older participants’ performance in the two exercises.<sup>3</sup> Our experimental findings concerning Castilian in the two age groups therefore reveal the following:

---

<sup>3</sup>When analysing these results, we must be careful not to simply say that there is a lower level of Castilian competence than Catalan competence across the board, basing our findings on table 5.1 (see page 204). While at first glance this appears to be the case, this lower level of Castilian actually represents very different things for the two age groups. While the level of identification of non-normative Castilian is indeed lower than the level for Catalan with younger participants, this is not the same for older participants, who display this comparatively ‘low’ level in both languages. Therefore, a conclusion which states that “there is a lower level of Castilian performance than Catalan

- All participants, regardless of age group and thus language of education, performed similarly when asked to identify non-normative linguistic convergence in Castilian (see table 5.2a, page 207).
- For younger participants, this represented a significantly lower rate of identification than in normative Catalan (see table 5.2b, page 208).
- Older participants showed no difference between their rate of identification of non-normative language use in the two languages (see table 5.2b, page 208).

These key findings concerning participant Castilian language competence can be summarised into the following conclusion:

### **Conclusion 2: CONTACT IN CASTILIAN**

**All participants, whatever their language of education, show similar levels of awareness of non-normative linguistic convergence in Castilian. For younger participants, this represents a significantly weaker awareness of normative Castilian than of normative Catalan. For older participants, this reflects no difference in awareness of non-normativity in the two languages.**

Note that, given the lack of cross-generational difference in performance, there is no need to invoke the apparent time hypothesis in order to speculate what such results could reflect, as regards Castilian language competence through time. Though this lack of difference between younger and older participants will prove interesting, and will be borne in mind, when examining the consequences for Catalan language policy in section 6.2.

As before, it would be interesting to see how these results reflect the self-report findings presented in section 5.3. Some parallelism can be observed, insofar as there was no significant difference between the two groups when they were asked to rate <sup>performance across the board</sup>” is misleading, since for older participants, the performance in the two exercises was the same.

their own Castilian linguistic competence. This is borne out in the fieldwork experiment results where there is no significant difference between the two age groups' rates of identification of non-normative linguistic convergence. Moreover, older participants did not self-report any difference in their own competence between Catalan and Castilian, which is also borne out by the results. However, younger people's self-reported perceptions differ from their actual performance in the exercises: they claimed the same level of competence in Catalan and Castilian, whereas the experiment shows that younger participants identified significantly fewer instances of non-normative linguistic convergence in Castilian than in Catalan. This is the first instance we have seen of self-report findings contradicting experimental performance. Either younger participants are overestimating their grasp of normative Castilian or underestimating their capabilities in normative Catalan, insofar as this can be related to the identification of non-normative occurrences of linguistic convergence. Future research may shed light on this issue.

When discussing awareness of normative Castilian, it would also be useful to see how the present findings fit in with existing work which addresses Castilian language competence in other Castilian-speaking areas. On the one hand, it may be seen as cause for concern that awareness of normative Castilian be lower than awareness of normative Catalan. However, a series of studies conducted by the Instituto de Evaluación (part of the Spanish Ministerio de Educación) revealed that Catalan children's competence in Castilian was not especially low; rather, it was decidedly average. In two rankings of all seventeen Spanish autonomous communities, Catalonia was placed in ninth position for a test on ten-year-olds published in 2010, and eleventh position for a test on thirteen-year-olds published in July 2011 (Gutiérrez 2011). The consequences of this for Catalonia's language-in-education policy will be discussed in section 6.2, but these findings are also interesting when examined alongside the present study's conclusions on language contact in Catalonia. Conclusion 2 states that younger Catalan participants were less adept at identifying non-normative Castilian than non-normative Catalan in the experiment. However, this is not necessarily a reflection on their Castilian language skills overall, since other young Catalans are

testing higher than those in more monolingual Castilian-speaking areas such as Extremadura and Andalucía. This could mean that a comparative lack of knowledge of contact-induced non-normativity in Castilian is not enough to dictate that students have ‘poor linguistic competence in Castilian’. Alternatively it could mean that, even if a comparative lack of knowledge of contact-induced non-normativity in Castilian *did* equate poor linguistic competence, members of other Castilian-speaking communities outside Catalonia are weaker still, possessing even *poorer* Castilian linguistic competence. Further research is required to solve these issues. However, based on the thesis’ language contact conclusions, we can determine that younger participants display a weaker command of normative Castilian than of normative Catalan, whether this is an indicator of poor Castilian linguistic competence or not, and whether the youth of other autonomous communities also lack certain skills in normative Castilian or not.

To directly answer the contact research question using conclusion 2, younger Catalan-Castilian bilinguals are not as aware of contact-induced linguistic convergence in Castilian as they are in Catalan. There is no difference in awareness between younger and older members of the community regarding instances of contact-induced linguistic convergence in Castilian. Therefore, in accordance with the apparent time hypothesis, we cannot postulate any development in awareness of non-normative Castilian in recent years.

\* \* \*

In short, the experiment findings and subsequent discussion have taught us a great deal about Catalan-Castilian language contact. Our research question asked about speaker awareness of non-normative linguistic convergence between the two languages and if this had developed over recent years. In response, we can say that:

- Based on conclusion 1 and the apparent time hypothesis, we can infer that a higher level of awareness of items of contact-induced linguistic convergence appearing in Catalan has developed in recent years.

- Based on conclusion 2, we can infer that no such development has occurred in the awareness of items of contact-induced linguistic convergence appearing in Castilian.
- Based on conclusion 2, younger participants have a higher awareness of non-normative contact-induced linguistic convergence in Catalan than in Castilian. Older participants showed a similar, comparatively low level of awareness of contact-induced linguistic convergence for both languages.

Thus far, the discussion has solely focused on language contact, and specifically, the awareness of non-normative CILP. These conclusions will be developed in the next section, in order to answer our research question regarding Catalan language policy.

## 6.2 Answering the language policy research question

<b><u>POLICY RESEARCH QUESTION</u></b>
--

To what extent is current Catalan language-in-education policy meeting its aims?
--

The present thesis not only offers comment on language contact phenomena in Barcelona (conclusions and discussion given above in section 6.1), but will also use these findings to provide an analysis of the effectiveness of current Catalan language-in-education policy. More specifically, this section will examine the experimental results and develop the existing conclusions in order to assess the consequences of the implementation of Catalan-medium education in the majority of primary and secondary schooling throughout Catalonia.

This section will provide three conclusions. The first (conclusion 3), develops some of the experimental findings in an attempt to establish a link between the fieldwork and current Catalan language policy initiatives. Conclusion 4 will then elaborate on conclusion 1, which discussed language contact appearing in Catalan, by discussing policy implications for the Catalan language. Finally, conclusion 5 will develop conclusion 2, which discussed language contact in Castilian, to talk about policy implications for the use of Castilian in the speech situation under examination.

\* \* \*

If we look at conclusion 1, given in full on page 265, we see that those younger experiment participants who had received Catalan-medium education were more adept at identifying non-normative Catalan than older participants. Invoking the apparent time hypothesis, we have already determined that this higher level of awareness of non-normative Catalan has developed in recent years. This can be elaborated in order to provide a more exact date for such a development. The development of awareness of non-normative linguistic convergence in Catalan witnessed in the experiment results must have occurred *after* the formative years of the older group and was largely in effect by the time the younger group reached this period. It would seem likely that the education system is a key vehicle for making speakers aware of instances of non-normative language use in Catalan.<sup>4</sup> As such, speakers were deemed to be in their formative period, at least for receiving formal instruction about non-normative language use, when they were in full time, compulsory education. The approximate age for the end of this period across groups is around 15 years of age. The date of the development in awareness of non-normative linguistic convergence in Catalan can thus be approximately identified. The lower limit of the older age group is 55 years of age; the average age of the end of compulsory education for them was 15, and so this occurred 40 years before the experiment ( $55 - 15 = 40$ ). Likewise the upper limit of the younger age group is 35; the average age of the end of compulsory education is still taken as 15 (for consistency), and so this point occurred 20 years before the experiment took place. The development in participant awareness of non-normative linguistic convergence in Catalan witnessed in the experiment (and stated in conclusion 1, page 265), thereby occurred between 20 and 40 years before the experiment took place, i.e. between 1970 and 1990.

The period from 1970 to 1990, as we know from chapters 1 and 2, saw a great deal of growth in institutional support of Catalan, encompassing not only the intro-

---

<sup>4</sup>This is to be discussed in much greater detail later in this section. The key role of the education system, and of language-in-education policies such as MOI choice, will be explained in the build-up to conclusion 4. For our present purposes, we shall rely on the common-sense assumption that the education system is an important (not even necessarily the *most* important) means of diffusing knowledge about what constitutes normative and non-normative language use.



duction of Catalan as the chief vehicular language of the education system, but also new media, advertising, signposting and administration, all conducted through the medium of Catalan, and all mandated by the Generalitat's language policies. These policies began in earnest with 1983's *Llei de Normalització Lingüística*, which then developed and engendered numerous other pieces of more specific linguistic legislation. Therefore, given the time frame provided above, based on experimental findings and the apparent time hypothesis, we can postulate the following:

**Conclusion 3: POLICY AND THE FIELDWORK**

***Normalització lingüística* policy initiatives and resultant legislation have, at least in part, led to the increase in participant awareness of non-normative linguistic convergence in Catalan witnessed in conclusion 1.**

\* \* \*

The present section will now develop existing findings in order to offer a conclusion as to how Catalan language-in-education policy is meeting its aims, with specific regard to the Catalan language. Much as conclusion 1 examined language contact phenomena as they relate to Catalan, the upcoming conclusion 4 will view the Catalan language situation through the lens of recent policy developments. In order for this to be achieved, firstly the role played by language-in-education policy, and specifically MOI choice, will be shown to be of key importance within the range of *normalització lingüística* policies implemented by the Generalitat (these policies being postulated in conclusion 3 as causes of the cross-generational difference in performance viewed in the experiment). MOI choice will be shown to be of importance not only on a theoretical level, but also through discussion of how these choices are actually implemented in Barcelona classrooms. Then, in order to answer the policy research question, Catalan language-in-education policy aims will be revisited, and it will be shown how the experimental findings can be employed to assess the degree to which these aims are being met. It will then be possible to formulate a conclusion which

uses the experimental findings concerning Catalan language use (i.e. those expressed in conclusion 1) in order to ascertain the effectiveness of Generalitat language-in-education policy in relation to the Catalan language.

Conclusion 3 has postulated that Catalan governmental *normalització lingüística* initiatives may be responsible for the cross-generational differences in awareness of non-normative linguistic convergence appearing in Catalan, and the importance of language-in-education policy has been cursorily mentioned. But while clearly insightful, this could be refined and made more specific. We are still not clear as to the exact nature of the language planning initiatives and resultant legislation mentioned above. This thesis focuses on language-in-education policy, and even more precisely, decisions made regarding the medium of instruction of a given school system. If this focus is to firmly tie in with the experimental findings, a link must be established between MOI choice and awareness of non-normative language use. However, the move to implement widespread Catalan-medium education was not the only language planning decision which has led to the current linguistic situation in Barcelona. Various government-led measures were introduced, resulting in an array of different constructs and processes which have benefited the Catalan language. The role of decisions regarding MOI choice must therefore be placed within this panoply, and its role in generating the current linguistic state of affairs ascertained. Only then will we be able to be more certain that the cross-generational difference in awareness of non-normative linguistic convergence in Catalan is attributable to MOI choice.

Firstly, we need to ascertain what the ‘panoply’ under discussion might be. That is to say, what else has been implemented as a result of top-down governmental decisions with the aim of bettering the position of the Catalan language? As mentioned above, the increase in Catalan-language media (television, radio and internet) has been supported by linguistic legislation passed down from the Generalitat. In some cases, the media even explicitly targets non-normative language use, such as in the popular TV3 game show *El Gran Dictat*, where one round asks contestants to replace Castilianisms with the normatively correct Catalan lexical item.<sup>5</sup>

---

<sup>5</sup>The drive to raise awareness of non-normative language use is also attested in the media of other linguistic communities, as in the Radio e Televisão de Portugal show *Bom Português*.

It is therefore clear that, when it comes to eradicating non-normative language use, the education system is not the only player. This can be said without even examining Generalitat linguistic input into other sectors such as signposting and private enterprise. However, for our purposes, it would be helpful to establish the education system (and more specifically the choice of Catalan as MOI) as a *key* player, if not the *only* player, in the drive to limit non-normative Catalan language use. In order to do this, let us revisit what it means to choose Catalan as MOI, and what ends this seeks to achieve. Arguably, given the lack of institutional support for Catalan under the Franco regime, and the consequently weakened position of Catalan by the time of the dictator's death, the aim was to strengthen the status of the language. The choice of a given language as MOI bolsters its standing within a community, and this view is supported by numerous theoretical concepts. The notion of *ethnolinguistic vitality*, coined by Howard Giles and colleagues in 1977, ascribes greater or lesser vitality to a language based on three criteria: *status*, *demography* and *institutional support*. The decision to use a given language as the primary vehicle for delivery of education, certainly constitutes an increase in institutional support, and arguably also in status. Moreover, the notion of *linguistic capital*, developed from Bourdieu's *cultural capital*, also supports this, and is defined as 'fluency in, and comfort with, a high-status, world-wide language which is used by groups who possess economic, social, cultural and political power and status' (Morrison and Lui 2000, 473). In making sure that all Catalan citizens have fluency in, and comfort with, the Catalan language, the Generalitat is arguably redefining linguistic capital in Catalan terms. If the language receives a greater degree of support and status, then the political power may well follow. One could argue that there is nowhere better to cement this support and status than through MOI decisions, since it is thanks to the education system that students acquire the requisite fluency and comfort with the language in question.<sup>6</sup>

In theory, we can see that the decision to make a particular language the MOI is of central importance in the nascent language policy a given community. This

---

<sup>6</sup>For an analogous case concerning the use of Chinese in Hong Kong, and cementing the theoretical ties between MOI choice and linguistic capital, see Morrison and Lui 2000.

now needs to be linked to the stigmatisation of non-normative language use, to show that the results of the fieldwork experiment are directly attributable, at least in part, to the introduction of Catalan-medium education. While the reduction of speaker instances of non-normative Catalan language use may not be an explicit aim of the introduction of Catalan-medium education, it is arguably a consequence. Teachers are required by law to demonstrate a certain level of competence in normative Catalan (see page 280). It is therefore presumed that teachers are in a position to correct non-normative Catalan language use with the requisite degree of accuracy. These teachers not only habitually highlight instances of non-normative language use (as indeed is the case in linguistic communities around the world), but also have a large window of opportunity in which to do so. The benefits of a system whereby Catalan is the vehicle of education in all classes (except Castilian language and literature) is that the student receives greater exposure to normative Catalan (from the teachers and by extension, from fellow pupils). In short, teachers can correct students' non-normative Catalan in a physics lesson, a history class, or a physical education session: they are not limited to the Catalan language and literature classroom.

In this section, a series of links has been established, between MOI choice, recent Catalan language policy initiatives and the increased awareness of non-normative Catalan language use, succinctly summarised in figure 6.1.<sup>7</sup>

Therefore, based on the links presented in this section, we can deduce that developments in language-in-education policy, particularly MOI choice, form an integral part of the recent Catalan language policy initiatives which have arguably led to an increase in participant awareness of non-normative linguistic convergence appearing in Catalan.

Having now demonstrated the central importance of MOI choice (see figure 6.1) as a

---

<sup>7</sup>It should be pointed out that even if MOI choice were not deemed to be of central importance within recent language policy developments (contra the arguments presented here), the effectiveness and relevance of the three-axis model would remain unchanged. This is because the results of all language policy developments, and not just MOI choice, fulfil the criteria of TDP. Therefore the result of developments in media, or private enterprise are manifestations of TDP. However, this thesis has deduced that MOI choice is arguably of the greatest importance in instigating the top-down phenomena in Catalan viewed in the experiment, and for this reason, our discussion is valid.

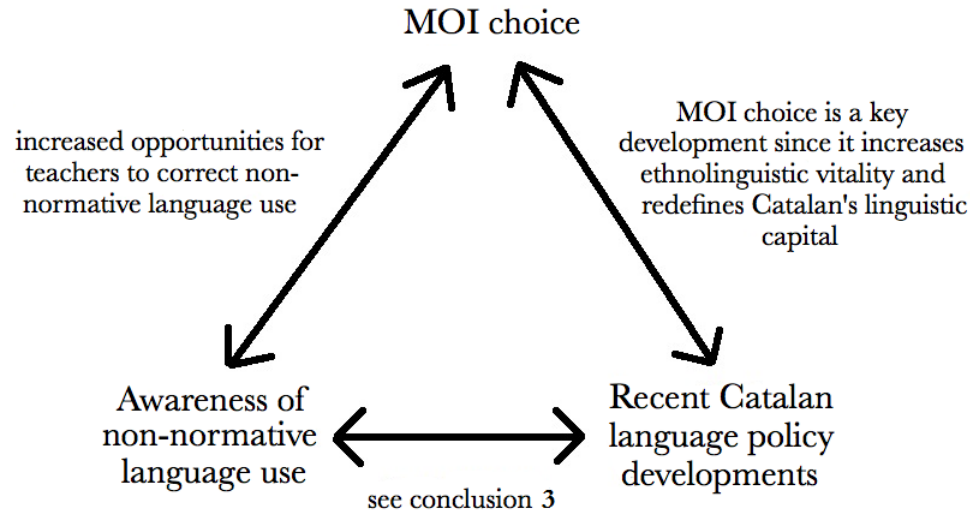


Figure 6.1: The links established between *MOI choice*, *Recent Catalan language policy developments* and *Awareness of non-normative language use*

key player within the wider array of language policies implemented by official bodies, it would be informative to examine how these decisions are concretely applied in Barcelona classrooms. This brief discussion of the practical concerns behind the implementation of Catalan as MOI will serve as a bridge between the analysis of the theoretical importance of MOI just undertaken, and its relevance within the Catalan community. With full knowledge of the theory behind MOI choice, as well as the practice regarding how this is achieved in Catalan schools, we will be able to progress to an assessment of the Catalan linguistic situation. The original Catalan language-in-education policy aims can then be evaluated, in order to provide an answer to our policy research question.

Barcelona residents who underwent Catalan-medium education<sup>8</sup> report that non-normative linguistic convergence appearing in Catalan is frequently corrected in numerous classes and scenarios. Secondary school Catalan language lessons consist of exercises where, as a class, pupils are required to identify non-normative language

<sup>8</sup>Due to a lack of published literature on the topic, the upcoming brief discussion is based on informal discussion with contacts aged between 25 and 35.

use. Moreover, these breaches of normative Catalan, when committed spontaneously by students, are corrected in front of their peers. Such correction of students' Catalan is not limited to Catalan class, with written work produced in other lessons, ranging from history to biology, heavily penalised for use of non-normative Catalan. In primary education (where several subjects are taught by one teacher), verbal, public corrections of written or oral violations of normative Catalan are commonplace, whatever the subject matter. Interestingly, those consulted had few memories of violations of normative Castilian being similarly addressed.<sup>9</sup>

In short, non-normative linguistic convergence between Catalan and Castilian (particularly when manifested as Castilianisms in Catalan) is openly tackled in schools in both primary and secondary education, with violations of normativity corrected openly in front of entire classes of students. This information about the implementation of MOI in the Catalan context helps set the scene to answer our policy research question. We now know that MOI is of central importance within the language policies applied (see figure 6.1), and that, in the Catalan context, this entails a specific drive towards ensuring linguistic normativity in classroom settings. Therefore, we can proceed to use our existing findings concerning speaker knowledge of non-normative language (conclusions 1 and 2) to answer our policy research question, to see if Catalan language policies are meeting their aims.

Thus far, it has been deduced that recent language policy developments, and particularly language-in-education policies such as MOI choice, have resulted in an increase in speaker awareness of non-normative Catalan language use. So, now this information needs to be directly applied to our policy research question so as to provide an answer as to how effective recent Catalan language policy decisions have been. In order to test the effectiveness of these policies, we must be reminded of their original aims. Then, the policies can be judged as more or less successful depending on the degree to which the original policy objectives have been addressed.

---

<sup>9</sup>This may go towards explaining the imbalance between the tackling of skills in Catalan and Castilian, to be summarised in conclusion 5. However, further research and development is necessary before such a claim can be made.

As already stated in section 2.2.1, the seminal piece of post-Franco Catalan language policy legislation is the *Llei de Normalització Lingüística* (LNL) of 1983. In article 1 of its *títol preliminar* the objectives are clearly identified as follows:

- To protect and encourage the use of Catalan for all citizens.
- To effectively implement the official use of Catalan.
- To normalise the use of Catalan in all media of social communication.
- To ensure the extension of Catalan language competence.

It is this final objective that bears the clearest relation to the content of this study. Speaker awareness of – and by extension, avoidance of – non-normative language use, is a facet of language competence; all the other objectives solely address language *use*, not language *competence*. Therefore, if the consequences of current language policies can be seen to ensure Catalan language competence, then they can be judged as successful, at least in part.

The above fragment of the LNL discusses the general aims of Catalan language policy. However, if we accept the central importance of language-in-education decisions (as highlighted in figure 6.1), then the specific aims of Generalitat policy regarding the school system must also be taken into account. An addendum to the 1983 LNL which specifically addresses aims regarding student language competence in Catalan and Castilian as a result of the introduction of new, Catalan-medium education, states that ‘the curriculum... should ensure that students possess an accurate and balanced mastery of the two official languages’ (un tractament curricular... que assegurí als alumnes un domini correcte i paritari de les dues llengües oficials, Generalitat de Catalunya 1983b, article 2).

Therefore the government’s most fundamental aims regarding Catalan language competence are that the knowledge of Catalan be sufficiently spread. On a specific, language-in-education level, it must be guaranteed that students possess sufficient skill in both Catalan and Castilian as a result of the schooling they receive. If the experiment findings show us that these objectives are being addressed, then the policies can be considered to be functioning correctly.

It has already been claimed that recent language policies (specifically, MOI choice) have led to an increase in awareness of non-normative language use in Catalan. But this does not necessarily guarantee that said policies are meeting their initial objectives. For that to be the case, the policies need to be spreading Catalan language competence and ensuring that students are sufficiently skilled in both languages upon leaving the education system, as per the aims and objectives of the 1983 legislation, laid out above. So, firstly we need to determine the following: can an increase in speaker awareness of non-normative Catalan language use be interpreted as a way of ensuring Catalan language competence, and thus meeting the policy objectives as defined?<sup>10</sup>

To answer this question, it must be determined how Catalan language competence is to be evaluated. It bears repeating that term ‘language competence’ in this case is not being used in a Chomskyan sense, but rather as a translation of the term in the legislation, ‘coneixement’ (Eng *knowledge*). The four components of language competence – reading, writing, speaking and listening – can be concretely assessed through testing. Such Catalan language testing is already in place, and consists of a number language exams provided by the *Institut Ramon Llull*, each different exam corresponding to a level of the CEFR. While students are currently not required to take these exams as part of compulsory education, the *Certificats de Llengua Catalana* are approved by the government, and indeed schoolteachers are required to provide evidence of having passed the *nivell de suficiència* exam (CEFR level C1) before they are allowed to teach (Generalitat de Catalunya 2002, article 12). Moreover, upon completing secondary education at eighteen years of age, students are presumed to have a level of Catalan equivalent to the *nivell de suficiència* (Generalitat de Catalunya 2002, primera disposició adicional). We may therefore assume that the Catalan language competence of those whose education has been in the medium of Catalan can be defined by the criteria of the *nivell de suficiència* exam. It is a foregone conclusion that those who have been schooled in Catalan will pass this exam (should they hypothetically be required to do so), and as such, its content should therefore

---

<sup>10</sup>The question of ‘balanced competence’ and discussion of parallel awareness of non-normative language use in Castilian will be discussed later in the section.



pose no problems. This is therefore a fitting way of judging what exactly makes up the presumed Catalan language competence of a person who has been through the Catalan education system. Unsurprisingly, this exam judges language use in all four skill areas (reading, writing, speaking and listening), and in order to pass, students need to be familiar with the ‘correct’ usage of lexical items and grammatical structures. In other words, in order to attain this level, one is presumed to have a firm grasp of what does, and does not, constitute normative Catalan language use. Therefore, in order to demonstrate Catalan language competence so as to work in an official capacity, one must pass a governmental language exam (the *nivell de suficiència*) or its equivalent (Catalan-medium education to the age of eighteen); and part of the skill set needed to pass this exam is an ability to identify non-normative language use in Catalan.

We are now in a position to bring our existing findings and conclusions together to form another, important conclusion that will provide an answer to our policy research question. We have thus far ascertained that recent language policy has potentially resulted in an increase in awareness of non-normative language use (see conclusion 3), and that MOI choice is a key component of this policy (see figure 6.1). Moreover, it has just been shown that awareness of normativity in Catalan is deemed to be a central part of language competence. Therefore, through increasing awareness of non-normative language use, these measures have increased language competence. It has been shown above, through examination of the original aims and objectives of the 1983 legislation, that if there has been a demonstrable increase in Catalan language competence, then the policies can be judged as successful, at least in part. The policies can thus be seen to be positively contributing towards their objectives, and the following conclusion can be drawn:

**Conclusion 4: POLICY AND THE CATALAN LANGUAGE**

**The Generalitat can be seen to be making successful strides towards its objectives of ‘ensuring Catalan language competence’ and ‘[guaranteeing that students possess] an accurate mastery of Catalan [as one of the two official languages]’.**

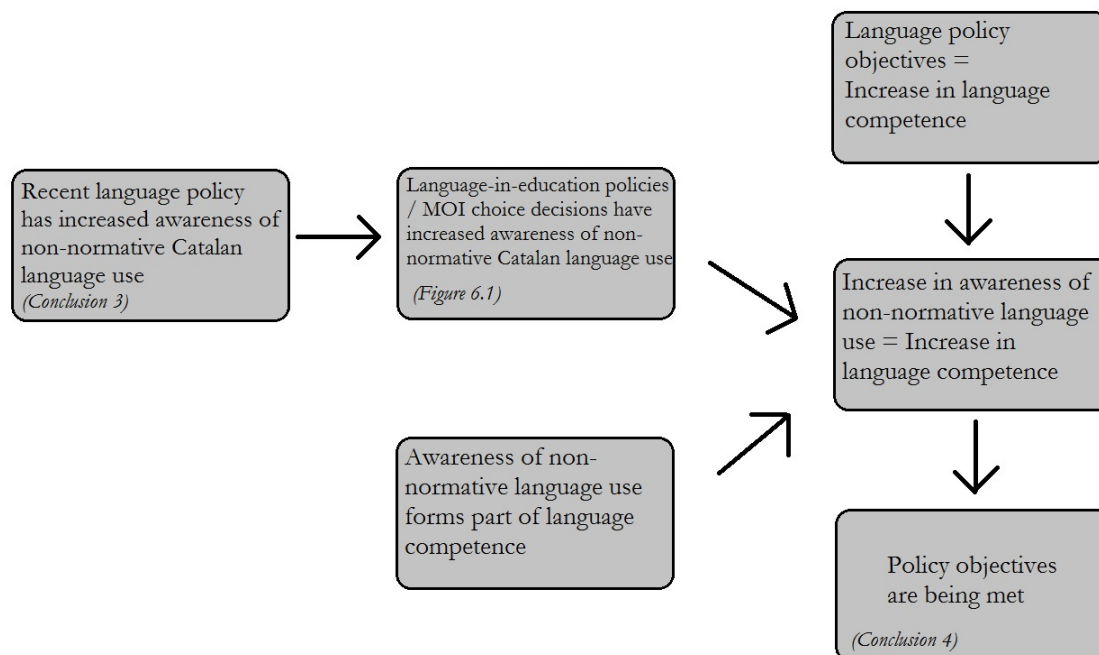


Figure 6.2: The validity of conclusion 4

This last conclusion was derived from a combination of previous conclusions and statements whose validity was proven in this section of the discussion, and is visually represented in figure 6.2.

\* \* \*

This section will finally develop present findings in order to offer a conclusion as to whether language-in-education policy in Barcelona is meeting its aims with specific focus on the Castilian language. The upcoming conclusion 5 will derive from the existing discussion in a similar way to conclusion 4. Conclusion 4 took the experimental findings of conclusion 1 concerning language contact in Catalan, and expressed the consequences for policy aims; conclusion 5 will do the same by using the experimental findings of conclusion 2 concerning language contact in Castilian, and expressing how this links to language policy. This will be achieved simply by examining the fieldwork results concerning Castilian and comparing the performance of the two generations of participants.

As well as identifying non-normative Catalan language use, participants were asked to comment on contact elements as they appear in Castilian. Almost half of the results given in chapter 5 are concerned with identification of Catalanisms in Castilian. Indeed, while the Castilian of language contact areas has been known to receive scholarly attention,<sup>11</sup> it is more frequent to examine the Catalan language practices of Catalan speaking areas.<sup>12</sup> Therefore, an examination of contact Castilian offers some perspective on comparatively underexamined linguistic practices.

Table 5.2a (page 207), tells us that participant age was a decisive factor in identifying non-normative linguistic convergence in Catalan, which we have seen constitutes a key component of Catalan language competence. Furthermore, it was revealed that participant age had no bearing upon ability to identify non-normative Castilian language use. So a primary question we may want to ask is *has the increased Catalan language competence of speakers resulted in a concomitant decrease in Castilian language competence?* In answering this, the results actually prove difficult to interpret.

Conclusion 2 stated that there were no cross-generational differences when examining levels of awareness of non-normative linguistic convergence in Castilian. At first glance, this seems to indicate that the increased Catalan language competence of younger participants is *not* accompanied by a decreased level of Castilian language competence, given the lack of difference between the groups. However, the mean level of education of each age group needs to be taken into account: the Pearson correlation (significant at the  $p < .01$  level) between the variables of *age group* and *level of education* (see table 5.23 on page 239) reveals that older participants have a significantly lower level of education than their younger counterparts. It can be seen in table 5.10 (page 225) that participants with lower levels of education identify fewer instances of non-normative language use across the board, and it has already been determined that older participants in the study have a lower level of formal education. Therefore, the lack of difference between the two groups in identifying non-normative Castilian

---

<sup>11</sup>Indeed, the *International Journal of the Sociology of Language* dedicated its entire double issue 193/194 to the sociolinguistic study of Castilian varieties in contact. Moreover, many papers on related topics can be found on [www.carstensinner.de](http://www.carstensinner.de).

<sup>12</sup>For example, issue 47 of the *International Journal of the Sociology of Language*, or indeed most of the output of the journal *Treballs de Sociolingüística Catalana*.

language use could actually be interpreted differently: older people *should* be scoring lower than younger people (given their lower level of education), so a parity between the groups might actually indicate a *higher* level of Castilian language competence among older participants in a situation where the level of education for both groups was in fact equal. Following this argument, young participants' increased language competence in Catalan is actually therefore at the expense of a diminished Castilian language competence, when compared to their older counterparts. However, this argument, aside from being largely speculative for want of more conclusive evidence, is not without its problems. Dependent sampling techniques performed on the sample divided along the axis of *level of education* (table 5.12, page 227) showed that there were no significant differences with regard to within-participant practices. This is to say, within each level of education group, members all exhibited the same patterns concerning identification of non-normative Castilian use relative to non-normative Catalan use: i.e. no pattern could be discerned. This prevents us from drawing a clear conclusion from our results since, if there is such a strong correlation between *age group* and *level of education*, then we might expect similar results in table 5.12 as found when dependent sampling techniques were administered to the sample along the axis of *age group* (table 5.3, page 209). This is not however the case, with the different age groups exhibiting different within-participant trends, although this could be related to the larger number of *level of education groups*. A certain element of doubt is nonetheless present and maybe the relationship between *age group* and *level of education* is not so strong after all.

In short, it is unclear whether the younger participants' higher performance in Catalan (and by extension the increase in Catalan language competence from older to younger participants) has been accompanied by a concomitant drop in Castilian competence or not. On the one hand, the equal performance of both age groups in the Castilian exercise may indicate a drop in Castilian language competence, since older participants are shown to have a lower level of education (table 5.23), and therefore should be identifying fewer instances than – not the same number of instances as – younger participants across the board (table 5.10). But then this correlation between

age group and level of education in table 5.23 may not be so important, based on dependent sampling results in tables 5.3 and 5.12, and therefore we would be able to take the results at face value, i.e. that an equal rate of identification of non-normative Castilian language use for both groups indicates equal language competence, and thus there has been no loss of competence. Given the complexity, the speculativeness of the arguments, and the general lack of clarity of the situation, no conclusions can therefore be made regarding a potential loss of Castilian language competence as a result of post-Franco Catalan language policies, and further investigation is necessary.

In order to draw definitive conclusions regarding Castilian language competence, unequivocal findings are required. Although we cannot be certain if there has been a cross-generational decrease in Castilian language competence, it is nonetheless clear that there is a discrepancy between the ability to identify non-normative Catalan and non-normative Castilian. Younger participants are less adept at identifying non-normative Castilian than they are non-normative Catalan (see table 5.2b, page 208). It has already been concluded that an increase in Catalan language competence is a goal of current Catalan language policy, and this has been deemed successful through an increased ability to identify non-normative Catalan language use (conclusion 6), but the same clearly cannot be said for Castilian. While it cannot be determined whether speakers are getting worse or simply remaining constant when it comes to identifying non-normative Castilian, this much is certain: they are getting no better, even with increased exposure to education. This therefore means that while the current education system is clearly addressing aims regarding Catalan language competence, it is not giving results that demonstrate that it is equally concerned with students' use of Castilian. This is potentially an important shortcoming when a key aim of current Catalan education is that students possess 'an accurate and *balanced* mastery of the two official languages' (*un domini correcte i paritari de les dues llengües oficials*, Generalitat de Catalunya 1983b, article 2). The experiment has revealed that younger participants (who have been through this Catalan-medium education system) do not possess a 'balanced' ability when it comes to identifying non-normative linguistic convergence. This is because they consistently identify more

instances of non-normativity in Catalan than in Castilian. Their ability to identify non-normative language use is thus stronger in Catalan and thereby ‘imbalanced’. Therefore, the following conclusion can be drawn:

**Conclusion 5: POLICY AND THE CASTILIAN LANGUAGE**

**The current Catalan education system is not fully achieving its aim of ensuring balanced levels of speaker competence in the two official languages, since less emphasis is placed on identifying non-normative language use in Castilian than in Catalan.**

\* \* \*

In short, developing the experiment findings and existing conclusions on language contact has allowed us a great deal of insight into the effectiveness of current Catalan language-in-education policy. Our research question asked to what extent was policy meeting its aims, and in response we can say that:

- Based on conclusion 3, language policy played a role in the difference in performance between older and younger participants in the fieldwork experiment.
- Based on conclusion 4, current Generalitat language-in-education policy measures can be evaluated as making successful progress towards meeting their aims, insofar as they are ensuring Catalan language competence.
- However, based on conclusion 5, current Generalitat language-in-education policy measures cannot be judged as successful on all fronts, since they are not adequately promoting Castilian and Catalan language competence on an equal level.

Having now provided satisfactory answers to both policy and contact research questions, this discussion chapter shall briefly address any related theoretical concerns that may open the way for interesting future research into language policy and contact, both inside and outside the Catalan-Castilian linguistic community of Barcelona.

### 6.3 Related issues and directions for further research

The fieldwork experiment results have now been analysed and discussed to provide answers to the research questions on language policy and language contact in Barcelona. As such, the primary aims of the thesis have been met. However, the theoretical background discussion, methodology and subsequent results give rise to issues for further research and development. While such concerns fall out of the remit of the thesis, some will be briefly detailed here. The potential of the present work as a springboard to other interesting research projects will thus be made clear.

This section will primarily examine the consequences of the experimental findings for the three-dimensional model of sociolinguistic phenomena presented in chapter 1. A final, supplementary conclusion will be offered in due course as to the potential future usefulness of such a descriptive device as my three-axis cube. After this, other potential directions for further research concerning language policy and language contact will be addressed.

\* \* \*

In chapter 1, I proposed a new, three-axis theoretical model which could be used to classify sociolinguistic phenomena present in linguistic communities. Language policy and language contact were defined as example background situations to phenomena which, according to this cube model, were maximally differentiated. As shown in chapter 1, existing discussions of sociolinguistic phenomena ascribe difference on the basis of one, or at most two, axes. My three-axis approach allows for maximal differentiation of phenomena. Moreover, their positioning on a three dimensional cube gives an idea of simultaneity, and indeed, more than two phenomena can be placed on the cube at once. The phenomena resulting from language policy scored positively on the three axes of **+/- planned**, **+/- metacommentary** and **+/- driven by overt prestige**, and were termed *top-down phenomena* (TDP). The phenomena resulting from language contact scored negatively on the same axes and were termed *bottom-up phenomena* (BUP). The Barcelona speech community has shown the presence of both these phenomena, with language contact resulting in instances of

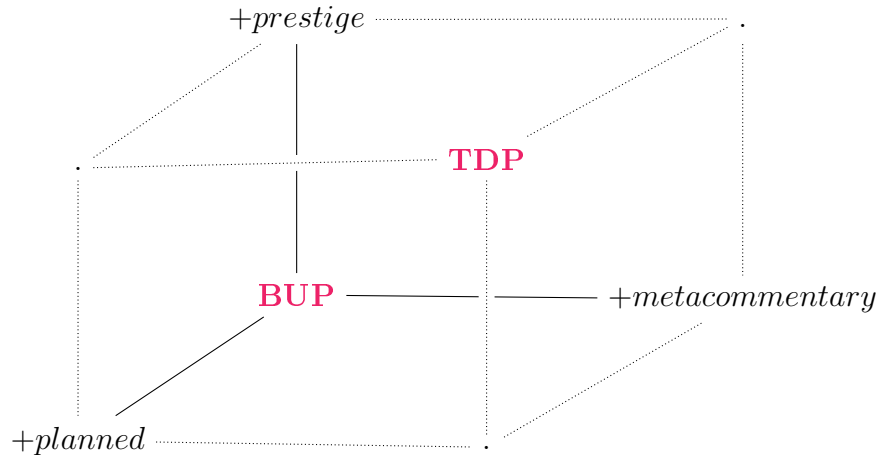


Figure 6.3: Reproduction of figure 1.3, showing where TDP and BUP appear on the cube

contact-induced linguistic convergence, and language policy arguably bringing about a cross-generational shift in awareness of such non-normative contact-induced phenomena. Therefore, both TDP and BUP are at work in the Barcelona linguistic community. Their position on the cube can be seen in figure 6.3 (reproduced from chapter 1).

The fact that both TDP and BUP are at work in this situation has potentially interesting consequences for the nature of the phenomena themselves. When formulating the research questions, one of the reasons that speaker awareness of tokens of non-normativity was deemed interesting was because it allowed for the testing of a specific competition between TDP and BUP. This was that if speakers could offer metacommentary on specific tokens of non-normativity that originally entered the language as instances of BUP, then surely these instances no longer fulfilled the criteria of BUP, since they were now *+ metacommentary*. The experiment shows a difference in performance between older and younger participants concerning the metacommentary axis. The significant difference between age groups means that younger people are consistently offering a greater deal of commentary on tokens of non-normative linguistic convergence in Catalan (i.e. Castilianisms) than their older counterparts. Thus, between these two groups, there has arguably been a shift concerning these Castil-



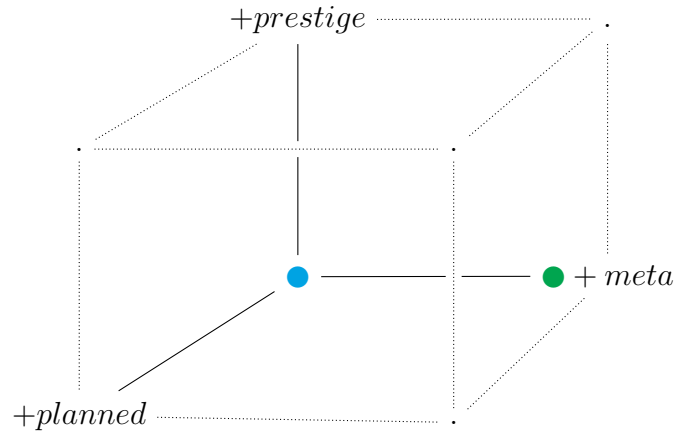


Figure 6.4: The positioning of Castilianisms on the three-axis cube for older participants (blue dot) and younger participants (green dot).

ianisms. For the older group they remain  $-metacommentary$ , whereas for younger participants, they are now  $+metacommentary$ . This is represented graphically in figure 6.4. In light of our existing conclusions, this metacommentary shift could be tied in with policy developments, since conclusion 4 (and the discussion leading to it) argued that the heightened ability of the younger generation to identify non-normative linguistic convergence in Catalan (and by extension their elevated Catalan language competence when compared to older participants) was due at least in part to Generalitat advances on fronts such as language-in-education policy and the establishment of Catalan as MOI. Therefore, this metacommentary shift in instances of BUP could be attributable to TDP (i.e. the policy developments). This would need to be examined further, in other situations to see if it is indeed feasible that two points on the model could interact with one another in such a way. All of this highlights the interesting potential of such a descriptive device of sociolinguistic phenomena as the three dimensional model, and argues for its inclusion and development in future work.

Other speech situations could theoretically be characterised by any number of simultaneous, distinct sociolinguistic phenomena, as in figure 6.5. Obviously, not all of the theoretical configurations outlined on the cubes in figure 6.5 are attested, or even likely. For example, if a phenomenon is  $+planned$ , it is not likely to be  $-driven\ by\ overt\ prestige$ . Similarly, if a phenomenon is  $+planned$ , it is not likely to

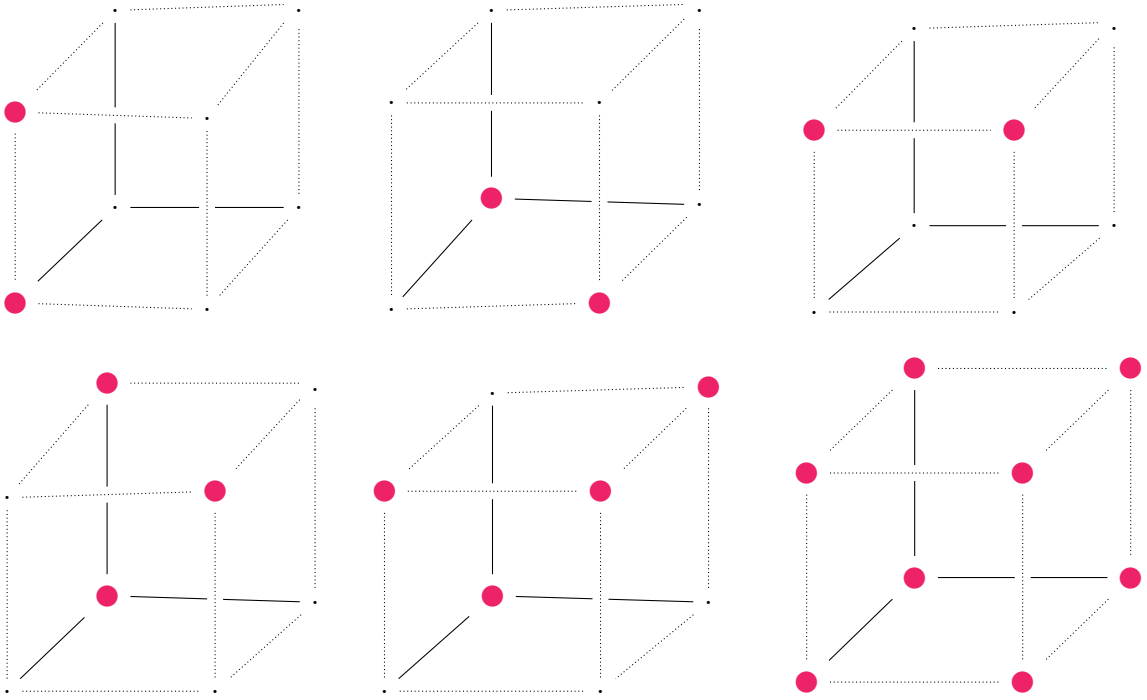


Figure 6.5: Examples of potential configurations of simultaneous, distinct sociolinguistic phenomena in hypothetical linguistic communities

be *–metacommentary*, at least not in the early stages of embedding in the linguistic community. However, the point is that such combinations are theoretically possible according to the model; whether they are likely or not is a matter for further study. Moreover, the benefit of such a construct is that there is no need to limit oneself to even two simultaneous phenomena. As seen in figure 6.5, if the axes are taken as binary, then eight potential distinct phenomena can occur at once. If the axes are interpreted as scalar (i.e. processes need not be + or – a particular property, but could lie somewhere in between), the number increases exponentially.

This model therefore has two strong benefits. Firstly, it classifies sociolinguistic phenomena along three axes, which is more than existing output attempts, allowing for maximal differentiation of processes. Secondly, it offers a clear visual representation of the simultaneity of phenomena that can occur within a community, whether that may refer to two, three or more distinct forces. The need for such a model is supported by the findings of my experiment, which has revealed that the two phenomena

which are maximally differentiated by the model can, and do, co-occur. If these two phenomena co-exist and interact, then what about other potential combinations? A model which accurately classifies sociolinguistic phenomena and allows for the simultaneity of processes is therefore a highly beneficial tool and this is expressed in the following conclusion:

**Conclusion 6: POTENTIAL FUTURE PROJECTS**

**The fact that multiple divergent sociolinguistic phenomena are shown to bear upon a single linguistic community means that a model is required which can deal with simultaneity and (potential) maximal distinctiveness of phenomena.**

\* \* \*

In addition to these potential theoretical advances, the primary foci of this work, i.e. language policy and language contact, could be developed, both in reference to the Catalan-Castilian linguistic situation and beyond.

Reflecting on our language policy findings, conclusions 4 and 5 have revealed that the Generalitat is making great strides towards ensuring students possess an ‘accurate’ knowledge of Catalan on leaving the education system, but that more could be done to achieve a ‘balanced’ competence of Catalan and Castilian. This overall assessment of the language-in-education policy decisions implemented in Catalonia could be compared with other Catalan-speaking regions which differ as regards their acquisition planning measures. There are currently studies which compare the different education measures in each Catalan-speaking territory and the impact this has on intergenerational Catalan language transmission in each case (cf. Vila i Moreno 2011 for a general overview, and Hawkey 2011 for a specific discussion of Catalunya Nord). Such cross-Catalan comparisons could be elaborated through recourse to the present work. Now that we are more familiar with the intricate consequences of Catalan-medium education in Catalonia (for example, the impact on contact varieties), similar questions could be asked of other Catalan-speaking regions. This becomes even

more revealing in the case of areas such as Catalunya Nord (in Southern France) and l'Alguer (in Sardinia), since these areas are governed by different polities, with arguably more linguistically assimilationist tendencies, and the varieties of Catalan are in contact with languages other than Castilian.

Another interesting field for exploration would be the contemporary state of Catalan language rights, particularly topical in light of the 2010 protests in Barcelona (see page 34). Given the different levels of provision for the Catalan language in the various regions where it is spoken (cf. Hawkey 2011 for a detailed comparison), it may be enlightening to investigate the degree to which Catalan speakers feel their rights are being supported or neglected. The findings of this present study compare the Catalan language competence of two groups, one whose rights have been arguably much more respected than the other. Such data is clearly germane to any study of language rights in the Catalan-speaking world.

Another potential direction would be to use this thesis as a starting point for a longitudinal study of language competence in a given linguistic community. The current findings could be supplemented with data collected at intervals in the future to track how Catalan and Castilian language competence in Barcelona may progress.

Outside of the Catalan-speaking world, other polities where the endemic linguistic variety is being revitalised, maintained or supported through processes of TDP, such as Wales, could offer a point of comparison with Catalonia, and constitute an interesting opportunity for interdisciplinary, collaborative work.

The present work's findings concerning language contact could also be taken in new and interesting directions. The fact that this work examines Catalanisms in Castilian, as well as the more frequently examined Castilianisms in Catalan, is a testament to the bidirectionality of contact-induced linguistic interference in the Barcelona linguistic community. This could be more fully explored in the Catalan-speaking world through a comparative study of different Catalan-speaking regions, which are characterised by different patterns of language use. In some regions, the presence of Catalan in the media is less pronounced, while in others Catalan holds little to no official status. Moreover, the languages with which Catalan co-exists are different in the various

regions. As such, a comparative study of Catalan language contact in various Catalan-speaking communities would allow scholars to draw interesting conclusions about the form, function and usage of Catalan throughout the areas where it is an autochthonous linguistic variety.

As with the findings on language policy, this can be generalised to examine other sites of language contact worldwide. This may include situations where both languages receive a degree of institutional support (such as Basque or Welsh), or indeed where one language is largely unsupported (as with many immigrant groups in France, for example). In any of these cases, an approach which bears in mind not only multidirectionality of interference, but also the simultaneity of maximally differentiated sociolinguistic phenomena, may prove beneficial.

## 6.4 Final conclusions

This thesis set out to offer insight into contact-induced linguistic phenomena in the Catalan-Castilian bilingual community of Barcelona, as well as assessing the effectiveness of recent Catalan language-in-education policies. The work made explicit reference to both the Catalan and Castilian languages, as they appear in this linguistic community.

Concerning Catalan, conclusion 1 showed how younger participants are more adept at identifying contact-induced linguistic phenomena than older participants, when these phenomena contravene the norms of standard Catalan. Conclusion 3 then revealed that Catalan *normalització lingüística* language policy efforts were at least in part responsible for this cross-generational difference in ability to identify non-normativity in Catalan. Then, after determining the central role of language-in-education policy with these normalisation efforts, conclusion 4 determined that the Generalitat is making successful strides towards its policy objectives of ‘ensuring Catalan language competence’.

As regards Castilian, conclusion 2 revealed that there is no difference cross-generationally when participants were asked to identify non-normativity in Castilian.

It remained unclear whether this reflected an inter-generational drop in Castilian competence, or simply no difference, but either way normative Castilian competence was not increasing, contrary to what is witnessed with Catalan. Therefore conclusion 5 determined that, in some respects, the Generalitat is falling short of its policy aims, since balanced competence in both Catalan and Castilian is not being achieved.

This work could also serve as a springboard to further study both inside and outside the Catalan-speaking world. Conclusion 6 demonstrated the future potential of the project in supporting the usefulness of the three-dimensional model of sociolinguistic phenomena. This is just one of many ways in which the present project could be developed.

In short, the methodology and findings of this thesis can be applied in further study in order to provide valuable insight concerning language policy developments and language contact phenomena. It is my hope that this thesis will prove useful to scholars in many different fields. To Catalanists, I hope this work offers a representation of the Catalan and Castilian languages in a period of great cultural, historical and political significance for Catalonia. To linguists, I hope that discussion of the effectiveness of language policy and the impact of contact phenomena proves enlightening. And to scholars in general, I hope that this examination of a dynamic and engaging linguistic situation inspires innovative and fruitful research in the future.

# Bibliography

- Ager, D. 1997. *Language, community and the state*. Exeter: Intellect.
- Ager, D. 2001. *Motivation in language planning and language policy*. Clevedon: Multilingual Matters.
- Ager, D. 2005. Image and prestige planning. *Current issues in language planning* 6 (1), 1–43.
- Ajuntament de Barcelona. Website. *Distribució territorial de la renda familiar a Barcelona*. <http://www.bcn.cat/estadistica/catala/dades/barris/renda/rdfamiliar/a2008.htm>. Accessed online 01/10/2010.
- Alexandre, V. 2006. *TV3 a traïció*. Barcelona: Edicions Proa.
- Amelang, J.S. 1986. *Honored citizens of Barcelona : Patrician culture and class relations 1490-1714*. Princeton, NJ: Princeton University Press.
- Arjona, M. 1979. Usos anómalos de la preposició *de* en el espanyol de México. *Anuario de Letras* 17, 167–84.
- Avui. 2009. Neix el gran grup en català. *Avui.cat*, 27/11/09 .
- Badia i Margarit, A. 1964. Llengua i cultura als Països Catalans. In F. Valverde (1980) *Aproximació crítica a la sociolingüística catalana*, Barcelona: Edicions 62.
- Baker, C. 2006. *Foundations of bilingual education and bilingualism*. Bristol: Multilingual Matters.
- Barber Casanovas, T. 2007. Lexicografia i traducció: l'argot juvenil. Accessed online at <http://www.carstensinner.de>, 23/06/2009.
- Barrera i Aguilera, M. 2000. Estudi sobre les interferències lingüístiques en l'obra de Josep Pla. Accessed online at <http://www.carstensinner.de>, 23/06/2009.
- Bastardas, J. 1995. *La llengua catalana mil anys enrere*. Barcelona: Curial.
- Bastardas i Boada, A. 1996. *Una ecologia de les llengües. Medi, contactes i dinàmica sociolingüística*. Barcelona: Edicions Proa.
- Batterham, M. 2000. The apparent merger of the front centering diphthongs – EAR and AIR – in New Zealand English. In A. Bell and K. Kuiper, eds., *New Zealand English*, Wellington: Victoria University Press.

- Bentivoglio, P. 1981. El dequeísmo en Venezuela: ¿un caso de ultracorrección? *Boletín de Filología de la Universidad de Chile*, 705–19.
- Blas Arroyo, J.L. 1991. Problemas teóricos en el estudio de la interferencia lingüística. *Revista Española de Lingüística* 21, 2 265–289.
- Boix i Fuster, E. and Vila i Moreno, F.X. 1998. *Sociolingüística de la llengua catalana*. Barcelona: Ariel.
- Bon català. Website. *Boncatala – La llengua catalana*. <http://boncatala.com>. Accessed online 27/07/2011.
- Bourdieu, P. 1976. The school as a conservative force: Scholastic and cultural inequalities. In Morrison, K. and Lui, I. (2000) Ideology, linguistic capital and the medium of instruction in Hong Kong, *Journal of Multilingualism and Multicultural Development* 21 (6), 471–86.
- Bourhis, R. 1984. *Conflict and language planning in Quebec*. Clevedon: Multilingual Matters.
- Branchadell, A. and Melià, J. 2011. Language policy in the world of business and consumer affairs: From non-existence to ineffectiveness. In M. Strubell and E. Boix-Fuster, eds., *Democratic Policies for Language Revitalisation: the Case of Catalan*, Basingstoke: Palgrave Macmillan.
- Bratt Paulston, C. and Heidemann, K. 2006. Language policies and the education of linguistic minorities. In T. Ricento, ed., *An introduction to Language Policy*, Oxford: Blackwell.
- Bugarski, R. 1992. Language situation and general policy. In R. Bugarski and C. Hawkesworth, eds., *Language planning in Yugoslavia*, Columbus, OH: Slavica.
- Calsamiglia, H. and Tusón, A. 1984. Use of languages and codeswitching in groups of youths in a *barri* of Barcelona: Communicative norms in spontaneous speech. *International Journal of the Sociology of Language* 47, 105–22.
- Calsamiglia, H. and Tusón, E. 1980. Ús i alternança de llengües en grups de joves d'un barri de Barcelona: Sant Andreu de Palomar. *Treballs de Sociolingüística Catalana* 3, 11–82.
- Castellsaguer, L. 2010. Barcelona atreu cada cop més a turistes de noves procedències. Accessed online at <http://www.bcn.es>, 22/03/2011.
- Caudmont, J. 1982. Contact de langues et évolution linguistique. In J. Caudmont, ed., *Sprachen in Kontakt - Langues en contact*, Tübingen: G. Narr.
- Cedergren, H. 1987. The spread of language change: Verifying inferences of linguistic diffusion. In P.H. Lowenberg, ed., *Language Spread and Language Policy: Issues, Implications and Case Studies. Georgetown University Round Table on Language and Linguistics*, Georgetown: Georgetown University Press.
- Chambers, J.K. 2002. *Sociolinguistic theory*. Oxford: Blackwell, 2nd edition.



- Chomsky, N. 1986. *Knowledge of language : its nature, origins and use*. New York: Praeger.
- Colón, G. 1989. *El español y el catalán: juntos y en contraste*. Barcelona: Ariel.
- Commission of the European Communities. 2005. Communication from the Commission to the European Parliament and the Council: The European Indicator of Language Competence. Brussels: European Union.
- Conversi, D. 1997. *The Basques, the Catalans and Spain*. London: Hurst & Co.
- Cooper, R.L. 1989. *Language planning and social change*. Cambridge: Cambridge University Press.
- Coromines, J. 1954. El que s'ha de saber de la llengua catalana. In Vallverdú, F. (1980) *Aproximació crítica a la sociolingüística catalana*, Barcelona: Edicions 62.
- Corporació Catalana de Mitjans Audiovisuals. 2005. Llibre d'estil de la CCMA. Accessed online at <http://esadir.cat>, 26/07/2011.
- Costa Carreras, J. 2009. *The Architect of Modern Catalan: Pompeu Fabra (1868-1948)*. Amsterdam: John Benjamins.
- Crystal, D. 1987. *The Cambridge Encyclopedia of Language*. Cambridge: Cambridge University Press.
- Crystal, D. 1997. *A dictionary of linguistics and phonetics*. Oxford: Blackwell, 4th edition.
- De Rosselló i Peralta, C. and Boix Fuster, E. 2004. El català, el castellà i l'anglès: tres llengües en joc: les ideologies lingüístiques dels estudiants de la Universitat de Barcelona. *Treballs de Sociolingüística Catalana 18*, 189–208.
- Deumert, A. 2004. *Language standardisation and language change: The dynamics of Cape Dutch*. Amsterdam: John Benjamins.
- Dewaele, J.M. et al. 2003. Introduction and overview. In J.M. Dewaele et al., eds. (2003) *Bilingualism: Beyond basic principles*, Clevedon: Multilingual Matters.
- Dreyfus, M. and Juillard, C. 2001. Le jeu de l'alternance dans la vie quotidienne des jeunes scolarisés à Dakar et à Ziguinchor (Sénégal). *Cahiers d'Études Africaines 163-4*, 667–96.
- Durfort i Coll, M. 2011. *Vocabulari de biologia de la reproducció*. Barcelona: IEC.
- Eastman, C.M. 1983. *Language planning: An introduction*. San Francisco: Chandler and Sharp.
- Echenique Elizondo, M.T. and Sánchez Méndez, J. 2005. *Las lenguas de un reino: Historia lingüística hispánica*. Madrid: Gredos.
- El català com cal. Website. *El català com cal. El parlem bé?* <http://elcatalacomcal.blogspot.com>. Accessed online 02/06/2011.

- Ellis, D. 1999. *From language to communication*. Mahwah, NJ: Taylor & Francis, 2nd edition.
- Enciclopèdia Catalana. Website. *Gran Enciclopèdia Catalana*. <http://www.enciclopedia.cat>. Accessed online 02/06/2011.
- Fabra, P. 1919a. Converses filològiques, num. 1. Per la puresa de la llengua: Els castellanismes. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Fabra, P. 1919b. Converses filològiques, num. 12. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Fabra, P. 1919c. Converses filològiques, num. 17. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Fabra, P. 1919d. Converses filològiques, num. 2. Mots catalans. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Fabra, P. 1920. Converses filològiques, num. 110. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Fabra, P. 1923a. Converses filològiques, num. 454. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Fabra, P. 1923b. Converses filològiques, num. 464. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Fabra, P. 1925a. Converses filològiques, num. 684. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Fabra, P. 1925b. Converses filològiques, num. 686. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Fasold, R. 1984. *The sociolinguistics of society*. Oxford: Basil Blackwell.
- Ferguson, C. 1953. Diglossia. *Word* 10, 25–39.
- Ferguson, C. 1987. Conventional conventionalization : Planned change in language. In T. Huebner (1996) *Sociolinguistic perspectives. Papers on Language in Society 1959–1994*, Oxford: Oxford University Press.
- Ferrando Francés, A. and Amorós, M.N. 2005. *Història de la llengua catalana*. Barcelona: UOC/Pòrtic.
- Fischer, J. L. 1958. Social influences on the choice of a linguistic variant. *Word* 14, 47–56.
- Fishman, J. et al. 1966. *Language Loyalty in the United States*. The Hague: Mouton.
- Galindo Solé, M. 2003. Language contact phenomena in Catalonia: The influence of Catalan in spoken Castilian. Accessed online at <http://www.lingref.com/cpp/wss/1/paper1004.pdf>, 21/07/2009.

- Galindo Solé, M. 2004. Les funcions discursives de l'alternança de codis als patis de les escoles de Catalunya. *Treballs de Sociolingüística Catalana* 18, 249–272.
- Generalitat de Catalunya. 1979. Estatut d'Autonomia de Catalunya. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1983a. Llei de Normalització Lingüística. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1983b. Ordre de 8 de setembre de 1983 (DOGC 362, del 9, i correcció d'errada en el DOGC 369, de 5 d'octubre), per la qual es desplega el Decret 362/1983, de 30 d'agost, sobre l'aplicació de la Llei /1983, de normalització lingüística a Catalunya, a l'àmbit de l'ensenyament no universitari, modificat per l'Ordre de 6 de desembre de 1983 (DOGC 415, de 9 de març de 1984). Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1983c. Reial decret 2568/1986, de 28 de novembre (BOE 305, de 22 de desembre i correcció d'errades, BOE 12, de 14 de gener de 1987), pel qual s'aprova el reglament d'organització, funcionament i règim jurídic de les entitats locals. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1990. Decret 214/1990, de 30 de juliol (DOGC 1348, de 28 de setembre, i correcció d'errades en el DOGC 1402, de 4 de febrer de 1991), pel qual s'aprova el reglament del personal al servei de les entitats locals. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1991. Llei 16/1991, de 10 de juliol (DOGC 1469, del 19, i BOE 190, de 9 d'agost, i correccions d'errades en els DOGC 1494, de 18 de setembre, DOGC 1507, de 18 d'octubre, i DOGC 1604, de 10 de juny de 1992), de les policies locals. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1994. Decret 14/1994, de 8 de febrer (DOGC 1859, de l'11, i correcció d'errades en el DOGC 1893, de 6 de maig), pel qual es regula l'exigència del coneixement del català als concursos per a la provisió de llocs de treball dels ens locals de Catalunya reservats a funcionaris amb habilitació de caràcter estatal. Annex derogat i apartat 1 de l'article 2 redactat d'acord amb el Decret 161/2002, d'11 de juny. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1996. Llei 8/1996, de 5 de juliol (DOGC 2232, del 19), de regulació de la programació audiovisual distribuïda per cable, modificada per la Llei 2/2000 de 4 de maig (DOGC 3133, del 5) del Consell Audiovisual de Catalunya. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1998a. Decret 269/1998, de 21 d'octubre (DOGC 2756, de 2 de novembre), del règim jurídic de les concessions per a la prestació del servei de radiodifusió sonora en ones mètriques amb modulació de freqüència per a emissores comercials, amb modificacions posteriors. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1998b. Llei de Política Lingüística. Barcelona: Generalitat de Catalunya.

- Generalitat de Catalunya. 1999a. Ordre de 16 de desembre de 1999 (DOGC 3046, del 30) per la qual es regula el funcionament dels establiments dedicats a la neteja tèxtil. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 1999b. Reial decret 1465/1999, de 17 de setembre (BOE 230, del 25), pel qual s'estableixen criteris d'imatge institucional i es regula la producció documental i el material imprès de l'Administració de l'Estat. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 2001a. Decret 152/2001, de 29 de maig, sobre l'avaluació i certificació de coneixements de català (DOGC 3406, d'11 de juny), modificat pel Decret 52/2003, de 21 de setembre (DOGC 4225, de 23 de setembre) i pel Decret 3/2006, de 17 de gener (DOGC 4554, de 19 de gener). Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 2001b. Decret 348/2001, de 4 de desembre (DOGC 3541, de 27 de desembre), pel qual es regulen els centres de bronzejat artificial. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 2002. Decret 161/2002, d'11 de juny, sobre l'acreditació del coneixement del català i l'aranès en els processos de selecció de personal i de provisió de llocs de treball de les administracions públiques de Catalunya. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 2003a. Decret 115/2003, de 8 de gener (DOGC 3801, del 16), pel qual es regula el règim jurídic transitori de les televisions locals per ones terrestres. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 2003b. Decret legislatiu 2/2003, de 28 d'abril (DOGC 3887, de 20 de maig), pel qual s'aprova el Text refós de la Llei municipal i de règim local de Catalunya. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 2006. Decret 108/2006, de 25 d'abril, pel qual s'aprova la modificació dels Estatuts del Consorci del Centre de Terminologia Termcat. Barcelona: Departament de la Presidència.
- Generalitat de Catalunya. 2009. Llei 12/2009, del 10 de juliol (DOGC 5422, del 16), d'educació. Barcelona: Generalitat de Catalunya.
- Generalitat de Catalunya. 2010. Llei 20/2010, del 7 de juliol (DOGC 5672, del 16), del cinema. Barcelona: Generalitat de Catalunya.
- Giles, H. et al. 1977. Toward a theory of language in ethnic group relations. In H. Giles, *Language, ethnicity and intergroup relations*, London: Academic Press.
- Gilliéron, J. and Edmont, E. 1902-1910. *Atlas Linguistique de la France*. Paris: Champion.
- Gobierno de España. 1978. La Constitución Española de 1978. Madrid: Gobierno de España.

- Golden, S. 2000. Projecte Sinibald de Mas: projecte pilot d'aculturació de la comunitat de parla xinesa a Catalunya: treball de camp. *Treballs de Sociolingüística Catalana* 14/15, 121–31.
- Gómez Torrego, L. 1991. Reflexiones sobre el dequeísmo y el queísmo en el español de España. *Español Actual* 55, 23–44.
- Grosjean, F. 2001. Individual bilingualism. In R. Mesthrie, ed., *Concise encyclopedia of sociolinguistics*, Oxford: Elsevier.
- Grupo Interalia. 1998. Lenguas peninsulares en contacto: Castellano y catalán en la producción de estudiantes universitarios bilingües. Accessed online at <http://www.carstensinner.de>, 23/06/2009.
- Gumperz, J. 1962. Types of linguistic communities. *Anthropological Linguistics* 4 (1), 28–40.
- Gutiérrez, M. 2011. El nivel de castellano de los alumnos de Catalunya está en la media del resto de las comunidades. *La Vanguardia*, 06/09/2011 .
- Gysels, M. 1992. French in urban Lubumbashi Swahili: Codeswitching, borrowing or both. In C. Eastman, ed., *Codeswitching*, Clevedon: Multilingual Matters.
- Haarmann, H. 1990. Language planning in the light of a general theory of language: A methodological framework. *International Journal of the Sociology of Language* 86, 103–26.
- Hamers, J.F. and Blanc, M.H.A. 2000. *Bilinguality and Bilingualism*. Cambridge: Cambridge University Press.
- Haugen, E. 1959. Planning for a standard language in Modern Norway. *Anthropological Linguistics* 1 (3), 8–21.
- Haugen, E. 1966. Dialect, Language, Nation. In E. Scherabon Firchow et al., eds., *Studies by Einar Haugen*, The Hague: Mouton.
- Haugen, E. and Weinreich, U. 1958. Reports. In *Proceedings of Eighth International Congress of Linguists*, 771–810, Washington: Linguistic Society of America.
- Hawkey, J. 2008. *Code-switching in Catalonia: The link between topic and identity*. Unpublished MA thesis.
- Hawkey, J. 2011. Towards an understanding of the contemporary sociolinguistic situation in the Pyrénées-Orientales. What questions need to be asked? *Cahiers AFLS* 15 (1), 36–52.
- Heine, B. and Kuteva, T. 2005. *Language contact and grammatical change*. Cambridge: CUP.
- Hernández García, C. 1994. Propuesta de variables sociolingüísticas para la obtención de los informantes para un corpus de la variedad del español de Barcelona y Área Metropolitana. Accessed online at <http://www.carstensinner.de>, 23/06/2009.

- Hoffmann, C. 1995. Monolingualism, bilingualism, cultural pluralism and national identity: Twenty years of language planning in contemporary Spain. *Current Issues in Language and Society* 2 (1), 59–89.
- Hornberger, N. 2006. Frameworks and models in language policy and planning. In T. Ricento, ed., *An introduction to Language Policy*, Oxford: Blackwell.
- Hughes, R. 1992. *Barcelona*. London: Harvill.
- IDESCAT. 2004. *Renda familiar disponible bruta (RFDB). Base 2000. Índex. Distribució per municipis. Any 2004*. <http://www.idescat.cat>. Accessed online 02/06/2011.
- IDESCAT. 2008a. *Població segons coneixement del català. Any 2008*. <http://www.idescat.cat>. Accessed online 09/05/2012.
- IDESCAT. 2008b. *Població segons coneixement del castellà. Any 2008*. <http://www.idescat.cat>. Accessed online 09/05/2012.
- IDESCAT. 2008c. *Població segons llengua d'identificació. Any 2008*. <http://www.idescat.cat>. Accessed online 09/05/2012.
- IDESCAT. 2008d. *Població segons llengua d'identificació per lloc de naixement. Any 2008*. <http://www.idescat.cat>. Accessed online 09/05/2012.
- IDESCAT. 2008e. *Població segons usos lingüístics amb la mare per grups d'edat* <http://www.idescat.cat>. Accessed online 09/05/2012.
- IDESCAT. 2008f. *Població segons usos lingüístics amb el pare per grups d'edat* <http://www.idescat.cat>. Accessed online 09/05/2012.
- Institut d'Estudis Catalans. 1997. *Diccionari de geologia*. Barcelona: IEC.
- Institut d'Estudis Catalans. 1998. *Vocabulari forestal, 2nd edition*. Barcelona: IEC.
- Institut d'Estudis Catalans. 1999a. *Proposta per a un estàndard oral de la llengua catalana I. Fonètica*. Barcelona: IEC.
- Institut d'Estudis Catalans. 1999b. *Proposta per a un estàndard oral de la llengua catalana II. Morfologia*. Barcelona: IEC.
- Institut d'Estudis Catalans. 2009. *Nomenclàtor oficial de toponímia major de Catalunya, 2nd edition*. Barcelona: IEC.
- Institut d'Estudis Catalans. 2010. *Diccionari jurídic català*. Barcelona: IEC.
- Institut d'Estudis Catalans. Website. *Gramàtica de la llengua catalana* <http://www.iecat.net>. Accessed online 19/05/2012.
- Instituto Nacional de Estadística. Website. <http://www.ine.es>. Accessed online 02/06/2011.
- Jané, A. 1976a. El bub bub. *Avui*, 26/05/1976 .
- Jané, A. 1976b. El dofí i el delfí. *Avui*, 09/06/1976 .
- Jané, A. 1976c. Les cometes. *Avui*, 01/06/1976 .

- Jané, A. 1976d. Noms del rosselló. *Avui*, 02/05/1976 .
- Jané, A. 1980. Prendre comiat. *Avui*, 16/05/1980 .
- Jespersen, O. 1946. Language: its nature, development and origin. In Chambers, J.K. (2002), *Sociolinguistic theory: second edition*, Oxford: Blackwell.
- Kaplan, R.B. and Baldauf Jr, R.B. 1997. *Language planning from practice to theory*. Clevedon: Multilingual Matters.
- Kloss, H. 1940. *Das Volksgruppenrecht in den Vereinigten Staaten von Amerika*. Essen: Essener Verlagsanstalt.
- Kloss, H. 1967. 'Abstand languages' and 'Ausbau languages'. *Anthropological Linguistics* 9 (7), 29–41.
- Kloss, H. 1969. *Research Possibilities on Group Bilingualism: A Report*. Quebec: International Centre for Research on Bilingualism.
- Kloss, H. 1977. *The American Bilingual Tradition*. Rowley, MA: Newbury House.
- Labov, W. 1972. *Sociolinguistic Patterns*. Oxford: Basil Blackwell.
- Labov, W. 2001. *Principles of linguistic change: Social factors*. Oxford: Blackwell.
- Lamuela, X. and Murgades, J. 1984. *Teoria de la llengua literària segons Fabra*. Barcelona: Quaderns Crema.
- Leclerc, J. Website. *L'aménagement linguistique dans le monde* Université Laval Website ([www.tflq.ulaval.ca](http://www.tflq.ulaval.ca)). Accessed online 12/04/2011.
- Levine, M.V. 1990. *The reconquest of Montreal: Language policy and social change in a bilingual city*. Philadelphia: Temple University Press.
- Llamas, C. et al. 2007. *The Routledge companion to sociolinguistics*. London: Routledge.
- Lodge, A. 2007. The history of French. In D. Ayoun, ed., *French Applied Linguistics*, Amsterdam: John Benjamins.
- Mackey, W.F. 2005. Multilingual cities / Mehrsprachige Städte. In U. Ammoth et al., eds., *Sociolinguistics: An international handbook of the science of language and society. Second completely revised and extended edition*, 1304–12, Berlin: Walter de Gruyter.
- Mahootian, S. 2006. Code switching and mixing. In *Encyclopedia of Language and Linguistics*. Accessed online at <http://www.sciencedirect.com>, 09/02/2008.
- Majoral, B. 2001. Cantar en català. In A. Colomines, ed., *És possible viure en català?*, Barcelona: Angle.
- Maurais, J. 2008. *Le vocabulaire des Québécois: Etude comparative (1983 et 2006)*. Montreal: Office québécois de la langue française.

- McLaughlin, F. 2001. Dakar Wolof and the configuration of urban identity. *Journal of African Cultural Studies* 14, 153–72.
- McRoberts, K. 2001. *Catalonia: Nation building without a state*. Oxford: Oxford University Press.
- Meyerhoff, M. 2006. *Introducing Sociolinguistics*. Abingdon: Routledge.
- Moravcsik, E. 1978. Language contact. In J.H. Greenberg, ed., *Universals of human language: Volume 1, method and theory*, Stanford, CA: Stanford University Press.
- Morrison, K. and Lui, I. 2000. Ideology, linguistic capital and the medium of instruction in Hong Kong. *Journal of Multilingualism and Multicultural Development* 21 (6), 471–86.
- Myers-Scotton, C. 1993. *Social motivations for code switching: Evidence from Africa*. Oxford: Oxford University Press.
- Nadal, J.M. and Prats, M. 1982. *Història de la llengua catalana: 1 - Dels orígens fins al segle XV*. Barcelona: Edicions 62.
- Nelde, P.H. 1997. Language conflict. In F. Coulmas, ed., *The handbook of sociolinguistics*, Oxford: Blackwell.
- Ninyoles, R.L. 1969. Conflicte lingüístic valencià. In Vallverdú, F. (1980) *Aproximació crítica a la sociolingüística catalana*, Barcelona: Edicions 62.
- Nymark, J. 2000. Diglòssia i polítiques lingüístiques. Anàlisi comparativa de tres experiències diferents: Catalunya, Noruega i el Paraguai. *Treballs de Sociolingüística Catalana* 14/15, 179–210.
- Office for National Statistics. Website. *SOC2010 Volume 3 NS-SEC (Rebased on SOC2010) User Manual*. <http://www.ons.gov.uk/about-statistics/classifications/current/soc2010/soc2010-volume-3-ns-sec-rebased-on-soc2010-user-manual/index.html>. Accessed online 29/09/2010.
- Ouakrim, O. 2000. Una aproximació sociolingüística del berber al Nord d'Àfrica i a Catalunya. *Treballs de Sociolingüística Catalana* 14/15, 89–97.
- Paloma i Sanllehí, D. and Rico i Busquets, A. 2008. *Diccionari de dubtes i barbarismes*. Barcelona: Edicions 62 / Proa.
- Paquot, A. 1988. *Les Québécois et leurs mots. Etude sémiologique et sociolinguistique des regionalismes lexicaux au Québec*. Quebec: Presses de l'Université Laval.
- Parlament de Catalunya. 2006. *Estatut d'Autonomia de Catalunya*. Barcelona: Generalitat de Catalunya.
- Pazos i Noguera, M.Ll. 1990. *L'amenaça del «català light»*. Barcelona: Tibidabo.
- Penny, R. 2002. *A History of the Spanish language*. Cambridge: Cambridge University Press, 2nd edition.
- Pericay, X. and Toutain, F. 1986. *Verinosa llengua*. Barcelona: Empúries.



- Pitarch i Almela, V. 2002. La llengua catalana i el sistema educatiu valencià: notes i reflexions. *Treballs de Sociolingüística Catalana* 16, 217–226.
- Poplack, S. 1980. Sometimes I'll start a sentence in Spanish y termino en español. In L. Wei, ed. (2000) *The Bilingualism Reader*, London: Routledge.
- Pountain, C.J. 2001. *A History of the Spanish language through texts*. London: Routledge.
- Pujolar, J. 1997. *De què vas, tio*. Barcelona: Empúries.
- Querol i Puig, E. 2009. L'ensenyament secundari al País Valencià i a Catalunya: Entre les paradoxes de la hipocresia i els mites en contacte. *Treballs de Sociolingüística Catalana* 20, 129–91.
- Rafel, J. 2010. Introducció a l'edició de les converses filològiques. In J. Mir and J. Solà, eds. (2010) *Pompeu Fabra: Obres Completes*, Barcelona: Edicions Proa.
- Real Academia Española. 2001. *Diccionario de la lengua española*. Madrid: Real Academia Española, 22nd edition.
- Ricento, T. 2006. Language policy: Theory and practice – An introduction. In T. Ricento, ed., *An introduction to Language Policy*, Oxford: Blackwell.
- Romaine, S. 2000. *Language in society: an introduction to sociolinguistics*. Oxford: Oxford University Press, 2nd edition.
- Romaine, S. 2003. Variation in language and gender. In J. Holmes and M. Meyerhoff, eds., *The handbook of language and gender*, Oxford: Blackwell.
- Rubió y Lluch, D.A. 1908. Discurs del vis-president del congrés. *Primer Congrés Internacional de la Llengua Catalana*, Barcelona.
- Sabater, E. 1984. An approach to the situation of the Catalan language: Social and educational use. *International Journal of the Sociology of Language* 47, 29–42.
- Sallabank, J. 2008. Prestige from the bottom up: A review of language planning in Guernsey. In A.J. Liddicoat and R.B. Baldauf Jr, eds., *Language planning in local contexts*, Bristol: Multilingual Matters.
- Schiffman, H. 1996. *Linguistic culture and language policy*. London: Routledge.
- Schmidt-Rohr, G. 1932. *Die Sprache als Bildnerin der Völker*. Jena: Eugen Diederichs.
- Segarra, M. 1985. *Història de la normativa catalana*. Tordera: Enciclopèdia Catalana.
- Segura i Ginard, L.J. 2004. Introducció al marc legal de la llengua catalana a les Illes Balears. *Treballs de Sociolingüística Catalana* 18, 117–127.
- Seib, J. 2001. La variedad bilingüe del español hablado en Cataluña caracterizada por interferencias y convergencias con el catalán. Accessed online at <http://www.carstensinner.de>, 23/06/2009.

- Sepa Bonaba, E. 2000. Afrosaharians a Catalunya : Des de quan i des d'on. *Treballs de Sociolingüística Catalana* 14/15, 103–20.
- Serrano, M.J. 1998. Estudio sociolingüístico de una variante sintáctica: el fenómeno *dequeísmo* en el español canario. *Hispania* 81 (2), 392–405.
- Siemund, P. and Kintana, N. 2008. *Language contact and contact languages*. Amsterdam: Benjamins.
- Silva-Corvalán, C. 1996. *Language contact and change: Spanish in Los Angeles*. Oxford: Oxford University Press.
- Sinner, C. 2002. *El castellano de Cataluña: Estudio empírico de aspectos léxicos, morfosintácticos, pragmáticos y metalingüísticos*. Tübingen: Max Niemeyer.
- Solé i Sabaté, J.M. and Villaroya, J. 1994. Cronologia de la repressió de la llengua i la cultura catalanes. Accessed online at <http://www.unitat.org/articles/repressio.html>, 03/06/2011.
- Spolsky, B. 2004. *Language Policy*. Cambridge: Cambridge University Press.
- Stolt, R. 2009. *The relationship between language and nation in the development of Austrian German*. Norderstedt: GRIN Verlag.
- Strubell, M. and Boix-Fuster, E. 2011. *Democratic Policies for Language Revitalisation: the Case of Catalan*. Basingstoke: Palgrave Macmillan.
- Szigetvari, M. 1994. Catalanismos en el español actual: Katalán elemek a mai spanyol nyelvben. Accessed online at <http://www.carstensinner.de>, 23/06/2009.
- Tagliamonte, S. 2005. So who? Like how? Just what? Discourse markers in the conversations of young Canadians. *Journal of Pragmatics* 37 1896–1915.
- Termcat. Website. *Termcat. Centre de terminologia*. <http://www.termcat.cat>. Accessed online 17/04/2011.
- Thomason, S.G. 2001. *Language contact: an introduction*. Edinburgh: Edinburgh University Press.
- Thomason, S.G. and Kaufman, T. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press.
- Thompson, W. and Hickey, J. 2005. *Society in focus*. Boston, MA: Pearson, Allyn & Bacon.
- Tonkin, H. 2005. Language planning. In M.J. Ball, ed., *Clinical sociolinguistics*, Oxford: Blackwell.
- Torres, J. 2003. L'ús oral familiar a Catalunya. *Treballs de Sociolingüística Catalana* 17, 47–76.
- Tovar, A. 1968. La lucha de lenguas en la Península Ibérica. In F. Vallverdú (1980) *Aproximació crítica a la sociolingüística catalana*, Barcelona: Edicions 62.
- Treffers-Daller, J. 2002. Language use and language contact in Brussels. *Journal of Multilingualism and Multicultural Development* 23 (1/2), 50–64.

- Trudgill, P. 1986. *Dialects in contact*. Oxford: Basil Blackwell.
- Trudgill, P. and Chambers, J.K. 1998. *Dialectology*. Cambridge: Cambridge University Press, 2nd edition.
- Turell Julià, M.T. 1994. Sociolingüística i contacte de llengües: el rastre del català en el discurs espanyol. *Treballs de Sociolingüística Catalana* 12, 51–62.
- Ubach Codina, N. 2009. Pròleg. *Treballs de Sociolingüística Catalana* 20, 15–17.
- UNESCO. 1953. *The use of vernacular languages in education*. Paris: United Nations.
- Vernet, J. and Pons, E. 2011. The legal systems of the Catalan language. In M. Strubell and E. Boix-Fuster, eds., *Democratic Policies for Language Revitalisation: the Case of Catalan*, Basingstoke: Palgrave Macmillan.
- Vila i Moreno, F.X. 2003. Els usos lingüístics interpersonals no familiars a Catalunya. Estat de la qüestió a començament del segle XXI. *Treballs de Sociolingüística Catalana* 17, 77–158.
- Vila i Moreno, F.X. 2008. Language-in-education policies in the Catalan language area. *AILA Review* 21, 31–48.
- Vila i Moreno, F.X. 2011. Language-in-education policy. In M. Strubell and E. Boix-Fuster, eds., *Democratic Policies for Language Revitalisation: the Case of Catalan*, Basingstoke: Palgrave Macmillan.
- Vila i Moreno, F.X. and Galindo Solé, M. 2009. El sistema de conjunció en català en l'educació primària a Catalunya: Impacte sobre els usos. *Treballs de Sociolingüística Catalana* 20, 21–69.
- Wei, L. 1998. The 'why' and 'how' questions in the analysis of conversational code-switching. In P. Auer, ed., *Code-switching in Conversation: Language, interaction and identity*, London: Routledge.
- Weinreich, U. 1953. *Languages in contact: Findings and problems*. The Hague: Mouton, seventh printing, 1970 edition.
- Wheeler, M. et al. 1999. *Catalan. A comprehensive grammar*. London: Routledge.
- Winford, D. 2003. *An introduction to contact linguistics*. Oxford: Blackwell.
- Wolfram, W. and Schilling-Estes, N. 1998. *American English*. Oxford: Blackwell, 2nd edition.
- Woolard, K. 1989. *Double talk: bilingualism and the politics of ethnicity in Catalonia*. Stanford, CA: Stanford University Press.

# Appendices

The appendices which follow contain information related to the fieldwork experiment. Appendix A consists of the pre-pilot questionnaire (detailed in section 4.1.2.2) which was used to gather information as to the nature of contemporary Catalan and Castilian speech practices in Barcelona. These findings went on to inform the content of the ultimate experiment. Appendix B contains the different stimuli used in the pilot study (explored in depth in section 4.2), which was then modified and refined. Appendices C, D and E contain material administered in the final fieldwork experiment: the stimulus texts, participant details form and consent form, respectively. Finally, appendix F consists of a diagram which represents the network of the fifty fieldwork participants.

## Appendix A: Pre-Pilot Questionnaire

The twenty pre-pilot participants only received the original Catalan questions which follow. No translation into Castilian was provided. Here, an English translation is offered afterwards for clarity.

### Catalan original

Com a part de les meves investigacions de doctorat, examinaré la situació sociolingüística a Barcelona, enfocant els canvis estructurals que existeixen en les llengües catalana i castellana. Per tant, li agrairia qualsevol comentari que em pogués fer, encara que sigui molt anecdòtic. En la mesura en què sigui possible, faci referència

a la situació de Barcelona. Tanmateix, si vostè coneix una situació semblant que em pugui ajudar, em faci una descripció aquí.

1. Com parla català la gent jove (és a dir, les persones que tenen menys de 35 anys)? Com parla català la gent més gran? Existeixen factors lingüístics que distingeixen els dos grups? Per exemple, si una persona gran digués que els joves catalans parlen d'una forma 'incorrecta', a quin tipus de costum lingüístic es referiria? Així mateix, si una persona jove digués que la gent gran parlen d'una forma 'antiquada', a qué es referiria? Li agrairia qualsevol comentari que em pogués fer, sigui de sintaxi, de fonologia o de vocabulari.
2. Llegeixi la pregunta 1, i torni a contestar, fent referència a la llengua castellana: Com parla castellà la gent gran / jove? Existeixen factors lingüístics que distingeixen els dos grups? Li agrairia qualsevol exemple que em pogués donar, encara que fos molt trivial.
3. Si vostè coneix alguna situació que em pugui ajudar (pel que fa el català, el castellà o qualsevol idioma), em faci una descripció aquí.

### **English translation**

As part of my PhD studies, I will be examining the sociolinguistic situation in Barcelona, focusing on the structural changes which exist in Catalan and Castilian. Therefore, I would be most grateful for any commentary that you could offer, however anecdotal. As far as possible, please make reference to the situation in Barcelona. However, if you are familiar with a similar situation that could help me, feel free to mention it here.

1. How do young people speak Catalan (that is to say, people under 35 years of age)? How do older people speak Catalan? Are there any linguistic factors which distinguish the two groups? For example, if an older person were to say that young people speak Catalan 'incorrectly', to which kind of linguistic practice would they be referring? Similarly, if a younger person were to say that

older people speak in an ‘old-fashioned’ way, to what would they be referring? I would be most grateful for any comments you could make, be they related to syntax, morphology or vocabulary.

2. Read question 1, and answer it again, this time making reference to Castilian. How do younger / older people speak Castilian? Do any linguistic factors exist which distinguish the two groups? I would appreciate any example you could offer, however trivial.
3. If you are familiar with any situation that could help me (with regard to Catalan, Castilian or any other language), please mention it here.

## Appendix B: Pilot Study Materials

The pilot study consisted of the identification of non-normative language use in two prepared texts (much as in the final experiment), followed by an exercise where vocabulary was elicited using a stimulus image (not included in the final experiment).

The following stimulus texts were offered to the four pilot study participants in a randomised order. Participants only saw the original (i.e. Catalan and Castilian) versions. The English translations are provided for clarity. Line numbers did not appear on the original stimuli, but are included here to aid reference when examining the instances of non-normative language use as detailed in section 4.3.

### Catalan pilot dialogue

- 1 - Quina il·lusió veure’t així! Fa estona que no xerrem de debò.
- 2 - Tens raó. Tens una casa molt maca.
- 3 - Moltes gràcies. Vols prendre alguna cosa? Tinc una munt de xocolata en aquell
- 4 armari, baix la televisió. En vols?
- 5 - No gràcies. Collons, fa molt fred, oi?!
- 6 - Sí però esclar nena, desembre és així. Per això no fa calor. Poso la calefacció.
- 7 - Doncs, què vas fer anit?

- 8 - Vaig veure aquella pel·lícula, saps? Sobre el tiburó. Com es diu?
- 9 - Ah sí sí. No recordo, però en anglès és Jaws.
- 10 - Això mateix.
- 11 - No hauria hagut de veure-la. La meva mare sempre em renya per veure pel·lícules  
12 que fan por.
- 13 - Potser és una tonteria, però quants anys tens?
- 14 - Tinc 17.
- 15 - Ai, perdona, no he pensat. Per això encara vius amb els teus pares.
- 16 - Això mateix. No és qüestió de falta de maduresa, sinó que sóc jove i prou! I tu?  
17 Què vas fer?
- 18 - Collons, ahir va ser un desastre!
- 19 - Per què? Què va passar?
- 20 - La meva germana va caure a la acera i es va quebrar una cama.
- 21 - Déu meu!
- 22 - Ja ho sé. Menys mal que només era la cama i no el cap.

### English translation of Catalan pilot dialogue

- What a nice surprise, seeing you like this! It's been ages since we really chatted.
- You're right. You've got a lovely house.
- Thanks. Do you want something to eat? I've got a load of chocolate in that  
cupboard over there, under the television. Do you want some?
- No thanks. Christ, it's cold, isn't it?!
- Yeah, but of course, that's December for you love. That's why it's not hot. I'll  
put the heating on.
- So, what did you do last night?
- I saw that film, you know? The one about the shark. What's it called?
- Oh yeah. I don't remember, but in English, it's Jaws.
- That's the one.
- I shouldn't have watched it. My mum always tells me off for watching scary  
films.

- I might be being stupid, but how old are you?  
- 17.  
- Ah, sorry of course, I wasn't thinking. That's why you still live with your parents.  
- Exactly. It's not a question of being immature, just that I'm young, full stop!  
And you? What did you do?  
- Christ, yesterday was a nightmare.  
- Why? What happened?  
- My sister fell over on the pavement and broke her leg.  
- My God!  
- I know. Thank God it was only her leg and not her head.

### Castilian pilot dialogue

1 - Venga tío, ¿por qué no te pones el abrigo? ¡Hace un frío que te mueres!  
2 - Tanto me da si me pongo malo. ¡Así me pillo una baja!  
3 - ¿Qué? ¿Qué dices? ¡No seas gilipollas! Ponte un abrigo, que estos días hace  
4 más frío que nunca. ¿No eres consciente que está cambiando el tiempo? En la tele,  
5 el Iñaki Gabilondo ha dicho que la mediana de hoy serían cinco grados.  
6 - ¡No seas mariquita! Hace un rato en el parque habían unas chicas con minifaldas,  
7 y no tenían nada de frío. Venga, vamos a entrar en calor. ¿Quieres tomar un café o  
8 algo? Y luego, cenamos, y nos vamos de juerga hasta que nos pongamos ciegos y nos  
9 hagan fuera. ¿Qué te parece?  
10 - ¿Con la crisis económica que estamos pasando? ¡Qué va! Vives en un mundo de  
11 fantasía. Te digo una cosa: la vida de ahora es algo diferente de la vida de hace dos  
12 años. Por cierto, ¿qué hora es?  
13 - Son las cuatro y media.  
14 - ¡Joder! Me tengo que ir. No sabía que fuera tan tarde... mi reloj debe de estar  
15 atrasado. Me voy porque me van a llevar una alfombra a las cinco, así que tengo que  
16 estar en casa cuando piquen a la puerta. Y encima, tengo que hacer una presentación  
17 mañana.  
18 - ¡Vaya excusa! La sabes muy larga, tú.



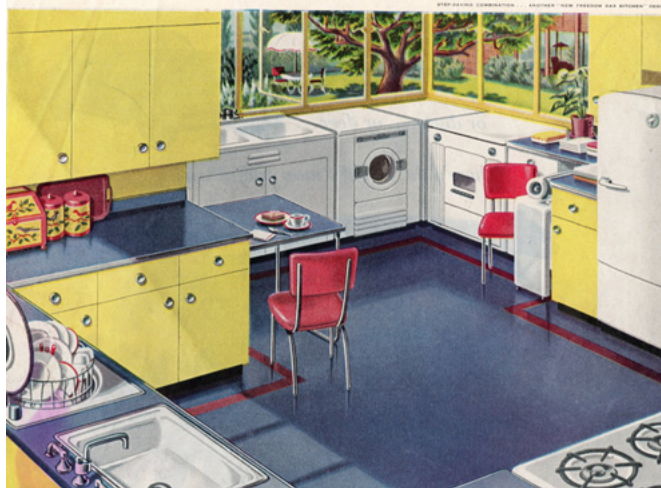
- 19 - Bueno, de todas formas mejor que me vaya.  
20 - Vale, pues muy bien. ¡Cuídate!

### English translation of Castilian pilot dialogue

- Come on mate, why don't you put your coat on? It's bloody freezing!
- I don't care if I get ill. That way, I get some time off work!
- What? What you on about? Don't be a twat! Put your coat on. These days it's colder than ever. You not aware that the climate's changing? On tele, Iñaki Gabilondo said that today's average would be five degrees.
- Don't be a little pussy! Just now in the park there were these girls in tiny mini skirts, and they didn't seem cold. Fine, let's go in the warm. You fancy a coffee or something? Then we can grab some dinner, and go out clubbing until we get twatted and they chuck us out. What do you think?
- What with the credit crunch and everything going on? No way! You live in a dream world, man. I'll tell you something: life is pretty different to the way it was two years ago. By the way, what time is it?
- Half four.
- Fuck! I've gotta go. I didn't know it was so late.... my watch must be slow. I'm off because I'm getting a rug delivered at five, so I've gotta be at home when they knock the door. On top of that, I've gotta give a presentation tomorrow.
- Wow, some excuse! You've got it all figured out.
- Well anyway, I'd best be off.
- Alright fine, take care.

### Image

After the text-based task, the pilot study participants were shown the following image and asked to name as many of the items which appear as possible. This was to gain insight into the availability of particular lexical items to each participant. If the interview was conducted in Catalan, they were asked to do this in Catalan. Likewise, if the interview was conducted in Castilian, this section was completed in Castilian.



## Appendix C: Final Fieldwork Experiment Texts

The following texts were offered to all fifty final experiment participants in a randomised order. Participants only saw the original (i.e. Catalan and Castilian) versions. The English translations are provided here for clarity. Again, line numbers did not appear on the original stimuli, but are included here to aid reference when examining the instances of non-normative language use as detailed in section 4.3.

### Catalan final dialogue

#### SAFAREIG

(dues persones a la feina)

- 1 - Escolta'm, Elena, has vist en Joan avui?
- 2 - Sí, és a l'hospital!!
- 3 - I això?
- 4 - Que no ho saps? La setmana passada era la festa de despedida de la Marina.
- 5 Te'n recordes?
- 6 - Ah sí, sí. No vaig anar-hi. Era el meu aniversari.
- 7 - Ai, és veritat. No he pensat. Tant se val, en Joan va tenir un accident. Va caure
- 8 a l'acera, quebrant-se una cama.
- 9 - Déu meu!

- 10 - Sí, sí. Però no em sorprèn gens. Hi havia un munt d'ampolles de jerès baix  
11 l'ordenador del seu despatx i se n'havia begut ben bé una de sencera!!
- 12 - Déu n'hi do!
- 13 - Sí, sí. Ben borratxo, guinyant l'ull a tothom, encara que la seva dona fos al seu  
14 costat tota la nit. Tonteries així.
- 15 - Menys mal que hi era si al final el van haver de dur a l'hospital. Quina vergonya!
- 16 - Doncs el que et vas perdre, perquè el seu cap va anar a parlar amb ells! Òbvia-  
17 ment els ha demanat el certificat metge... però no m'hauria agradat estar en la seva  
18 situació!
- 19 - Per què? Què lis va dir?
- 20 - No ho sé. Però ja és prou grandet, no creus?
- 21 - Quants anys té, per cert?
- 22 - Té 35.
- 23 - Què dius?! Sembla més gran.
- 24 - Ai, és veritat, m'he liat, ara pensava en una altra persona.
- 25 - Deu tenir com uns 45 anys. Potser és la crisi dels 40.
- 26 - Potser. Però, de persones de la seva edat, n'hi ha a milions, i no cal beure tant!
- 27 - Ja tens raó. De totes maneres, no cal fer-ne safareig.

### English translation of Catalan final dialogue

#### GOSSIP

(two people at work)

- Hey Elena, you seen John today?
- Yeah, he's in hospital!!
- How come?
- You mean you don't know? Last week was Marina's leaving do. You remember?
- Ah yeah. I didn't go. It was my birthday.
- Ah true. I wasn't thinking. Anyway, John had an accident. He fell over on the pavement and broke his leg.
- Oh God!

- Yeah yeah. But I'm not at all surprised. There was a load of bottles of sherry under the computer in his office, and he must have necked a whole one!!

- Bloody hell!

- Yeah yeah. He was well drunk, winking at everyone even though his wife was right next to him all night. Stuff like that.

- Just as well she was there if at the end he was taken to hospital. How embarrassing!

- That's not all, because his boss went to speak with him. Obviously he asked them for the doctor's certificate... but I wouldn't have liked to have been in his shoes!

- Why? What did he say to them?

- I don't know, but he's a big boy to be doing that sort of thing, don't you think?

- How old is he, by the way?

- 35.

- What are you on about? He seems older.

- Oh yeah, you're right. I got confused there and was thinking about someone else.

- He must be about 45. Maybe it's a mid-life crisis.

- Maybe. But there are millions of people his age, and they all don't go around drinking that much!

- Yeah you're right. Anyway, we shouldn't be gossiping!

### Castilian final dialogue

#### DE NUEVO TÚ

(una conversación por teléfono)

1 - ¿Dígame?

2 - Hola, soy la Meritxell.

3 - Hola, guapa, ¡qué ilusión! Hace muchísimo tiempo que no nos vemos. Dos meses,  
4 me parece, ¿no? Tenemos que ponernos al día. Tengo muchas noticias: ¡mi vida es  
5 completamente diferente de mi vida de hace dos meses!

6 - ¡No me digas! ¿Quedamos mañana y me cuentas?

- 7 - Mañana me viene muy bien. ¿A qué hora?
- 8 - Pues yo plego a las siete, así que podemos quedar a las siete y media.
- 9 - Perfecto. ¿Qué quieres hacer?
- 10 - No sé, tú misma, pero prefiero salir en plan tranquilito, que tú eres muy juerguista
- 11 y no quiero acabar como la última vez, que estuvimos bebiendo hasta que nos hicieron
- 12 fuera y al día siguiente tuve una resaca... Además, tuve gripe la semana pasada y
- 13 aún me duelen las orejas y el cuello.
- 14 - ¡Vaya excusa! La sabes muy larga tú. Pues nada, ¿que vamos a tomar un café?
- 15 - Sí, muy bien, ¿conoces el café Virreina?
- 16 - No, pero será en la plaza Virreina, supongo.
- 17 - Exacto, con Encarnació. Reconocerás la cafetería fácilmente porque enfrente hay
- 18 una farola que tiene una estatua de la Sagrada Familia sobre.
- 19 - ¡Sí, ya, ya, un sitio un poco cutre!, pero de todas formas prefiero tomar un café
- 20 ahí, porque la verdad es que no tengo mucha pasta. Mi compañero de piso se fue de
- 21 vacaciones la semana pasada sin pagar el alquiler. Se ha ido con todo el dinero y me
- 22 ha dejado sin.
- 23 - Jo, ¡qué fuerte! ¡Ahora eres consciente que todos los tíos son unos gilipollas!
- 24 - ¡Jajá, pues sí! Nos vemos mañana, entonces. Y escúchame, reina, abrígate bien.
- 25 En la tele han dicho que la mediana de mañana será cinco grados y no quiero que te
- 26 pongas mala otra vez, después de la gripe que has tenido, ¿vale?
- 27 - Vale, niña, tranquila. Escucha, tengo que irme, alguien está picando a la puerta.
- 28 - Venga, hasta mañana, entonces.
- 29 - Hasta mañana, cuídate.
- 30 - Un beso, chao.
- 31 - Chao.

### English translation of Castilian final dialogue

**YOU AGAIN**

(a telephone conversation)

- Hello?

- Hello, it's Meritxell.

- Hi love, what a nice surprise! It's been ages since we've seen each other. Two months, right? We've got to catch up. I've got so much news for you: my life right now is totally different to my life two months ago!

- No way! Let's meet up tomorrow and you can tell me all about it.

- Tomorrow's fine by me. What time?

- Well I finish work at seven, so we could say seven thirty.

- Perfect. What do you want to do?

- I don't know, it's up to you, but I'd rather just have a quiet one. You like getting wasted, and I don't wanna end up like last time. Drinking until they kicked us out, and then I had a hangover the next day... Also, I had the flu last week and my ears and throat still hurt.

- Wow, some excuse! You've got it all figured out. Never mind, what about a coffee?

- Cool, do you know the Virreina café?

- No, but I'm guessing it's in Virreina square.

- Exactly, on the corner with Encarnació street. You'll easily recognise the café because there's a lamppost opposite with a statue of the Sagrada Familia on top.

- Yeah yeah, I know, a shitty little place. But either way, I'd rather have a coffee there, because the truth is, I don't have much money. My flatmate went on holiday without paying the rent. He went with all the money and left me without.

- Bloody hell, that's a bit much! Now you're aware that all men are bastards!

- Haha, well quite! So we'll see each other tomorrow then. And listen, love, wrap up warm. On the tele they said that the average temperature tomorrow would be five degrees, and I don't want you to get ill again, after that flu you had, okay?

- Alright babe, fine. Listen, I've gotta go, someone's knocking at the door.

- Cool, see you tomorrow then.

- See you tomorrow, take care.

- Ta-ra.

- Bye.

## Appendix D: Participant Personal Details Form

Participants were asked whether they wished to fill in the Catalan or Castilian version of the form. An English translation is provided for clarity.

### Catalan

Referència:

Edat:

Lloc de naixement:

Professió:

Quant temps fa que vius a Barcelona?:

A quina escola / quin col·legi anàves? (educació primària):

A quin institut anàves? (educació secundària):

Llengua que parles amb la família:

Castellà / Català / Ambdues / Altra

Llengua que parles amb els amics:

Castellà / Català / Ambdues / Altra

Nivell d'estudis que tens / estàs realitzant (indica el nivell més alt):

ESO

Batxillerat (o equivalent: Grau Mitjà de FP, de disseny etc.)

Grau Superior de FP, de disseny etc.

Educació universitària: Estudis de 1r cicle

Educació universitària: Estudis de 2n cicle

Educació universitària: Estudis de 3r cicle (doctorat)

Durant els teus estudis, en quina llengua es feien normalment les classes?:

Castellà / Català / Ambdues / Altra

Lloc de naixement del pare:

Professió del pare:

Lloc de naixement de la mare:

Professió de la mare:

Llengua	Quants anys tenies quan vas aprendre aquesta llengua?	On vas aprendre aquesta llengua? ( <i>a casa, a l'escola etc.</i> )
Castellà		
Català		
Altra		

*Indica la resposta adequada:*

	Sí, perfectament			No gens	
	1	2	3	4	5
<b>Castellà</b>					
El sé parlar.	1	2	3	4	5
El sé escriure.	1	2	3	4	5
El sé llegir.	1	2	3	4	5
L'entenc.	1	2	3	4	5
<b>Català</b>					
El sé parlar.	1	2	3	4	5
El sé escriure.	1	2	3	4	5
El sé llegir.	1	2	3	4	5
L'entenc.	1	2	3	4	5
<b>Altre idioma</b>					
El sé parlar.	1	2	3	4	5
El sé escriure.	1	2	3	4	5
El sé llegir.	1	2	3	4	5
L'entenc.	1	2	3	4	5



## Castilian

Referencia:

Edad:

Lugar de nacimiento:

Profesión:

¿Cuánto tiempo llevas en Barcelona?:

¿A qué escuela / colegio ibas? (educación primaria):

¿A qué instituto ibas? (educación secundaria):

Lengua que hablas con la familia:

Castellano / Catalán / Las dos / Otra

Lengua que hablas con los amigos:

Castellano / Catalán / Las dos / Otra

Nivel de estudios que tienes / estás realizando (indica el nivel más alto):

ESO

Bachillerato (o equivalente: Grado Medio de FP, de diseño etc.)

Grado Superior de FP, de diseño etc.

Enseñanza universitaria: Estudios de 1r ciclo

Enseñanza universitaria: Estudios de 2o ciclo

Enseñanza universitaria: Estudios de 3r ciclo (doctorado)

¿Durante tus estudios en qué lengua se impartían normalmente las clases?:

Castellano / Catalán / Las dos / Otra

Lugar de nacimiento del padre:

Profesión del padre:

Lugar de nacimiento de la madre:

Profesión de la madre:

Lengua	¿Cuántos años tenías cuando aprendiste esta lengua?	¿Dónde aprendiste esta lengua? ( <i>en casa, en la escuela etc.</i> )
Castellano		
Catalán		
Otra		

*Rodea la opción adecuada:*

	Sí, perfectamente			No, para nada	
	1	2	3	4	5
<b>Castellano</b>					
Lo sé hablar.	1	2	3	4	5
Lo sé escribir.	1	2	3	4	5
Lo sé leer.	1	2	3	4	5
Lo entiendo.	1	2	3	4	5
<b>Catalán</b>					
Lo sé hablar.	1	2	3	4	5
Lo sé escribir.	1	2	3	4	5
Lo sé leer.	1	2	3	4	5
Lo entiendo.	1	2	3	4	5
<b>Otro idioma</b>					
Lo sé hablar.	1	2	3	4	5
Lo sé escribir.	1	2	3	4	5
Lo sé leer.	1	2	3	4	5
Lo entiendo.	1	2	3	4	5

## English

Reference:

Age:

Place of birth:

Occupation:

How long have you lived in Barcelona?:

What primary school did you attend?:

What secondary school did you attend?:

Language spoken with family:

Castilian / Catalan / Both / Other

Language spoken with friends:

Castilian / Catalan / Both / Other

Level of studies you possess / are currently undertaking (indicate the highest level):

GCSE equivalent

A-Level equivalent (or corresponding vocational qualification)

Higher vocational qualification

University: 1st cycle (three years of university study)

University: 2nd cycle (four years of university study)

University: 3rd cycle (doctorate achieved)

During your studies, what language were classes normally delivered in?:

Castilian / Catalan / Both / Other

Father's place of birth:

Father's occupation:

Mother's place of birth

Mother's occupation:

Language	How old were you when learned this language?	Where did you learn this language? ( <i>at home, at school etc.</i> )
Castilian		
Catalan		
Other		

*Circle the appropriate response:*

Yes, perfectly

No, not at all

1

2

3

4

5

**Castilian**

I can speak it. 1 2 3 4 5

I can write it. 1 2 3 4 5

I can read it. 1 2 3 4 5

I understand it. 1 2 3 4 5

**Catalan**

I can speak it. 1 2 3 4 5

I can write it. 1 2 3 4 5

I can read it. 1 2 3 4 5

I understand it. 1 2 3 4 5

**Other language**

I can speak it. 1 2 3 4 5

I can write it. 1 2 3 4 5

I can read it. 1 2 3 4 5

I understand it. 1 2 3 4 5

## Appendix E: Participant Consent Form



### Formulari de consentiment

Signeu aquest formulari si l'investigador us ha explicat els objectius de la investigació.

Títol de l'estudi: 'Change from below' and 'change from above': their impact on contemporary speech practices in Barcelona.

Queen Mary Research Ethics Committee Ref: QMREC2009/0158

- Moltes gràcies per haver participat en una entrevista. L'investigador us explicarà els objectius del projecte abans que signeu aquest formulari.
- Si teniu preguntes o dubtes sobre la investigació, demaneu-les a l'investigador abans de signar. L'investigador us donarà una còpia personal d'aquest formulari, que podeu consultar en qualsevol moment.
- En qualsevol moment, podeu retirar-vos del projecte si no voleu participar-hi per qualsevol motiu. En aquest cas, només cal notificar a l'investigador i la vostra contribució serà retirada immediatament.

#### **Declaració del participant:**

Jo, ....., declaro que l'investigador m'ha explicat aquest projecte i jo vull participar-hi. No trobo cap problema amb la gravació de la meva entrevista.

Signatura:

Data:

#### **Declaració de l'investigador:**

Jo, James Hawkey, declaro que he explicat bé el contingut i els objectius de la investigació.

Signatura:

Data:

Participants only received the above Catalan form. No translation into Castilian was provided. The following English translation is offered here for clarity.



**Consent form**

Please sign the following form if the researcher has explained the aims of the project.

Title of study: 'Change from below' and 'change from above': their impact on contemporary speech practices in Barcelona.

Queen Mary Research Ethics Committee Ref: QMREC2009/0158

- Many thanks for having participated in an interview. The researcher will explain the objectives of the project before you sign the present form.
- If you have questions or doubts about the project, please ask the researcher before signing. You will also receive your own copy of this form, which you can consult at any time.
- At any point, you can withdraw from the project if you no longer wish to participate for whatever reason. In this case, merely contact the researcher and your contribution will be immediately withdrawn.

**Participant declaration:**

I, ....., declare that the researcher has explained this project, and I wish to participate. I have no problem with my interview being recorded.

Signed:

Date:

**Researcher declaration:**

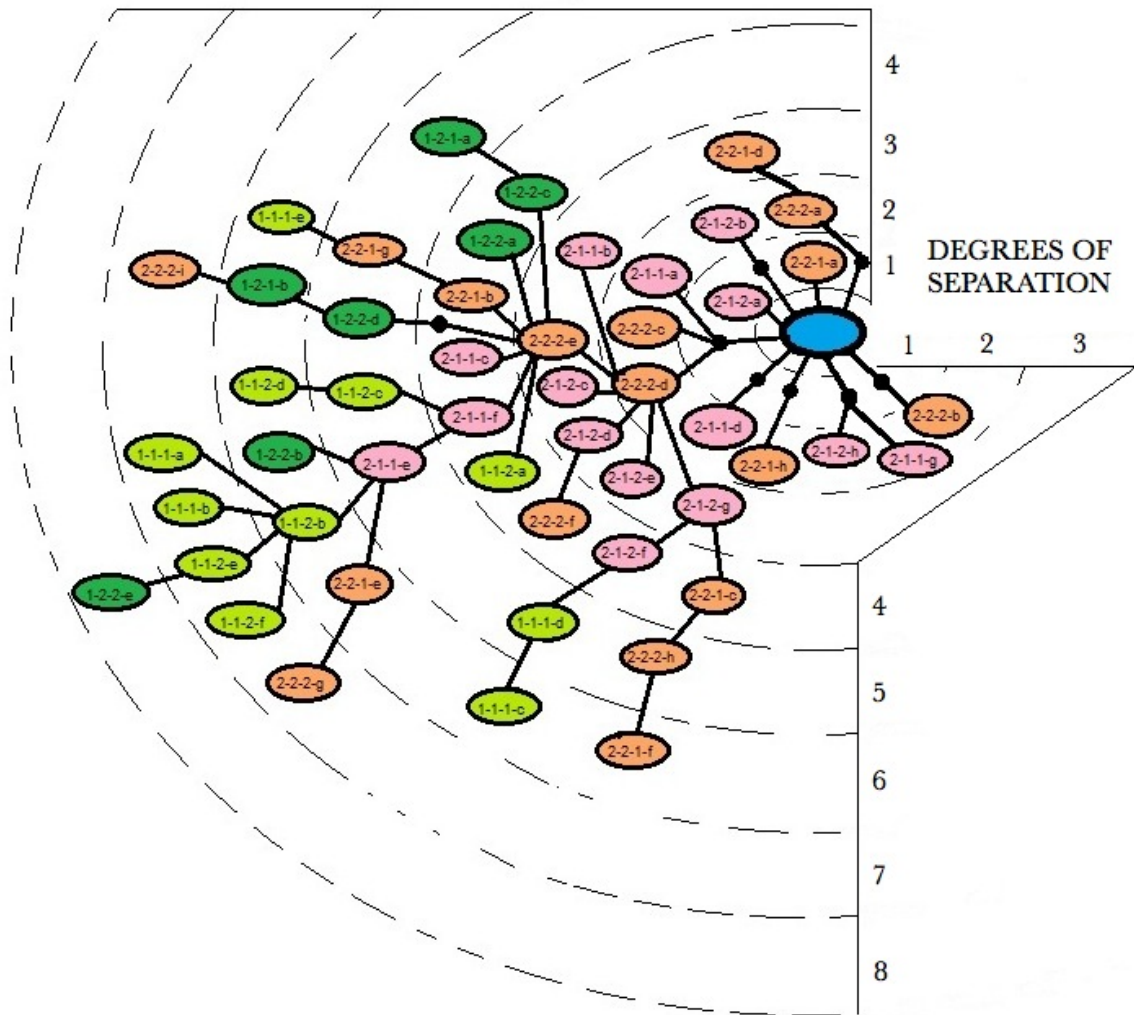
I, James Hawkey, declare that I have fully explained the content and aims of the study.

Signed:






Date:

## Appendix F: Network of participants

Participants' anonymised references are used in the diagram, which is also colour-coded for ease of use. The differently coloured circles represent participants of different sex and age. Thick black lines indicate that one participant introduced the next to the study, while black dots stand for people who were ineligible or unable to participate, but nevertheless introduced subsequent participants to the experiment.



Key

Older Female		Younger Female	
Older Male		Younger Male	
Researcher		Non-participant	