

Teaching and Learning Arabic Variation Through Vocabulary

A mixed-methods study on diglossic vocabulary building for higher-education students of Arabic as a foreign language

Submitted by Giorgia Ferrari to the University of Exeter as a PhD dissertation towards the degree of Doctor of Philosophy in Arab and Islamic Studies,
January 2018

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

I certify that all material in this dissertation which is not my own work has been identified with appropriate acknowledgement and referencing and I also certify that no material is included for which a degree has previously been conferred upon me.

Signature

Abstract

The field of Teaching Arabic as a Foreign Language (TAFL) has seen in recent decades a growing interest in portraying and teaching one of the most salient and intrinsic features of Arabic: language variation. This thesis takes a position in contrast to approaches that portray the two varieties as being distinct and well-defined dichotomic units, in favour of an approach that interprets them as two heterogeneous language varieties within one singular linguistic system. The two language varieties are embodied by Standard and Colloquial Arabic and it is argued here for the teaching of both varieties to students of Arabic as a foreign language. In this light, this thesis sets out to investigate the development of two language skills, vocabulary knowledge and language awareness, in a diglossic learning environment. Moreover, it explores the attitudes and perceptions of the students towards Arabic variation.

Two experimental methods based on focus-on-form instruction are used in this research to teach Colloquial Arabic to students of Arabic as a foreign language at higher-education level, and the empirical research is conducted within a semi-embedded research design in which qualitative and quantitative data are collected. Students from three universities participate in this research: the Universities of Exeter, Genoa and Milan. This allows for the comparison of results from students of different mother tongues. The main research question that this thesis sets out to answer is: does focus-on-form instruction lead to vocabulary development in two diglossic varieties, namely Standard and Colloquial Arabic, more effectively when it focuses on the two varieties separately or when it links their forms? Two sub-questions investigate which of the two methods of focus-on-form instruction lead more efficiently to the development of language awareness, and the impact they have on students' attitudes towards Arabic variation. The last sub-question asks to what extent the development of the diglossic language skills and attitudes is a consequence of the method of instruction received. The results of this study suggest that the answer lies in focusing predominantly on one variety at a time with additional consolidation exercises that compare the forms of the two varieties.

The main contributions of this thesis are both theoretical, to the literature of TAFL, and empirical, regarding the development of the language skills and attitudes measured.

Table of content

1	CHAPTER 1 – Introduction	11
1.1	Introductory words	11
1.2	Research questions.....	12
1.3	Organisation of the study.....	14
2	CHAPTER 2 – Literature Review	18
2.1	Introduction.....	18
2.2	Arabic varieties	19
2.2.1	Classical Arabic.....	20
2.2.2	Dichotomy between spoken varieties and Classical Arabic	21
2.2.3	Development of Standard Arabic	23
2.2.4	Standard Arabic and Colloquial Arabic	25
2.3	Diglossia	27
2.3.1	Criticism of Ferguson’s model of diglossia	31
2.3.2	Standard and prestige	34
2.3.3	Variationist approaches to Arabic variation	37
2.3.4	Educated Spoken Arabic.....	39
2.4	Approaches of this research to Arabic diglossia.....	42
2.4.1	The reference-packagings model.....	43
2.4.2	Diglossic code-switching	47
2.5	Implications of Arabic variation for TAFL.....	50
2.6	Conclusion.....	56
3	CHAPTER 3 – Theoretical Framework.....	60
3.1	Introduction.....	60
3.2	The systematic nature of variation.....	63
3.2.1	Intrinsic variation of SA	63
3.2.2	Intrinsic variation of CA	65

3.3	Diglossic code-switching use in real-life communication.....	68
3.3.1	Intra-dialect conversations	70
3.3.2	Cross-Dialect Interactions	71
3.3.3	Written productions	73
3.4	Arabic proficiency	75
3.5	Focus of investigation.....	77
3.5.1	Diglossic vocabulary development.....	78
3.5.2	Language awareness.....	84
3.5.3	Students' perceptions and attitudes of Arabic variation	86
3.6	Conclusion.....	86
4	CHAPTER 4 – Methodology.....	89
4.1	Introduction.....	89
4.2	Review of focus of investigation	91
4.3	Focus-on-form instruction.....	95
4.4	Research method	100
4.4.1	Mixed-Methods Research Design	100
4.4.2	Data Collection and Analysis Methodology	105
4.5	Selection of participants	116
4.5.1	Selection of the Universities.....	118
4.5.2	Pilot study.....	119
4.6	Language material.....	123
4.6.1	Orthography of language material.....	124
4.6.2	Activities and exercises of language material	127
4.7	Conclusion.....	129
5	CHAPTER 5 – Diglossic vocabulary development	132
5.1	Introduction.....	132
5.2	Statistical formulae	134
5.3	Language tests.....	136
5.3.1	Design of pre-test.....	136
5.3.2	Design of post-test and delayed post-test.....	140
5.4	Results on language tests of participant universities.....	147

5.5	Results of subgroups on diglossic vocabulary development.....	155
5.5.1	University of Milan.....	156
5.5.2	University of Genoa.....	161
5.5.3	University of Exeter.....	163
5.6	Comparative analysis of results on diglossic vocabulary building	167
5.6.1	Low-average proficiency subgroups	168
5.6.2	Average proficiency subgroups.....	170
5.6.3	High-average proficiency subgroups.....	172
5.7	Conclusion.....	173
6	CHAPTER 6 – Language awareness	176
6.1	Introduction.....	176
6.2	Data collection on language awareness.....	178
6.3	Results on language awareness of participant universities.....	181
6.4	Results of subgroups on language awareness.....	188
6.4.1	University of Milan.....	189
6.4.2	University of Genoa.....	193
6.4.3	University of Exeter	196
6.5	Comparative analysis of results on language awareness	199
6.5.1	Low-average proficiency subgroups	199
6.5.2	Average proficiency subgroups.....	201
6.5.3	High-average proficiency subgroups.....	203
6.6	Conclusion.....	205
7	CHAPTER 7 – Perceptions and attitudes towards variation	207
7.1	Introduction.....	207
7.2	Questionnaire administered before language course.....	208
7.2.1	Reasons for learning Arabic.....	208
7.2.2	Language skills development and teaching methodologies	210
7.2.3	Students’ attitudes towards SA	213
7.3	Questionnaire administered after language course.....	214
7.3.1	Overall experience of studying Arabic simultaneously.....	214
7.3.2	Integration of CA into a SA programme	217
7.3.3	Influence of CA on SA passive and active skills	219

7.4	Conclusion.....	221
8	CHAPTER 8 – Conclusion.....	224
8.1	Introduction.....	224
8.2	Position of the research towards TAFL	227
8.3	Main findings and contribution.....	229
	8.3.1 Diglossic vocabulary development.....	229
	8.3.2 Diglossic language awareness.....	232
	8.3.3 Students’ perceptions of variation	236
8.4	Limitations	238
8.5	Conclusion.....	239
	Appendices.....	243
I)	Questionnaires	243
	APPENDIX A Questionnaire administered before language test	243
	APPENDIX B Questionnaire administered after language test	245
II)	Language Material for Experimental Group.....	246
	APPENDIX C Experimental Group - Lesson 1	246
	APPENDIX D Experimental Group - Lesson 2	250
	APPENDIX E Experimental Group - Lesson 3	257
	APPENDIX F Experimental Group - Lesson 4.....	260
	APPENDIX G Experimental Group - Lesson 5	263
	APPENDIX H Experimental Group - Lesson 6	267
III)	Language Material for Control Group.....	269
	APPENDIX I Control Group - Lesson 1	269
	APPENDIX J Control Group - Lesson 2	273
	APPENDIX K Control Group - Lesson 3.....	279
	APPENDIX L Control Group - Lesson 4	282
	APPENDIX M Control Group - Lesson 5	285
	APPENDIX N Control Group - Lesson 6.....	289
	Bibliography	290

Table List

Table 1	Qualitative and quantitative phases of the research	103
Table 2	Components of the empirical research.....	113
Table 3	Matching exercise (pre-test).....	138
Table 4	Identification of groups of vocabulary words as CA or SA items...	142
Table 5	Language tests, University of Milan	149
Table 6	Language tests, University of Genoa	152
Table 7	Language tests, University of Exeter.....	154
Table 8	Diglossic vocabulary in low-average control group, Milan.....	156
Table 9	Diglossic vocabulary in low-average experimental group, Milan	157
Table 10	Diglossic vocabulary in average control group, Milan	158
Table 11	Diglossic vocabulary in average experimental group, Milan	158
Table 12	Diglossic vocabulary in high-average control group, Milan	159
Table 13	Diglossic vocabulary in high-average experimental group, Milan ..	160
Table 14	Diglossic vocabulary in low-average control group, Genoa	161
Table 15	Diglossic vocabulary in low-average experimental group, Genoa..	161
Table 16	Diglossic vocabulary in average control group, Genoa	162
Table 17	Diglossic vocabulary in average experimental group, Genoa	162
Table 18	Diglossic vocabulary in low-average control group, Exeter.....	163
Table 19	Diglossic vocabulary in low-average experimental group, Exeter ..	163
Table 20	Diglossic vocabulary in average control group, Exeter	164
Table 21	Diglossic vocabulary in average experimental group, Exeter.....	165
Table 22	Diglossic vocabulary in high-average control group, Exeter	166
Table 23	Diglossic vocabulary in high-average experimental group, Exeter.	166
Table 24	Diglossic vocabulary across low-average control groups.....	168
Table 25	Diglossic vocabulary across low-average experimental groups	168
Table 26	SD in diglossic vocabulary of low-average control groups	169
Table 27	SD in diglossic vocabulary of low-average experimental groups ...	169
Table 28	Diglossic vocabulary across average control groups	170
Table 29	Diglossic vocabulary across average experimental groups	170
Table 30	SD in diglossic vocabulary of average control groups.....	171
Table 31	SD in diglossic vocabulary of average experimental groups.....	171
Table 32	Diglossic vocabulary across high-average control groups	172
Table 33	Diglossic vocabulary across high-average experimental groups ...	172
Table 34	SD in diglossic vocabulary of high-average control groups.....	173
Table 35	SD in diglossic vocabulary of high-average experimental groups..	173
Table 36	Language awareness in control group, Milan.....	181
Table 37	Language awareness in experimental group, Milan	182
Table 38	Language awareness in control group, Genoa	184

Table 39	Language awareness in experimental group, Genoa.....	184
Table 40	Language awareness in control group, Exeter.....	186
Table 41	Language awareness in experimental group, Exeter.....	186
Table 42	Language awareness in low-average control group, Milan.....	189
Table 43	Language awareness in low-average experimental group, Milan ..	189
Table 44	Language awareness in average control group, Milan.....	190
Table 45	Language awareness in average experimental group, Milan.....	191
Table 46	Language awareness in high-average control group, Milan.....	192
Table 47	Language awareness in high-average experimental group, Milan .	192
Table 48	Language awareness in low-average control group, Genoa.....	193
Table 49	Language awareness in low-average experimental group, Genoa	193
Table 50	Language awareness in average control group, Genoa.....	194
Table 51	Language awareness in average experimental group, Genoa.....	195
Table 52	Language awareness in low-average control group, Exeter.....	196
Table 53	Language awareness in low-average experimental group, Exeter	196
Table 54	Language awareness in average control group, Exeter.....	197
Table 55	Language awareness in average experimental group, Exeter.....	197
Table 56	Language awareness in high-average control group, Exeter.....	198
Table 57	Language awareness in high-average experimental group, Exeter	198
Table 58	Language awareness across low-average control groups.....	199
Table 59	Language awareness across low-average experimental groups ...	200
Table 60	SD in language awareness of low-average control groups.....	200
Table 61	SD in language awareness of low-average experimental groups ..	201
Table 62	Language awareness across average control groups.....	201
Table 63	Language awareness across average experimental groups.....	202
Table 64	SD in language awareness of average control groups.....	202
Table 65	SD in language awareness of average experimental groups.....	203
Table 66	Language awareness across high-average control groups.....	203
Table 67	Language awareness across high-average experimental groups..	204
Table 68	SD in language awareness of high-average control groups.....	204
Table 69	SD in language awareness of high-average experimental groups .	205

Figure List

Figure 1	Anderson and Krathwohl's revision of Bloom's taxonomy.....	80
----------	--	----

Abbreviation and Acronyms

AFL	Arabic as a foreign language
CA	Colloquial Arabic
CLT	Communicative language teaching
COV	Coefficient of variation
ESA	Educated spoken Arabic
FSA	Formal spoken Arabic
FFI	Form-focused instruction
H	High variety
L	Low variety
MLF	Matrix Language Frame
MSA	Modern Standard Arabic
SA	Standard Arabic
SD	Standard Deviation
TAFL	Teaching Arabic as a foreign language

Acknowledgements

I would like to thank all who have contributed to the completion of this project with their intellectual, or emotional support. Firstly, I would like to thank the Institute of Arab and Islamic Studies of the University of Exeter, for providing me with administrative and academic support during the PhD process. I would like to thank my supervisor John Peter Cooper for his unwavering support across the different phases of this thesis. I would like also to thank my second supervisor Manuela Elisa Bibiana Giolfo for her constant and informative support and advices. In addition, the academics of the University of Exeter I collaborated with during the process have had key impact on my work thanks to their inestimable intellectual support. I would like in particular to thank Fatima Rawan for her invaluable support and advice both personal and professional.

My heartfelt gratitude goes to Ermanna, Laura and Norma for their continuous support over the years and for pushing me to embrace challenging professional decisions.

My parents and my entire family are to be acknowledged for believing in me and supporting each of the decisions I have made through this process. The completion of this thesis would have not been possible without the support of Riaz, who embarked on this journey with me and remained by my side despite my erratic behaviour.

Finally, I wish to thank all the friends who have accompanied me along the PhD process, particularly Ana, for having shared every day of this path with me and for showing me what *home* truly means.

There is not sufficient space to name each of the friends separately that have supported me in Exeter, Italy, Morocco, and Oxford. Many thanks to the many friends, and to the Unicorns PhD community, for their unconditional support, motivation and encouragement during these years.

1 CHAPTER 1 – Introduction

1.1 Introductory words

The field of Arabic-as-a-foreign-language acquisition has been characterised until recently by a great focus on the teaching of grammar, with limited attention to active and communicative language skills such as speaking. However, as Zeinab Taha (1995: 175) puts it, “with the emergence of the proficiency movement, the rising awareness of the learners’ needs, and the shifting focus towards communicative competence, two major camps have arisen. One camp emphasizes the role of grammar in the classroom and maintains that if the students are not taught grammar formally, their linguistic skills will not develop properly. The other camp, however, emphasizes fluency and communication, and deemphasizes the formal teaching of grammar in the classroom.” On this dichotomic basis, further methodologies of second language acquisition developed within the teaching of Arabic as a foreign language, with the aim to combining communicative language teaching approaches with grammar-based activities and instruction.

Until recently, grammar-based methodologies have focused only on teaching the formal variety of Arabic under three main assumptions. First, it is considered to be the only proper variety to be used in higher-education environments; second, it is the only variety shared by the different regions of the Arab worlds and therefore the key to establishing communication with native Arabic speakers independently from their country of origin; and finally, on the supposition that knowledge of the formal variety facilitates the acquisition of the colloquial variety without the need for formal instruction in the latter. Nowadays, this approach to Arabic teaching is still the most widespread, as I elucidate in section 2.5. It is nevertheless challenged by an increasing number of approaches that argue for the teaching of both the formal and the colloquial varieties of Arabic, as both varieties constitute Arabic. An in-depth explanation of Arabic varieties is provided in section 2.2. Arabic varieties are within skilled-based programmes, in which SA is used for reading and writing, and it overlaps with CA in speaking and listening skills. Since these approaches have developed in recent decades, little research has been done on the impact that they have on language skills development. This is particularly true for language

skills developed within a diglossic environment and it entails the need for studies that focus on their acquisition and improvement. It is in the hope to respond to this need that my research is planned and carried out. The additional purpose of this research is to offer new insight into the advantages and disadvantages of combining the teaching of Colloquial and Standard Arabic, and this is achieved by focussing on focus-on-form methods of instruction. For purpose of clarity I would like to define the use of “diglossic vocabulary development” that I make in this thesis. By diglossic vocabulary development I mean vocabulary development of corresponding vocabulary in two diglossic varieties, namely Standard and Colloquial Arabic. This research aims at contributing to the field of diglossic vocabulary development within mainstream academic programmes and therefore it studies the effects of integrating the teaching of Colloquial Arabic into Standard Arabic-based academic programmes. Finally, this thesis aims at providing new observations and understanding of the effects the teaching of two varieties of Arabic has on the students’ perceptions and motivation towards variation. Students from the Universities of Exeter, Genoa and Milan participate in this study. This allows me to assess whether the two different mother tongues of the students have an impact on their acquisition of diglossic skills. The selection of the participant universities is explained in detail in section 4.5.1. The variables measured in this study are three, two of quantitative nature and one of qualitative nature. I outline them in the following section, together with the research questions.

1.2 Research questions

The research questions reflect the duality of the nature of the variables studied: quantitative on one side and qualitative on the other. This thesis investigates quantitatively the development of diglossic language skills, which are interpreted as being composed of (i) vocabulary retention of correspondent items and forms in two language varieties; (ii) ability to code-switch and code mix between the two varieties; and (iii) the reported presence of diglossic language awareness. The students’ perceptions of language variation in Arabic are also observed, and the data collected to inform the observation are qualitative in nature.

The four research questions are outlined as follows:

1 - Does focus-on-form instruction lead to diglossic vocabulary development more effectively, when it links forms in Standard and Colloquial Arabic, or when it focuses only on one variety?

2 - Is focus-on-form instruction more effective for diglossic language awareness, when it links vocabulary forms in Standard and Colloquial Arabic, or when it focuses only on one variety?

3 - What impact does focus-on-form instruction have on students' perceptions of Arabic variation, when it links vocabulary forms in Standard and Colloquial Arabic and when it focuses only on one variety?

4 - To what extent is the development of diglossic vocabulary knowledge a function of the method of focus-on-form instruction received?

The mixed-methods approach is reflected in the nature of the research questions: the first two questions are quantitative, the third is qualitative, and the fourth includes both quantitative and qualitative data. Quantitative and qualitative methods are integrated with the aim to achieving complementarity and obtaining an exhaustive understanding of the variables studied. I place the mixed-methods approach within an embedded quasi-experimental design. Embedded quasi-experimental designs are primarily based on one data set. In this research this is the case of the quantitative data set. The second data set, which is here the qualitative data set, has a secondary but complementary role. The research questions reflect the design of the study in that they call for both type of data in a proportion that represents the structure of the design, i.e. two questions are quantitative in nature, since quantitative data are at the core of this study, while only one research question determines the collection of qualitative data, since these have a secondary role. The last question entails the merging of all the data collected in the empirical research. The process of merging the data provides the research with the opportunity to gain in breadth and depth of understanding of the variables studied, as it provides information of different nature and collected with different methodologies. Moreover, mixing quantitative and qualitative data offsets the weaknesses inherent to using each approach by itself. The research design, the data collection process and the

methodological tools employed to analyse the data are provided in detail in chapter four.

1.3 Organisation of the study

Chapter 1 – Introduction

The first chapter of this thesis presents the key building blocks of this thesis. It also aims at defining the design of this thesis. To frame this project, it first outlines Arabic language variation and its impact on the teaching of Arabic as a foreign language. It elucidates one of the areas in this field in which research is most needed, and it explains how this thesis aims at contributing to that area of investigation. It therefore identifies the aim of study of this thesis and its research questions. The first chapter also mentions the mixed-methods research design that is employed in this thesis and the rationale for using it. It finally mentions how the research questions lead to the collection of different types of quantitative and qualitative data. The details of the research methodology and of the types of data to be collected are provided in the methodology chapter.

Chapter 2 – Literature Review

The literature review chapter is based on literature on Arabic language and, more specifically, on Arabic variation. The chapter also discusses the implications that Arabic variation has for the teaching of Arabic as a foreign language. The chapter is divided into three main sections. It firstly outlines the historical development of Arabic variation, with the aim to outlining the roles that Arabic varieties have within the Arabic-speaking community. The second section defines the approach of this research to Arabic variation, which is interpreted in this thesis as intrinsically heterogeneous in nature, rather than characterised by distinct linguistic entities. The last section identifies the impact that Arabic variation has on TAFL, with a specific focus on higher-education instruction.

Chapter 3 – Conceptual and Theoretical Framework

This chapter explains the theoretical and conceptual framework which is going to be followed in the thesis. It also explains the main features of the threefold focus of investigation of this thesis, namely diglossic vocabulary development; diglossic language awareness; and students' perceptions of variation in Arabic language reality and learning. This is based on the literature on Arabic outlined in chapter two, which is used to ground the framework. This thesis interprets linguistic and metalinguistic abilities and attitudes of native speakers as being the main catalysts for variation in language use. This thesis adopts the 'reference-packagings' model, developed by Giolfo and Sinatora (2011) to frame its understanding of Arabic variation. The chapter is divided into three main sections. The first section describes in detail its interpretation of variation as an integral feature of all Arabic varieties, and it shows that diglossic code-switching and code-mixing are used at almost all levels and communication. The second section is based on the theoretical basis of the previous section and it identifies how this thesis interprets native-like proficiency and its main components: the knowledge of at least two language varieties and the ability to code-switch between them. The last section identifies the learning outcomes that this research adopts to measure the development of language knowledge in this thesis.

Chapter 4 – Methodology

The methodology chapter is divided into five sections and it defines all the methodological elements of this research. The first section recapitulates the main features of the focus of investigation of this thesis. The second section identifies the teaching approach used in this research to conduct the empirical language research, which is based on focus-on-form instruction. The third section describes the methods of data collection assessment, and how the data are instrumental to answering the research questions. The fourth part describes the selection of the participants for this research and the logistics of the

empirical research. The description of the language material and activities I designed and developed for the research are provided in the last section.

Chapter 5 – Diglossic vocabulary development

The fifth chapter is the first empirical chapter. It clarifies the observation of diglossic vocabulary development in the three universities observed in this research. It is divided into five sections. The first section defines the statistical formulae used to measure the development of diglossic vocabulary. The second section delineates the exercises of the language tests and how they verify and assess the knowledge of diglossic vocabulary and the ability to code-switch. The following sections analyse the data collected on diglossic vocabulary development. They focus respectively on: (i) comparing the results scored within the three universities; (ii) comparing the results scored by groups with similar proficiency in Standard Arabic within universities; and (iii) comparing the results scored by groups with similar proficiency in Standard Arabic across universities.

Chapter 6 – Diglossic language awareness

The sixth chapter is the second empirical chapter and it focuses on the development of diglossic language awareness. Its structure repeats the structure of the previous chapter. It firstly explains diglossic language awareness is measured, and it defines how the exercises of the language tests assess its knowledge. In the last three sections I report the data collected with the aim to: (ii) comparing the results scored by groups with similar proficiency in Standard Arabic within universities; and (iii) comparing the results scored by groups with similar proficiency in Standard Arabic across universities.

Chapter 7– Perceptions and attitudes towards variation

The sixth chapter is the last chapter of the empirical section and it focuses on observing the perceptions of the students towards Arabic variation. The chapter is divided into two sections. First, the answers of the participants to the

questionnaire conducted before the beginning of the language course are analysed. Second, I analyse the answers of the participants to the questionnaire conducted after the completion of the language course.

Chapter 8 – Conclusion

This last chapter presents the discussion on the key findings of this thesis. This chapter firstly reviews the contribution, innovation and findings of each of the three empirical chapters. It proceeds by answering the main research question and three subquestions. By doing so, it addresses both the hypotheses and assumptions. The first section reflects on each one of the empirical chapters of this thesis, which correspond to each aspect of the focus of investigation of this thesis: diglossic vocabulary building, which corresponds to chapter five; diglossic language awareness, which corresponds to chapter six; and students' motivations towards Arabic variation, which is chapter seven. The next section then gives the limitations of the thesis. The last section of the chapter finally merges and discusses the findings and observation together.

2 CHAPTER 2 – Literature Review

2.1 Introduction

This chapter sets out to define language variation in Arabic and discusses the implications of this variation for TAFL. The thesis is positioned within the contemporary discussion of Arabic as a multi-faceted language, characterised by the presence of multiple language varieties. The features of these varieties are distinct and are interpreted by every member of the speech community in a unique manner. This research regards Arabic language varieties as fulfilling different linguistic functions and reflecting distinct metalinguistic perspectives held by native speakers. It also interprets the choices determining language variety use as being determined by numerous linguistic, sociolinguistic and behavioural factors. While Arabic varieties alternate through diglossia, they are treated in this thesis as the realisation of a single unit. This is because despite their distinct characteristics, Arabic varieties are perceived by native speakers as being part of one cultural system. The following chapter is divided into three main sections.

The first section outlines the historical development of Arabic variation from the pre-Islamic period, before the seventh century CE, to the present day. Within this context I identify the distinctive elements of the three main language varieties: Classical Arabic, Standard Arabic and Colloquial Arabic. Understanding their historical development is crucial in elucidating their roles within the Arabic-speaking community, their differences and their overlapping features.

The second section analyses various linguistic interpretations of Arabic as a discretely stratified language characterised by clearly defined language levels and varieties. I critically analyse these approaches from a dynamic perspective which views Arabic language varieties as being intrinsically heterogeneous in nature, rather than homogeneous and distinct entities. Crucially, the latter approach interprets native speakers' language attitudes and linguistic proficiency as being the main catalysts for observed variation in language use. This thesis adopts the 'reference-packagings' model, developed by Giolfo and

Sinatora (2011) to elucidate my position towards the nature of Arabic variation and its use among native speakers.

The third, and last, section identifies the impact that the varied nature of Arabic has on TAFL, focussing specifically on higher-education instruction. It outlines the range of positions among academics and educators on the appropriate varieties of Arabic that ought to be taught in formal education. It also evaluates these stances against students' needs, expectations and preferences. Moreover, it investigates the effect that language choices within TAFL can have on the representation of Arabic linguistic nature. Traditional approaches to TAFL may portray Arabic as being composed of discrete segments which hold unequal levels of privilege, each of which is capable of being studied as a separate unit. Alternatively, they can paint a rich picture of a dynamic language in which linguistic varieties are nuances of the same entity. This thesis recognises that most existing teaching methodologies fall within the former model of TAFL and tend to privilege the teaching of Standard Arabic. The second model is currently less widespread, being comprised of experimental methodologies and a few well-established teaching approaches, and it includes the teaching of both Standard and Colloquial Arabic. The process of outlining both models of reference is fundamental to this thesis as the teaching methodology proffered in later chapters is based on innovative approaches to TAFL, with a view to incorporating them into mainstream practice. This study thereby aims to contribute to the existing research in the field of Arabic Applied Linguistics, and specifically the assessment of potential techniques to incorporate the teaching of Colloquial Arabic vocabulary within a Standard Arabic-based programme.

2.2 Arabic varieties

This section outlines the historical development and key features of the three main varieties of Arabic: Classical Arabic, Standard Arabic (hereafter SA) and Colloquial Arabic (hereafter CA). It highlights the role of Classical Arabic as the marker of pan-Arab identity and its role within the religious sphere. It also identifies its development into SA and their divergences from the vernacular varieties. CA is not a specific variety of spoken Arabic in itself, but I use it as a broad term that encompasses the characteristics of the colloquial varieties of

Arabic. The historical development of Arabic language varieties is key to understanding the reasons why, in modern usage, these varieties are characterised by certain specific, and in some cases unique, features.

2.2.1 Classical Arabic

In his seminal work, 'The Arabic Language and National Identity', Suleiman (2003:69) states that: "Arabic has many of the ingredients which make it eminently suitable to play the role of one of the primary markers of national identity in the modern period." He describes Arabic as a symbol of collective identity, by virtue of its bond with religion and Islamic theology, as well as Arab nationalism (Suleiman, 2003:66-7). The kind of Arabic at issue here is an elevated form of Arabic, also labelled *al-'Arabiyya*, Classical Arabic, or *fuṣḥā* (Corriente, 1976). It is an elevated language form that emerged in the sixth and seventh century CE and that corresponds to the language of pre-Islamic poetry and the Qur'an (Holes, 1995:8).

The nomadic tribes that lived in the Arabian Peninsula before and in the early days of the Islamic Revelation had a flourishing oral poetic tradition expressed through the linguistic eloquence and sophisticated grammar structure of Classical Arabic. According to Abu-Absi (1986:345) Classical Arabic represents the highest literary realisation of the language and the bond that unites all Arabs within a rich, shared heritage, due to its eloquence and refined structure. In Giolfo and Sinatora's (2011:117) words, it denotes "a mythical, abstract, reconstructed or constructed essence of Arabness" at three different levels: metalinguistic, socio-political, and anthropological. Classical Arabic is also the language which is believed by Muslims to have transmitted the word of God in the Qur'an and it is therefore considered divine and sacred (Parkinson, 1991:38). Its role is not only to act as the chief vehicle and instrument of the Islamic religion, but it is also the essential key to interpreting its religious message (Zughoul, 1980:203). Thus, it acted as the main vehicle for the dissemination of Islam during the expansion of the Islamic Empire. This led to its development as a divine language to be used for religious and, later on, for intellectual and legal purposes (Mazraani, 1997:9). The first attempts at formal standardisation date back to the eighth and ninth century CE (Corriente, 1976).

Beside the eloquent language of their poetic composition, the nomadic tribes that were dwelling in the Arabian Peninsula employed regional dialects as a means of daily communication. The different vernaculars spoken among the tribes were accepted by grammarians as linguistically correct, although they were considered a non-prestigious category of Arabic (Versteegh, 2001:132). It is widely agreed that Classical Arabic was used orally by poets and orators who were native speakers of different dialects (Ferguson, 1959b:616) and that it differed from any spoken dialect. However, so far there is no unanimity among Arabic linguists on the specific point of origin for the dichotomy between spoken varieties related to everyday life and the prestigious variety of poetry and eloquence. The various theories on the genesis of this dichotomy and the related position of this research are explained in the following section.

2.2.2 Dichotomy between spoken varieties and Classical Arabic

There is no overall agreement so far about the origin of the separation between the everyday spoken varieties and the eloquent form of Arabic. However, my position is in line with Weninger (2011) and, while it acknowledges that it is unclear how, and in which circumstances, these varieties initially evolved, it also considers the dichotomy between them to already have been in place in the pre-Islamic period. This interpretation implies that language variation is deeply and historically rooted in the nature of Arabic and explains why it has been challenging for Arabic linguists to locate a standard variety that represents uniformity and unity. My position is detailed in this sub-chapter but I firstly provide a general overview of the different theories on the origin of the dichotomy between the two varieties.

The prevailing theory argues that Classical Arabic developed from two dialects that were spoken in the Arabian Peninsula, namely Hijazi and Nejdī, which respectively spread throughout the Western region of the peninsula and the inland desert area (O'leary, 1923:19). Another group of scholars hold that Classical Arabic was the progenitor of spoken vernaculars, yet view them as being a corruption of it (Owens, 2001:424). A third group of scholars claim that modern vernaculars share a heritage as old as Classical Arabic, with both varieties deriving from a common ancestral form of proto-Arabic (Owens, 2001:424-5). Finally, there exists a debate around whether dialects arose from

an older *koine* (Mitchell, 1975:72) or supradialect. Using the term *koine*, Ferguson argues that the spoken dialects developed from a supradialect variety, whereas Classical Arabic followed an independent path. This variety “was not identical with any of the earlier dialects and [...] differed in many significant respects from Classical Arabic, but was used side by side with the Classical language during early centuries of the Muslim era” (Ferguson, 1959b:616).

Contrary to the theories mentioned above, two scholars, Blau (1977) and Versteegh (2001), argue that Classical Arabic and Pre-Islamic spoken dialects were in essence the same language and that the dual linguistic reality that characterises modern Arabic developed after the spread of the Islamic Empire. Their theory is based on the evaluations and conclusions of Arab grammarians, who compared Classical Arabic and pre-Islamic poetry with spoken vernaculars. It should be noted that the analysis of Arab grammarians did not comprise the entirety of spoken dialects and it focussed only on a marginalised group of pre-Islamic Bedouin dialects (Al-Sharkawi, 2010:40).

Several obstacles prevent researchers from gathering reliable and uncontested information about the nature of pre-Islamic Arabic (Holes, 1995; Versteegh, 2001). Firstly, the Arabic script was borrowed from Nabatean at the expense of clarity, as the latter lacked some consonants that existed in Arabic. Moreover, the diacritic marks had not been standardised at the time of the first scripts and therefore the dots, which differentiate between the letters and the short vowels, were missing. Finally, dwellers of the Arabian Peninsula used to write on perishable items such as palm leaves or tree bark, meaning that very few pre-Islamic Arabic writing samples have survived the passage of time (Versteegh, 2001). It is for this reason that the debate on the origin of the dichotomy between spoken and Pre-Islamic elevated oral varieties remains open and strongly contested.

Some evidence of different features between the two varieties was, however, already found in the pre-Islamic period and for this reason “most Arabists share the assumption that Arabic was characterised by diglossia before Islam or [...]”

that it occurred in less than a century after its spread" (Larcher, 2001:604-605)¹. As noted at the beginning of this section, my position is in line with Larcher (2001) and Weninger (2011) in positing that language variation is deeply rooted in Arabic and was even to be found before the spread of Islam. This position considers Classical Arabic as disconnected from any specific dialect spoken throughout the Arabian Peninsula and it implies that the ancestors of modern dialects existed alongside Classical Arabic and evolved into present-day varieties of CA. It recognises that Classical Arabic did indeed serve as a basis for the standardisation of SA, and that the Qur'an represents its highest and most eloquent form. The development of SA on the basis of Classical Arabic allowed the codified language variety to continually evolve and remain in use as the standard form. It also enabled SA's inheritance of the peculiar features of Classical Arabic, as the symbolic embodiment of a mythical Arabness. This function is discussed in more detail in the following sub-section.

2.2.3 Development of Standard Arabic

Understanding the development of Classical Arabic into SA is important because it underscores SA's role as a standard and systematised variety that developed on the basis of the codified form, to mitigate its lack of vocabulary and modernity. On the one hand, it is more dynamic than Classical Arabic and, being widely used in contexts such as contemporary literary and media broadcasting, educated native speakers use it more frequently than Classical Arabic in their everyday lives. On the other hand, it inherited from Classical its pan-Arab role and symbolism within the Arabic-speaking community and therefore it fulfils a unifying role. In the researcher's opinion, this provides an insight into the reasons behind the continuation of the codified variety as a symbol of linguistic standard as well as unity and purity.

According to Kaye (1994:53), the Qur'an was the main factor resulting in "the preservation of the classical language in a frozen state, while its contemporary spoken dialects continued to change, as all living languages do." The Islamic conquests that started in the seventh century CE entailed a linguistic contact

¹ Translated from French by the researcher.

with non-Arabic-speaking communities. A general fear that this contact could lead to the adulteration of Arabic, and therefore to the language of Islam, vehicle of understanding of the divine revelation and symbol of its conquests and triumphs, triggered a process of standardisation aimed at safeguarding Arabic purity against the risk of being corrupted by foreign varieties (Eisele, 2002:7). The process of codification into fixed forms dates to as early as the ninth century (Van Mol, 2003). It is within this process that the Bedouin began to be regarded by sedentary populations as the ideal type of Arab and their language as a focus for the preservation of the purity of the pre-Islamic period (Versteegh, 2001:37). This idealisation of the Bedouin was linked to a perception that linguistic purity was best attained by way of strict geographical restriction; accordingly, the more isolated and self-contained a tribe or social group was, the more likely it was to preserve its language unadulterated (Suleiman, 2012:203).

The process of systematisation of the language gave birth to a strong cultural tradition of grammar study, rhetoric, and literary criticism (Eisele, 2002:7). However, despite the strong tendency towards the codification and dissemination of a unified and pan-Arab language, the linguistic situation within the Islamic Empire was extremely fragmented. Holes (1995:34) reports that at the beginning of the decline of the Islamic Empire, the linguistic situation was

“one of considerable fragmentation and complexity. In the urban areas across the Muslim Empire, spoken Arabic showed variation along both the horizontal (geographical) and vertical (social) axes. In the countryside, the substrate languages continue to survive, though by now challenged by Arabic [...]. The spoken Arabic of inner Arabia, [instead], remained structurally close to the original [tribal dialects] of six centuries before, not having been subject to prolonged contact with other languages.” (brackets in original)

Moreover, in the last centuries of the Islamic Empire, around the thirteenth century CE, other languages, such as Turkish and Persian prevailed over Arabic in many communities. Their political and economic power overshadowed Arabic as a channel for culture influence for six centuries, until the so-called Arab renaissance (Eisele, 2002:7). The Napoleonic campaign in Egypt, which occurred between the end of the eighteenth century and the beginning of the nineteenth, is historically considered to be the beginning of this renaissance, which was triggered in Egypt and spread throughout other Arabic-speaking countries (Holes 1995; Versteegh 2001). This period was

characterised by a revival of the classical language and a vision for the creation of separate, independent Arab nations in opposition to the colonial powers (Walters, 1996:163).

Haeri (2003:308) argues that through western colonisation, “western political, economic, and scientific ideas proliferated and ramified through the Arab world.” This led to the translation of numerous foreign words, especially in the sciences, and these translations began to pour into written Arabic. As a result, a new form of Arabic emerged which became widely dispersed through its deployment in broadcast media, nationalised school curricula and textbooks (Walters, 1996:163). This new form of Arabic, namely SA, has remained the modern iteration of Classical Arabic up to the present day. SA shares most of its morphosyntactic rules with Classical Arabic and canonical Islamic literature, with only small simplifications which enhance its prestige as a model of eloquence and excellence (Mitchell, 1986:9). However, it tends to be more consistently different from Classical Arabic in style and lexicon (Ryding, 2005:9). This is because SA is used in modern contexts such as news broadcasting, contemporary literature and academic writing, and therefore it needs expressions and terminology that do not exist in Classical Arabic. Moreover, stylistic changes are necessary due to the journalistic and lively nature of SA, as opposed to the static nature of Classical Arabic (Bentahila and Davies, 1991:372). It is interesting to note that the distinction between Classical Arabic and SA exists only in the analysis of western research. Native Arab speakers conceive only one standard form, which is called *fuṣḥā* and represents the literary, eloquent and standard form as opposed to mere dialects (Bassiouney, 2009:27). I focus in the following sub-chapter on the comparison between the characteristics and functions of SA and CA.

2.2.4 Standard Arabic and Colloquial Arabic

During the Arab renaissance, European linguists who had been previously concerned exclusively with Classical Arabic and considered vernacular languages corrupt, started looking at them from a different perspective as they discovered that “dialects often contained forms that were much older than the corresponding forms in the standard language” (Versteegh, 2001:5). Linguist scholars thereon started to record and analyse the vernacular forms and treat

them as fully developed languages. From these analyses it emerges that there are several key linguistic differences between SA and CA. Firstly, they have different phonological systems (Ibrahim, 1986); secondly their lexicon is divergent; and thirdly, CA is characterised by a simplified morpho-syntax, particularly in the case-marking system which is preserved in SA and absent in CA (Holes, 1995; Mansouri, 2000).

CA varieties can be geographically identified as follows. Fisher and Jastrow (quoted by Van Mol, 2003:22) divide the contemporary Arabic dialectal regions into five large groups based on differences based mainly based on phonology, morphology and lexicon. These are: the Arabian Peninsula, Mesopotamia, the Maghreb, Egypt and the Syro-Palestinese area. As this research focuses on the development of diglossic vocabulary, I focus on the relationship that exists between SA and CA vocabulary. In all the afore mentioned regions, there are three patterns that describe the difference between CA and SA vocabulary forms (Abdel-Malek, 1972: 138-139). Firstly, there is vocabulary variation between SA and CA. Secondly, there are phonological differences. Shared vocabulary items can show phonetic differences, distinct vocalisation or different pronunciation. Abdel-Malek states (1972:138) that the differences between SA and the CA of a given region can be specified by a general rule. For example, in some of the afore mentioned areas, where /q/ occurs in a SA form, a glottal stop usually occurs in the corresponding CA form. Thirdly, there are vocabulary items that are not shared by SA and CA. The distinction between vocabulary items as proposed by Abdel-Malek is essential to this study as this differentiation is adopted in the researcher's approach to diglossic vocabulary building. This becomes clear in chapter four.

The linguistic trait that most distinguishes CA and SA is the medium of their performance: the former is written and the latter is spoken. This means that, although SA has users with strong reading and listening comprehension and even occasional speaking and writing use (Parkinson, 1991:39), it is not "a *spoken* language and therefore is nobody's native language" (Mitchell, 1975:70, emphasis in the original). This leads to a situation in which SA synchronically has no native speakers, while CA represents the actual native language and native speakers convey messages related to their personal or intimate spheres and opinions using CA. Diachronically, in contrast, Arabic language speakers

imbue the two varieties with a moral dimension, which elevates the status of SA: SA is perceived as preserving a pure state of the language whereas CA is often seen a corruption of it (Grande, 2012:125). The term ‘dialect’ as opposed to SA, in fact, often carries a pejorative connotation in the Arabic tradition (Owens, 2006:12). This highlights a dichotomy that affects modern Arabic-speaking communities, in which SA represents the quintessence of Arabness although vernaculars are the authentic native languages. This being said, it is not uncommon to find CA forms and colloquialisms inserted within SA. According to Parkinson (1991:38-9),

“Arabs, even highly educated ones, find it difficult and unnatural to use *fusha* spontaneously without referring to a prepared text which is then partially or entirely read. [Thus,] we can observe that the oral intervention of colloquial Arabic is becoming more and more frequent in these situations for purposes of better communication.”

The features listed above might suggest Arabic linguistic reality to be that of two markedly distinct varieties with specific linguistic and moral dimensions, and idealised interpretations attributed to the two language forms. Instead, and in agreement with Giolfo and Sinatora (2011:115), this research posits that these features are artificially constructed by native speakers at a cognitive level. The existence of this dualism at a cognitive level is crucial to language variation and results in a linguistic performance that switches between language varieties. In practical terms this means that the shift between language forms can be required by the nature of the language performance as we have seen above (SA is written whereas CA is spoken), but it can also express individual linguistic and metalinguistic choices. This model is explained in detail in section (2.4.1) and I now focus on the description of Arabic linguistic reality according to the diglossic model proposed by Ferguson. Analysis of Ferguson’s model and the criticisms directed towards it allow me to better elucidate the model of reference that this thesis adopts.

2.3 Diglossia

We have seen in the previous section the varieties that linguistically characterise Arabic: standard varieties on the one side and vernaculars on the other. Ferguson labelled this linguistic situation as “diglossia” in 1959 (1971[1959]:1), by modelling the term on the French word *diglossie*, which had

been previously used by the French linguist Marçais (1930) in his studies of North African vernaculars. Diglossia is a linguistic situation in which two distinct, codified and stable varieties of the same language “exist side by side throughout the community, with each having a definite role to play” (Ferguson, 1971[1959]:1). More specifically, they consist of a "High" (hereafter H) or prestigious variety and a "Low" (hereafter L) or inferior variety. In Ferguson's words (1971[1959]:16), diglossia is

“a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned by formal education and is used for most written and spoken purposes but is not used by any sector of the community for ordinary conversation.”

Within a diglossic speech community speakers select the variety to be used in accordance with the distinct function it fulfils. Thereby, in one environment the high superimposed form is the only appropriate variety, whereas the ordinary form alone must be used in the other set of circumstances, with infrequent overlap (Ferguson, 1971[1959]:5). Ferguson does not distinguish between Classical Arabic and SA in his representation of H. He identifies H as being related to the religious sphere as applies to Classical Arabic, but he also links it to situations where SA is the usual means of communication, such as politics, news and modern literature. For consistency with his model, Ferguson's H and L is used in this section without the distinction being made between Classical and Standard Arabic.

Ferguson claims (1971 [1959]:6) that “the importance of using the right variety in the right situation can hardly be overestimated.” He goes on to distinguish between the typical circumstances in which H or L are adopted. The former is appropriate for contexts such as sermons in church or mosque, political speeches, university lectures, news broadcasts, personal letters, newspaper editorials and poetry. It is also the variety used in formal schooling and higher education, as the language of instruction (Maamouri, 1998:31). In contrast, L is appropriate for conversations with family, friends and colleagues, radio soap operas and folk literature (Ferguson, 1971[1959]: 5).

Diglossic varieties are considered by Ferguson as belonging to the same language. However, diglossia has been subsequently extended to bilingual

speech communities, such as Spanish and Guaraní in Paraguay, and multilingual situations where different languages carry distinct functions and are appropriate for separate circumstances (Fishman, 1967). In my view, and in agreement with Hudson (1980), this approach does not consider multiple clear-cut differences between diglossia and bilingualism, i.e. the genetic relation between the language varieties and the functional roles of the language varieties. Firstly, in a diglossic community only one language variety is the mother tongue, and the second variety is learned through formal schooling. In bilingual communities both language varieties are mother tongues. Secondly, diglossia implies the existence of a complementary distribution of language functions within the same linguistic community that does not exist in bilingual communities (Bassiouney, 2009; Boussofara-Omar, 2005; Fasold, 1984; Holes, 1995; Maamouri, 1998; Versteegh, 2001).

Diglossia comes into being where there is an extensive body of written literature composed in H; where education is limited to a small elite within the speech community and a period of several centuries elapses between the establishment of the former and the latter (Ferguson, 1971[1959]:8). It is generally agreed that diglossia can persist stably for centuries (Snow, 2013:65). However, it eventually vanishes due to social changes brought about by modernisation, with L replacing H in almost all domains (Ferguson, 1971[1959]; Hudson 2002). Ferguson lists the social developments that lead to the decline of diglossia as follows: firstly, the spread of literacy on a large scale; secondly, the rise of communication among distinct social and geographical segments of the community; and finally, a widespread desire for a national language that can serve as the symbol of independence and autonomy (1971[1959]:18-19). Together with mass literacy in the Arab world, mass media has helped considerably “to spread the knowledge of SA and the urban forms of spoken Arabic” (Abdulaziz, 1986:15 quoted by Hudson, 2002:37). According to Amer, Adaleh and Rakhieh (2011:21), there is an expectation among many Arab scholars that Arabic in the future will be a unified and standardised language used for both speaking and writing. As Ferguson puts it (1971 [1959]:19),

“When these trends appear, leaders in the community begin to call for unification of the language [...] The proponents of H argue that H must be adopted because it connects the community with its glorious past or with the world community and because it is a naturally unifying factor as opposed to the divisive nature of the L

dialects. [...] The proponents of L argue that some variety of L must be adopted because it is closer to the real thinking and feeling of the people; it eases the educational problem since people have already acquired a basic knowledge of it in early childhood; and it is a more effective instrument of communication at all levels.”

Ferguson believes that H can succeed in establishing itself as a standard variety only if it is already serving as a standard language in at least one language community and then prevails within the diglossic community (1971[1959]: 20). Otherwise H tends to remain a liturgical language associated with religion that is studied only by scholars or specialists but that is not actively used in the community. In this scenario, L acts as the basis of the new standard, whether it is a relatively pure L or a form considerably mixed with H. Ferguson suggests (1971[1959]: 21) that within the Arab world, this could lead to the development of three standard varieties: first “Maghrebi”, based on the variety spoken in Rabat or Tunis; second “Egyptian”, based on the Cairene SA; and finally, “Eastern” based on Baghdadi dialects. He suggests moreover that other varieties could develop such as Syrian and Sudanese (1971[1959]:21). There are nine features that sociolinguistically identify the presence of diglossia within a community according to Amer, Adaileh and Rakhieh (2011:20), summarised as follows:

- i. [H] and [L] are in a strict complementary functional distribution;
- ii. H is uniformly held in higher esteem than L by members of a speech community;
- iii. H has associated with it a substantial and highly regarded body of written literature;
- iv. Proficiency in H is typically attained as a result of formal schooling, whereas proficiency in L is attained through the natural process of mother-tongue acquisition;
- v. The pronunciation, grammar, and vocabulary of H are standardized and tolerate only limited variation, whereas there is wide variation in the pronunciation, grammar, and vocabulary of L;
- vi. Diglossic situations are extremely stable and typically persist for several centuries at least;
- vii. There are always extensive differences between the grammatical structure of H and L;
- viii. There exists a series of phonologically unrelated lexical doublets for concepts frequently expressed in H and L; and
- ix. The sound system of H and L constitutes a single phonological structure of which L is the basic system.

Whilst it is difficult to predict whether the spoken varieties of Arabic will take over from H, it is likely that L forms will increasingly be adopted and accepted over time (Belnap and Bishop, 2003). In fact, the use of L forms in formal and semi-formal domains has been increasingly accepted within Arabic-speaking countries (Kaye, 1994; Versteegh, 2001). Diglossic patterns are however still strong in the Arab world due to two main factors delineated in the previous sections. Firstly, H plays an irreplaceable role within Islam and as a pan-Arab nationalist symbol (Versteegh, 2001). Secondly, mass education and literacy is a relatively recent phenomenon, resulting in the prolonged conservation of diglossic patterns (Snow, 2013:65). In fact, although L is the native speaker's mother tongue, it is also linked to low access to education and therefore often seen as inferior to H (Abu-Absi, 1990:34; Bentahila and Davies, 1991:371), which is normally acquired through higher education (Versteegh, 2001:195). Consequently, the pre-eminence of the written form makes it the basis for the creation of fixed models and grammar rules (Haugen, 1972:246). It is for this reason that Arabic linguistic norms, grammar and orthographic rules are based on H, not the mother tongue L. There are several criticisms of Ferguson's theory of Arabic diglossia, which are delineated in the following sub-section. These criticisms are pivotal to the articulation of the position of this research on Arabic variation.

2.3.1 Criticism of Ferguson's model of diglossia

The theory of diglossia has drawn several criticisms. Firstly, Ferguson does not specify the exact distance that is required between L and H for a language situation to be considered diglossic (Bassiouney, 2009:11). This omission occurs because Ferguson intended specifically to draw an idealised picture of a diglossic reality (Bassiouney, 2009:11). He claims that H and L reflect variation on the basis of occasions of use and are influenced by the linguistic environment, their addressee(s), the topic(s) and setting(s), and so forth (Ferguson, 1996:56). Ferguson does not consider social factors in determining language choice, supposedly because they were not considered sufficiently scientific when he first approached diglossia (Bassiouney, 2009:12). He consequently places more emphasis on linguistic repertoires and appropriateness as influences on language variation (Ferguson, 1996:60). In my opinion it is not possible to define a precise distance between L and H since

their deployment is highly subjective and individualised. While agreeing that external factors do determine language variation, I believe that it is also induced by subjective and individual elements. This point aligns with Bassiouney in asserting that the role of individuals is fundamental in negotiating “socially agreed patterns of language choice” (2009:12).

Secondly, Ferguson’s model addresses H and L as if they were homogeneous and static systems (Meiseles, 1980:121). It is therefore so categorical that its application to dynamic language varieties is problematic (Kaye, 1970; Mitchell 1986; Holes 1993). Moreover, Ferguson considers every intermediate linguistic form between H and L to be lacking a regular structure and to be unsystematic, and therefore he offers no theoretical framework to describe them (Wahba, 1996:104). His view of diglossic realities is that of language communities where two main varieties occur alongside each other and whose use is strictly determined by appropriateness of use. My position is in contrast with Ferguson in not only considering H and L to be dynamic, but in also regarding variety mixing as being present in the same linguistic context. Kaye (1994:53) reports that although a university professor giving a lecture uses SA, it is not uncommon to find the material interspersed with some L. Ferguson himself asserted that the reality of the Arabic language is more complex than the ideal description that he depicted (1971[1959]:6):

“In the Arab world [...] formal university lectures are given in H, but drills, explanation, and section meetings may be in large part conducted in L. [...] Often a considerable part of the teacher’s time is taken up with explaining in L the meaning of material in H which has been presented in books for lectures.”

The employment of diglossia in religious and political speeches has also been recently studied. In this context, the speaker may seek both to reach a greater audience and also to be perceived as a member of the community (Holes, 2004; Soliman, 2008). In such instances the boundaries of the rigid dichotomy drawn by Ferguson are blurred, and can be criss-crossed by speakers in order to convey specific messages. As far as literary production is concerned, nowadays it is not uncommon to come across literary texts that make partial use of CA to convey messages related to everyday life or personal and intimate issues. Amer, Adaileh and Rakhieh (2011:21) argue that the spread of literacy allows speakers to employ a greater stylistic range when writing in Arabic, since many young Arabs are beginning to see unmitigated SA as too formal for

personal correspondence with peers. They prefer to use a variety of Arabic that is at least somewhat closer to CA, if not primarily CA. Also, diglossia can also be a consequence of “an author’s incomplete control of the prescribed variety” (Amer, Adaileh and Rakhieh, 2011:20). Not all educated native speakers can produce grammatically correct SA despite hearing and reading it every day, because they have little opportunity to practise it (Parkinson, 1996:92). In his study of variability in SA grammar skills, Parkinson (1996:99) reports that although most educated speakers appear to be aware of the basics of the case-marking system when it is brought directly to their attention, most cannot produce correct case-marked sentences, nor complex grammatical structures.

Finally, Ferguson perceives that the H variety possesses both the quality of prestige and the role of the standard language variety. This doesn’t reflect the fact that these two characteristics do not always and compulsorily coexist in the same Arabic variety, and that a variety can be prestigious without being standard. For this reason, Van Mol suggests that when tackling the issues of prestige and standard, it is of fundamental importance to specify whether one is considering the written or the oral language (2003:11). H carries both features simultaneously (Van Mol, 2003:46), whereas SA can be prestigious but, as it is uncodified, cannot function as a standard language variety. The issue of standard and prestige in Arabic varieties is tackled in detail in the following section 2.3.2.

To summarise, although Ferguson is persuaded that speakers in a diglossic context have no freedom in choosing which language variety to use, and that they passively abide by superimposed and socially determined boundaries (Amer, Adaileh and Rakhieh, 2011:20), my view is that language variety is determined by speakers’ subjective choices, sociolinguistic factors and their command of language varieties. In this thesis, I interpret linguistic choices to be determined by a combination of the speakers’ attitudes and beliefs about language varieties, in that their use of language variation is functional to “identify themselves and to mark social contexts” (Wahba, 1996:103), and their language proficiency in the formal variety. This means that native speakers do perform Ferguson’s abstract dichotomous model but they actively decide which variety to employ and how to mix them. As discussed in section 2.2.4, I adopt Giolfo and Sinatora’s (2011) model to describe language variation and the

representation of varieties in native speakers' minds. CA and SA exist in the mind of the native speaker both with a linguistic and a metalinguistic dimension. The crucial aspect of Giolfo and Sinatora's model is that it portrays Arabic as unitary body, which is conceptually divided into two complementary parts, CA and SA, and each one has both a linguistic and a metalinguistic nature as explained above. I now proceed to outline in detail the significance of the concepts of Standard and Prestige in Arabic.

2.3.2 Standard and prestige

The Arabic diglossic community can be looked at from an international and an intranational perspective (Angheliescu, 1974:82). From an international perspective, H is represented by both Classical Arabic and SA. It is equally shared by all Arab countries and plays a unifying role between them. L is represented by the colloquial varieties and serves as the distinctive mother tongue for each country. From an intranational perspective, within each country, diglossia is found in every speech community. In this regard, Bentahila and Davies (1991:381) state that "the duality between written and spoken language exists within each Arab nation and within the repertoire of each individual educated speaker."

The prestigious language is defined at the international level on a horizontal dimension, as it is shared by all Arab countries. Similarly, the abstract label 'colloquial variety' is applied horizontally across all Arab nations, portraying the spoken variety used as mother tongue in each country. At an intranational level there is a potential overlap of the functions of language varieties. Classical and Standard Arabic fulfil the role of symbols of Arabness, religion and purity. However, the marker of national identity is each country's mother tongue. For instance, Haeri (2003:37) claims that Classical and Standard Arabic are often perceived by Egyptian native speakers as varieties whose musical qualities move listeners, inspiring feelings of spirituality, nostalgia and community. The marker of Egyptian identity and national culture is nevertheless Egyptian Arabic (Haeri, 2003:37). Among the numerous dialects that are included within Egyptian Arabic there exists a prestigious form that is regarded as the symbol of national identity. This prestigious spoken variety differs from Standard and Classical Arabic. This contradicts Ferguson's model of diglossia, in which

standard and prestige are only found in SA. In reality, it is only the former quality that is unique to SA while the latter may also exist in CA. Bassiouney (2009:18) claims that the equivalence between the standard and the prestigious variety is the result of applying western research to the Arab world.

The prestigious and standard forms of Arabic can coexist without necessarily overlapping. SA has undergone a process of standardisation leading to the development of fixed grammatical rules (Van Mol, 2003:12) and it also simultaneously fulfils a prestigious function because it is the model of erudition and civilization (Van Mol, 2003:21). CA cannot function as a standard language because of its lack of codification but it evolves spontaneously to reach the point where a specific spoken variety enjoys more prestige than the others (Van Mol, 2003:12). This is usually the variant that is perceived by the members of the community as a model and an example worth imitating (Al, 1987:35 quoted by Van Mol, 2003:15) and often coincides with the variety spoken in the capital or in major urban centres. In the above example we have seen that there is one variety of Egyptian Arabic, Cairene Arabic, which is considered prestigious and acts as a symbol of national identity. According to Bassiouney (2009:18):

“Many studies have shown that for most speakers, there is a prestige variety of L, the identity of which depends on many geographical, political and social factors within each country, and which may in certain circumstances influence speech”

The issue in Arabic diglossic speech communities is that although there is only one codified standard variety, there are two prestigious varieties, these being SA and a local vernacular variety “which represents the prestigious variety of common use for the speakers of all dialects in a given Arab country” (Wahba, 1996:121). The spoken variety of greater prestige will tend to use the vernacular as the basis and borrows lexicon, set expressions, and discourse markers from SA (Walters, 1996:169). For example, urban dialects of Jordanian cities are treated as prestigious varieties and non-Cairenes consider Cairene dialect the most prestigious among Egyptian Arabic vernaculars (As Abdel-Jawad, 1986; Bassiouney, 2009). Similarly, in his studies on linguistic variation in Bahrain, Hole shows that there exists a locally recognized prestigious variety which clearly differs from H (Holes, 1983; 1986). In his study of variation in Alexandrian Arabic, Wahba analyses variation within educated and non-educated social groups, and states that in Arabic diglossic reality (1996:120):

“It is common practice to regard the Classical variety of the language as the one that carries prestige, in comparison with the local variety. Since the Classical variety is mainly restricted to religious usage, the prestige value of Classical Arabic has been transferred to Modern Standard Arabic (MSA). In fact, MSA constitutes a standard variety across the international community of educated individuals in the Arab world, but within each Arab country there is a regional variety of the language that functions as the standard. In Egypt, for example, the Cairene Colloquial dialect—as opposed to the Upper Egyptian, for instance, functions as the standard variety for all Egyptian speakers. [...] This is because it is the dialect of the country's capital, and the one most prevalent in television and radio programs that do not use MSA”.

Eisele's (2002:20) theory of “regimes of domination” draws a connection between linguistic, metalinguistic and socio–psychological levels within language varieties and claims that individuals consciously express their own individual identity through them. Eisele introduces three regimes of domination: intranational or localistic; pan-Arab; and international. As he (2002:20) puts it,

“localistic regimes of domination involve the valorisation of the speech of local centres of linguistic autonomy, often that of the capital cities but sometimes that of a specific class. This is reflected in a dialect being termed as ‘prestige’ dialect [...]. The pan-Arab regime of domination valorises Classical Arabic (or its modern reflex, MSA) and the cultural heritage, and is the most dominant and authoritative of these regimes in terms of institutional support and cultural weight. [Finally] there is an “international” regime of linguistic authority or domination, which reflects the dominant economic, political, and technological status of European languages, primary among them being English. This regime is reflected in the conflict or competition between Arabic and these colonial and postcolonial foreign languages, as measured in the amount of borrowings from them, the extent of bilingualism, and the perceived necessity of foreign language education in these languages”.

Prestigious spoken varieties can be employed to enable clear communication when two dialects come into contact with each other. Manfredi's research (2012) shows that in such circumstances, speakers can either import features from SA or from the mutually recognised prestigious variety. The first strategy is termed ‘classicising’ and the second ‘levelling’. As Van Mol (2003:50) puts it:

“The levelling strategies consisted of adaptations of the language using their higher dialectal forms which did not necessarily belong to the dialect of the speaker. [...] The aim of the levelling was to suppress local dialectal expressions and characteristics in favour of characteristics that were more common cross-dialectally. The ‘classicising’ strategies consisted mainly of borrowings from Classical Arabic”.

Manfredi's (2012:150) findings show that social contact and urbanization are two of the main factors that accelerate dialect levelling in contemporary Arabic. This shows us that some varieties of CA are in fact considered to be

prestigious. They are employed in day to day language use, as well as in contact-induced code switching, to facilitate clear communication.

The reinterpretation of prestige in Arabic varieties has a direct consequence for TAFL. It supports the view that students need to develop strong skills in CA, in addition to mastering SA. CA is not an isolated variety only used in informal and familiar contexts. It can be the prestigious variety employed in a semi-formal context or the shared language form that permits clear communication. It also shows that the prestigious variety often borrows a large amount of SA vocabulary, within CA morphosyntactic structures. This in turn means that mastery of CA doesn't simply provide students with the vocabulary of everyday life, but that it is crucial to achieving higher levels of oral proficiency. In the next sub-section I identify the approaches to Arabic variation proposed in response to the limits of Ferguson's diglossia.

2.3.3 Variationist approaches to Arabic variation

Several scholars who have criticised Ferguson's model for being too categorical subsequently proposed approaches to Arabic variation that posit intermediate language levels, ranging between the H and L poles (Meiseles, 1980). The first approaches I identify are those of Blanc (1960) and Badawi (1973) who pinpoint how the dichotomy of H and L is too simplistic. However, they do not present dynamic models as the possible alternatives. On the one side, Blanc distinguishes five different language levels that range from a pure form of Classical Arabic to plain colloquial (1960:85). Between these two unadulterated forms are posited three hybrid levels of colloquial form that possess a decreasing proportion of CA linguistic features mixed with Classical Arabic traits. The division that Blanc proposes is based on linguistic criteria that leave the boundaries among the intermediate levels highly unclear and it is therefore hard, if not impossible to verify the three middle levels (El-Hassan, 1977:119). In response to Blanc's distinction, Badawi adds a sociolinguistic dimension to the linguistic factors that determine language levels, including factors such as education level and social class (Eisele, 2002: 16). The levels he identifies range from *fuṣḥā al-turāth* (Classical Arabic) to *'āmmiyyat al-'ummiyyīn* (Illiterate spoken Arabic) gradually shading into *fuṣḥā al-'aṣr* (MSA),

'āmmiyyat al-muthaqqafīn (Educated Spoken Arabic) and *'āmmiyyat al-mutanawwiriin* (Semi-literate Spoken Arabic) (Badawi, 1973:89).

Badawi's model has limitations similar to those outlined in Blanc's approach: it lacks definitional clarity regarding the dividing lines between variation levels (Hary, 1996:71). It is unclear where one level ends and the next begins because Badawi's levels are proposed with a view to simply using them as heuristic devices in the study of sociolinguistic and linguistic features. The five levels range from purely classical to purely colloquial at the ends of the scale and, as Wahba (1996:105) summarises it, "the degree of occurrence of the linguistic features of either classical or colloquial across the other three levels depends on the proximity of any given level to one end or the other." In this model, the third level, *'āmmiyyat al-muthaqqafīn*, is placed in an abstract position that is equally distant from the two extremes and is close to representing a standard language. This level is identified by a set of rules that fulfil a vision of linguistic unity and uniformity which in turn proves the existence of, and linguistically represents, a unified society according to Labov (1972:143). With respect to the extreme poles, these are pure forms and ideal varieties which do not exist as language realities according to Hary (1996:72). Badawi's emphasis on a form of Arabic that is heterogeneous and equally distant to both poles reveals an intention to identify a unified language (Giolfo and Sinatora, 2011:107).

Elgibali (1996) and Hary (1996) criticise the afore mentioned attempts to frame language levels in favour of a more fluid model. Such critiques of Blanc and Badawi's approaches can be divided into variationist and language continuum models. The variationist model portrays Arabic-speaking communities as being characterised by a constant mixing of language varieties. Speakers move along a language continuum wherein no one strictly uses only one variety, not even in the same sentence (Elgibali, 1996:105), and they constantly change their position due to linguistic or sociolinguistic factors, such as language proficiency, the mix of participants in the discussion, the purpose and function of the conversation and the personal relationship between speakers (Hary, 1996: 76). The theories of language continuum claim that Arabic varieties are to be placed on a continuum without the need to establish a fixed numbers of levels. This owes to the impossibility and impracticality of determining a discrete number of varieties in the multiglossic situation of Arabic (Hary, 1996:71). These critical

approaches to the existence of discrete language levels have been seminal in the development of research into language varieties, as they suggest that sociolinguistic factors trigger the shift between language varieties (Eisele 2002:16).

My position agrees with the variationist approaches in so far as they criticise the existence of two dichotomous and defined language forms. However, these approaches suggest the existence of multiple language varieties without succeeded in providing a clear description of their boundaries and distinctive features. They also erroneously imply that the mixing process follows rigorous rules that result from defined hierarchies of language varieties. The approaches that suggest the existence of a language continuum avoid these pitfalls, by positioning the continuum in place of the discrete language varieties and asserting that the blending is unpredictable and undefinable. In my opinion, language continuum theories also have a crucial weakness: they do not discard the interpretation of Arabic as a set of multiple language varieties where SA and CA are blended. I propose to reinterpret language mixing in view of the nature of Arabic language varieties (SA is written and CA is oral), and in view of the speakers' linguistic and metalinguistic attitudes towards the above mentioned 'reference-packagings model'. Finally, I suggest that the use of variation needs to be interpreted in light of the existence of different degrees of prestige within both SA and CA, as previously explained. I now move on to describe the theorisation of a middle variety that lies idealistically between SA and CA. My position towards it allows me to identify the crucial features of my approach to Arabic variation, which is described in section 2.4.

2.3.4 Educated Spoken Arabic

Blanc's model claims the existence of a supra dialectal Educated Spoken Arabic (hereafter ESA), a third dimension of stylistic difference that is equally distant from SA and CA and serves as a "medium for intellectual and cultured discourse" (Mejdell, 2012:159). ESA embodies a standard form of reference, which is not identified within the diglossic model. Despite the inadequacy of the diglossic model to pinpoint a standard form used by educated native speakers in their everyday language, Ferguson's theorised the development of an intermediate form which could emerge as a language between H and L. He

describes this middle language, or “luġah al-wuṣṭā” (Ferguson 1971[1959]:10), as a relatively uncodified and unstable form, which resembles spoken Arabic but it is used in semiformal or cross-dialectal situations. Therefore, it borrows items from both H and L. Although it generously allows the use of spoken vocabulary, it is mainly characterised by standard vocabulary with few or any inflectional endings and with certain features of classical syntax on a fundamentally colloquial base in morphology and syntax (Ferguson, 1971 [1959]:10-11).

Contrary to Ferguson’s interpretation of the middle variety as uncodified and unstable, some scholars believe ESA to undergo a process of standardisation. Ryding (1991:212) describes it as a mixed variety that carries a high degree of prestige and formality and can be used in intercommunications throughout the entire Arabic-speaking world. In Ryding’s (1991:211) words, ESA is used “for inter-dialectal conversation by educated native speakers, for semiformal discussions, and on other social occasions when the colloquial is deemed too informal, and the literary, too stilted.” Abu-Absi (1986:342) also defines it as the variety “used among educated Arabs who come from various dialectal backgrounds and who find it cumbersome or artificial to use the literary language, [and] it is characterized by the tendency to use literary vocabulary and colloquial grammar.” The claim that ESA goes through a process of codification is a functional approach that allows for it to be taught to students of AFL. Under this approach, ESA serves the specific purpose of functioning as language standard independent from SA and CA although it retains some of their characteristics: it is a standard variety that can be analysed and codified, and its grammatical structure can be formally taught (Eisele, 2002:16). The theorisation of ESA recalls the unresolved issue of finding a standard, prestigious and codified language that is used by native speakers in their everyday life and that is equally distant from CA and SA.

However, as of today, there is no agreed-on definition of ESA in the literature; there are no fixed morphosyntactic and phonological features that define it; and the role of personal and regional elements in the outcome of its mixing is still unclear (Albirini, 2016:22). I therefore agree with Albirini that “even if a number of rules of language mixing were to be specified, they would not be themselves sufficient to posit the existence of a new language variety” (2015:22). The

drawback to the conceptualisation of ESA is that, although it helps in creating a model of reference that combines SA and CA, it does not represent an existing variety of Arabic that is defined by clear features. We have seen above that the patterns of mixing and switching between SA and CA can be unpredictable. Also, they have been proven to be persistently strong (Snow, 2013:65). Thus, the theorisation of a variety based on their mixing does not seem to be representative of the authentic use of language varieties. In fact, authenticity in Arabic is expressed through a division within the use of language varieties based on their different functions, and it also sees a constant blend of these varieties in the speech of native speakers, "with different degrees depending on many social variables, such as situation and education" (Wahba, 2006:144). Other social variables can be the participants, the topic and the formality of the situation. As for education, Parkinson (1992:226) claims that speakers of a language "vary considerably in their control of particular language forms." He also adds (1992:226) that in language situations "with a relatively codified, formal, probably mainly written end of the linguistic spectrum, one is likely to find extreme variability among speakers or users both in their knowledge of the formal variety and their ability to use it effectively." In his research among Egyptian native speakers, he also found that they perform receptive skills (reading and listening) better than productive skills (writing and speaking) (Parkinson, 1992:242).

My position is that there are three main flaws in the ESA model. Firstly, it has not yet been proved whether ESA, either in the form of "the main Arabic urban vernaculars" (Miller, 2004:184) or as a supraregional form (Ryding, 1991), has "succeeded in becoming regional and/or national standard" (Miller, 2004:184). Secondly, further research claims that other language levels exist but their limits are too blurred to be defined (Parkinson, 1991). Finally, Mitchell's original theorisation of ESA described it as a stylistically-controlled variation. However, such variation needs to be approached "in terms of socio-demographic categories [since it is necessary] to define what significant types of variation occur and with which sociological entities they are associated" (Owens, 2001: 440). As Holes (1995:279) points out, "the problems of description and explanation are inextricably bound up: in any passage of Arabic speech, whether monologue or conversation, one cannot track [...] the moment by

moment, unpredictable changes in language form unless one is also aware of concurrent changes in the ideational content of the discourse and the interpersonal relationships of the participants, as perceived by the participants themselves.” Educated Arab speakers spontaneously code-switch between SA and CA and they following an almost unpredictable pattern of choices (El-Hassan, 1977:114-115). As it is not possible to delimit the extent to which the two varieties are mixed and to establish a standard pattern, I argue that it is important to focus on the cognitive processes that can potentially trigger the switch. The model I adopt to describe native speakers’ language behaviour is the ‘reference-packagings model’ (Giolfo and Sinatora, 2011) and I look at it in the following section.

2.4 Approaches of this research to Arabic diglossia

So far this chapter has tackled the discussion around the nature and development of Arabic varieties and diglossia, with the view to elucidating my position and perspectives within this debate. These viewpoints consider variation as intrinsic to the Arabic language both in terms of linguistic essence and language use, and they are critical toward a strict dichotomy that describes Arabic as composed of discrete units, whether two as in Ferguson’s diglossia, or more in the multiglossic models. However, it also rejects the suggestion of an indefinite linguistic continuum composed of multiple discrete language varieties, along which native speakers move at their own discretion. The stance of this thesis is that Arabic is to be considered one language as part of one cultural system, which holds multiple varieties at its core. These are employed through code-switching by native speakers in response to their linguistic and metalinguistic attitudes as well as their control of language forms. Each individual within the speech community, however, has a peculiar linguistic behaviour that can vary from the behaviour of other individuals and “it is shaped by his/her cultural specificity” (Giolfo and Sinatora, 2011:121). As anticipated in section 2.2.4, I adopt Giolfo and Sinatora’s (2011) model to describe language variation. According to it, CA and SA exist in the mind of the native speaker both with a linguistic and a metalinguistic dimension. The linguistic features of the colloquial variety form the basis of the syntax of CA. The ideal features of the standard variety, instead, correspond to SA (Giolfo and Sinatora, 2011:115).

As for the metalinguistic dimension of the two varieties, they comprise the ideological components that are associated with the spoken and the standard varieties and are used by native speakers to convey their cultural and linguistic identities through them. I describe in the following section the features of this model in detail. I firstly focus on contemporary attitudes hold by educated native speakers towards Arabic varieties, and especially towards CA, and I then describe the model I employ to describe Arabic linguistic and metalinguistic reality.

2.4.1 The reference-packagings model

Albirini states (2016:78) that language attitudes are key to understanding the prevailing assumptions about the language varieties used by speakers in a given linguistic community, and therefore the sociolinguistic context that characterises its language. He describes them as follows: “in the Arab context, language attitudes have been found to be a major vehicle of sound change; a basis of interdialectal (un)intelligibility; a determinant of language resilience and maintenance; a catalyst of transformations in language use and function; a yardstick of linguistic convergence and divergence; a marker of functional status and code choice, an index of tribal, local, national, religious, ethnic, and global identities; and a predictor of second language achievement” (Albirini, 2016:79). Saville-Troike (1989:181) adds that individuals can seldom choose what attitudes to have towards a language since attitudes are acquired as a factor of group membership and as part of the process of gradual acquisition of the characteristics and norms of a culture or group in a particular speech community. Research studies have reported a positive change of native speakers’ attitudes towards SA in terms of status, which seems to comparably increase with their higher education levels (Hussein & El-Ali, 1989; Haeri, 2003; Ennaji, 2007; Saidat, 2010).

Albirini’s study on language attitudes among university students from Egypt, Jordan, Morocco and Saudi Arabia confirms a positive disposition towards SA, but it also shows that its status is increasingly challenged by English. SA is identified as the favoured language with regard to religion, official status, Arab belonging, literacy, media and education, whereas English seems to exemplify professional accomplishment, science and technology, and general usefulness

in everyday life (Albirini, 2016:92). Together with positive attitudes towards English, the analysis of which is beyond the scope of my research, Albirini sheds light on increasingly favourable attitudes towards the mixing of SA and CA in social spheres that were traditionally assigned to SA (2016:118). His finding seems to reveal a possible change in the generalised attitudes of many educated Arabic speakers towards CA that might no longer be seen as a challenger competing against SA but, instead, as a complement to it. This translates into a new viewpoint by which CA is not perceived as threatening SA's unifying power but as a different variety of the same language. The fact that nowadays both varieties are needed to represent language unity and identity might explain why diglossia is not viewed as a social problem by many educated Arabic speakers (Albirini, 2016:95). It also emerges that CA fulfils a role of identity representation. Finally, this shows that the linguistic behaviour of educated speakers of Arabic is influenced, among other things, by linguistic as well as metalinguistic factors, such as linguistic attitudes towards language varieties.

Giolfo and Sinatora (2011:115) bring together the afore mentioned linguistic and metalinguistic factors in their approach to diglossia. In their words, diglossia does not separate two language varieties but rather it is "a line that divides an abstract *unity* into two complementary models" (italics in original). The two complementary models represent CA and SA at an abstract level, in which each of them has "a *linguistic* and a *metalinguistic* nature" (Giolfo and Sinatora, 2011:115) (italics in original). The linguistic nature of the two models comprises the ideal features associated with the spoken variety, on one side, and the ideal features associated with the standard variety on the other. The metalinguistic nature of the two models, instead, includes the metalinguistic features that are associated with each of the two linguistic abstractions. The two models (CA and SA) have an identical structure, i.e. they both have a linguistic and a metalinguistic nature, and they are identical in how they function. Giolfo and Sinatora (2011:121) describe them as follows:

"The contemporary situation of the Arabic-speaking world is characterised by vast linguistic variability [which] manifests itself as a systemic variation within the speech community between two 'reference-packagings' which are two distinct linguistic and metalinguistic models of reference. These are not part of the linguistic spectrum, they do not correspond to any variety of the language, but are present in the mind of the native speaker as distinct interiorised abstractions which serve as reference

frames for two distinct functions of linguistic expression, namely speaking and writing, within one cultural system. The objective distance between the two 'reference-packagings' appears reduced in the mind of native speakers, who perceive the two models as being part of one Arab(ic) cultural system."

The structure of the models that Giolfo and Sinatora introduce can be interpreted as the motor that triggers code-switching and mixing within the speech community and that, by being shaped by the individuals' cultural peculiarities, determines each speaker's speech features. Interestingly, Khamis-Dakwar and Froud conducted a neurolinguistics research project with native speakers of Palestinian CA that lead them to a delineation of diglossia that is remarkably in line with the conceptualisation developed by Giolfo and Sinatora. Their research aimed at determining how diglossia is represented in the speakers' brain, and they investigated (i) whether SA and vernacular varieties of Arabic are neurally represented as separate or interrelated linguistic systems (Khamis-Dakwar & Froud, 2007:2); and (ii) whether diglossic switching is rule-governed and can be predicted. By focussing on brain responses to lexical switching, they pinpointed that SA and Palestinian CA are represented in the brain as two distinct languages with distinct lexical stores in which Palestinian CA is stored as a first language and SA as a second (2007:163). Native Palestinian speakers process and practice lexical code-switching between SA and CA similarly to bilingual speakers and they show evidence that switching into or out of SA constitutes switching between two language variants of different linguistic systems (Khamis-Dakwar & Froud, 2014:294). This is in line with Giolfo and Sinatora's model of reference-packagings, which sees a dichotomy between the features pertaining to the spoken and the written channels and two distinct models of reference for these channels. These channels are well-defined at the abstract level in the speakers' minds and the two language varieties are separated at the neuro cognitive level, but at the practical linguistic level they are performed as one language, which is part of one Arab(ic) cultural system.

Similarly, experiments conducted by Ibrahim and Aharon-Peretz support the assumption that SA and CA have the status of two separate languages in the cognitive systems of Arabic native adults and that the representation of SA is that of the second language (2005:65). By considering Arabic as inherently structured around the abstract combination of two different linguistic poles, it

appears clear that proficiency and fluency in CA do not guarantee equivalent proficiency skills in SA. Khamis-Dakwar and Froud also investigated whether diglossia can be predicted, as Boussofara-Omar argues (2003). If it were possible to predict the switching between Arabic language varieties, then the existence of one or multiple varieties that result from diglossic switching could be theorised. However, empirical data do not support this assumption and, as Giolfo and Sinatora state, code-switching and mixing is triggered by the homomorphic nature of the two reference models native speakers have in their minds and it is empirically realised through language varieties by its combination with native speakers' cultural specificities (2011:121).

To sum up, we have seen in this section that metalinguistic features, such as language attitudes, are intrinsic components of language varieties. According to Giolfo and Sinatora's reference-packagings model, metalinguistic and linguistic features complementarily shape abstract models of reference of CA and SA that native speakers have in their minds for the spoken and the written channels. These abstractions are not language varieties, but models of reference. Language varieties are instead empirically realised by combining the reference-packagings model with cultural specificities and social variables. Aharon-Peretz (2005) and Khamis-Dakwar and Froud's (2007) data support the assumption that the two language varieties CA and SA refer to two abstract models that are separate in the speaker's mind. They also seem to show that code switching patterns do not follow specific and superimposed guidelines. This supports the reference-packagings theory that sees the two models of reference as part of one Arab(ic) cultural system, but their realisation into language varieties depends on the subjectivity of native speakers rather than on fixed rules. This led me to discuss authenticity in Arabic and to state that it is a consistent blend of language varieties with different degrees of use depending on metalinguistic features as well as social variables. Among social variables we can identify the topic, the formality of the situation and the speakers' knowledge of SA. Considering the afore mentioned reconceptualisation of diglossia based on the reference-packagings model, I argue that learners of Arabic should be exposed to real-life language use and language variation. This comprises focusing on both SA and CA, on diglossic code-switching, and on metalinguistic features and social variables that trigger such switching. I also argue that students of

Arabic, whose goal is to communicate with Arab speakers (Palmer, 2007:116) and whose model is the “educated native speaker” (Younes, 2015:25), require developing communicative competence through linguistic and sociolinguistic training as we will see in chapter four. I will now proceed to explain diglossic code-switching.

2.4.2 Diglossic code-switching

The term “code-switching is normally used when mixing between two different languages” (Van Mol, 2003:87), but it can also happen in diglossic speech communities as speakers tend to shift from the two language varieties through a process that is subject to certain fixed structures (Eid, 1988). Within this framework, code-switching in diglossic communities has been extensively approached in recent years (e.g. Abdel-Malek, 1972; Abu-Melhim, 1991; Holes 1993; Saeed, 1997; Mejdell, 2006; Bassiouney, 2006, 2009; Soliman, 2008). Within Arabic diglossia, linguistic changes “ultimately depend on the social contexts in which diglossic varieties are used” (Walters, 1996:173). Moreover, according to Myers-Scotton (1993) members of the speech community have developed a communicative competence that allows them to move within a framework where they know the social values of any variety. They have also developed “an evaluation metric for the likely social cost and potential social benefits of any code choice” (Walters, 1996:185). For this reason, “while the object of previous multiglossic variationist studies was based on sociolinguistic variables such as function, situation, topic and education, [studies on diglossic code-switching] shifted the attention from sociolinguistic variables towards the speakers’ behaviour” (Giolfo and Sinatora, 2011:110). This does not mean that the speakers’ behaviour is independent of sociolinguistic factors, but it means that it is not solely influenced by them. As we have seen above, it is influenced by linguistic, sociolinguistic and metalinguistic factors. Albirini’s study on code-switching patterns employed by educated speakers of Egyptian, Gulf, and Levantine dialects of Arabic in specific language domains, namely religious discussions, lectures, political debates, and soccer commentaries, shows that there are precise reasons why speakers switch from CA to SA and/or Classical Arabic (2011:541). These are:

- i. to introduce formulaic expressions;

- ii. to highlight the importance of a segment of discourse;
- iii. to mark emphasis;
- iv. to introduce direct quotations;
- v. to signal a shift in tone from comic to serious;
- vi. to produce rhyming stretches of discourse;
- vii. to take a pedantic stand; and
- viii. to indicate pan-Arab or Muslim identity

Thus, Standard Arabic is the code employed for expressing pan-Arab concepts and the motivational patterns can be linked with prestige, importance, eloquence, and seriousness. On the contrary, patterns of switching from SA to CA can be summarised as follows (Albirini, 2011:547):

- i. to induce parenthetical phrases and fillers;
- ii. to downplay a particular segment of the discourse;
- iii. to signal indirect quotes;
- iv. to simplify a preceding idea;
- v. to exemplify;
- vi. to mark a shift in tone from serious to comic;
- vii. to discuss taboo or derogatory issues;
- viii. to introduce daily-life sayings; and
- ix. to scold, insult, or personally attack.

As Albirini (2011:547) shows, one of the most transparent patterns of code-switching to CA relates to the use of parenthetical phrases and fillers, which are implemented to introduce a point that is not completely part of an utterance, but adds to its message. Fillers are sometimes used to interrupt the flow of SA and introducing CA discourse. Generally, speaking, CA is related to low prestige and everyday topics. Although this could seem a regression to Ferguson's diglossia, within diglossic code-switching speakers alternate between the two codes based on their perceptions of the statuses of the codes themselves and the specific sociolinguistic functions associated with them. These functions are set in advance of their use rather than emergent during the interactional process. In other words, speakers enter the communicative exchange with preconceived notions about the code polarity both in terms of language prestige and about the role the code polarity in indicating sociolinguistic functions of varying prestige and complexity (Albirini, 2011:547).

Myers-Scotton's model, the Matrix Language Frame (MLF), allows structuring different kinds of code choice, both in bilingual and in diglossic realities. She lists four different kind of codeswitching: 'codeswitching as marked choice'; 'codeswitching as exploratory choice', 'codeswitching as sequential unmarked choice' and finally 'codeswitching as unmarked choice'. Boussofara-Omar (2005) applies this MLF model to Arabic and argues that middle varieties are in fact diglossic switching, and that there is no variety conventionalised as a third form or ESA. The switch takes place within two languages in contact, or language varieties in contact. One of them is the Matrix Language, which determines the morphosyntactical nature of the speech and provide for some morphemes, and the second is the Embedded Language, which provides for other morphemes within the code-switching process. "What is being conventionalized are patterns of switching between two varieties, with the dialect as the matrix variety into which constituents (i.e. content morphemes) from CA/MSA are embedded" (Boussofara-Omar, 2005:55). According to Boussofara-Omar (2005:77), "the reconsideration of the third language and middle varieties in light of the model of the grammatical structures of code switching demonstrates that diglossic switching is non-random, systematic and predictable." Thus, Boussofara-Omar triggers a process of "re-conceptualisation of Ferguson's diglossia, as she perceives that it is in diglossia that the justification for a non-randomness at the linguistic level (linguistic constraints) as well as the motivation of variation (social motivation) should be found" (Giolfo and Sinatora, 2011:111). Boussofara-Omar defines patterns of switching between two varieties as conventionalised. On the contrary, as explain above, this thesis adopts the reference-packagings model and it argues that patterns of code-switching are based on native speaker's individual choices and variation lies in the Arabic native speaker's mind. The difference between this approach and diglossia is that the two reference-packagings cover complementary linguistic and metalinguistic skills. This complementarity make them part of one cultural system. Within diglossia, instead, both varieties encompass all language skills but they are used in different situations and are therefore mutually exclusive. Within diglossia, language variety use derives only from sociolinguistic circumstances and situations. In the reference-packagings model, in contrast, although sociolinguistic circumstances and pragmatic

reasons play an important role in influencing choices of language varieties, the key to language switching is the native speaker. I explore now how Arabic diglossia affects TAFL as well as the implications that Giolfo and Sinatora's model has on it.

2.5 Implications of Arabic variation for TAFL

In the previous sections, we have seen that the domain of SA use is predominantly that of reading and writing, whereas it overlaps with CA in speaking and listening. Due to the spread of CA on social media, written productions can occur in CA but this lacks a fixed codification and it is therefore not included in the skills to be taught within an AFL programme. Diglossia affects TAFL under four main perspectives. The first concern is related to which language variety should be taught. The answer ranges from one variety, either SA or CA, or both varieties. In the second case, it is necessary to establish which emphasis is to be allowed for each variety and in which order the two they should be introduced. Finally, I believe it is important to add a query about the availability of time and investigate whether there is there enough time within an AFL programme to cover the teaching of both varieties.

Most university Arabic classes focus on SA and usually offer optional classes of CA for advanced students or in study abroad programmes (Amara, 2006; Hee-Man & El-Khazindar, 2006; Ryding, 2006). In these cases, the teaching of CA is concise and frequently separated from SA (Palmer, 2007:114). Historically, the approach to Arabic teaching that favours SA "flourished after the Second World War and, as of today, it is still the most common approach" (Alosh, 1997:88). As Al-Batal (1992) reports, reasons given for opposing the teaching of CA include lack of materials for teaching vernaculars, limited class time and impediments to the decision of which vernacular to teach. Moreover, teachers of AFL often share the prevalent attitude that SA is the authentic Arabic, and that the vernaculars are its corrupted and ungrammatical diversions (Ryding, 2008:16). The reason for not incorporating CA within SA is threefold: first, the development of oral skills in SA is seen as a tool for students who can use it as a springboard to acquire later any spoken dialect in its social context, ideally by travelling to the target country; second, the use of SA within an environment as academic as a university classroom is considerate appropriate from a

sociolinguistic perspective; and finally, students of intermediate and advance levels are expected to perform at higher levels of abstraction, which makes the use of SA appropriate (Alosh, 1997: 89-90). Alosh (2009:54) doubts that the goal of training students to perform like Arabic native speakers, i.e. read a text in SA and discuss it in CA, solely through classroom instruction, is realistic and achievable. He (2009:54) claims that “this ability can be developed as learners gain fluency in a dialect learned during study abroad, not during the few contact hours available in state-side programs”. Moreover, he (2009:57) suggests that Arabic instruction:

“should focus on building a solid foundation in MSA for at least the first four semesters. The field should not be concerned with the so-called “linguistic reality.” It is the situation that exists for the native speakers of the language and also for advanced non-native learners. For beginners and intermediate learners, the task should be the learning of the Standard code first. I would follow two guiding principles. The first one would be the attainment of a certain level of proficiency in Standard Arabic (say Intermediate Mid), by the end of the fourth semester. As teachers, we should focus on developing communicative skills although no dialect is involved. These skills are transferable and would be useful when the learners eventually had the opportunity to travel to one of the Arab countries and learn its specific dialect. We should not deceive ourselves by thinking that we can replicate the process of native-speaker linguistic development. The differences between a native speaker and a non-native learner are so numerous and significant that they defy any comparison.”

Nevertheless, as Palmer (2007:112) states, students who study only SA face difficulties in communicating with native speakers, since they “are kept outside the in-groups and often experience frustration and embarrassment when trying to [interact] with Arabic speakers.” Language teaching models that largely focus on SA do not reflect the multifaceted nature of Arabic, and do not expose learners to its linguistic variety. By not being exposed to real-life language input, students are less likely to come into contact with native speakers’ authentic use of the language. While the main goal of AFL students in the 1960s and 1970s centred around reading old Arabic texts, all indications are that students now want to learn Arabic with the goal of functioning in all the language skills, namely understanding, speaking, reading and writing the target language in the way it is used by its native users (Belnap, 1987; Younes, 2006). Therefore, responding to the learners’ needs is equivalent to providing them with effective tools to perform in real-life situations. Since interaction is “somewhat off-limits to those who do not speak the appropriate code” (Palmer, 2007:112), both varieties –SA and one vernacular – ought to be taught. In line with this

assumption, new materials that incorporate vernaculars alongside SA are being developed (Ryding, 2006:15).

This leads us to the second main concern, which inquiries about the emphasis that should be allowed for each variety in an AFL programme. I agree with Eisele (2006:219) that usually “the choice is [...] a matter of the value systems of the learners and pedagogues involved.” A great emphasis is placed on CA by language teachers and curriculum developers who “value speech over other forms of language-based communication” (Eisele, 2006:219), whereas prevalence for text-based activities is to be found in Arabic academic studies that mainly focus on literature, politics and media. In my opinion the emphasis is to be functionally placed according to the relationship between language varieties and linguistic skills. Within a diglossic language curriculum, the ultimate goal needs to be centred on tasks, and the curriculum should be designed towards achieving task-based goals: two language skills, namely reading and writing, need to be performed in SA, whereas speaking and listening involve both SA and CA.

This is followed by the third of the afore mentioned perspectives: the order in which the two varieties should be studied. At the present time, there exist different approaches to Arabic teaching in which both SA and CA are included. I start by exploring the simultaneous method (Al-Batal, 1992; Fakhri, 1995; Al-Batal and Belnap, 2006; Wahba, 2006; Palmer, 2007, 2008) and I subsequently proceed to outline integrated approaches (Younes, 2009; Giofio and Salvaggio, 2018). Within simultaneous approaches SA is taught concurrently with CA and the two teaching tracks complement each other rather than competing (Ryding, 2006:17). There are specific reading and speaking goals for each track. According to this method, “the Arabic classroom can and should be a place [where] multiple registers co-exist, as they do in real life” (Al-Batal and Belnap, 2006:397). Al-Batal argues (1992:298) that the simultaneous approach seems to adequately deal with Arabic diglossia as it reflects the diglossic situation, as it exists nowadays in the Arab world, in the classroom. This is achieved by introducing SA as a written variety alongside one spoken dialect for communication. In line with this approach, Fakhri (1995) proposes to simultaneously teach SA and one vernacular and to develop reading and writing skills in SA, while introducing a vernacular in the listening and speaking skills.

Wahba (2006), suggests presenting SA and CA as separate entities at the early stages of learning, followed by mixed texts at the intermediate level and integration at advanced levels. Three major shortcomings have been identified within the simultaneous approach: first, it “does not account for variation in speaking brought about by situational, contextual, personal, and other factors” (Alosh, 1997:94). Second, speaking is restricted to CA and reading and writing to SA, thus creating an artificial dichotomy. Third, students run the risk of confusing the two varieties (Ellis, 2001).

Integrated approaches, instead, start with a language variety and gradually integrate the other. Younes (1990) and Giolfo and Salvaggio (2018) propose to start with CA and integrate SA (I define it here a CA-based integrated approach). In these approaches, the emphasis is placed on familiar and informal vocabulary and contexts at the beginning of the programme, for which CA is particularly appropriate. Reading and writing activities are also introduced at an early stage, and they are conducted in SA through a clear separation of tasks performed in the two varieties. This separation becomes “more fluid” as the level of proficiency of the learners increases (Giolfo and Salvaggio, 2018:7). As Younes (2009: 60) states, SA:

“occupies an increasingly more (sic.) prominent role in the curriculum with the move towards the less familiar, less concrete and more formal, but integration remains an important feature of the whole program. An attempt is made to develop the four language skills simultaneously. Speaking activities are conducted in [CA] throughout the course, while reading and writing are conducted in Fusha. One lesson typically involves work on more than one language skill, which results in a continuous and spontaneous movement from Fusha to [CA] and vice versa as a function of the linguistic situation and the language material that are being replicated. Following common practice by native speakers, material presented in Fusha is discussed in [CA], which contributes to the continuous movement between the two language varieties.”

Data show that the opportunity to learn a variety of Arabic that can be used in daily conversations with native speakers, greatly motivates the students. By analysing phonology and grammar, Haddad (2006) concludes that cognitively it is preferable to learn a vernacular before SA, but he argues that further research is needed in this area. “Typically, two objections are raised against the integrated approach: the fear of confusing students, and the difficulty or cultural/political sensitivity of deciding which [CA] should be introduced in the program to the exclusion of others” (Younes, 2009:63). According to Younes,

(2009:63), confusion between SA and CA is minimised in the integrated programme because the two varieties are introduced in the classroom through language materials that keep their skills separate: SA is presented through reading passages to be read and understood but not to be actively spoken (Younes, 2009:63). On the contrary, CA materials are introduced and regularly used as a foundation for speaking activities. The opportunity to develop the necessary skills to use in their proper contexts enables the students to develop a correct approach to the sociolinguistic realities of Arabic (Younes, 2009:63). According to Ryding (2009:50):

“The challenges to our field lie in integrating authentic spoken discourse skills and strategies into traditional MSA curricula to the extent that they are necessary for communicative competence at any proficiency level. These challenges include the materials, sequencing, design, and teaching of primary discourse skills. [...] What our curricula need is restructured access to both the primary and secondary discourses of Arabic. The new architecture of Arabic as a foreign language – including curricular goals, sequencing, and text-type – needs to be constructed with full respect to issues of discourse type, interactive functional skills, the building of firm foundations, and expanded definitions of linguistic, cultural, and social norms and appropriateness. It includes written Arabic as the cornerstone of literacy, and it includes spoken forms of Arabic, both colloquial and educated, as cornerstones of spoken fluency.”

Ryding (2009:51) adds: “This is not meant to imply that Arabic programmes should not teach individual dialects; but it does mean that, in the real world of Arabic usage, students need more than an acrolect [i.e. the H form] and a basilect [i.e. the L form]. They need to learn how to calibrate the formality of their speech, and how to distinguish and adjust to particular situations and regionalisms. [...] Learners need instruction, not just exposure. And part of that instruction incorporates the cultural and linguistic pragmatics of interactive discourse focusing on contextualized uses of language.”

A SA-based integrated approach, in which students begin with SA and are introduced to CA only at a more later stage, “has dominated Arabic pedagogy up to now, whether it is proficiency-oriented or not” (Eisele, 2006:218). Ryding (2006:16) calls it “reverse privileging”, as the vernacular language of the primary discourses of familiarity is postponed or minimalised, while the language of secondary and formal discourse is made central. She claims that it is discouraging and limiting for students to be denied early access to the vernacular skills with which they could informally interact with Arabic speakers. I agree with Ryding and I share the assumptions of the CA-based integrated

approach. However, as we have previously seen, the main pedagogical approach that encompasses instruction in both varieties begins by teaching SA and introduces CA only at a later stage. In order to reproduce the most widespread AFL teaching reality, this study follows the same order and introduces CA in groups that have already undergone formal teaching of SA. I highlight here five assumptions of the SA-based blended approach as presented by Alish (1997) that I undertake in this research:

- i. “The process of learning and teaching should be reflected by a rigorous methodology”: course materials should be achievement-oriented and students should be explicitly taught learning strategies (Alish, 1997:96).
- ii. “A foreign language course in an academic institution should not be designated with the purpose of turning out native or near-native speakers”. In light of this, foreign language programmes should set realistic and feasible goals (Alish, 1997:97). I agree with Alish that the goal of reaching near-native proficiency might be difficult to achieve. However, I believe that students should be helped and guided to become autonomous and independent learners, and to adopt effective approaches to Arabic acquisition. This should happen by engaging them in developing language learning and vocabulary building strategies. Although it is unlikely that students reach native-like proficiency levels through language courses, if the instruction aims at supporting them to become skilled and independent learners, it is also likely to promote, in turn, their language knowledge development until very high levels of proficiency.
- iii. “Arabic language varieties should not be taught as discrete systems”. Learners should be aware that the SA and CA forms they learn are perceived by native speakers as part of one single language (Alish, 1997:98);
- iv. In an Arabic programme there should be a separation between SA and CA based on specialisation of function in order to preclude any conflict between the objectives specific to SA (i.e. reading and writing) and those common to both (i.e. listening and speaking) (Alish 1997:100).
- v. “Teaching rules of use is as crucial as teaching rules of grammar” (Alish, 1997:97): students must be aware of when and how elements of SA and CA are used and both the reason and the way they alternate. In order to

reach linguistic proficiency in Arabic, learners also need to develop skills which allow them to perform native-like diglossic differentiation, thus to understand not only sociolinguistic factors that can trigger language variation, but also the reasons that lead native speakers to switch between them.

The final question I introduced at the beginning of this section asks whether there is enough time to efficiently cover both varieties. The answer lies, I believe, within point (b) among Alish's assumptions. TAFL comprises more than one language variety as well as sociolinguistic and code-switching skills to be developed. The workload that students need to undertake is undoubtedly challenging and extremely demanding. However, research can help teachers and instructors to support students in developing learning strategies that enhance autonomy. One of the main aims of this study is in fact to investigate teaching methodologies that enhance learners' autonomy, promote the efficient development of code switching skills and provide the students with the necessary tools to become interactive users in both language varieties.

2.6 Conclusion

The aim of this chapter was to identify and review the literature on Arabic variation upon which this research is based. Its purpose was also to analyse the impact that different approaches to variation have on TAFL.

The conceptual framework of the research is grounded mostly on linguistic analysis of Arabic diglossia, but other literatures were also drawn from, including Arabic Sociolinguistic and Applied Linguistics. The chapter reiterated the key features of Arabic varieties because they have a central role in this research. It also details their historical development, and it outlines my position towards their essential nature and their roles in the language community.

Based on a view that the varieties are capable of fulfilling complementary linguistic, sociolinguistic and ideological roles, I have elaborated my agreement or disagreement with the main theoretical models used to describe their nature and functions. Arabic varieties are regarded in this thesis as partly responding to over-imposed linguistic and sociolinguistic rules of use, and partly being a product of native speakers' ideologies. Moreover, although complementarity is a hallmark of language varieties, this does not entail mutual exclusivity of

functions and significances. For example, concepts such as prestige, identity and belonging can be expressed by both varieties.

The chapter was divided into four sections.

The first section gave a historical overview of the development of Arabic varieties and identified their differential features and their commonalities. This historical synopsis provided a detailed picture of variation as an intrinsic feature of Arabic. Moreover, I identified the historical changes in language varieties, with the aim of portraying how, and to what degree, their main characteristics have evolved. Language varieties' structures and nuances of significance were highlighted through their historical evolution. Arabic used to be, and still is, commonly referred to as a language composed of discrete varieties that fulfil specific functions. My position disagreed with this analysis and it defined it as being too restrictive. I proposed, instead, that language varieties can be dichotomous and respond to different contexts of use, but they can also fulfil similar functions. For example, only SA is appropriate in formal written communication media whereas only CA is appropriate in spontaneous oral conversations. However, both of them may be employed to define identity and express the speaker's position within the speech community. The difference among them is, in this case, the meaning of their intrinsic message. While SA conveys a sense of belonging to a pan-Arab(ic) reality, CA can express group identity or national identity.

The next section was structurally and theoretically based on the analysis of language varieties introduced in the previous part. While the focus of the first section was placed on Arabic varieties as singular realities, it prefaced the analysis of their combination for the purposes of communicating within the speech community, which was the scope of the second section. This section analysed how Arabic variation is performed within Arabic speech communities. This linguistic performance is referred to as 'Arabic diglossia'. Ferguson's definition of diglossia was key to explaining an interpretation of linguistic functions that has remained almost undisputed for decades and that reinforces the ideological stances of prestige and codification for SA, and linguistic corruption and absence of rules for CA. My theoretical approach disagreed with that of Ferguson and was founded on elements selected from other theoretical models. I also focussed in this section on emphasising an aspect of CA and SA

that is, in my opinion, greatly undervalued: I argued that while only SA can fulfil the role of the standard and codified language, both SA and CA can assume a prestigious role.

The third section of the chapter defined the model I adopt to describe language variation, namely the 'reference-packagings' model. The model was depicted as having two different levels. On the one hand, it reflects the dimension that language variation occupies in the native speakers' brains. On the other, I explained its use in communication acts. The reference-packagings model describes how the use of both language varieties occurs by means of code-switching. This allows speakers to abide by linguistic rules that require using appropriate forms in specific sociolinguistic contexts. Code-switching also allows for shifts between language varieties to implicitly express subjective language ideologies that are related to use of one language variety or the other. In this thesis, I refer to the switch between language varieties as diglossic code-switching, and I explained its characteristics in detail in this section.

The fourth and final section described the impact that the mainstream interpretations of diglossia, based on Ferguson's model described in section 2.3, have had on the field of TAFL. This field is characterised by the predominance of SA teaching at the expense of CA. Moreover, the emphasis placed on the two varieties is often perceived as necessarily exclusive. I conversely argued for the importance of methodologies that include both language varieties and that provide adequate emphasis to their roles in real-life use.

My theoretical position favours integrated methodologies that introduce CA and SA almost simultaneously, and that start with the former. However, my research applies to TAFL programmes that exclusively, or mainly, teach SA and where the introduction of CA occurs, if at all, only at a very advanced stage of instruction. The purpose of this investigation is to explore methodologies to teach diglossic vocabulary items and forms, and to analyse how the teaching of diglossic vocabulary affects students' language awareness and motivations. I propose to investigate teaching methodologies within the field of TAFL that do not deviate significantly from prevailing approaches for two reasons. Firstly, if the methodologies investigated are shown to be effective, they could be feasibly introduced within programmes of TAFL that already exist without altering them

drastically. Moreover, such an analysis could also provide insights into the limitation of the methodologies that are currently used.

3 CHAPTER 3 – Theoretical Framework

3.1 Introduction

In the previous chapter I sought to position myself within the wider field of analysis of Arabic diglossia and its impact on the field of TAFL. I also posited that the ability to master both SA and CA is of paramount importance to achieving language proficiency. That chapter has shown that the dichotomy between SA and CA has generated significant debate. SA is often perceived to be in conflict with CA and to challenge its unifying role provoking fragmentation, alienation, and the loss of shared cultural values (Ryding, 2013). The predominant approach to diglossia interprets it as having an unsettling effect on communication among Arabic speakers as well as a negative and disruptive influence on the education of Arab children. It also views it as having negative implications for socioeconomic, cultural and political advancement. This approach towards diglossia is the one that receives most attention.

This research criticises the implications of focussing solely on the negative aspects of diglossia and it argues that these negative readings overshadow positive attitudes towards CA and Arabic variation, and that they portray a distorted and misrepresentative picture of Arabic reality. It therefore highlights more positive readings and interpretations of diglossia.

First, the propagation of an unfavourable view of diglossia as being a disruptive element hinders the recognition of its positive features and use. Interestingly, researchers have observed instances of countervailing attitudes among educated Arab speakers, who are increasingly using a prestigious form of CA and expressing favourable views about the cultural values that it embodies (Albirini, 2016). Native speakers also seem to appreciate the role of prestigious spoken forms in representing national identities, and are amenable to the use of such forms in semi-formal contexts. Acknowledgement of these positive views, alongside the negative perceptions of diglossia, sheds a more accurate light on real-life language use. It also supports one of the assumptions of this thesis: the rejection of a dichotomous view of the roles and functions attributed to language varieties, in favour of a more flexible understanding of their codes and of the shades of opinion that native speakers have towards them. Based on this view, this research emphasises the significance of native speakers' ideologies and

linguistic choices, together with sociolinguistic factors, in understanding language variation and diglossic code-switching.

Second, the strident defence of an essentialised form of Arabic against the threat of colloquial influences perpetuates the idea of Arabic as a monolithic entity by denying its multifaceted nature, creating a deep cleavage between reality and theory. In the field of TAFL it also encourages language teaching models that focus solely or predominantly on SA, which is therefore the dominant instructional medium and instructional goal of most AFL programmes (Ryding, 2013). Pedagogical models entirely built on SA do not reflect Arabic linguistic variety and they lead to a discontinuity between language reality and teaching. This produces “linguistic uncertainty among native and non-native users of Arabic” (Palmer, 2008:84). On the one hand, native speakers face the challenge in schools to “unlearn the language used at home and among friends to relearn what is supposed to be their native language” (Palmer, 2008:84). On the other, students of AFL are not exposed to spoken varieties and their use in real-life situations (Farghaly, 2005:29). It would be of great interest to investigate pedagogical approaches aimed at supporting Arabic native speakers to overcome the above-mentioned challenges, but this is beyond the scope of this research, as I focus here on AFL learners and programmes.

The great stress I place on the role of both linguistic and metalinguistic factors in diglossic code-switching led me to adopt the reference-packaging model for language variation, as explained in the previous chapter. By attributing equal prominence to both elements, the model acknowledges the importance of understanding not only sociolinguistic features that are external to the speaker, but also pragmatics and metalinguistic factors that are subjective and therefore dependent entirely upon each speaker’s personal choices. For example, we have seen in the previous chapter that, as prestige applies to both language varieties, the choice of which variety to use in a formal and semi-formal context is often subject to native speakers’ perceptions of language roles and the metalinguistic messages they want to convey.

Considering this interpretation of Arabic linguistic reality, this thesis aims at identifying teaching approaches that include SA and CA, and that highlight the importance of sociolinguistic as well as ideological and subjective factors in prompting diglossic code-switching. The main objective of the afore mentioned

teaching approaches is to help learners of Arabic to efficiently develop diglossic vocabulary and to become proficient users of diglossic code-switching. This research also argues for an approach to the teaching of Arabic diglossia that exhibits CA prominence in the spoken language across a rich and diverse range of situations and levels of formality. This does not endanger the role of SA as the irreplaceable variety used in the two language skills of reading and writing. However, it encourages the development of speaking and listening skills in SA and CA, and that CA be used across different ranges of prestigious contexts.

The bases of the theoretical framework of this thesis are:

- i. Variation as an integral feature of Arabic;
- ii. The interpretation of Arabic variation and its representation through the reference-packagings model;
- iii. The application of the reference-packagings model to TAFL through the development of diglossic code-switching skills.

I explained the above three points in detail in chapter two with the aim of defining the pillars upon which the theoretical framework of this thesis is built. This chapter will outline the key features of the theoretical framework and is divided into three sections.

The first section describes an approach to variation as an integral feature of all Arabic varieties, which are intrinsically heterogeneous. It shows that diglossic code-switching is used at almost all levels and communication acts. This reinforces my position that students of AFL need to become skilled users of this linguistic practice. Moreover, it justifies my assertion that learners of Arabic ought to also develop noticing and interpretation skills that help to develop a facility in diglossic code-switching use and in intuiting the motivations that trigger it. In the same section I illustrate the reasoning behind my selection of the variables employed in this study.

The second section defines native-like proficiency. Its definition is based on the interpretation of Arabic variation in the reference-packagings model and its application to TAFL. I describe two Arabic proficiency benchmarks that I use as measures of expertise in both SA and CA. The first is the ability to “move” (Younes, 2006:162) between language varieties and the second is the ability to

use the language appropriately in order to convey meaningful messages, that is “socio-pragmatic competence” (Giolfo and Salvaggio, 2018:6).

The last section identifies the learning outcomes that this research seeks to measure for the teaching of Arabic diglossic vocabulary and code-switching. The learning outcomes and pedagogical methodology adopted in this research are based on the definition of proficiency outlined in the second section. While the pedagogical methodology is outlined in chapter four, I focus here on the definition and measurement of diglossic vocabulary development, language awareness and students’ perceptions and attitudes.

3.2 The systematic nature of variation

Although Arabic is referred to as a diglossic language, research has shown this to be an imprecise definition, as we have seen in the previous chapter. Numerous studies (for example Badawi, 1973; Meiseles, 1980 and Walters, 2003) criticise a binary approach to Arabic varieties and they claim the existence of multiple levels of mixing that lead to the creation of intermediate and mixed language levels between the diglossic and polarised varieties CA and SA. I have criticised the variationist approaches to Arabic in section 2.3.3 and I claimed that Arabic varieties are heterogeneous in their nature and therefore intrinsically various. We will see below that SA and CA varieties are of no exception.

3.2.1 Intrinsic variation of SA

Although SA is relatively homogeneous across the Arabic-speaking world, there are substantial and consistent divergences in vocabulary between different regions and countries. For example, according to Holes (2004:47), there are differences in SA “vocabulary that differentiate the Maghreb countries – principally Tunisia, Algeria, and Morocco – from those of the Mashreq, or eastern Arab world. [...] These often reflect underlying east-west dialect differences or [...] the different foreign languages with which each area has historically had most contact.” Wilmsen’s research analyses SA in newspaper texts printed in Egypt and Lebanon and has found vocabulary differences due to interferences from the colloquial Arabic spoken in the two countries (2010:122-

3). His research reveals that features are sometimes transferred directly from writers' vernaculars into SA, whereas in other cases SA syntax, such as sentence structure or the ordering of direct and indirect objects, is influenced by CA. Holes (2004:48) reports that spoken broadcasting is also characterised by the use of a plethora of language varieties. These range from pure SA to pure dialect and they show "a greater or lesser mixture of MSA and dialectal elements, depending on the speaker's (or writer's) perception of the formality of the context. [...] In some circumstances speakers switch between using a "colloquialized" MSA and a "standardized" colloquial as they perceive the demands of the speech context to change" (Holes, 2004:48) (quotation marks in original).

In his study, Magidow (2013) analyses the choices of Syrian native speakers in producing a written text in SA and how they identify a standard register in a given written text. He does so by investigating the elements that native speakers find inappropriate for the targeted register level and that are therefore filtered out from a SA text. The main results show a general avoidance of negative interference from CA, i.e. speakers avoid forms that occur in colloquial speech. Although they do not always agree on the best possible alternatives in SA, they are "united in their dispreference" (Magidow, 2013:151) for what is perceived to be too close to the colloquial forms. The existence of numerous colloquial varieties and the consequent differences in vocabulary items among them infer that it is not possible to reach an agreed-upon pan-Arabic perception of negative interferences from CA. The variations between vernacular forms are therefore likely to lead to "differences in speakers' avoiding strategies, thus leading to further variation in the terms regularly used in formal MSA" (Magidow, 2013:161). Similar results have also been reported by Parkinson (1992:237) in his study conducted with native Egyptian speakers.

Finally, Ibrahim's (1986) research shows that the boundaries of (un)acceptable forms of SA change within regional variation, and for example that Lebanese newspapers seem to allow more room for colloquial or overlapping forms than their Syrian counterparts. It is therefore clear that the perception of what is SA is not totally homogeneous among Arabic speakers and that "the boundaries between colloquial and MSA are largely maintained [and adjusted] by the speakers themselves" (Magidow, 2013:161). In his study about language ability

and ideology in the Egyptian Arabic speech community, Parkinson (1992:250) states that “Egyptian native speakers of Arabic are not of one mind when it comes to their formal language.” This further supports the reference-packagings model which argues for the role of native speakers’ metalinguistic models in their linguistic choices and performances. It also supports the claim that authenticity in Arabic is variation and that foreign learners of Arabic need to develop sociolinguistic expertise to pinpoint and decipher it. SA internal diversity is identified here for its role in supporting the my claims that variation affects every variety of Arabic and for this reason it represents authenticity. This does not exclude nor diminish its role as the codified and written variety, and it is in light of these functions that it is adopted in this study. SA is employed here as it is described in reference grammars (Haywood, 1998; Ryding, 2005) as well as dictionaries (Wehr, 2012).

3.2.2 Intrinsic variation of CA

Moving away from SA, we enter the domain of CA to find that this umbrella term comprises countless local vernaculars showing differences in their lexica, phonology and morphosyntax (Versteegh, 2001:107). The dialectological literature has suggested various taxonomies based on different criteria to classify the wide range of vernacular varieties and three of them are outlined here: geographical, anthropological and sociocultural. As I have said earlier, a “popular” (Albirini, 2016:31) classification of Arabic vernaculars is based on geographical factors and has identified five regional groups. According to Versteegh (2001:145) these are: the Maghreb, Egypt, the Levant, Iraq/Mesopotamia and the Gulf (Versteegh includes Yemen in this last group). This classification is geographical in that it considers linguistic characteristics that are largely shared by the speakers of the afore mentioned zones. Mutual intelligibility between speakers of different regional vernaculars may not always be achieved and “the greater the distance between any two points of comparison, by and large, the greater will be the differences between the ordinary vernaculars spoken in them’ (Holes, 2004:3). Thus, the varieties of CA spoken at the margins of any given area “differ from each other considerably, and certainly to the point of mutual unintelligibility, if we were to compare what might be called the plain uneducated vernaculars” (Holes, 2004:3). However,

these varieties are often considered “to share a collective history and a wide range of phonological, syntactic, morphological, and lexical features that justify their subsumability within the family of Arabic varieties” (Albirini, 2015:9).

There is a large linguistic microvariation within regional groups and vernaculars can be further subdivided on the basis of the countries in which they are spoken, such as “Omani Arabic” or “Tunisian Arabic” (Theodoropoulou and Tyler, 2014:23). At a further level of taxonomy, country-specific CA varieties can be divided on the basis of their human geography: urban, rural or Bedouin (Palva, 2006). This overlaps with anthropological criteria that differentiate CA varieties based on the main lifestyle of the societies in which they are spoken, i.e. nomadic or sedentary (Bassiouney, 2009:19). Due to a large interaction among nomadic and sedentary communities, the boundaries between their spoken colloquial varieties are not always straightforward, with rural and urban dialects also containing Bedouin features and vice versa (Versteegh, 2001:149). Various studies have suggested that urban varieties are considered to hold a prestigious status among urban and non-urban speakers. This is due to fact that cities often enjoy a higher socioeconomic status, more modern infrastructure and communication systems and their inhabitants have access to higher educational institutions than the rural and Bedouin populations (Abdel-Jawad, 1986; Al-Wer, 2002, 2007; Wahba, 1996).

As we have seen in the previous chapter, it is important to note that within each modern Arab nation-state the colloquial of the capital city is usually considered the most prestigious and acts as a local common spoken variety (Altoma, 1969; Ferguson, 1959; Haeri, 1991; Ibrahim, 1986; Miller, 2006; Versteegh, 2001) despite the fact that it is not insusceptible to language variation. Amer, Adaileh and Rakhieh (2011:23) report the example of the vernacular spoken in Amman, the capital of Jordan. They call it ‘Madani Arabic’ and describe it as being different from local dialects as well as being very diverse in itself. On the one hand, it is linguistically distant from the vernaculars of both the south and the north regions of Jordan. On the other hand, it is not “uniform throughout the city” (Amer, Adaileh and Rakhieh, 2011:23) either. For example, there seem to be a relevant difference between the variety used by the population who settled in Amman but is not originally from the capital, and the population that was born in the capital instead. Together with geographical and anthropological factors at

play in influencing intra-country vernacular variation, we have previously seen that several sociocultural aspects are also included. Sociocultural factors can be gender, religion, family education and economical background (Al-Wer, 2014; Haeri, 1991; Miller, 2007 among others). This shows us that what makes the varieties spoken in the capital cities act as common local vernaculars is independent from the fact that they have not undergone a standardisation process and that are not exempt from language variation. It rather seems to stem from language attitudes of native speakers that transfer the supremacy status of a given capital city onto linguistic features considered to be distinctive characteristics of the varieties spoken in the capital.

The overall picture is of a kaleidoscopic variation, and hence identifying specific and unambiguous definitions of CA is not a straightforward process. It also emerges that language attitudes play a fundamental role in defining language varieties and prestige. In fact, we have seen that most of the criteria associated with standard varieties, such as prestige and formality, are based on native speakers' assessments and perceptions to form their judgement on the varieties and are therefore "evaluative" (Milroy, 2001: 533). These criteria have influenced the choice of the colloquial variety to be employed in this research. I will briefly explain here the reasons that led me to this choice.

Variety of CA used in this study

When the process of selecting a CA variety for this research started, I decided to firstly apply the "popular" (Albirini, 2016:31) classification of Arabic vernaculars based on geographical factors, and I selected one of the five regions identified by Versteegh, namely the Levant, comprising the vernaculars of Jordan, Lebanon Palestine and Syria. The reason for selecting this specific area is familiarity: having spent three years living in Jordan and Lebanon between 2006 and 2010, this is the Arabic regional vernacular I am most familiar with in. Although I would not claim a native-like knowledge of the vernaculars spoken in the region, the exposure received to these varieties is significantly greater to those of any other area. Similarly, living in the region gave me the opportunity to internalise metalinguistic information which are unfamiliar to me elsewhere. I then went on to narrow down the selection and focus on one specific area within the whole region. This decision was

underpinned by the belief that CA instruction should be informed by sociolinguistic research as it sheds light on how the language is used in everyday life, and the vast majority of publications in sociolinguistics adopt a country-specific or group-specific division (Albirini, 2016:31).

From this perspective, I decided to focus on the Lebanese/Central Syrian vernaculars (Versteegh, 2001), and specifically on the vernaculars spoken in the capital cities, due to their role as prestigious colloquial forms and to my familiarity with them. As the participants in this study are beginner students of Arabic, I decided to employ only one capital-based CA to avoid confusing the learners and producing unreliable results. I favoured Damascene CA over Beirut CA as I am more accustomed with resources that teach the former to non-native speakers, other than the latter. The approach to Damascene CA in this study replicates that of Magidow's (2013:148), which sees it as the variety defined by Cowell's grammar (1964) and dictionary (1965). Cowell treats the colloquials of the whole Levantine area but it is largely based on Syrian vernaculars, and specifically Damascene CA. The structure of the language material used for this research is explained in chapter four.

3.3 Diglossic code-switching use in real-life communication

The previous sub-sections show that diglossic code-switching is regularly and steadily used in communicative situations. I focus below on three communication settings: diglossic code-switching in intra-dialect communication; in inter-dialects contexts; and written productions. The analysis of the way in which this linguistic practice is used by native speakers in linguistic interactions, provides crucial details on the use of diglossic code-switching in real-life communication.

For the purpose of accuracy, it is necessary to clarify here that diglossic code-switching comprises two different forms of shifting between language variations in this thesis. One form refers to the mixing of two distinct codes in native speakers' oral interactions. The other form is a shift between different language varieties that are used in relation to their media of communication: oral or written. This switch refers to the alternation of Arabic varieties that native speakers employ according to their specific domains and functions: SA "is used

for reading, writing, and formal, scripted (as opposed to spontaneous) speaking, while [CA] is used for ordinary conversation” (Younes, 2015:17). Although it is true that the use of CA for reading and writing has increased in recent years with the availability of the internet, particularly in the form of text messages, Facebook comments, and comments on newspaper articles written in SA, the general pattern of usage remains that described above: SA for reading, writing, listening and scripted speech, and CA for listening and spontaneous conversation. Educated native speakers of Arabic, i.e. those who master a CA variety and are educated in SA, move from SA to CA and vice versa depending on the linguistic situation or task in that, for example, they read the news in SA but they discuss it in CA.

As seen previously, Arabic varieties are not used by educated native speakers as independent from each other, but they complement one another to form one system of communication. We have also seen that they are suited for specific functions and both “are necessary for functioning in the full range of situations where an educated native speaker is expected to function. Without one or the other the proficiency of such a speaker is incomplete” (Younes, 2006:159). Therefore, the two forms of diglossic code-switching can happen in the same speech or linguistic performance and they intrinsically constitute Arabic variation. This is of great importance as it informs the definition of native Arabic proficiency I adopt, which is in “diglossic native proficiency” as defined by Wahba (2006:146). Learners therefore need to become “diglossic language user” (2006:146). The factors that determine this linguistic performance in Arabic can be of manifold nature such as linguistic, social and metalinguistic as previously elucidated. In the following section I focus on diglossic code-switching and its performance in oral conversations. This gives me the opportunity to explore the nature of the language that native speakers use in oral interactions and to inform my pedagogical approach to diglossic vocabulary teaching. I have individuated three different forms of diglossic switching that are relevant to the students of Arabic as they represent real-life situations in which the students are likely to find themselves in: (i) intra-dialect communication, i.e. oral interactions with native speakers of the same vernacular they have learned; (ii) inter-dialect or cross-dialect communication, i.e. oral interactions with speakers of different vernaculars than the one they have learned, and finally (iii)

written productions. Code-switching in written production is not as common as oral interactions, but we will see below here that it is not uncommon and therefore students should be aware of it.

3.3.1 Intra-dialect conversations

Intra-dialect conversations in Arabic are “constructs produced by the patterns of simultaneous choices that speakers in a community make” (Holes, 2004:345). The descriptive difficulty is that these language choices are “probabilistic, not absolute” and they are only partly linguistically predictable (Holes, 2004:345).

One aspect that researchers agree on is that SA is not used for everyday-life conversations by any linguistic community (Younes, 2015; Holes, 2004). In a relaxed conversation, the participants of any local speech community use their local dialect. As Holes (2004:354) puts it, “it is unlikely that any group [from any single speech community] would deviate markedly from the local linguistic common denominator, that is, the dialect features that they all share. This means that the phonology, morphology, and sentence syntax would be dialectal virtually whatever they were talking about; choice of vocabulary, however, which depends much more directly on topic, would be more variable.” This makes clear that words, idiomatic expressions, “linguistically frozen elements” (Holes, 2004:355), and complete sentences may be borrowed from SA, but they will be placed in a morphosyntactic CA structure. If there were significant and frequent shifts in the language structure itself, they would likely “signal changes in the role a speaker is playing or claiming for him/herself in a conversation” (Holes, 2004:358). Younes (2015:18) adds that this style mixing happens “for the most part subconsciously and effortlessly. It is an essential part of the educated speaker’s linguistic competence.”

Farghaly (2005:29) describes the use of diglossic code-switching in intra-dialect communication as being regulated by exact rules. According to him, native speakers “have grammars of both varieties [and] internalize the rules that govern the switch from one variety to the other” (Farghaly, 2005:29). If this proved to be true, programmes and courses of AFL could provide the students with the keys to mastering these rules and therefore diglossic switching. However, I disagree with Farghaly’s position, and I argue that there is no

evidence of clear-cut guidelines that regulate diglossic shifting so far. Moreover, the degree of variation is such that no single set of grammatical rules can account for it and a certain degree of flexibility is necessary to capture the full complexity of code-switching between Arabic varieties (Gardner-Chloros, 2009:106). In my opinion, it is only legitimate to describe code-switching in terms of the grammatical regularities which characterise it in a given context (Gardner-Chloros, 2009:106). In the context of intra-dialect conversations, we have seen above that sentences are morphosyntactically placed on CA structures and they follow CA phonology. The choice of vocabulary items is instead less predictable and can include SA frozen elements or SA items integrated in CA phonological and morphological structure, making it a function of pragmatics rather than syntax. This supports one of my claims in that, as the context and implicature are the drivers, then switching is driven by general reasoning processes rather than grammatical rules generated in the language faculty of the brain. This has a twofold implication for TAFL. On the one hand, it enhances the argument for the teaching of both SA and CA varieties with a major focus on developing speaking skills in CA, as native-like use of the spoken language entails a strong command of CA morphosyntactical and phonological structures. On the other hand, it highlights the importance of developing the ability to move between the two codes in the same language performance and particularly in the case of vocabulary and phonology. The language material developed for this research aims at including both these teaching and methodological implications for TAFL as we see in chapter four.

3.3.2 Cross-Dialect Interactions

Several studies have shown that cross-dialect interactions among native speakers of Arabic can take place in SA, but that this is not a widespread pattern and it is not their only option. They may also switch towards a European language, or one of them could use the CA spoken by the other interlocutor if he or she is familiar with it (Abu-Melhim,1991; Mitchell, 1986; Shiri, 2002; Walters, 2003). Soliman's study reports that native speakers seem to make few and marginal modifications to their native CA, yet achieving a high degree of comprehensibility in their conversations (2014:129). The results of the study show that native speakers speak primarily in their own vernaculars in informal

cross-dialectal situations with minimal borrowings from other spoken varieties and SA. This is partly due to an increased familiarity of native speakers with other Arabic varieties, as a result of their expanded exposure to CA varieties through the media, travel experiences and greater interaction among speakers of different vernaculars. SA is shown in the study to be occasionally used as a frame of reference, particularly through its root system and SA cognates, in aiding comprehension of unfamiliar utterances in CA. There are however limited borrowings from SA and more dependence on the vernaculars.

Trentman's (2011:44) study also argues that the ability to understand an unfamiliar colloquial is aided by the knowledge of a close colloquial more than by the knowledge of SA. This is based on "historical developments in the Arabic language that have led to the dialects sharing a number of phonological, morpho-syntactic, and lexical features against MSA" (Trentman, 2011:44). When speakers of unfamiliar vernaculars interact, they are therefore likely to use their own colloquial variety, or to borrow consistently from a widespread colloquial and only rarely resort to SA. For example, speakers of North African colloquial varieties are likely to make considerable modifications to their vernacular mother tongue and borrow consistently from Syrian CA due to its prestigious and widespread role in the Arab world (Soliman, 2014:127). SA and dialectal borrowings are reported by the participants of Trentman's study as being instigated by different factors and motives such as the speakers' exposure and attitude to SA and other vernaculars. Soliman (2014:129) reports an extract of a conversation among three participants as follows:

"[one participant] switched into MSA for elevating (classicizing) the conversation when the topic became slightly formal; nevertheless, this modification was not applied by [another participant] in the same conversation, who expressed her attitude towards using MSA in conversations as sounding unnatural. In [other] examples [she] was also observed to borrow from other dialects and she explained that she grew up in an Arab country different from [her native country] where she used to speak with other dialect speakers and, therefore, she was comfortable switching to another dialect in order to help comprehension. [The third participant] in the same conversation said that, although she understands [a few dialects], she feels shy speaking in any dialect other than hers."

This overview of educated native speakers' language behaviour in cross-dialect conversations can be used to inform AFL teaching pedagogy. The above study shows that except for speakers of North African vernaculars, native speakers tend to use their own colloquial varieties with minimal borrowing from SA. They

use SA, their local CA and their knowledge of other colloquial varieties to aid comprehension. Within the field of TAFL this informs my position as follows: both CA and SA ought to be employed to enhance mutual comprehension and can be used as a frame of reference. Students of AFL can develop such ability by being trained in switching their lexicon between varieties and they can also practice developing their comprehension skills by using SA to decipher cognates in CA and vice versa, and by referring to SA roots to guess the correct meaning of CA vocabulary. These skills are useful both in their exposure to new vernacular varieties and in the acquisition of new CA vocabulary.

3.3.3 Written productions

Belnap and Bishop (2003) carried out a study in which they analysed written texts in Arabic and showed that hybrid forms are used, which are neither pure SA nor pure vernacular. They showed that speakers are concerned about making some types of errors in SA such as marking case endings, while other errors are perceived as acceptable for reasons of clarity. The study suggests that there must be a conscious process in which speakers try to conform their written language productions to both SA prescriptive norms and social pressures in order to be clear and reach a wide audience. Meanwhile, as we have previously seen, Wilmsen (2010) found evidence of interference of the writers' colloquial Arabic in formal SA newspaper-style texts. Research has shown that native speakers generally agree when asked to rank blended speech extracts and define their level of SA. However, more wide-ranging research on what makes texts seem more colloquial or more standard to native speakers, and what significance they attribute to different kinds and sizes of difference, has yet to be done (Holes, 2004:363). This implies that, when exposed to real-life written language material, students of AFL can encounter a certain level of mixing or interferences from local colloquials into the written production.

Authentic written material that reflects a mixture of language varieties can be used in class to train the students in pinpointing colloquial influences in SA lexicon and morpho-syntax. It can also be used to raise awareness among students about the role of CA and SA; to discuss the authors' attitudes towards the two language varieties, and their linguistic choices. These types of texts can

be used to create exercises of diglossic switching, in which students are asked to substitute vocabulary or linguistic structures that occur in the text as too colloquial, with more standard ones. By doing so, learners develop skills that are standard practice for educated speakers of Arabic when communicating with each other and with educated speakers from other parts of the Arab world: they spontaneously suppress features that have a particularly local flavour, or that are stigmatised, in favour of linguistic choices that are “more common or socially prestigious” (Younes, 2006:159). In this research students are made aware that diglossic switches can occur in written productions. As the participants are learners at beginner-level, I do not expose them to hybrid texts which could cause great confusion. Instead, I resort to groups of vocabulary words that are either SA only or that contain words in CA as well, and I ask the students to identify the linguistic nature of the vocabulary groups. This exercise is possible as many vocabulary items can be used both in SA and CA, whereas some words are part of one variety only. It is also particularly useful as it trains students’ noticing skills in pinpointing and detecting interferences of CA in the written language.

It is clear from the above-mentioned studies that the amount of SA and CA that a speaker uses depends on subjective and metalinguistic factors such as their attitudes towards the language, their command of SA and the metalinguistic message they want to convey in the communication process. We have also seen that SA is rarely employed in both intra-dialect and cross-dialect oral communications, but it is still used as a frame of reference in aiding cross-dialect comprehension and it is used to add frozen SA elements to the speech. Moreover, it remains the only standardised vehicle of written communication, although its purity is likely to be contaminated by the influence of writers’ spoken vocabulary. This supports the strand of research undertaken in this research and according to which authenticity in Arabic is reflected in variation and linguistic and “socio-pragmatic competence” (Giolfo and Salvaggio, 2018:6) are both essential to reach language proficiency. Language proficiency and the approach of this thesis to TAFL are tackled in the next section.

3.4 Arabic proficiency

Lightbown and Spada (2006:196) define communicative competence as “the ability to use language in a variety of settings, taking into account relationships between speakers and differences in situations.” Canale and Swain individuate four components of foreign language knowledge that are central to reaching communicative competence: linguistic or grammatical competence, which is knowledge of the target language’s linguistic forms and structures; sociolinguistic competence, that is knowledge of language use in context; strategic competence, or knowledge of how to succeed in communicative situations and ability to compensate for limited language resources; and finally discourse competence, i.e. knowledge of how written and spoken languages are combined grammatically and meaningfully in different genres (1980:28-9). They emphasise that learners should be provided with the information and experience needed in the communication process and that it is of essential importance to expose them to authentic communicative situations (Canale and Swain, 1980:28). They also state that the achievement of communicative competence in a foreign language involves all the components simultaneously without overemphasizing one of them (Canale and Swain, 1980:27).

The emphasis that Canale and Swain place on the importance of providing learners with sufficient sociolinguistic information and to expose them to authentic communicative situations supports my view that the development of communicative competence in Arabic is achieved through the development of linguistic, sociolinguistic and metalinguistic knowledge of language variation. Furthermore, the application of the features of communicative competence to Arabic entails that authentic spoken language must be part of discourse competence. This raises “an issue that needs to be decided within Arabic language teaching programs: the relationship between literacy and spoken language competence” (Ryding, 2013:70). My suggestion is that both are needed within AFL programmes, and both competences leads to proficiency in Arabic.

Alongside the development of discrete linguistic abilities and in addition to the four language skills for complete communication, i.e. speaking, listening, reading and writing, this research claims that learners of Arabic need to develop the ability to move between language varieties. As Ryding (2013:4) puts it, “the

key to being a functional Arabic speaker is flexibility and interconnectedness: the ability to operate at all levels and, even more important, to be able to navigate between them as required by different social contexts. Those who learn Arabic as a foreign language, therefore, face a daunting challenge: competence in a full spectrum of language varieties.” As the aim of learners of Arabic is to develop native-like language skills, the diglossic behaviour of native speakers ought to inform TAFL pedagogy. I report here two research studies that analyse native speakers’ diglossic development. The former suggests that Arab children develop “sensitivity to language arbitrariness as well as phonological detection and manipulation abilities similar to those of the bilingual children” (Eviatar and Ibrahim, 2000:453). The latter focuses on children’s morphosyntactic development in a diglossic environment, and analyses their ability to compare grammatical structures realised in the two different language varieties - SA learned in school, and CA acquired in everyday-life domains (Khamis-Dakwar, Froud and Gordon, 2011). Khamis-Dakwar, Froud and Gordon suggest that “skilful readers are those who can effectively shift between the skills and knowledge gained in their oral language to reading and writing in MSA” (2011:85).

Guidelines from both the American Council on the Teaching of Foreign Languages (ACTFL) (2012) and the Common European Framework of Reference for Languages (CEFR) (2001) indicate language proficiency as the ability to efficiently communicate socio-linguistically in real-life circumstances. Most importantly, they emphasise the concept of different stages and competences within language proficiency, which is not “an end result but a stage in a trajectory of development towards a specific outcome that more and more closely resembles the competence of an educated native speaker of a foreign language” (Ryding, 2013:6). The concept of stages allows for the differentiation between ‘proficiency’ and ‘performance’. The former refers to the level of knowledge that the learner has of the target language. The latter reveals the communicative competence that the learner has at different stages of their learning process. The CEFR combines an approach to mastering a foreign language based on six different levels of proficiency, and it also focuses on using the language in authentic situations. This leads to the description of language use in authentic situations as provided by Younes (2015:25): SA used

in reading and writing, whereas both CA and SA are used in listening and speaking. The difference in the latter group is that learners of Arabic at beginner levels need to master the skills of speaking and listening in CA for use in ordinary conversation, whereas students at advanced levels need to be able to use SA for delivering speeches or for formal interviews, and to adjust their CA to produce a form that is suitable for semi-formal conversations.

As Giolfo and Salvaggio (2018:6-7) point out, this

“...implies that the language varieties used by language learners should be the ones to which native speakers would normally resort to in [real-life] situations. [...] When considering the six CEF levels, which include both ‘ordinary conversations’ and ‘written and formal spoken purposes,’ we should then be able to differentiate for each of the socio-communicative tasks the language variety normally associated with it in that particular context. [...] Basic levels mainly refer to daily situations in which learners have to cope with different basic tasks (fulfilment of needs of concrete types, introducing themselves, shopping, traveling, etc.). In Arabic these domains are not normally covered by SA but are instead predominantly associated with CA. However, more advanced levels involve, alongside more complex listening/speaking skills (which involve both CA and SA), the comprehension and production of increasingly challenging written texts. These despite being normally associated with SA can well consist of a mix of CA and SA or even CA only.”

As previously explained, I consider proficiency not only as the ability to develop knowledge of language varieties in the language skills that distinguish their use, but also as the ability to move between language varieties. The switch between them is a response to different sociolinguistic requirements, as well as metalinguistic stances. Therefore, proficiency in Arabic comprises the ability to master two language varieties and to perform diglossic code-switching. The focus of this research is specifically on the development of diglossic vocabulary and diglossic code-switching skills, together with the investigation of language awareness and students’ attitudes and motivations developed while focussing on building diglossic vocabulary. I explain the focus on this research’s investigation in detail in the following section.

3.5 Focus of investigation

This section details the threefold focus of this research: the development of diglossic vocabulary and language awareness; and students’ attitudes towards Arabic varieties and their study. This section is organised in three parts and each one describes one specific area of focus of the research.

3.5.1 Diglossic vocabulary development

I explain here how the development of diglossic vocabulary knowledge and code-switching metalinguistic and sociolinguistic skills are defined and measured in this research. The evaluation of the learning outcomes, instead, is explained in detail in the methodology chapter (Sections 4.3 and 4.4). Vocabulary knowledge is a multi-faceted phenomenon about which there is still limited knowledge. However, one of the most commonly agreed-on views is that it occurs along a continuum of development. The fundamental idea is that non-native learners move along a continuum in which receptive and productive vocabulary are placed at either end of a continuum of knowledge and that the former comes before the latter and is an essential requirement of productive knowledge (Melka, 1997; Pigott, 1981; Palmberg, 1987). Henriksen (1996) adds to the receptive-productive continuum the depth-of-knowledge continuum in which degrees of word knowledge are operationalised at different levels of understanding or comprehension, from partial to precise and exhaustive. In order to apply this to Arabic and its diglossic reality, it is necessary to resort to Giolfo and Sinatora's approach to diglossia and Khamis-Dakwar and Froud's (2007:163) findings on the neurolinguistics representation of Arabic vocabulary as I described in the previous chapter. Khamis-Dakwar and Froud's data shows that although the two language varieties are perceived to be parts of one linguistic system, CA and SA vocabulary are stored separately in the brain and code-switching between them resembles bilingual switching.

The combination of Khamis-Dakwar and Froud's findings with previously-mentioned approaches to vocabulary knowledge leads to the interpretation of Arabic word knowledge that I suggest here. Arabic vocabulary knowledge encompasses two language varieties and there exist two separate receptive-productive knowledge continua, one for SA and one for CA. Also, different levels of depth-of-knowledge correspond to every word along each continuum. Depth of vocabulary knowledge refers to the accumulation of knowledge developed through encountering and using vocabulary items in a variety of different contexts that lead to learning their forms, meanings, and uses. I define 'diglossic code-switching skills' as the skills that comprise the speaker's ability to create semantic, phonological and grammatical links between corresponding words in the two varieties. Diglossic vocabulary knowledge therefore takes

place at two levels: linguistic and metalinguistic. The linguistic level is the word knowledge that I define here as the command of corresponding CA and SA vocabulary along the receptive-productive and depth-of-knowledge continua. Diglossic code-switching skills are, instead, developed at a metalinguistic level. They are based on metalinguistic knowledge and the ability to create semantic, phonological and grammatical relationships between SA and CA and on sociolinguistic abilities to functionally switch between the two varieties. Arabic word knowledge can thus be summarised as the combination of pragmatic competence (Taha, 2006:360-1), diglossic vocabulary knowledge and diglossic code-switching skills of corresponding words in SA and CA.

I focus in this study on two aspects of diglossic vocabulary development: CA vocabulary retention along the receptive-productive continuum and diglossic code-switching skills. I investigate vocabulary retention only along the receptive-productive continuum because the amount of time available for language instruction within this research project was of twelve hours, and in my view this is not sufficient to also develop and assess the depth-of-knowledge continuum. The duration of the course is defined by multiple reasons, which are explained in section 4.5.2. The rationale behind this selection is twofold. First, I introduce CA vocabulary that the students are already familiar with in SA. In light of this, my assumption is that building CA lexis-knowledge should result in developing diglossic knowledge of corresponding SA-CA vocabulary. Second, my supposition is that by focussing on CA forms (phonology and grammar) and content (lexis and semantics) that are already known in SA, students are likely to develop a framework of reference that can facilitate the development of diglossic code-switching skills. For example, I focus on everyday-life verbs in the present tense; on the pronunciation of specific consonants; and on common beginner-level vocabulary that the students already know in SA. The details of the topics covered are provided in section 4.6.

I measure CA vocabulary-retention and diglossic code switching skills using Bloom's taxonomy (1956) of Educational Objectives, which introduces a set of six educational learning objectives organised into levels of complexity and specificity. His aim was to structure and better understand the learning process, and he proposed that learning engages three psychological spheres: cognitive, affective and psychomotor. He focused on the cognitive domain, which he

identified as overseeing processing information to develop knowledge. In 2001 Anderson and Krathwohl revised Bloom’s taxonomy. They introduced verbs for each of the categories and a rearrangement of the sequence of learning. Their revision of Bloom’s taxonomy is represented in the table in Fig. 1 below.

Figure 1 Anderson and Krathwohl’s revision of Bloom’s taxonomy.

lower order thinking skills			higher order thinking skills		
remember	understand	apply	analyze	evaluate	create
recognizing <ul style="list-style-type: none"> identifying recalling <ul style="list-style-type: none"> retrieving 	interpreting <ul style="list-style-type: none"> clarifying paraphrasing representing translating exemplifying <ul style="list-style-type: none"> illustrating instantiating classifying <ul style="list-style-type: none"> categorizing subsuming summarizing <ul style="list-style-type: none"> abstracting generalizing inferring <ul style="list-style-type: none"> concluding extrapolating interpolating predicting comparing <ul style="list-style-type: none"> contrasting mapping matching explaining <ul style="list-style-type: none"> constructing models 	executing <ul style="list-style-type: none"> carrying out implementing <ul style="list-style-type: none"> using 	differentiating <ul style="list-style-type: none"> discriminating distinguishing focusing selecting organizing <ul style="list-style-type: none"> finding coherence integrating outlining parsing structuring attributing <ul style="list-style-type: none"> deconstructing 	checking <ul style="list-style-type: none"> coordinating detecting monitoring testing critiquing <ul style="list-style-type: none"> judging 	generating <ul style="list-style-type: none"> hypothesizing planning <ul style="list-style-type: none"> designing producing <ul style="list-style-type: none"> constructing

Source: Table developed by Iowa State University on Anderson and Krathwohl’s revision of Bloom’s taxonomy²

The six levels of classification, which identify the process of learning, are placed from lower-order to higher-order thinking skills. They are: remember, understand, apply, analyze, evaluate and create. Each level is described by specific verbs that not only define its main characteristics, but can be easily translated into learning objectives and outcomes and can therefore help in developing an effective assessment plan. The application of this taxonomy to my research allows me to determine the language objectives and outcomes against which I measure the development of CA receptive-productive knowledge and diglossic code switching skills. I explain below how the six cognitive levels of Bloom’s taxonomy can be used to achieve these purposes.

² Table available at: <http://www.celt.iastate.edu/wp-content/uploads/2015/09/RevisedBloomsHandout-1.pdf> (last accessed date: 13/12/2017)

Diglossic vocabulary learning outcomes

I outline the learning outcomes of both CA vocabulary retention and diglossic code-switching skills based on the six cognitive levels of Bloom's taxonomy. Within each of the six levels I have selected the key verbs that allow me to delineate the expected outcomes. It is important to mention that I have delineated learning outcomes that can be applied to the participants of my study, i.e. beginner students of SA at higher-education levels who have never been exposed to any variety of CA. However, their flexibility allows for the adaptation of the six level to students with different levels of language proficiency. Similarly, they can also be adapted to students that are familiar with CA but that have never been exposed to SA, or to learners that have a certain degree of knowledge of both varieties. I list below the key verbs I selected for each level, and I explain how they allow me to delineate the expected outcomes.

For the first cognitive level, 'remember', I selected the following verbs: arrange, define, recall and recognise. Students are able to recognise and recall newly learned CA vocabulary, as part of the learning outcomes expected within their CA vocabulary receptive-productive knowledge. They are also able to identify isolated words, for example from a list that is not placed in context, as CA or SA vocabulary, as part of the learning outcomes within their diglossic code-switching skills.

The second cognitive level is 'understand' and I choose the following verbs within this group: compare, contrast, describe, explain, extrapolate, match, paraphrase (retell), predict and translate. Students are able to translate from CA into their mother tongue and vice versa, and they can also explain, describe and extrapolate phonological and grammatical patterns that apply to CA, as part of the learning outcomes expected within their CA vocabulary receptive-productive knowledge. As per their diglossic code-switching skills, learners are expected to be able to linguistically and semantically contrast and match newly learned CA vocabulary against correspondent already known vocabulary in SA. They can retell SA texts in CA. They can also extrapolate phonological and grammatical patterns that link SA and CA and use these patterns to predict how to form new vocabulary.

Within the third cognitive level, 'apply', I selected the following verbs: apply, execute and role play. I identify the learning outcomes as follows: firstly, students are able to apply their linguistic knowledge of CA to form correct sentences, as part of their CA vocabulary receptive-productive knowledge. Secondly, by virtue of their diglossic code-switching skills they are also able to functionally apply their sociolinguistic knowledge to use CA and SA in authentic contexts through role plays.

The fourth cognitive level is 'analyse'. The key verbs I choose within this level are: contrast (discriminate), examine and identify. The expected learning outcomes can be listed as follows: within their CA vocabulary receptive-productive knowledge, students are able to break down language material into constituent parts, determining how the parts relate to one another and to an overall structure or purpose. They are also able to analyse complex language information to identify the most appropriate and correct variety to use, as part of their diglossic code-switching skills.

The fifth cognitive level is 'evaluate' and the verbs I opted for are: assess and estimate. The expected learning outcomes of CA vocabulary receptive-productive knowledge and diglossic code-switching skills in this level coincide. Students can assess their language use based on their own judgement. They are aware of their progress over time.

The last level is 'create', and I suggest that key verbs are: compose, construct, create, design, generate, invent, plan, produce, rewrite. Within this level students are expected to be able to reorganise elements into a new pattern, to create and give novel commands. The expected learning outcomes that I suggest within the CA vocabulary receptive-productive knowledge list learners' ability to invent new details for a story, generate and formulate answers to hypothetical questions. Within this category the students use language varieties not only in response to sociolinguistic appropriateness, but they are also able to make subjective choices as to how to use diglossic code switching to better achieve their communication goals. The diglossic code-switching learning outcomes of this level, instead, describe learners as being able to extrapolate phonological and grammatical patterns that apply to CA and that link SA and CA, they can develop their comprehension skills using SA to decipher cognates in CA (and vice versa) and refer to SA roots to guess the correct meaning of CA

vocabulary. These skills are useful both in their exposure to new vernacular varieties as well as to acquire new vocabulary in the varieties of CA they are already familiar with. In my view, the limited amount of instruction the participant receive does not allow one to set learning objectives and investigate learning outcomes across the whole range of categories included in the taxonomy and I therefore focus only on the first five categories. Further research is thus needed to investigate the learning outcomes of this category. The evaluation of the learning outcomes included in the first five cognitive levels is carried out through quantitative analysis, which is explained in detail in the methodology chapter below (Sections 4.3 and 4.4).

My adoption of Bloom's taxonomy is not carried out without reservation in this research, and I partially differ from its hierarchical and sequential arrangement of cognitive skills. I regard the skills identified by the taxonomy as being interrelated and dependent on each other to function most efficiently and effectively. For example, the three high-order cognitive skills (Analyse, Evaluate and Create) can occur concurrently at a cognitive level in the execution of an exercise. Let us consider a task in which learners are asked to link written SA nouns, verbs, adjectives and discourse markers to their CA counterparts. The task demands the learner to analyse how the vocabulary in the two varieties relates syntactically and semantically. Every time an inference is made about each word, they will be creating meaning, and the correctness of every inference needs to be evaluated. In this research, there is no assumption that cognitive skills are placed along a sequential order as Bloom's taxonomy suggests; rather, this categorisation of cognitive skills is employed to describe different processes that take place in diglossic vocabulary development. For example, it includes phases related to creating hypotheses about how the target language works, to creatively seeking opportunities to test those hypotheses, to producing communication strategies that creatively compensate for lack of knowledge of foreign language words, and finally to autonomously investigating efficient metacognitive strategies to learn. Bloom's taxonomy also offers the opportunity to develop learning objectives and expected outcomes based on the descriptions it provides of the cognitive learning phases.

In conclusion, Bloom's model, especially in Anderson and Krathwohl's (2001) adaptation, is used here as a holistic classification of the different objectives

that can be set for students in the cognitive domain of learning. It also allows me to define competences within language proficiency as emphasised in the guidelines from both the American Council on the Teaching of Foreign Languages (ACTFL) (2012) and the Common European Framework of Reference for Languages (CEFR) (2001). They highlight the existence of different stages and competences within language proficiency, which, as previously seen, is not an end result but a stage of development (Ryding, 2013:6). The concept of stages allows me to differentiate between 'proficiency' and 'performance' as previously delineated. The former refers to the performance of the learner in the target language, whereas the latter reveals students' communicative competence at different stages of their learning process.

I have explained in this section my approach to diglossic vocabulary development through CA vocabulary retention and diglossic code-switching skills, and I now move to the second focus of investigation: language awareness.

3.5.2 Language awareness

Language awareness is defined in this thesis as cognizance and explicit knowledge about diglossic code-switching and conscious perception of language use. It is also involved in the ability to switch between language varieties. Learners are cognitively aware of the substitution that they perform during diglossic code-switching; they report being conscious of their experience; and they are able to provide a description of the rule they have applied to realise such switch. This thesis defines awareness on the basis of the criteria identified by Allport (1988): a display of diglossic code-switching realised through verbal or written production of the targeted form, and a description of the underlying rule employed to realise such switching. The description of the rule is fundamental to demonstrate awareness of the linguistic experience. Based on Allport's criteria, Leow (1997) conducted a study to quantitatively and qualitatively address the role of awareness in foreign language behaviour. His target item was a Spanish morphological form, and he analysed students' recognition and reproduction of the target form as well as think-aloud protocols produced by the same learners. In think-aloud protocols, participants think

aloud as they are performing a set of indicated tasks. This provides the researcher observing the process with insight into the participant's cognitive processes and it allows for thought processes to be as explicit as possible during the performance of the task. Leow (1997:126) identified three levels of awareness:

- i. presence of cognitive code-switching; absence of meta-awareness; absence of morphological rule formation (where participants did not provide a report of their subjective experience, nor did they verbalise any rule);
- ii. presence of cognitive code-switching; presence of meta-awareness; absence of morphological rule formation (where participants did report their subjective experience but did not provide any verbalisation of the rule);
- iii. presence of cognitive code-switching; presence of meta-awareness; presence of morphological rule formation (where participants provided both a report and a verbalisation of rule formation).

Leow put forward three conclusions based on the data of his study. First, different levels of awareness lead to differences in processing diglossic code-switching. More specifically, meta-awareness appears to correlate with an increased use of hypothesis-testing and morphological-rule formation, while the absence of meta-awareness correlates to an absence of such processing. Second, the findings indicate that greater awareness contributed to recognition and accurate production of the noticed forms. Finally, the outcomes of the research provide empirical support for the facilitative effects of awareness on foreign-language behaviour.

Language awareness is measured in this thesis using Leow's levels of awareness. As I explain in detail in the methodology section in chapter four, the language post-test and delayed post-test that the participants take after the end of the language course are designed to test their oral accuracy (cognitive code-switching) in performing diglossic code-switching. Their scores are analysed against retrospective think-aloud protocols. In retrospective think-aloud protocols, participants describe their cognitive experience of performing diglossic code-switching immediately after having performed it. More

specifically, in the retrospective think-aloud protocols, the participants describe their mental processes (meta-awareness) and the steps (morphological rule formation) they took while carrying out code-switching tasks.

I have identified here my approach to language awareness and I now move to the third focus of investigation: students' perceptions and attitudes.

3.5.3 Students' perceptions and attitudes of Arabic variation

As mentioned in the previous chapters, all indications are that students want to learn Arabic with the goal of functioning in all the language skills in the way they are mastered by native speakers (Belnap, 1987, Younes, 2006). Therefore, responding to the learners' needs is equivalent to teaching both Arabic varieties and providing the students with effective tools to perform in real-life situations. This research tries to respond to the learners' needs by integrating CA teaching into a SA programme, and it does so experimentally through a combination of communicative-based teaching and focus-on-form instruction (FFI). Students' perceptions and attitudes and responses to the teaching approaches they are exposed to, are informed through qualitative data collected via open-ended questionnaires. This thesis does not use standardised attitude or perception scales, and it relies on self-reporting questionnaires that investigate students' perception of CA and SA in relation to their language performance, usefulness of the instructions in improving vocabulary and autonomy and, finally, their enjoyment of language learning.

3.6 Conclusion

In this chapter I have outlined the theoretical framework employed in this research. This framework is based on three theoretical perspectives that I outlined in the previous chapter: the understanding of language authenticity in Arabic as a consistent performance of language variation; the definition of language variation through the reference-packagings model; and the definition of language proficiency within TAFL as a combination of two abilities. These being, firstly, the ability to master two language varieties, and, secondly, to code-switch between them as native speakers do. There are two forms of code-switching that are performed between Arabic varieties: the shift between

varieties during oral interactions, and the switch of varieties as a function of the communication medium, which can be broadly summarised as attributing SA to the written medium and CA to the oral medium. Key to understanding language variation in Arabic is that, at a pragmatic level, the two varieties are performed as part of one Arab(ic) cultural system.

The chapter is divided into three sections that collectively contribute to building the theoretical framework of this thesis. In the first section I analyse variation as an integral feature of all Arabic varieties. This allows me to demonstrate that every Arabic variety is intrinsically heterogeneous and that diglossic code-switching is a fundamental feature of language use. Not only do native speakers switch from SA to CA and vice versa, as is claimed by approaches to Arabic variety asserting its dualistic entity, but they also shift between different forms of CA. This feature is important because of the consequences that it has for TAFL. The acknowledgement of diglossic code-switching as a daily and standard practice in Arabic language use entails the need to help learners to develop code-switching skills and to integrate a specific focus on its development within AFL programmes.

Alongside becoming skilled users of diglossic code-switching, learners also need to develop interpretative skills with which to analyse the reasons behind its use by native speakers in different situations. Moreover, the reconceptualisation of Arabic code-switching is key to the definition of language proficiency adopted in this thesis. Proficiency in Arabic implies not only possessing knowledge of SA and at least one variety of CA as the two main language varieties, but it also requires the ability to notice and perform diglossic code-switching. The former is the skill to understand the reasons that trigger the shifts between varieties as performed by native speakers, and the latter is the ability to actively perform code-switching in its two forms as described above.

Research is greatly in favour of the teaching of both SA and CA in AFL classrooms. I agree with Ryding (2013: 3) that CA should be included within the skeleton of the AFL programme because the key to being a functional Arabic speaker is the ability to employ and combine the diglossic, multifaceted components of the language to achieve functional and socially appropriate communication. It is also of relevance in support of my argument that students of Arabic strongly agree that they are able to integrate more into the culture in

an Arabic-speaking country when they communicate in a spoken variety of Arabic (Palmer, 2008: 91).

However, I argue in this chapter for, and I base my theoretical framework on, the importance of adding training in diglossic code-switching alongside the teaching of both varieties in AFL programmes. On this premise, a specific aspect of SA gains relevance in TAFL: the fact that its roots and linguistic patterns have an important role in the recognition of CA cognates. For example, learners of Arabic can use the root and pattern system of SA in order to correctly guess the meanings of unfamiliar utterances in CA, as native speakers do. I regard this aspect as beneficial and therefore valuable to explore, and for this reason I include it in my approach to the teaching of CA. I explain this in the methodological chapter, in the section on language instruction.

On the basis of the afore mentioned theoretical premises, I finally identify in this chapter the threefold focus of investigation of this research and I define each aspect in detail. Firstly, I explore diglossic vocabulary building. I investigate it through the analysis of diglossic-vocabulary and code-switching skills development. Diglossic-vocabulary building is prompted by means of developing CA vocabulary that the learners already know in SA. Secondly, I focus on the effect that CA vocabulary instruction aimed at developing diglossic skills and abilities has on language awareness. Finally, I investigate the impact of the afore mentioned instruction on students' motivation and attitudes towards Arabic variation, its study and their language performance.

I now move on to the methodology chapter, which explains in detail the approach I use to carry out the experimental language instruction; to collect data; and to analyse them.

4 CHAPTER 4 – Methodology

4.1 Introduction

This chapter details the research method used to conduct my empirical study and to answer my research questions. This research investigates the effects that integrating the teaching of CA into a SA programme at higher-education level has on the development of diglossic vocabulary and language awareness. It also inquires into the corollary effects on students' perceptions of variation in Arabic that result from the afore mentioned integration. The investigation is preceded by a pilot study and it involves students of three different universities. In chapter three I provided the detailed description of the focus of investigation of this research. First, I explained that by diglossic vocabulary development, I have in view a twofold definition that includes the acquisition of corresponding SA and CA vocabulary, and the development of diglossic metalinguistic and sociolinguistic code-switching skills to shift between language varieties. Second, I elucidated that diglossic language awareness is seen here as the ability to perform a linguistic switch that converts forms and vocabulary from a language variety into another, and it is combined with the cognitive awareness of the switch performed. The switch between varieties can happen in the same communication act during oral interaction or it can occur as a function of the medium of communication used: SA is used for written texts, formal presentations and formal oral contexts, whereas CA is used for everyday life oral interactions. Finally, the participants study CA and SA simultaneously as they attend CA lessons in the experimental language course organised as part of this research, and SA lessons in their academic programme. Students' views on Arabic variation are analysed to discern the linguistic challenge that studying two varieties simultaneously can pose and whether the approach to variation changes among students in the experimental and control groups.

This chapter is divided into five sections. The first summarises the main features of the threefold focus of investigation as elucidated in chapter three, namely diglossic vocabulary development; language awareness; and students' perceptions and attitudes towards variation in Arabic language reality and learning (Section 4.2). The second section identifies the pedagogical approach of FFI (Section 4.3). Focus-on-form refers to instruction that draws the attention

of learners to linguistic structures within a meaningful context. The section subsequently focuses on the two experimental methods of FFI I adopt in the language instruction of the empirical research. The two methods are distinguished by the nature of their focus: one exclusively focuses on the teaching of CA vocabulary and forms, whereas the other offers a constant and explicit comparison between corresponding vocabulary and forms in CA and SA.

In the third section I describe how I collect, measure and assess the data that inform this study, and how they are instrumental to answering the research questions (Section 4.4). I introduced the research questions in chapter one and I recapitulate them here. The first investigates the impact that two different and experimental methods of FFI have on diglossic vocabulary development. Two subsidiary research questions explore the role that the two experimental methods of FFI play in the development of language awareness, and in the shaping of students' opinions and attitudes towards Arabic variation. The last and main question sets out to explore the relationship that exists between the first three research questions. More precisely, it investigates whether and how the development of diglossic vocabulary knowledge and language awareness are interconnected with the shaping of views on Arabic variation as a linguistic factor and its acquisition in the AFL classroom. In the third section I also explain the design of this investigation and the use of case studies. I employ an embedded quasi-experimental design in which data are analysed both qualitatively and quantitatively through a mixed-methods framework. As this is a small-scale language education research project, quantitative data are analysed by means of non-parametric statistics. Qualitative data are analysed instead through NVivo software. I developed the research design and methods of inquiry specifically for this research as, to the best of my knowledge, a research method that investigates the relationship between the three afore mentioned components has not yet been developed.

In the fourth part I provide an account of the sources of this research and explain their selection followed by considerations of coding, research ethics and data storage (Section 4.5). I collected the data from 63 beginner-level Arabic students enrolled in the Universities of Milan and Genoa in Italy, and Exeter in the U.K.. The data-collection period ranged from September 2014 to June 2015;

prior to these studies I carried out a pilot study with 10 students between May and July 2014 at Exeter University. The reason why I have chosen these three universities is my familiarity with their courses, their teaching methodologies and facilities. I have completed my Undergraduate at the University of Genoa, my Master's Degree at the University of Milan and I am currently a PhD candidate at the University of Exeter. Due to my familiarity with these universities' systems I decided to confine my research to these three case-studies. Carrying out my research in any other university where I have less familiarity, would have raised risks of bias. The selected universities are not equally divided among their nationalities: two of them are Italian and only one is British. However, this does not impact the results of my research as I do not compare the results divided by nationalities but by students' learning outcomes.

In the last section I describe the language material and activities I designed and developed for this research. I outline the orthographic differences between the language material used with the experimental and control groups and I identify the differences of focus on forms between the instruction in the two groups (Section 4.6).

4.2 Review of focus of investigation

This section summarises the three focal points of this research, which I have extensively tackled in the previous chapter. I briefly recapitulate them here with the aim of restating their features clearly before explaining the methodological procedure through which they are measured and analysed. I start with diglossic vocabulary development in its main components, i.e. diglossic vocabulary building and diglossic metalinguistic and sociolinguistic code-switching skills development. I subsequently restate my interpretation of diglossic language awareness as well as students' perceptions and attitudes towards variation in Arabic language reality and learning.

Diglossic Vocabulary Development

This study focuses on two features of diglossic vocabulary development: CA vocabulary retention (along the receptive-productive continuum of word knowledge) and diglossic code-switching skills. The development of CA

vocabulary is a key element to building diglossic lexicon in this thesis on one condition: that the vocabulary items to be learned in CA are already known in SA. My assumption is that this ought to result in building knowledge of corresponding SA-CA items and therefore of diglossic vocabulary. Moreover, I hold the supposition that the focus on CA forms (phonology and grammar) already mastered in SA, encourages learners to develop a framework of reference that includes correspondent grammatical and phonological rules in both varieties. The two methods of FFI I employ, permit a comparison to be made between the effectiveness of teaching CA vocabulary and forms using two different methodologies: by means of explicit references to their SA equivalents, or without the creation of links to their SA equivalents. Beside their effectiveness, I explore the strengths and weaknesses of the two methods in leading to the development of diglossic knowledge and on shaping positive attitudes towards Arabic variation. I regard diglossic code-switching skills as having a sociolinguistic and a metalinguistic nature and drawing on the knowledge of diglossic vocabulary and forms. More specifically, these are the skills that allow learners to perform the shift between language varieties by combining linguistic knowledge with sociolinguistic and metalinguistic competence.

I have set out the theoretical ground upon which I develop the interpretation of CA vocabulary retention (of items already known in SA) and diglossic code-switching skills as structural factors of diglossic knowledge, and whereupon I outline the features of diglossic vocabulary development. I assess and analyse the students' learning outcomes in diglossic vocabulary development using Bloom's taxonomy of educational objectives, which identifies six learning objectives arranged into levels of difficulty. As explained in chapter three, I focus on the first five cognitive levels of the taxonomy. This is because my experimental language course covers twelve hours and I do not consider this a sufficient time to cover as far as the last cognitive level. The duration of the course is defined by multiple reasons, which are explained in section 4.5.2. I list below how I adapted the five learning objectives of Bloom's taxonomy to diglossic vocabulary development. I firstly detail how they apply to CA vocabulary retention, and I subsequently focus on diglossic code-switching skills.

The five cognitive levels of Bloom's taxonomy that apply to CA vocabulary receptive-productive knowledge are identified as follows:

- i. Students recognise and recall newly learned CA vocabulary;
- ii. Learners translate from CA into their mother tongue. Also, they extrapolate and explain phonological and grammatical patterns that apply to CA vocabulary items;
- iii. Students form correct sentences in CA;
- iv. Learners can infer meaning from a text by analysing how words relates to each another syntactically and semantically;
- v. The expected learning outcomes of the fifth learning objective for CA vocabulary receptive-productive knowledge and diglossic code-switching skills coincide. Students can assess their language use based on their own judgement.

The five cognitive levels of Bloom's taxonomy apply to diglossic code-switching skills as follows:

- i. Students can distinguish groups of words as being CA or SA;
- ii. They can link linguistically and semantically newly learned CA vocabulary against correspondent already known vocabulary in SA; students can retell SA texts in CA; extrapolate phonological and grammatical patterns that link SA and CA; and use these patterns to predict how to form new vocabulary;
- iii. Learners apply their sociolinguistic knowledge to use CA and SA in authentic contexts through role plays;
- iv. They analyse complex language information to identify the most appropriate and correct variety to use;
- v. The last expected learning outcome coincides with that of CA vocabulary receptive-productive knowledge. Students can assess their language use based on their own judgement.

The first four learning outcomes are measured through a language test that I administered to the students twice. Participants sat the test at the end of the experimental language course, and they resat it after a three-week period. Details regarding the administration of the test are provided in in chapters five and six, in which they are also analysed. The last learning outcome, which is

shared between CA vocabulary receptive-productive knowledge and diglossic code-switching skills, was evaluated by measuring student's language awareness, which is defined through Leow's levels of awareness. This is explained in the following section.

Diglossic Language Awareness

Language awareness is defined in this research as the explicit knowledge of diglossic code-switching. I assess this knowledge by evaluating code-switching tasks that the students perform in their language tests. Students are not only requested to carry out several diglossic code-switching tasks, but also to verbalise the grammatical rules they apply to realise such tasks, and their cognitive experience of the switching between language varieties. This entails the existence of three levels of language awareness, as outlined by Leow and explained in chapter Three in section 3.5.2. Leow's levels of awareness are identified as follows: presence of cognitive code-switching; presence of meta awareness (awareness of the subjective experience of diglossic code-switching); and presence of morphological rule formation (verbalisation of the rule applied to realise the switch). The evaluation of language awareness takes place in three different stages. First, I assess the presence of cognitive code-switching by evaluating diglossic code-switching tasks that students are requested to perform in the language tests (The language tests therefore measure CA vocabulary-retention; diglossic code-switching skills; and presence of cognitive code-switching in diglossic code-switching). The other two levels of language awareness are measured using information provided by the students and collected via retrospective think-aloud protocols. These are retrospective moments of reflections, in which the participants describe the mental processes they experienced while performing the code-switching tasks, and they explain the rules applied to perform the said tasks. Their level of awareness related to their use of diglossic switching also informs the fifth learning outcome of Bloom's taxonomy.

Students' perceptions of Arabic variation

As explained above, I use two different methods of FFI to introduce CA vocabulary and forms. One method draws explicitly on the students' previous knowledge of SA, whereas the second method focuses exclusively on CA. I employ the two methods with groups of students that participate in this research during the academic year. This entails that the attendance of CA lessons that are part of this research, occur over the same period of time as the attendance of SA classes at the students' Academic Institutions. Thus, all the participant experience studying CA and SA concurrently during this research. After the conclusion of the experimental language course, students answer an open-ended questionnaire which allows me to analyse the impact that the experimental teaching methods employed in this research have on their perceptions of Arabic variation; to assess the students' evaluation of the effectiveness of the instruction received, and to gather in-depth reports on their experience of studying CA and SA simultaneously.

I compare the information gathered here with data that I collect before the experimental language course begins. These data provide information on students' attitudes and motivations towards SA and towards its study as a foreign language. Students also share their learning strategies and the methodologies they have been exposed to in their learning experience; their assessment of Arabic complexity; and the main challenges perceived in improving in the different language skills. The comparison between the data collected before and after the language course allows me to better understand the impact that this research's experimental methods of instruction have on the development of students' views and motivations towards Arabic language and learning.

I have recapitulated the threefold focus of this research and the variables studied. I now move on to describe the methods of instruction used in the language courses of this research.

4.3 Focus-on-form instruction

In this section I define the teaching methodology employed in the empirical classroom research of this study: a combination of communicative language

teaching (CLT) and FFI. I explain how I integrate the two methodologies, the difference in the teaching approach adopted in the control and experimental groups, and how I apply these approaches to the teaching of Arabic.

Meaning and Form

Academics and educators agree that foreign language instruction is most effective when it includes attention to both semantic and structural aspects of the language taught (Lightbown and Spada, 2006: 184; Ellis, 2001; 2006; 2000, Norris and Ortega, 2000). On the one hand, learners that are exposed and engaged in meaning-based language teaching are likely to develop comprehension skills, oral fluency, communicative abilities and self-confidence. On the other, cognitive instruction and particularly focus on form seems to help learners to improve pronunciation and acquire morphological and syntactic features of the target language (Lyster, 1994). Doughty and Williams (1998:212) describe the term form as including all aspects of the target language, ranging from grammar to vocabulary, and according to them, it is highly likely that lexical acquisition can be enhanced through FFI. It is important to underline that, although there seems to be a dichotomy between meaning and form, I do not imply that form does not have meaning or that meaningful language use does not have form. I only stress a distinction between form and meaning in this research with the aim to differentiate a meaning-focused approach from a structured focus that primarily involves attention to linguistic features.

Characteristics of FFI

Since the teaching of CA in this research focuses both on meaning and structure, it combines CLT and FFI. First, CLT is aimed at helping learners to develop comprehension skills, oral fluency and communicative abilities. Explanation and examples of communicative-based exercises and activities are provided in detail in the sub-chapter on the teaching material developed for this research (Section 4.6). Second, FFI concurrently adds cognitive activities to CLT to enhance vocabulary and metalinguistic skills development. The cognitive approach to language learning is applied on the assumptions that cognitive theories make about the nature of language, language learning and

the process of learning. Although it does not generate fluency-facilitator knowledge, this approach encourages learners to cognitively engage in understanding how the target language works (Ellis, 1997). For this reason, I am interested in exploring the impact that it has on the teaching of Arabic varieties. It is plausible to expect, for example, that it encourages learners to wonder about the nature of CA forms and vocabulary and to relate them to their SA equivalents. Also, it is likely that cognitive instruction encourages learners to explore how words change phonetically and morphologically from one variety to the other, and which are the patterns of switch that link the two varieties.

The reason that led me to choose FFI as the cognitive methodology to employ in this research is twofold. Firstly, Lightbown and Spada (2006:181) show an increasing agreement among researchers that FFI helps learners in content-based or communicative instruction to acquire language structures that are unlikely to be learned without training. This applies to my research because I focus on the development of diglossic vocabulary and metalinguistic skills at the beginner level of language proficiency, both of which require formal instruction and are burdensome to be learned without. Secondly, students that start studying a foreign language when they are beyond childhood, such as at higher-education level, and whose exposure to the target language takes place primarily within the classroom, seem to benefit from FFI in that it helps them to efficiently use their limited exposure to the language they are learning (Lightbown and Spada, 2006). Various definitions of FFI exist, and I use what Lightbown and Spada label as isolated and integrated FFI. Both consist in attention to form within a programme that is primarily communicative, but the former is provided in activities that are isolated from content-based interaction, whereas the latter occurs during CLT.

Isolated FFI occurs in discrete units of metalinguistic instruction that are dedicated to attention to form. It can be taught both in preparation for communicative activities and after tasks in which the students have encountered some difficulties with a specific language form. Integrated FFI, instead, is incorporated within activities where the primary emphasis is placed on meaning, and learners' attention is drawn to language forms during communicative tasks to help them express meaning more accurately. Focus on form in integrated FFI can be incidental or planned. It can be incidentally raised

during language interactions or anticipated and planned by the teacher. It includes explicit elicitation of correct forms, brief explanations, corrective feedback, and input enhancement. Both isolated and integrated FFI can encompass the statement of rules through metalinguistic terminology and explicit feedback. In Lightbown and Spada's perspectives, a comprehensive curriculum or syllabus should not make a choice between integrated and isolated FFI, but rather, it should include both. They believe that the challenge lies in discovering the conditions according to which isolated and integrated FFI are most appropriate.

Tomlin and Villa (1994) suggest that drawing attention to some aspects of language form independently from engaging in meaningful communication as isolated FFI does, may be particularly useful for beginning learners. In his research, he shows how isolating and focusing input on specific features of the target language can help learners detect and understand form-meaning relationships (Van Patten, 2003, 2004; Van Patten and Cadierno, 1993). This applies particularly to language features that develop very slowly in the absence of attention and that have low salience, low frequency, or low communicative value (Norris and Ortega, 2000). Examples of features that develop at a slow pace if learners' attention is not drawn to them are: grammatical cues, morphemes and prepositions. This is because foreign language learners tend to focus their attention predominantly on open-class words (nouns, verbs, adjectives, and adverbs) rather than morphosyntactical features. Once the form-meaning connections have been established through isolated FFI, the development of greater fluency is likely to be favoured by integrated FFI. I employ isolated FFI in my research and I explain the reasons for this in the following section. I also explain how I include two different methods of FFI within a CLT approach. The students are divided in two groups, identified as control and experimental group, according to the method they are exposed to.

The language instruction used with the experimental and control groups has the same communicative approach but it employs two different forms of FFI with the two groups as mentioned above. The difference between the language instruction lies in the nature of FFI: it draws the attention of the students only to CA forms in the control group, whereas it focuses on CA forms and their SA

equivalents in the experimental groups. The motivations that led me to use isolated FFI are manifold and are as follows:

- i. Isolated FFI is particularly useful for beginner levels and my study focuses on beginner students of Arabic in higher education;
- ii. Isolated FFI draws attention to language features that have low communicative value, i.e. that are not particularly salient in oral language and that learners might not have heard or noticed in the first input. This is particularly applicable, for example, to features such as links between CA and SA forms. Isolated FFI can be useful for creating the necessary salience to help learners notice links between diglossic forms that are useful for the development of their language skills but that are not directly relevant to the oral input;
- iii. Isolated FFI is likely to provide the students with opportunities to learn forms they would not be explicitly exposed to in the CLT classroom, such as sociolinguistic forms. The social dynamics of the classroom do not provide the opportunities to observe real-life situations in which different forms can be used to perform different sociolinguistic language functions. However, isolated FFI draws students' attention to sociolinguistic functions of language use and it is therefore likely to prepare them to develop sociolinguistic abilities (Lyster, 1994).

I decided not to use integrated FFI because I was hesitant that drawing attention to CA and SA linguistic and metalinguistic forms within communicative activities at beginner level could cause confusion and overwhelm the students. However, at higher proficiency levels, more fluent and accurate use of that feature may best be encouraged through integrated FFI. On the contrary, several studies on FFI have reported that foreign language learners benefit most from isolated FFI when they are at beginner's level in their language acquisition (Mackey and Philp, 1998; Lightbown and Spada, 1999). Explanation and examples of isolated FFI activities developed specifically for the focus and experimental group will be explained in detail in the section on teaching material (Section 4.6) and I detail now the research method I propose to accomplish the research objectives, which is a mixed-methods approach within an embedded quasi-experimental design.

4.4 Research method

I explain here the research method I apply in this study and I identify the reasons for using this approach. I specify the types of data collection and analysis used, and I describe how the method chosen inform the research questions.

4.4.1 Mixed-Methods Research Design

The students who participate in this research are beginner-level university students of Arabic who have never been exposed to CA. As part of this experimental research, they undergo a twelve-hour introductory course to CA, and they are divided into control and experimental groups. The study is set within an inductive, bottom-up approach to research that starts from an empirical inspection of the data and seeks meaningful patterns that allow hypothesis generation and abstraction to describe a picture of the phenomenon studied. The independent variable is the explicit link between SA and CA forms, which is present only in the experimental group. The dependent variables of this study, i.e. the variables being tested, are the three aspects of AFL acquisitions mentioned above: diglossic vocabulary development; diglossic language awareness, and students' perceptions and attitudes towards language variation. The phenomenon studied has a threefold nature and it is multi-layered. For this reason, both quantitative and qualitative data are necessary to analyse it, as quantitative and qualitative components are used to address different aspects; focus on different variables; and therefore, their combination allows to achieve a more accurate understanding of the variables analysed.

The collection of both quantitative and qualitative data is conducted through a mixed-methods approach to research. Quantitative and qualitative methods are integrated to achieve complementarity; to gain a more exhaustive understanding of the variables studied; and to pursue new perspectives and frameworks. As Wesely (2013:300) puts it, the integration of the two sets of data not only leads to a deeper grasp of the phenomena studied, but it can also reveal contradictions between them and help raising new assumptions, suppositions and perspectives. I place the mixed-methods approach in this research within an embedded quasi-experimental design. In embedded quasi-

experimental designs, one data set provides a complementary, but secondary, role in a study based essentially on the other data type (Creswell et al., 2003). This research is based substantially on a quantitative design and it embeds a qualitative element within it. For this reason, I identify it as a 'QUAN + qual' design, where I mark the former in capital letters ('QUAN') as it is the predominant form of data collection, whereas the latter is marked in lowercase ('qual') because it is the subsidiary method employed. Comparative analysis within and across sets of experimental and control groups is used to measure quantitatively and qualitatively the threefold focus of investigation of this research within and across the participant universities. I list below the stages of my empirical research and I elucidate the nature of the data collected (quantitative or qualitative) in the different research phases. The details of my methodological approach are explained in sub-section 4.4.2.

Quantitative and Qualitative stages of the empirical research

Before the beginning of the empirical research, qualitative data are gathered through an open-ended questionnaire to understand the perceptions of the students towards SA and their experiences as students of AFL. This is followed by a language test that provides quantitative figures on SA vocabulary knowledge. As I explain in section 4.3.2, the results of the pre-test provide precise information on the students' actual proficiency levels and this is used to better understand their language outcomes after the experimental language course. After the end of the course I collect data on students' perceptions of Arabic variation, through the second open-ended questionnaire administered in this research. I then collect data on language knowledge by means of two processes: the first data collection process (hereafter the post-test) takes place a week after the end of the language course and the second occurs after a period of three weeks (hereafter the delayed post-test). The steps followed in the two data gathering procedures are identical. Firstly, I collect quantitative data through a language test that focuses on CA vocabulary retention; diglossic code-switching; and on the first level of language awareness (cognitive diglossic code-switching). The test is followed by the collection of qualitative data on the remaining two levels of language awareness (meta-awareness of diglossic code-switching and verbalisation of the rules applied to realise the switch). This

is carried out through retrospective think-aloud protocols, which consist in the verbalisation of the cognitive processes involved in performing a task and they are provided by the learners retrospectively, immediately after completing the diglossic code-switching tasks that are included in the language test (Godfroid, Housen and Boers, 2010:174). The data collected through retrospective think-aloud protocols are analysed twice. Firstly, they are turned into quantitative figures and are used as scores to be added to the language outcomes on CA vocabulary retention; diglossic code-switching skills and cognitive code-switching. By doing so, I obtain figures on diglossic knowledge that include diglossic vocabulary development and language awareness. Secondly, the data are analysed qualitatively in the last phase of the research. In this phase I, merge the whole databases of qualitative and quantitative data gathered in all the phases of the research. I analyse them together to seek information that can provide new perspectives and far-reaching insights to better understand students' cognitive experiences of diglossic code-switching and their development of diglossic knowledge. The steps of the research are summarised as follows:

- i. Qualitative (qual) prior to language course:
Participants respond to an open-ended questionnaire before the language intervention starts;
- ii. QUANTITATIVE (QUAN) prior to experimental language course (pre-test):
Participants sit a language test before the experimental language course starts;
- iii. QUANTITATIVE (QUAN) a week after the end of the experimental language course (post-test):
Participants sit a language test immediately after the experimental language course finishes;
- iv. qualitative (qual) immediately after the post-test:
After having taken the language test at the end of the experimental language course, participants respond to an open-ended questionnaire. They also share orally their cognitive experience of doing diglossic-code switching tasks in the test. I collect data on their cognitive experiences through retrospective think-aloud protocols. These data represent students' metacognitive awareness of performing diglossic-code switching. Data on

metacognitive awareness are analysed twice. I firstly convert them into quantitative figures and I add them to the language outcomes collected in phase (iii). I also analyse them qualitatively in the last phase of the research;

- v. QUANTITATIVE (QUAN) three weeks after the post-test (delayed post-test):

This part repeats phase (iii) and collect the same quantitative data after a period of three weeks;

- vi. qualitative (qual) immediately after the delayed post-test:

In this phase I collect data through retrospective think aloud protocols. The procedure to use data extracted from the protocols is identical to the one followed in phase (iv).

- vii. Integration of qualitative and quantitative data:

The last stage occurs during the analysis of the data. I integrate the data gathered in the qualitative and quantitative phases with the aim of finding key insights into the relationships between the results obtained.

The phases are visually represented in the table below for clarity of interpretation.

Table 1 Qualitative and quantitative phases of the research

Prior to experimental language course	(qual) Open-ended questionnaire	(QUAN) Pre-test: SA language test
A week after the experimental language course		(QUAN) Post-test on diglossic vocabulary knowledge and cognitive code-switching
Immediately after the post-test	(qual) Retrospective think-aloud protocols used to collect data on meta-awareness and presence of morphological rule formation + (qual) Open-ended questionnaires	(QUAN) Retrospective think-aloud protocols converted into quantitative figures

Three weeks after the post-test		(QUAN) Delayed post-test on diglossic vocabulary building and cognitive code-switching
Immediately after the delayed post-test	(qual) Retrospective think-aloud protocols used to collect data on meta-awareness and presence of morphological rule formation	(QUAN) Retrospective think-aloud protocols converted into quantitative figures
During data discussion	(QUAN) + (qual) Integration of the data gathered in the qualitative and quantitative phases	

Source: Author.

The mixed-methods approach is reflected in my research questions, which are quantitative (1 and 2), qualitative (3) and mixed-methods (4) in nature. As I illustrated in the introduction chapter, they are as follows:

1 - Does focus-on-form instruction lead to diglossic vocabulary development more effectively, when it links forms in Standard and Colloquial Arabic, or when it focuses only on one variety?

2 - Is focus-on-form instruction more effective for diglossic language awareness, when it links vocabulary forms in Standard and Colloquial Arabic, or when it focuses only on one variety?

3 - What impact does focus-on-form instruction have on students' perceptions of Arabic variation, when it links vocabulary forms in Standard and Colloquial Arabic and when it focuses only on one variety?

4 - To what extent is the development of diglossic vocabulary knowledge a function of the method of focus-on-form instruction received?

I apply the research design to the data collected from the three participant universities. This allows me to conduct cross-case analysis. I illustrate the selection of the sources and the ethics-approval process in the following section.

4.4.2 Data Collection and Analysis Methodology

The seven steps of the study as listed in the previous section are explained in detail below.

i. qual – before the experimental language course

The qualitative tool I use to collect data in the first phase of the empirical research is a questionnaire with open-ended questions, which is given to the experimental and control groups before the experimental language course starts. The questions allow me to obtain in-depth information on students' attitudes, opinions, and experience regarding learning Arabic as a foreign language. This tool is used to best learn from participants as it allows them to provide information without constraints, and it is also used on the assumption that the nature of the questions can provide different insights, illustrations, and interpretations. I combine them for nodes, i.e. themes and categories, using the qualitative software NVivo. This software supports the coding of the data into a selection of adequate nodes, for comparison and evaluation, in the data analysis stage. Examples of questions about students' attitudes towards studying Arabic are, 'Why do you study Arabic', 'What are, in your opinion, the most exciting and/or discouraging aspects of studying Arabic?', 'What do you think is the hardest aspect to be acquired in Arabic?'. Moreover, they are asked if they are planning to follow a language course, in which Arab country and why. Students are also asked to describe the teaching methodology or methodologies they have been exposed to in their experience as learners of Arabic, as well as the language learning strategies they use to memorise vocabulary. Finally, in order to understand their exposure to and knowledge of Arabic culture, they are asked about their knowledge and opinions of Arabic music and films.

ii. QUAN – pre-test before the experimental language course

Two premises of this study are that CA vocabulary and forms introduced in the language course are familiar to the students in their SA equivalents, and that participants are students of SA that have never been exposed to CA. The link between CA and SA is explicit in the experimental group and it is avoided in the control group, which focuses only on CA. On this basis, I investigate whether

and how the two methods used with the two groups, lead to building CA meaning and structural knowledge, and whether and how they lead to developing diglossic knowledge and language awareness of corresponding SA-CA words.

I verify the familiarity of the students with the selected SA vocabulary and forms with their Arabic teachers at the three participant Universities, as well as against their language programmes. In order to have an in-depth view of learners' actual familiarity with the afore mentioned vocabulary items and forms, and to differentiate between their proficiency levels, I administer a SA written and oral pre-test before the beginning of the language treatment. The results of the pre-test are not used with the aim of comparing the students' language improvement after the experimental language course as they have no knowledge of CA before the course starts. The results are collected on the basis that they provide a clear insight into the students' language knowledge in SA before the beginning of the language experiment. The comparison between these results and the data collected after the language treatment allow for a better understanding of the impact that the language treatment has on the participants. This is particularly true for the exercises in the post-test that involve SA, i.e. exercises to measure diglossic vocabulary development and language awareness. The comparison between pre-test and post-test scores helps to analyse the impact of the language treatments on students with similar levels of proficiency in SA at the beginning of the treatment, as well as among learners that have different levels of language proficiency in SA. In my view, the results of the pre-test also allow for a more detailed understanding of the learners' attitudes towards Arabic language reality and learning as their views can be analysed in light of their proficiency levels. Finally, the data collected quantitatively in the pre-test further support the interpretation of qualitative data collected before and after the treatment by adding information on learners' language knowledge at the beginning of the experiment. The vocabulary tested in the pre-test covers the following topics: greetings, introducing oneself, classroom objects, numbers, days of the week and months, telling the time. It also includes a few everyday-life verbs in the present tense, adjectives, Wh-questions, prepositions and compass points.

Data are collected through a language test composed of a written and an oral part, and I analyse the results of the test by measuring the mean scores, the standard deviation, and the coefficient of variation. The mean score, or average, of a certain set of data is equal to the sum of all the values in the data set divided by the total number of values. The standard deviation measures the variation within a set of data. It indicates how much the values of a certain data set differ from its mean and so how widely the data is spread from the mean. The coefficient of variation is useful to compare the dispersion in different sets of data and particularly data which differ in their means. It is calculated by dividing the standard deviation by the mean. It expresses the variation as a percentage of the mean and it helps to understand how much the data vary within a set (relative variability) and therefore the relative magnitude of the standard deviation. The standard deviation is a measure of absolute variation or dispersion, whereas the coefficient of variation is a measure of relative variation or dispersion. If the data has coefficient of variation lower than 1, or 100%, its distribution is considered as having low-variance, while if it is higher than 1, or 100%, the distribution is considered as high-variance. The equal distribution of coefficient of variation equal to 1, is considered a medium-variance distribution. Mean, standard variation and coefficient of variation provide an insight into the main differences in language performance among the participants group, and the internal variation within each group.

iii. QUAN – a week after the end of the experimental language course (post-test)

The data gathered in this phase is used to answer multiple research questions. They inform the development of diglossic-vocabulary knowledge and of diglossic language awareness; they allow one to compare the results among sets of control and experimental groups as well as among groups of similar nature; and they facilitate the understanding of the relationship between diglossic language awareness and the development of diglossic vocabulary knowledge.

The post-test comprises a written and an oral test that assess CA vocabulary retention, diglossic code-switching skills, and the first level of diglossic language awareness, i.e. cognitive diglossic code-switching. As in the pre-test, I analyse

the results of the test by measuring the mean scores; the standard deviation; and the coefficient of variation. I evaluate the mean, standard deviation and coefficient of variation of each group and I compare it with that of other groups. I compare and evaluate the mean scores of sets of control and experimental groups that have participated in this research at the same time and within the same university. I then compare these results among the three universities that participate in this experiment. This allows me to establish whether the differences or similarities of results among pairs of control-experimental groups are analogous among universities or not.

Moreover, the figures of the post-test establish whether there is a statistically significant difference between the results of experimental and control groups. This in turn sheds light onto the impact that the two forms of instructions employed with the two groups have on the threefold focus of investigation of this research. Two sets of scores are analysed: results within groups and comparative scores between experimental and control groups. The analysis is carried out by means of non-parametric statistics. Non-parametric statistics are formulae that accommodate the features of small-scale language education research, as is the case of this research. These features include an approximatively small number of participants who are not recruited anonymously from a vast population of learners. This applies to my study in multiple perspectives. First, the participants are not randomly selected but they voluntarily take part. Second, the selection of the participants is not as broad as to reach the entire population of learners, but it is constrained within the Universities of Exeter, Genoa and Milan that provided support and access to their facilities for the conduct of the empirical research. It is important to highlight here that very often educational research tries to gain a deeper or broader understanding of a specific phenomenon that occurs in a defined learning environment and it rarely aims at generalising its findings to a population.

The non-parametric formula I employ here is 'Wilcoxon-Mann-Whitney Test'. This allows me to analyse two sets of dependent variable data, i.e. the sums of the post-tests scored by the experimental and control groups. It also allows me to establish whether there is a statistically significant difference between them. This formula is a nonparametric test used for comparing two independent

groups, and it requires that the independent variable has two levels. The independent variable of this research is type of instruction and it has two levels: first, the instruction in the control group combines CLT with FFI that focuses only on CA; and second, the instruction in the experimental group combines CLT with FFI that focuses on SA and CA simultaneously.

As anticipated above, I analyse in this phase the learning outcomes for CA vocabulary retention, diglossic code-switching skills and the first level of diglossic language awareness (diglossic code-switching). The first two are evaluated according to the categories of Bloom's taxonomy, whereas language awareness is measured using Leow's levels of awareness. The data collected to measure Leow's three levels of language awareness are dual in nature: quantitative and qualitative. Data on the first level (presence of cognitive code-switching) are collected quantitatively through language scores attributed to code-switching tasks that the students perform in the language tests. The remaining two levels (meta-awareness and verbalisation of rules for diglossic code-switching) are measured through retrospective think-aloud protocols. These provide qualitative data, which is converted in this phase into quantitative scores and ranking.

I attribute a score to the two components of the retrospective think-aloud protocols of each student, and the conversion is explained in chapter six. The reason behind converting the data is twofold. On the one hand, it allows me to add information on language awareness to the quantitative data collected in the post-test. The summation of the quantitative data collected is of great relevance for the research since it provides an exhaustive and manifold picture of the language knowledge and awareness developed by the students through the experimental language course of this research. On the other hand, I can join data on the three levels of language awareness and analyse them as a whole for each participant and every group. I can then compare the results between sets of control and experimental groups.

I compare the obtained quantitative figures on language awareness with CA language retention and diglossic code-switching knowledge. I analyse the scores obtained by sets of control and experimental groups. Intragroup analysis is used to measure the relative performances of each student within the group they are allocated to, and of each group as a whole. Intergroup analysis, i.e. the

comparison of the obtained figures among different sets of groups, allows me to reveal differences and similarities that exist between them. I compare the mean-variance of single groups among each other in the intergroup analysis. I firstly analyse the differences between sets of corresponding control and experimental groups within the same universities. I then compare the results among the Universities of Exeter, Genoa, and Milan. I additionally analyse the differences among groups that have identical nature: control groups and experimental groups are compared only among each other. This allows me to compare the results among similar groups within different universities and to identify patterns of similarities and differences among them. Finally, qualitative data collected through the retrospective think-aloud protocols are used in their qualitative and informative nature in the last stage of the research, when qualitative data is merged with quantitative data to better inform the results of statistical scores and rankings.

iv. qual – after the post-test

Two kinds of qualitative data are gathered after the language intervention. First, learners describe their experience of performing diglossic code-switching tasks and I collect this information through retrospective think-aloud protocols. The data gathered via retrospective think-aloud protocols provide qualitative information on the cognitive experience that each learner has of performing diglossic code-switching. I convert this data into quantitative figures in phase (iii), but I also keep it in its qualitative format and use it at the last stage of data analysis. In that stage, quantitative and qualitative data are merged to form an overall and detailed picture of the impact that this study has on the development of diglossic vocabulary and language awareness. I assess this impact on both language knowledge and subjective responses of participants to Arabic language variation; diglossic code-switching; and the study of CA and SA simultaneously through the different experimental methodologies of this research.

Second, I collect qualitative data through questionnaires with open-ended questions. The questions allow me to obtain in-depth information on learners' experiences of attending the language course while they study SA in their academic institutions, and their views on variation in Arabic language and

learning. I analyse these data in chapter seven. They are of fundamental importance to answer the fourth research question: 'What results emerge by comparing quantitative data on diglossic learning outcomes and accuracy with qualitative data on students' awareness and perceptions of Arabic reality and learning?'. Examples of questions in the questionnaire are; 'Describe the experience of studying simultaneously Colloquial and Standard Arabic'; 'Describe the main difficulties, challenges and positive aspects of integrating Colloquial Arabic with Standard Arabic'.

These two qualitative data sets (retrospective think-aloud protocols on language awareness and post-test questionnaires) are compared in chapter eight, in which I discuss whether there exists a connection between their perceptions of Arabic variation with their cognitive experiences of performing diglossic code-switching. Also, I seek to identify answers to whether there is a link between the students' experiences of performing diglossic code-switching and their motivations towards studying two Arabic varieties simultaneously. I explore whether there is a link between students' evaluations of the challenges of integrating CA and SA and their experiences of performing diglossic code-switching. Finally, I try to establish whether there is a connection between the participants' self-evaluation on the impact that the integration of CA instruction into their SA programme has, and their experiences of performing diglossic code-switching.

v. QUAN – three weeks after the post-test (delayed post-test)

The delayed post-test occurs three weeks after the post-test. It is a replication of the post-test (phase iii) and it has the same focus. The process of data gathering is identical and so is the data analysis. First, the participants sit a language test aimed at collecting figures on diglossic vocabulary knowledge (the test measures CA language retention; diglossic code-switching tasks; and presence of diglossic code-switching). I use the results of the tests to measure the mean, standard variation and coefficient of variation for every group. I compare the results between control and experimental groups using the Wilcoxon-Mann-Whitney test to determine whether there is a statistically significant difference between them. Second, I measure the two remaining levels of language awareness: presence of meta-awareness and awareness of

morphosyntactical rule applied to perform the diglossic code-switching. The data is collected and analysed by converting qualitative data from retrospective think-aloud protocols into quantitative figures. The qualitative data gathered is analysed, instead, in the last phase of the data analysis.

The analysis of the data collected through the delayed post-test adds one element that is not present in the analysis of the data collected through the post-test: the comparison between the results of the post-test and the delayed post-test among results of paired samples of the same participants. I use here the 'Wilcoxon Signed Rank Statistic' test to compare the results of paired samples and identify whether there is a statistically significant difference between the results of the post-test and the results of the delayed post-test. I evaluate the differences among corresponding control and experimental groups, and I compare the results obtained among the three universities participating in this study.

vi. qual – after the delayed post-test

This phase replicates phase (iv). Participants share their experiences of performing diglossic code-switching during the delayed language test and I collect these data through retrospective think-aloud protocols. The participants are asked to make the thought processes that occurred to them during the task performance as explicit as possible, and they are asked to elucidate the rules they apply to switch between language varieties. As I do in phase (iv), I use this information both quantitatively and qualitatively. Quantitative figures are analysed at this stage, whereas qualitative information is analysed at the last stage of the data analysis, when qualitative and quantitative data are merged. The data gathered through retrospective think-aloud protocols after the delayed post-test are useful to: (i) analyse more comprehensively learners' scores on language knowledge in the delayed post-test by adding the element of language awareness to the scores; (ii) assess long-term learners' awareness of diglossic code-switching changes; (iii) compare and analyse how learners' awareness change over a three-week period.

vii. integration of qualitative and quantitative data

This stage occurs during the data analysis and its purpose is to merge quantitative and qualitative results to achieve a better understanding of the two sets of data (Creswell, 2006:83). Data are merged to compare the results from the two datasets through a side-by side comparison technique (Creswell and Plano Clark, 2011:223), which involves presenting qualitative findings and quantitative results together in the discussion chapter. I present quantitative results followed by qualitative results in the form of quotes and I specify how the qualitative quotes support, or contradict, the quantitative results.

I summarise in the table below the components of my empirical research: (i) the research phases; (ii) the nature of the data gathered; (iii) the tools employed to collect them and (iv) to measure and analyse them; and finally, (V) the information provided by the data collected.

Table 2 Components of the empirical research

Phase 1 (qual) Prior to the experimental language course

Data collected on	Students' attitudes towards SA language and learning
Data collection tool	Open-ended questionnaires
Data analysis tool	NVivo
Data provide information on	Students' views towards SA and their opinions and experiences regarding learning AFL Teaching methodologies that students are exposed to in their AFL programmes Effective vocabulary language learning strategies employed by the students

Phase 2 (QUAN) Prior to the experimental language course

Data collected on	SA proficiency
Data collection tool	Language test (pre-test)
Data analysis tool	Language test scores for each participant
Data provide information on	Learners' actual proficiency in SA forms and vocabulary that will be taught in CA in the experimental language course. This allows to analyse the impact that the language experiment has on the development of diglossic language knowledge and awareness for the participants, based on their levels of proficiency at the beginning of the course.

Phase 3 (QUAN) A week after the conclusion of the experimental language course

Data collected on	Diglossic vocabulary-building. This is composed of: CA vocabulary retention; diglossic code-switching skills; presence of cognitive code-switching (first level of diglossic language awareness)
--------------------------	--

Data collection tool	Language test (post-test)
Data analysis tool	Language test scores used to analyse the mean scores; standard variation; coefficient of variation; and Wilcoxon-Mann-Whitney test
Data provide information on	<p>Development of diglossic language knowledge for each participant. This is measured against the first five categories of Bloom's taxonomy as listed in sections 3.5.1 and 3.5.1.1. It is also measured using the first level of language awareness as indicated by Leow (Leow's levels of awareness are listed in section 3.5.2).</p> <p>Establish whether the difference in language instruction between control and experimental groups leads to a statistically significant difference in diglossic vocabulary building between the two groups.</p> <p>Difference in diglossic vocabulary-building results between sets of control and experimental groups within the same Universities and also among the participant Universities.</p> <p>Insight into the main differences in language performance and mean scores within each single group, between sets of control and experimental groups within the same Universities and among the participant Universities.</p> <p>Analyse the similarities and differences in the results among groups that have identical nature, i.e. control groups and experimental groups are compared among each other. This allows to identify patterns of similarities and differences of groups that have the same nature among the participant Universities.</p>

Phase 4 (qual) Immediately after the post-test

Data collected on	Meta-awareness of the subjective experience of diglossic code-switching, and presence of morphological rule formation (second and third levels of diglossic language awareness)
Data collection tool	Retrospective think-aloud protocols
Data analysis tool	NVivo
Data provide information on	Whilst it is collected immediately after the conclusion of the empirical language research, this data inform the last phase of data analysis. In the last phase quantitative and qualitative data are merged to provide a detailed picture of the impact that this study has on the development of diglossic vocabulary and language awareness, and on the students' attitudes and motivations towards Arabic language reality and learning. I assess this impact both as language knowledge and as the students' subjective responses to Arabic language variation.

Data collected on	Students' attitudes towards variation in Arabic reality and learning
Data collection tool	Open-ended questionnaires
Data analysis tool	NVivo
Data provide information on	<p>Insight into students' perceptions of variation in Arabic and their attitudes towards studying, together with their cognitive experiences of performing diglossic code-switching.</p> <p>Answer questions that seek to identify whether there is a link between the students' experiences of performing diglossic code-switching and their motivations towards studying two Arabic varieties simultaneously; their evaluations of the challenges of integrating CA and SA; and their</p>

	self-evaluation on the impact that the integration of CA instruction into their SA programme has.
--	---

Phase 5 QUAN) three weeks after the post-test

Data collected on	Long-term diglossic vocabulary-building and language awareness
Data collection tool	Language test (QUAN) and retrospective think-aloud protocols (qual converted into QUAN)
Data analysis tool	Language test scores used to analyse the mean scores; standard variation; coefficient of variation; Wilcoxon-Mann-Whitney test; and Wilcoxon Signed Rank Statistic
Data provide information on	<p>Evaluate the long-term development of diglossic-vocabulary knowledge and language awareness for each participant (CA vocabulary knowledge; diglossic-code switching and diglossic language awareness) and for each group. The comparison is carried out within the same Universities and among the participant Universities.</p> <p>Compare the results of delayed post-tests between sets of control and experimental groups and determine whether there is a significant difference between them and if the difference is affected by the types of instruction received in the experimental language courses.</p> <p>Analyse the similarities and differences in the results of groups that have identical nature, i.e. control groups and experimental groups. The results of control groups and experimental groups are compared among each other with the aim to identifying patterns of similarities and differences of groups that have the same nature among the participant Universities.</p> <p>Establish the long-term relationship that each participant develops between language awareness and the other two components of diglossic vocabulary building and allow to investigate whether</p> <p>Compare paired samples to determine whether there is a statistically significant difference between two sets of data from the same group of people, as in between the post-test and the delayed post-test scores.</p> <p>Shed light on the impact that the experimental language instruction of this research has on long-term diglossic language retention.</p> <p>Evaluate sets of control and experimental groups within the same Universities and among the three universities participating in this study.</p>

Phase 6 (qual) Immediately after the conclusion of the post-test

Data collected on	Meta-awareness of the subjective experience of diglossic code-switching, and presence of morphological rule formation (second and third levels of diglossic language awareness)
Data collection tool	Retrospective think-aloud protocols
Data analysis tool	NVivo
Data provide information on	Provide an insight into the participants' cognitive processes and their experiences of performing diglossic code-switching during the delayed post-test. These data are collected in this phase but analysed in the last stage of the data analysis, when qualitative and quantitative data are merged

Phase 7 (QUAN and qual)	
Data collected on	The data collected in the first six phases are merged here.
Data collection tool	-
Data analysis tool	Side-by side comparison technique
Data provide information on	Quantitative and qualitative results are merged to achieve a better understanding of the two sets of data.

Source: Author.

So far, I have focused in detail on the design of the research calling attention to each phase and research tool used. I focus now on the selection of the participants, the coding system used, and the ethics approval process.

4.5 Selection of participants

In this research an embedded quasi-experimental mixed-methods research design is employed. A quasi-experimental design has in common with experimental designs that it builds its empirical study on a target population divided into experimental and control groups. However, it lacks the element of random assignment to the groups that is typical of experimental designs. In this thesis, the target population is represented by students of SA in higher education who have never been exposed to a variety of CA. As we have seen previously, they are enrolled in one of the following universities: Exeter in the U.K. and Genoa or Milan in Italy. All the participants are at beginner's levels. Since students from three universities participated in this study, the length of their exposure to Arabic instruction is comparatively different although they are at the same language levels. Thus, with the aim of selecting a consistent set of participants, I identify a threshold in the selection as follows: by the start of the language intervention all students must have had between 100 and 150 hours of SA instruction. The participation is supported by the above-mentioned universities but it is not compulsory for the students to join the study and they can decline to take part. This entails that participants are not randomly selected from a vast population of learners. Finally, the classes for both experimental and control groups are created to accommodate the students' needs and schedules, and they can choose to join in the group that suits their needs best. These circumstances occur very often in language-education research and lead to a dataset that does not come from a population following a probability

distribution. However, I explain in the next section how this sets of data can be quantitatively analysed using non-parametric statistics.

Several ethical considerations needed to be taken into consideration during the preparation of the research, in order to conduct the empirical study under appropriate conditions. An ethics-approval form in English for the students of the University of Exeter was prepared and approved by the Ethics Committee of the University before the language interventions took place, and it was translated into Italian for the universities of Milan and Genoa. The ethics-approval process included the need to ensure the voluntary and informed nature of participation of the students. A consent form was given to the students before the beginning of the empirical research explaining the main aims of the study and the objective of the data gathering, hence being part of the empirical research of a PhD thesis. Written consent was obtained in every case. The consent forms were signed before the first questionnaire was handed out, which gave me the possibility of explaining the aims of the interviews and managing the expectations of the participants.

In this research, codes are used order to protect confidentiality and ensure anonymity. Participants are de-identified by using codes instead of their names. The codes identify the University (MI for Milan, GE for Genoa and EX for Exeter) and the nature of FFI instruction received (CA for instruction that focuses only on CA and CASA for instruction that links CA and SA forms). I finally use numbers to differentiate between participants. For example, "MICASA1" is a participant enrolled at the University of Milan who has been allocated to the experimental group CASA and has been randomly attributed the number 1 to be distinguished from the other participants.

Regarding data storage, the transcription of the tests and the tests results, the answers to the open-ended questionnaires and list of participant names were kept separate in two different folders on the University of Exeter secured central data storage facility, namely the U: drive. Furthermore, the ethical-approval process required an assessment of possible harm caused by the research. It was concluded that the interviews were not likely to cause harm to the participants. Lastly, a declaration of interest ensured that there were no further commercial or other interests involved in the project. The use of qualitative software NVivo supported the qualitative coding of the questionnaires into a

selection of adequate nodes. It also represented an additional archival support for the gathered data. The use of qualitative software permits the exploitation of hypertext and hypermedia techniques and the possibility of creating flexible nodes through a variety of supports and texts (Atkinson and Delamont, 1996). Finally, the use of NVivo supported the retrieval of information and analysis of the qualitative comparative case studies and permitted the horizontal grouping of the interviews along selected topics and facilitated their comparison.

In the following section I explain the selection of the universities in which I conducted my language research and the information gathered through the pilot study, which helped me to structure the language interventions.

4.5.1 Selection of the Universities

The language intervention consists of a twelve-hour course and is organised in six lessons of two hours each. It ran at the universities that participate in my research and that allow me to use their facilities to carry out the language instruction as well as the language tests. Pre- and post-treatment questionnaires were also conducted using rooms provided by the universities. I conduct my empirical research in two different countries because I am interested in collecting and comparing data from speakers of different mother tongues. My universities of choice fall within countries in which I am equally comfortable in speaking the official language, namely Italy and the United Kingdom, and with using it in formal instruction. This is of crucial importance because, although the main language of instruction is colloquial Arabic, students at beginner levels very often need support in their mother tongue during second-language lessons, especially if the activities are based on focus-on-form. For this reason, I chose countries in which my ability to use the official language beside Arabic during classroom instruction is the same. I perceive this to be a fundamental factor since it avoids risks of bias in data collection and interpretation.

Within the afore mentioned countries, the selection of the three universities was also based on criteria drawn up to avoid bias in data collection. On the one hand, in all three universities the Arabic programme is based on SA without formal instruction in CA. On the other, I am equally as familiar with their

facilities, structures, organisation and instruction as I completed my bachelor's degree at the University of Genoa, earned a master's degree at the University of Milan and I have taught Arabic alongside my PhD programme at the University of Exeter.

The research took place in the academic year 2014/2015. I carried it out in Milan and Genoa from October until December 2014, and I conducted my research in Exeter in the second term of the same academic year, more precisely from January until March 2015. As Genoa and Milan are not far from each other, I combined the research in the two universities within the same weeks as follows: from Monday until Wednesday in Milan, and Thursday and Friday in Genoa.

Before starting the research, I first contacted the Universities and set a formal agreement in which I was accepted as visiting researcher at their Arab and Islamic studies or language institutes. This allowed me to make use of the universities' facilities and to deliver the lessons, tests and questionnaires at the universities' building rooms. I organised the schedules of the lessons through the timetable services of the three Universities, which allowed me to combine the free slots in the students' timetables with the rooms available in the buildings. The research has a compulsory attendance policy and the attendance to all six lessons was mandatory, in order to make sure that all the participants were exposed to same amount of classroom time. Finally, so as to minimise the impact of the course on the students' workloads, I did not request the participants to study the language content covered in the lessons for the tests. This allowed them to decide freely whether they wanted to study in preparation for the test. I now identify and explain the content and the language material developed for the language lessons of this research.

4.5.2 Pilot study

I conducted the pilot study at the University of Exeter during the second term of the academic year 2013/2014. Through it I acquired best practices, lessons learned and students' advice that helped me to make sure that the language interventions met the students' needs and interests without becoming burdensome. I also assessed whether the amount of vocabulary and language

forms I taught was appropriate for the students' time and availability to study between lessons. Moreover, the conduction of the pilot study helped me to understand how to efficiently organise the language intervention of this study and the phases of data collection. I decided to carry out the study during the second term of the 2013/2014 academic year as my empirical language research was formally scheduled to take place over the following academic year. The timeframe set for the pilot study left me enough time to make the necessary changes to the organisation and structure of language interventions and language tests before the beginning of the following academic year. I was also able to share the finalised version of my empirical research programme with the universities participating in this study with adequate notice.

Before carrying out the pilot study, I was uncertain about two structural aspects of the empirical research: whether to organise it as an intense language course or spread it over a certain number of weeks, and the exact length of the research. First, I needed to assess whether the students preferred the language treatment to be organised over a prolonged period of time or if they favoured an intensive course. In the pilot study, it emerged that they had no definite and strong preference between the two options. However, they expressed a preference for the study to be carried out during term time, as it is often the case that many of them do not live in their university town outside of term. This entailed spreading the empirical research over several weeks, because it was not possible to combine an intensive course with their university lessons' timetables. I therefore decided to organise the research over a prolonged period.

Second, the pilot study was helpful in determining the exact length of time of the empirical research, i.e. the overall time that the language course and the data collection phases (language tests and questionnaires) should take. The students expressed a preference for one lesson a week of two hours and they were willing to take as many lessons as possible since they were eager to learn a form of vernacular Arabic. I explain below in detail the process that led me to establish the exact number of lessons of this research.

On account of the research premises, students participate in the study only after having received between 100 and 150 hours of SA formal instruction, and

before they are exposed to any variety of CA. This applied to the students of the three universities as follows:

- i. The students of the University of Exeter spend their second year in Jordan within the year abroad programme organised by the University. This programme gives them the opportunity to live for almost a year in Amman; study intensively SA; and learn CA alongside. As they are exposed to CA in their second year, only first year students can participate in this research. By the end of the first term of their first year, the students reach 100 hours of formal instruction, which is the threshold for the participation in this research study. This narrowed down the amount of time available for Exeter students to the second term of the first year. In the academic year 2014/2015 this corresponded to eleven weeks.
- ii. The bachelor's degree courses of the Universities of Milan and Genoa did not offer study abroad year options at the time of this research. Neither Milan nor Genoa Universities had compulsory CA courses in their Arabic programmes. This meant that, unless the students had been exposed to a variety of CA for personal reasons, or had taken optional CA lessons organised by the universities, they could participate. It was often the case that students enrolled in the first year had not received enough instruction in SA to be able to join this research. In fact, the study saw the participation of second and third year students only. In order to be consistent with the length of the research in Exeter, I organised the research to take place over eleven weeks in both universities.

To sum up, through the pilot study I established that the empirical research should cover eleven weeks inclusive of pre-test, language instruction, post-test and delayed post-test. I also determined that the lessons should be two hours long, and that they should occur during term time.

Before the beginning of the pilot study I was concerned about arranging the empirical research during term time, as I was concerned that its combination with curricular lessons could prove to be an obstacle. I was partly concerned the three tests that comprise my research (the pre-test, post-test and delayed post-test) could have created a stressful environment and that this would have a negative impact on the students' well-being and academic performance. I was also concerned about the impact of the research on the students' workload.

Whilst I considered this a good opportunity to assess the effects of combining CA and SA instruction simultaneously, the experimental, and therefore untested, nature of the language instruction made me conscious of the potential downsides of integrating it within SA instruction. For example, it could have caused confusion between SA and CA or unexpected and negative outcomes. The teachers of Arabic of the Universities of Exeter, Genoa and Milan, as well as the students participating in the pilot study, helped me to tackle these concerns. The former accepted to meet me before the beginning of the study and address the potential negative outcomes of conducting my research during term time. I queried the latter after the completion of the pilot study, as to whether they had found the combination between this research and their standard curriculum lessons cumbersome, confusing or excessively demanding, and whether the tests had proven to be highly stressful. The teachers expressed a favourable opinion on the basis that my research involves the teaching of CA (with links to SA forms in the experimental group) at elementary-level, whereas the university lessons in SA that the students attend are high-beginner level or low-intermediate levels. The teachers were also supportive of the idea that the tests included in the research would not present an obstacle as they are anonymous; they are not evaluated or shared with the teachers; and therefore, they have no impact on the students' academic performance. Moreover, the students were made aware before the beginning of the pilot study that their withdrawal from the study would not be penalised and they were encouraged to act should perceive that the research was having a negative impact on their well-being. The students confirmed the responses of the teachers: they reported that the integration of the course with their University timetable had not been overwhelming and that they did not feel it caused confusion either in terms of their passive knowledge (understanding CA and SA) or active knowledge (recall and producing sentences in CA and SA correctly). Only one student dropped out the pilot study, and more specifically the experimental group, on the grounds of the excessive amount of information and the difficulty in combining it with the other subjects studied. I have explained in detail the pedagogical and contextual reasons that led me to define the organisation of my empirical research and I now move on to delineate the experimental language courses.

4.6 Language material

In keeping with the role and function of CA, which is confined to everyday oral interaction, teaching this variety should be restricted to oral skills. As Alish (1997:99) argues, “learning a colloquial variety is limited to the oral skill[s] only. Nevertheless, pedagogical constraints, such as the need for reinforcement of learning, dictate that some written [CA] be provided.” I therefore rely on CA written material for pedagogical purposes.

The material used in the language course was developed by the researcher. I drafted the first version before the pilot study and I trialled it in that occasion to test it before using it systematically in the experimental language course. In the pilot study, I tested, adjusted and refined the language material. I initially consulted with the participants to identify their topics of preference and then refined the content of the language resources in the pilot study. I discussed and examined with the students involved in the pilot study their vocabulary choices in order to identify which elementary topics they wished to learn in CA. This is because, as the students’ participation was voluntary and the attendance to all six lessons was required, I assumed that they would more likely be engaged if the topics presented were in line within their needs and requirements. The students expressed a particular interest in learning CA related to daily life topics to ameliorate their skills in everyday life interactions. They individuated areas in which they believed that their knowledge of SA did not provide them with the necessary tools to communicate with native speakers and therefore areas in which they felt hindered in having social interactions. Among the topics suggested I selected the most popular elementary areas: greetings, introducing oneself, numbers, days of the week and months, and telling the time. I added to these topics some grammatical forms and nouns that allow learners to make simple sentences in the present tense (a few everyday life verbs in the present), adjectives, Wh-questions, prepositions and compass points. The course was delivered over six weeks from the second week until the seventh of the language intervention and the topics were organised as follows:

Week 1: Greetings

Week 2: Revision of greetings, introducing oneself, numbers (until 19);

Week 3: Revision of introducing oneself, numbers (20-30), Wh-questions;

Week 4: Everyday life (includes an introduction of how to tell the time);

Week 5: Revision of greetings, numbers, everyday life and wh-questions. Telling the time. Days of the week;
Week 6: Describe position and direction using selected vocabulary from prepositions, classroom objects, and compass points. Revision.

4.6.1 Orthography of language material

The material was presented in Arabic script. The practice among some developers of CA materials has been to present the texts transliterated in Roman script but this practice is questionable since materials in CA are generally designed for students with a background in SA, which is normally taught with Arabic script (Alosh, 1997:99). A challenge this decision entails is that there does not exist a standardised orthography for CA. Moreover, spelling choices have the potential to take on social meaning. The field of the sociolinguistics of spelling has pointed out that in vernacular writing, both the choice of a spelling variant that is supplied by the standard variety and the choice of a spelling that departs from an existing standard version in some way constitute a social action, i.e., they transport social meaning by either complying with or breaking existing norms (Hinrichs, 2012:326). In this research the selection of the orthography for the language material is carried out with the aim of reflecting the main features of the FFI used in the two groups of the experimental language course. I firstly identified a shared linguistic area of vocabulary overlap that exists between CA and SA. I adopt here two definitions introduced by Bassiouney (2006:36) that define shared lexical items between SA and CA: “neutral” and “mixed”. The former refers to vocabulary in which the alternation between one lexical item in CA and the other in SA “is only phonological” (Bassiouney, 2006:30) and therefore it does not appear in the written word. This difference could be attributed to divergent “vowel pattern” or “realisation of consonants” (Bassiouney, 2006:33). For example the verb قال /qa:la/ (to say) can be realised phonologically as SA by pronouncing it /qa:la/ as well as CA by uttering it /ʔa:l/ or /ga:l/. The phonological variations at play here involve the pronunciation of the consonant ق /q/ that is realised in CA as a glottal stop /ʔ/, i.e. by hindering the airflow in the glottis (Watson, 2002:17), or as a voiced velar stop /g/; (ii) the dropping of the last short vowel in the pronunciation of the verb. Neutral words also include words which seem like SA

words because they may have some MSA features, but have no equivalent in CA and are therefore used in both codes.

Mixed items instead merge features of the two varieties and Bassiouney (2006:37) divides them into two forms:

- “1. Mixed forms that are mixed morpho-phonologically and lexically (within a word). [For example] the noun [e:ʔ (thing)]. Lexically the noun is MSA, the ECA equivalent would be ha:ga. The glottal stop at the end also seems like an MSA phonological feature. However, the realisation of the vowel as e: rather than the MSA diphthong ay, is an ECA feature.
2. Mixed forms that are mixed by blending a bound morpheme from one code and a free morpheme from another. This category includes MSA verbs, for example, which are saliently MSA with MSA morpho-phonological features, but which have an ECA variable attached to them. The passive verb bi-tunaffadh (it is implemented) is a clear example.”

By exploring mixed and neutral items, Bassiouney gives us a significant insight into code mixing, or its absence, at word level. The rationale that led to my orthographic choices is based on reflecting Bassiouney’s distinctions of CA and SA vocabulary. I firstly explain the orthographic choices I made for the material used with the control group and I subsequently tackle the orthographic choices that I apply to the material used with the experimental group.

In the control group I employ the word spelling adopted in “Syrian Colloquial Arabic, a functional course” (Liddicoat, Lenanne, and Abdul Rahim, 2012). My choice is motivated by the fact that I used various communicative exercises from this textbook and I found it consistent to adhere to the CA spelling its author employs. I followed throughout different guidelines for the experimental group instead.

First, I start by explaining the spelling of neutral lexical items, as per the definition provided by Bassiouney. These are written in SA. For example, بيت /bajt/ and كثير /kaθi:r/ are written according to their SA spelling, although their pronunciation in Levantine CA is phonologically different. This is rendered into CA by two means. First, the participants are provided with audio files for all the lexical items of the course so that they are able to listen to the CA pronunciation of the vocabulary covered in the course. Second, the experimental group is exposed to a comparison between SA and Levantine CA phonological forms that explicitly explains the variations of pronunciation. The variation of pronunciation of the word بيت /bajt/, which is pronounced /be:t/ in CA, is explained by elucidating how diphthongs change from SA to Levantine CA.

Diphthongs are monosyllabic sounds that begin with one short vowel and glide into a long vowel that has a different sound (ex. خَيْر /ħajr/ and يَوْم /jawm/). A common example in English is the sound at the end of the word ‘toy.’ Diphthongs in Spoken Arabic tend to eliminate the short vowel and the sound of the long vowel prevails. The sound of the long vowels can remain the same as it is in SA, or take a Levantine CA inflection. In the case of بَيْت /bajt/, the long vowel is pronounced in CA as /e:/ and therefore the word بَيْت is pronounced /be:t/ in Levantine CA. The change in the pronunciation of the word كَثِير /kaθi:r/, instead, is explained through two common rules as follows. The students are taught that the consonant ث /θ/ is pronounced in Levantine CA as /t/ and in some words, as /s/, and that very often the sequence ‘consonant-consonant-long vowel’ is characterised by the absence of a vowel on the first consonant. The word كَثِير is therefore pronounced /kti:r/ in Levantine CA.

As per mixed forms, I focus only the forms that are mixed by blending a bound morpheme from one code and a free morpheme from another. I employ a specific colour coding for this forms: black for SA morpho-phonological features and light blue for the CA variables attached to them. For example, the verb يَرْجِع /byərdʒaʕ/ is written using black for the SA feature رَجَعَ /radʒaʕa/ and using light-blue for the CA feature ب /b/. My aim is to visually facilitate the students to recognise the features typical of Levantine CA and SA. However, although visual tools are of help for the students, in my view the audio files that accompany the vocabulary are of fundamental importance to render the nuances of the CA pronunciation. As previously mentioned, the participants are provided with audio files for all the lexical items of the course so that they are able to listen to the CA pronunciation of the vocabulary covered in the course. This is because it is not possible to reproduce the entire range of CA sounds using the Arabic script. If we go back to يَرْجِع /byərdʒaʕ/, we see that, although the colour coding helps the students to differentiate between the suffix in CA and the verb in SA, the sound /ə/ is not represented. The students would therefore not be able to utter the verb in CA without an audio support. Similarly, we have seen that the long vowel of the word بَيْت /bajt/ is in fact pronounced /e:/ in CA and there is no correspondent to that sound in the Arabic alphabet. Finally, I write lexical items that exist only in CA using the colour coding previously explained. My colour of choice for CA is light-blue and therefore vocabulary in

CA is written in this colour. As per the spelling of these lexical items, I follow the spelling used by Liddicoat, Lenanne, and Abdul Rahim (2012). This is the same book I use as reference to prepare the material for the control group and I believe it is useful and consistent to adhere to the spelling of CA words it provides.

4.6.2 Activities and exercises of language material

The theoretical approach to the development of activities and exercises I adopt for both the experimental and control groups recalls that of Nation (2011:445). According to him (Nation, 2011:445), there needs to be a balance of opportunities for learning across four strands: meaning-focus input, meaning-focus output, language-focused learning, and fluency development. Thus, a well-planned speaking course builds on learning through meaning-focused input; provides substantial opportunities for meaning-focused output through speaking; gives deliberate attention to pronunciation and the learning of vocabulary; and develops fluency in speaking. All these elements are present in both the experimental and the control group. Moreover, with both groups I employ interactive activities and games. Games are particularly important in speaking activities as understanding is a key to achieving the goal of the activity. Interactive activities are also based on students' negotiation of meaning.

With both groups, I combine exercises to encourage speaking as a meaning-focused activity, but I also pay deliberate attention to pronunciation to improve the quality of spoken output and to stress the phonological characteristics of CA. I also focus on the development of sociolinguistic competence by working on topics such as greetings and by giving a particular focus to how to address people in appropriate ways. Most importantly, I focus on how to keep a conversation going, which is a particular useful strategy for learners to gain control in the early stages of speaking (Nation, 2011:451). There exists a range of simple ways in which learners can be helped to manage conversations. For example, a useful way to remain involved in a conversation is to answer questions with a short answer and add extra related information (Holmes and Brown, 1976). Finally, I focus on developing fluency in speaking in CA. Speaking occurs under time constraints and thus it is very important that

learners quickly become fluent in using the language items that they already know. For example, knowing the numbers in another language is of little value in spoken use unless these numbers can be accessed fluently (Nation, 2011:452). For this reason, for example, students in both groups practice numbers from 1 to 39 through the game 'Battleship.' This is a game that students play in pairs. Each player has a ruled grid on which they mark the location of their fleets of ships. The individual squares in the grid are identified by numbers. The locations of the fleet are concealed from the other player and they alternate turns trying to identify the location of the other player's ships by calling out numbers of the grid. Thus, not only do the numbers need to be learnt, they also need to be practised until they can be quickly recognised when they are heard and quickly produces when they are needed. Exercises to develop fluency are carried out once the vocabulary has become familiar to the learners, as unfamiliar vocabulary, grammatical constructions or discourse features hinder fluency development. The focus of the activities is on conveying messages and thus the tasks are communicative. Finally, there tends to be an encouragement to perform faster than usual speed and a great amount of repetition.

The afore mentioned communicative activities are combined with FFI. In this research FFI in the experimental group draws students' attention to the following equivalent forms and vocabulary in Levantine CA and SA:

- i. How the pronunciation of diphthongs changes from SA into Levantine CA;
- ii. How the combination of three letters 'consonant-consonant-long vowel' is characterised by the absence of a vowel on the first consonant;
- iii. Phonetic changes from SA to Levantine CA in the pronunciation of the following consonants ث /θ/ is pronounced /t/ in Levantine CA. In some words it can be pronounced /s/; ق /θ/ is pronounced /ʔ/ in Levantine CA. It can also be pronounced /g/; ء /ʔ/ is often not pronounced in Levantine CA; ة /a/ is often pronounced /e/ in Levantine CA;
- iv. Phonetic changes in the pronunciation of numbers;
- v. How Wh-questions change from SA into Levantine CA;
- vi. How the conjugation of the present tense changes from SA into Levantine CA.

The exercises in the language material provided for the experimental group is written exclusively in Arabic script and it specifies the equivalent in SA and CA of all the lexical items covered. As per the control group, FFI introduces the same afore mentioned forms through meaning-focused instruction, without explicitly explain their equivalents in SA. Students are provided with groups of words in CA that stress the phonological or lexical forms covered. The material is written in CA and the explanation of the CA phonological rules are provided in English.

4.7 Conclusion

This methodological chapter outlines the different methods, participants and measurements used to advance the empirical research. Not only have I explained the research method used to conduct the empirical study, but also how it supports the collection of data that answer my research questions. The research method is built on the theoretical perspectives that I have outlined through the first two theoretical chapters, namely the literature review and the theoretical framework. Language authenticity in Arabic is interpreted here as constant performance of language variation. Language variation is cognitively represented through the reference-packagings model and it is performed through diglossic code-switching. In order to represent the reality of Arabic, TAFL programmes should include, in my opinion, guidelines that support the development of two abilities, these being the ability to master two language varieties, and, secondly, the ability to code-switch between them as native speakers do. This thesis contributes to this field by researching whether and how two instruction methods based on the combination of FFI and CLT, can support the development of diglossic knowledge of vocabulary and code-switching skills at higher-education level. In doing so, it also investigates the effects that the two methods have on students' perceptions of variation in Arabic and their development of code-switching awareness.

The chapter is divided into five sections that contribute to building the methodological framework of this thesis as well as to describing the features of the language resources developed for this research. In the first section I summarised the threefold focus of investigation of the research. I also recapitulated the learning outcomes against which I assess the students'

development of diglossic vocabulary. These are based on the first five levels of Bloom's taxonomy of educational objectives. Diglossic language awareness is measured, instead, using Leow's level of language awareness. Finally, I also collect students' perceptions of variation in the reality and learning of Arabic. In the second section I explained how I combine FFI and CLT using two distinct methods of instruction, with the aim of comparing their impact and effects on the development of diglossic vocabulary and code-switching skills. The two methods of instruction are used with groups of participants that are divided into experimental and control groups. The delineation of the characteristic of the two instructional methods employed in the research has led to the subsequent section, in which I provided the details of the research methods and the data collection. This research is based on a mixed-methods research design, in which quantitative and qualitative data are collected and are integrated to achieve a more exhaustive understanding of the variables studied. The mixed-methods methodology is applied through an embedded quasi-experimental design. In the fourth section I explained the selection of the sources; the instruments used for data analysis; and the coding system used. I also delineated the pilot study carried out before the beginning of the empirical research and I explained how it provided fundamental information for the delineation of the empirical language research. Finally, I outlined the material used in the language lessons and the features of their language activities in the last section.

This methodological section also shows how the research questions are to be analysed. It presents the methods used to gather the data and gives arguments supporting these choices. This chapter shows that quantitative and qualitative methods are necessary to account for the heterogeneity of cases. However, it also shows that methodological bias can exist in educational research, due to the need to accommodate the features of small-scale language education research, as is the case of this research. As shown in section 4.4.2, these features include a small number of participants who are not recruited anonymously from a vast population of learners. In order to overcome these limitations, the quantitative analysis of the data is carried out by means of non-parametric statistics. The theoretical chapters (chapter 2, 3 and 4) of this thesis have now presented in detail the research background, framework and methods

of the thesis. The next three chapters comprise of the empirical part of this thesis. Each of the three chapters focuses on one aspect of the threefold focus of investigation, and shows the data collected to answer one research question. The first empirical chapter focuses on the development of diglossic vocabulary (chapter 5), the second on the development of diglossic language awareness (chapter 6) and the last on the students' perceptions of variation in Arabic (chapter 7). In the last chapter (chapter 8) I aim at answering the last research question by discussing the analysis of the empirical chapters, and by merging quantitative and qualitative data to analyse the results on the development of diglossic knowledge with information on students' language awareness and their perceptions of variation in Arabic.

5 CHAPTER 5 – Diglossic vocabulary development

5.1 Introduction

This chapter is the first of three empirical chapters presenting the analysis of the threefold focus of investigation as detailed in the previous sections. This first empirical chapter aims to observe diglossic vocabulary development. It answers the research question: does focus-on-form instruction lead to diglossic vocabulary development more effectively, when (a) it links forms in Standard and Colloquial Arabic, or when (b) it focuses only on one variety? Scores of language tests on diglossic vocabulary knowledge from beginner-level students are analysed in detail. The students are enrolled in one of the universities that participate in this study. The scores are analysed both within and across universities. The analysis in this chapter is based on test scores of the experimental and control groups. To observe whether diglossic vocabulary development has taken place and how two different methods of FFI, one that accentuates the links between corresponding forms in Standard and Colloquial Arabic, and one that concentrates only on CA, influence it, this thesis argues that different methodological tools are useful. Firstly, two aspects must be jointly observed. Both CA vocabulary retention and diglossic code-switching skills need to be observed. Furthermore, to more specifically examine the development of diglossic vocabulary knowledge, the students' ability to notice and perform diglossic code-switching is also assessed. This is useful to perceive the development of students' ability to code-switch precisely.

An in-depth view of each student's proficiency within the beginner-level of SA was provided through a language test, i.e. the pre-test, before the start of the empirical language research. Two further tests (the post-test and the delayed post-test) measured the language outcomes on CA vocabulary retention; diglossic code-switching skills; and the ability to notice and perform diglossic code-switching. The comparison between the scores of the two tests is instrumental to understanding the impact of the two different methods of instruction used, on long-term diglossic vocabulary knowledge. The three tests are scored on a scale of 0 to 100 and this study applies five statistical formulae that use the scores. The formulae are: average score; Wilcoxon-Mann-Whitney test; standard deviation; coefficient of standard deviation; and Wilcoxon 2-

sample Rank Sum. A detailed explanation of each formula is provided in section 5.2.

This chapter is structured as follows. The objective of the first section is to explain how the formulae used are essential for comparing the impact of the two forms of FFI on diglossic vocabulary building. More precisely, the first section explains how the result of each formula helps to ascertain whether the impact of the two forms of instruction on diglossic vocabulary building is significantly different, and how it differs (Section 5.2). The second section delineates how the exercises of the language test verify and assess the students' knowledge of diglossic vocabulary and their ability to code-switch and perform code-switching (Section 5.3). First, I describe the structure of the pre-test, which is based on SA vocabulary items and forms that are introduced in the CA language course afterwards. Second, the post-test and delayed post-test, which are identical tests on diglossic vocabulary development, are illustrated. The third section analyses the correlation between the results of the pre-test with those of the post-test and the delayed post-test among experimental and control groups (Section 5.4). I carry out the comparison between the results of the experimental and control groups within the three participant universities. This comparison allows me to analyse the effects that the two different methods of FFI have on the students' development of diglossic vocabulary knowledge, and to have an overview of the long-term effect of the two methods of instructions. This facilitates a deep understand of how diglossic vocabulary knowledge varies in the long term among the participants of the same groups, and to compare its variation among pairs of experimental and control groups. The results of the post-test and delayed post-test are analysed in view of the students' results on the pre-test. The results of the pre-test allow me to analyse the development of students' knowledge considering their actual proficiency levels in SA at the beginning of the course. This section is followed by the fourth section, in which the results are organised in subsections in which the experimental and control groups are clustered on the basis of the students' scores in the pre-test and labelled as follows: low-average proficiency; average proficiency; and high-average proficiency. I finally compare the scores on CA vocabulary retention; diglossic code-switching; and presence of language chance of the post-test and delayed post-test across the participant Universities (Section 5.6).

5.2 Statistical formulae

This section illustrates the formulae used in this research to compare the impact of the two forms of FFI on diglossic vocabulary building.

The average score refers to the arithmetic mean average. For simplicity, I refer to the arithmetic mean average here as the 'average.' I calculate the average for each group of students using the values of their test scores, which are summed up and divided by the number of student of each group. The average score is calculated for each language test.

In order to determine whether there exists a significant difference between the average scores of the experimental and control groups, I employ the Wilcoxon-Mann-Whitney test. This test is used to measure the non-parametric distributions of data of this thesis, as explained in the previous chapter in section 4.4.2. If there is a significant difference between the average scores of the experimental and control groups, it is possible to argue that one form of FFI instruction is likely to lead to diglossic vocabulary development more effectively than the other. A significant difference between two groups means that there is a measurable difference between the groups and that, statistically, the probability of obtaining that difference by chance is very small. Thus, it is safe to assume that the difference is due to experimental manipulation, which is, in this research, the employment of two different teaching methods of FFI to teach CA vocabulary and forms. The Mann-Whitney test operates as follows. The null hypothesis is that there is no significant difference between the scores of the observed groups of students, and that any difference that is observed is due to sampling or experimental error. If there is no difference between the distributions of both groups, there is a 50% probability that an observation from a value randomly selected from one set of values is equal to an observation randomly selected from the other set. Similarly, there is a 50% probability that two randomly selected values, each one from one of the two sets of values, are different. Therefore the p-value is set at ($p \leq 0.05$). The Wilcoxon-Mann-Whitney test determines whether the null hypothesis is rejected or fails to be rejected. If the p-value is below or equals 0.05, the null hypothesis can be rejected and the difference between the experimental and control groups is statistically significant. If the p-value is above 0.05, the decision is not to reject the null

hypothesis and the difference between the experimental and control groups is not statistically significant.

Establishing whether there exists a statistically significant difference between the average scores of two groups does not provide us with information regarding the size of the difference between the groups. On the one hand, the results between two groups might be statistically significant but the difference between them could lack a meaningful effect. On the other hand, the results might be statistically not significant but the difference between them could carry revealing and enlightening information. With a view to including an evaluation of the difference between the average scores of sets of groups, I measure the standard deviation and the coefficient of standard deviation. If the data are normally distributed, most of the values in the set of data are clustered around the average, and only few examples tend to one of the two extremes. If the data are not normally distributed, there exist numerous values that are not close to the average. The standard deviation is a value that expresses whether and how tightly the examples are close to the average in a set of data, or if they are spread out from the average. Due to the fact that it provides information on the distance of the values from the average within a set of values, the standard deviation presents a picture of the distribution of the values within the set that is more precise than the one given by the average. The standard deviation is therefore of fundamental importance in comparing two sets of data to gain a better understanding about the position and heterogeneity of the values with respect to the average.

The coefficient of standard deviation is useful in comparing the degree of variation from one data series to another, as it shows whether two data sets have a similar or different probability of distribution of the data. The coefficient of standard deviation shows the extent of variability in relation to the average and it represents the ratio of the standard deviation to the average. It considers the average as the total variation and it expresses in percentage the frequency or probability of distribution. The data series with the smaller coefficient of standard deviation is less dispersed than the one with the larger coefficient.

Finally, I use the Wilcoxon 2-sample Rank Sum test to compare two sets of scores that are collected from the same participants. As per the Wilcoxon-Mann-Whitney test, this test is used to measure the non-parametric distributions

of data of this thesis, as illustrated in section 4.4.2 of the previous chapter. The Wilcoxon 2-sample Rank Sum test compares and analyses in this research two sets of scores collected from the same participants from one point in time to another, as is the case of data from the post-test and the delayed post-test. The null hypothesis is that there is no significant change between the scores of the pre-test and the scores of the post-test of the same group of participant. If the result of the test rejects the null hypothesis, then there is a significant difference between the scores. If the result of the test does not reject the null hypothesis, the difference between the scores is not significantly different. The p-value is set at 0.05.

I have described here the formulae used to analyse the empirical data collected through the language tests, and I proceed to illustrate the features and exercises that characterise the tests. The following section delineates how the exercises of the language test assess knowledge SA (through the pre-test) and of diglossic vocabulary; the ability to code-switch between language variety; and to perform diglossic code-switching (through the post-test and delayed post-test). I firstly illustrate the exercises of the pre-test, and I subsequently describe the tests on diglossic knowledge.

5.3 Language tests

The following section delineates how the exercises of the language test assess knowledge SA (through the pre-test) and of diglossic vocabulary; the ability to code-switch between language variety; and to perform diglossic code-switching (through the post-test and delayed post-test). I illustrate the exercises of the pre-test, and I subsequently describe the tests on diglossic knowledge.

5.3.1 Design of pre-test

The exercises of the pre-test are aimed at testing vocabulary within the topics that are covered in the experimental language course. The topics are as follows: greetings, introducing oneself, classroom objects, numbers, days of the week and months, telling the time. It also includes few everyday verbs in the present tense, adjectives, Wh-questions, prepositions and compass points. The test divides the topics in two groups that are respectively assessed via written

or oral part of the test. I firstly illustrate the written exercises and I afterward explain the oral questions.

Written test

The written test covers the following topics: introducing oneself; numbers; days of the week and months. It also includes few everyday verbs in the present tense, adjectives, prepositions and compass points. The exercises are presented here. The first is a reading exercise in which a student describes himself and his family in Arabic. The participants are asked to fill in a table with information on the student in English. I ask them to provide the information in their mother tongues (English or Italian) to ensure it is well understood. The passage reads as follows:

إِسْمِي سَامِي حَسَنٌ وَأَنَا طَالِبٌ جَدِيدٌ فِي هَذِهِ الْجَامِعَةِ. أَنَا مِنَ الْمَغْرِبِ، وَلِدْتُ وَكَبِرْتُ فِي مَدِينَةِ
الرَّبَاطِ وَلَكِنِّي أُسْكُنُ فِي مَدِينَةِ إِكْزِيْتِرِ الْآنَ. عُمْرِي تِسْعَ عَشْرَةَ سَنَةً. فِي الْحَقِيقَةِ، سَأُبْلِغُ عَشْرِينَ
سَنَةً فِي الشَّهْرِ الْقَادِمِ: عِيدَ مِيلَادِي فِي يَوْمِ ٢٠ مَارْسِ/ آدَار. لِي أَخٌ وَأُخْتُ:
أَخِي يَسْكُنُ هُنَا فِي أَنْكَلْتِرَا بَيْنَمَا أُخْتِي تَسْكُنُ فِي الْمَغْرِبِ مَعَ أَهْلِهَا، زَوْجُهَا وَبِنْتُهَا. ذَهَبْتُ إِلَى
الرَّبَاطِ السَّنَةَ الْمَاضِيَةَ وَزُرْتُ أَبْنَاءَ أَعْمَامِي وَكُلَّ أَقَارِبِي.

My name is Sami Hassan and I am a new student in this University. I am Moroccan, I was born and raised in the city of Rabat but I live in Exeter now. I am 19 years old. Actually, I will turn 20 next month: my birthday is on the 20th of March. I have a brother and a sister: my brother lives in Exeter whereas my sister lives in Morocco with her family (her husband and her daughter). I went to Rabat last year and I visited my cousins and all my relatives.

The students are asked to provide the following information on the text: the name of the student; his surname, nationality, age, and birthday; and information about the members of his. The second exercise consists of a matching activity with vocabulary related to numbers, days of the week and months.

Table 3 Matching exercise (pre-test)

English	Arabic
May	تسعة
Twenty-eight	يوم الاثنين
Thirty-six	يناير/ كانون الثاني
Saturday	خمسة عشرة
Nine	يوم الثلاثاء
Tuesday	أيار/ مايو
Monday	ستة وثلاثون
Fifteen	يوم السبت
January	ثمانية وعشرون

Source: author.

The third exercise requires the students to provide the opposites of a list of nouns and adjectives. The lexical items provided are: قريب (close); طويلة (long or tall); سهل (easy); كبيرة (big); جديد (new); غداً (tomorrow); يمين (right). Students are asked to maintain the gender of the adjectives provided, i.e. 'long' and 'big' are feminine and therefore should take a feminine opposite; whereas 'new', 'close' and 'easy' are masculine and should take a masculine opposite. In the fourth exercise, students are provided with the list of letters comprising the words of the four compass points and are asked to join the letters together and form the correct spelling of the four compass points شمال (North); جنوب (South); غرب (West); and شرق (East). The fifth exercise tests knowledge on prepositions of place. Several illustrations of the position of a ball in relation to a box and a table are provided. The students are asked to write the correct prepositions in Arabic that describe the relationship of the ball with the box and the table. The prepositions are: وراء (behind); أمام (in front); تحت (under); على (on); فوق (above); بجانب (next to); and بين (between). Finally, the last exercise assesses knowledge of

everyday-life verbs conjugated in the present tense. The verbs included in the exercise are: درس (study); لعب (play); عمل (work); قرأ (read); شرب (drink); أكل (eat); قال (tell); and سافر (travel). The exercise consists in conjugating each verb according to a personal pronoun that is written next to the verb in the exercise, and to translate the conjugated verb.

Oral test

The oral questions cover greetings, introducing oneself (this part complements the written questions on this topic), classroom objects, telling the time, and wh-questions. Students are shown an illustration of day and night. I point at one of them and I expect them to say 'Good morning' or 'Good evening.' I respond accordingly and I ask them 'How are you?'. I then show them an illustration of two individuals greeting each other and I utter 'Hello.' The students are expected to respond appropriately, i.e. by using the expression 'Hello to you'. I then ask them their name, their age, where are they from, their date of birth and what is the time. I show them a photo portraying classroom objects, such as a door; a desk; a chair; a window; a whiteboard; notebooks and pens; a bag; and a small plastic bottle of water. Finally, I test the knowledge of Wh-words by asking the participants to listen to recorded questions, individuate the Wh-words they hear and translate them in their mother tongues. The questions are: 'What is your name?', 'How old are you?', 'How many languages do you speak?', 'Where are you from?', 'Which day is today?', 'Why?'

The oral test is scored on a scale of 0 to 25, whereas the written test is scored on a scale of 0 to 75. The overall test is scored on a scale of 0 to 100. The students are made aware that the nature of the topics and the vocabulary items assessed in the test are covered in the CA language course. They are provided with the keys to the questions and the correct answers to the exercises for reference.

I have illustrated the design and content of the pre-test and I now explain the post-test and delayed post-test in the following section.

5.3.2 Design of post-test and delayed post-test

The post-test and the delayed post-test are identical as we have seen in the previous chapter in section 4.4.1. This is because the aim of the delayed post-test is to shed light on how the students' knowledge of diglossic vocabulary differs within a period of three weeks. I planned and designed the exercises in order to be able to assess the cognitive levels of Bloom's taxonomy, against which I measure CA vocabulary receptive-productive knowledge and diglossic code-switching. The language test is composed of a written and an oral part. I firstly illustrate the written part and I subsequently proceed to delineate the oral part.

Written test

The written test assesses the cognitive levels of Bloom's taxonomy as they apply to CA vocabulary receptive-productive knowledge and diglossic code-switching. I firstly explain the exercises designed with the aim of assessing CA vocabulary receptive-productive knowledge. The first cognitive level is interpreted in this thesis as the ability of the students to recognise and recall newly learned CA vocabulary. This is assessed through a listening exercise. The students listen to a recording in SA in which everyday life activities are described in the present tense. After listening to the recording, they are asked to write the verbs they have recognised and understood, in CA. The verbs are *بيدرس* (study); *بيلاعب* (play); *بيشتغل* (work); *بيقرأ* (read); and *بياكل* (eat).

According to the second cognitive level, learners are able to translate from CA into their mother tongue. As with the first cognitive level, the second is also assessed through a listening exercise. I play twelve pre-recorded sentences in CA during the test and the students are required to translate them into their mother-tongue. Examples of sentences to be translated are: How old are you?; It's 10:45; Today is Tuesday; I usually study until 6 pm; The bottle is behind the door. According to the second cognitive level, students are also able to extrapolate and explain phonological and grammatical patterns that apply to CA. This ability is assessed in the oral test as I elucidate in the following section. The third cognitive level of Bloom's taxonomy, the ability to form correct sentences in CA, is also assessed through the oral test.

Within the fourth cognitive level, learners can infer meaning from a text by analysing how words related to each other syntactically and semantically. This is assessed in the written test through a multiple-choice exercise. The text consists of a dialogue between two friends, Sarah and Mohammad, about their plans for the following afternoon. The dialogue is meant to simulate a real-life situation and therefore it is written in CA. Each sentence of the dialogue has a gap and the students are given three choices from which to choose the appropriate lexical item to insert in the gap. I illustrate below an extract of the exercise:

سارة: صحيح! _____ هلق، بشوفك!

محمد: بشوفك!

(أ) أذهب (ب) بروح (ت) بذهب

(Sarah: you are right. I am leaving now, see you soon!

Mohammad: See you soon!)

The three alternatives provided are selected among the following categories: the correct answer (as is the example of answer ب here, which is the verb 'I go' in CA); an answer that seems correct but it is in the wrong variety (answer أ, which is the verb 'I go' in SA); and a wrong answer (answer ت, which is a non-existent word that combines the CA present-tense prefix with the SA root of the verb 'to go'). The wrong answer can be an irrelevant word, or a non-existent word created by casually blending a bound morpheme from one code and a free morpheme from the other code. The nature of the alternatives provided and the relevance of their forms to the meaning of the text, elucidate whether the students are able to analyse how words relates to each other syntactically and semantically and to infer the correct meaning of a text on the basis of their analysis.

As per the fifth learning objective for CA vocabulary receptive-productive knowledge, the expected learning outcome is that students can assess their language use based on their own judgement and that they are aware of their progress over time. This ability is assessed in the oral test as I explain in the

following section. As we have seen in the previous chapter, the sixth and last level of the taxonomy is not assessed in this research due to time constraints.

I now proceed to outline the written exercises designed with the aim of assessing diglossic code-switching skills. Diglossic code-switching skills within the first cognitive level of Bloom's taxonomy, are interpreted in this research as the ability of the students to distinguish groups of words as being CA or SA. This is assessed in the written test by providing two lists of words, and both lists need to be labelled by the students as being CA or SA groups of words. The two lists are reported below.

Table 4 Identification of groups of vocabulary words as CA or SA items

List A	List B
شمال	على
أمام	يمين
في	سهل
كم	عمل
قنينة	ما (?)
حلو	مشغول

Source: author.

The lists are composed of ambiguous words that are used in both varieties, but both lists contain one word that belongs to one variety only. This is the case of the word

أمام (in front of) among the vocabulary of 'list A', and the interrogative word ما (what) among the vocabulary of 'list B'. Both words are SA vocabulary items and therefore both lists are SA list of words.

According to the second cognitive level, students can link linguistically and semantically newly learned CA vocabulary against correspondent already known vocabulary in SA; they can retell SA texts in CA; extrapolate phonological and grammatical patterns that link SA and CA; and use these

patterns to predict how to form new vocabulary. I assess through the written test the first of the expected learning outcomes of this level. The remaining outcomes are tested through the oral test. To test the ability of the students to link newly learned CA vocabulary against correspondent already known SA lexical items, they are asked to match words in CA with their equivalents in SA.

The word pairs CA-SA are as follows: حلوة - جميلة (beautiful); عمل - شغل (work);

لماذا - ليش (why); يسار - شمال (left); أنتو - أنتم (you plural); and نذهب - منروح (we go).

As reported in the previous section and in the previous chapter, the sixth and last level of the taxonomy is not assessed in this research due to time constraints.

With regard to the third, fourth and fifth cognitive levels, the third and the fifth are assessed in the oral test and the fourth is measured in the written test. The expected outcome of the fourth cognitive level is that students analyse complex language information to identify the most appropriate and correct variety to use. The assessment of this level is integrated within the multiple-choice exercise based on the dialogue between Sarah and Mohammad as explained above. The dialogue is aimed at depicting the use of the language in a real-life situation and therefore is written in CA, bar one part that is written in SA. This occurs because in one segment of the conversation, SA is the appropriate variety to use. In this part one of the two friends asks the other's opinion on a presentation aimed at an academic audience. The segment of the dialogue in which the presentation is repeated out loud by one of the interlocutors is written in SA, to simulate the language variety properly employed in formal contexts. The students participating in this study are not aware that both CA and SA are used in the dialogue, but they are made aware, before the beginning of the test, that the dialogue aims at portraying an authentic use of the language. I illustrate below an extract of the exercise:

سارة: تقع المدينة في _____ إيطاليا

(أ) جَنُوب (ب) جَنُوب (ت) الجَنُوب

وهي مشهورة _____.

(أ) كثير (ب) كثيرة (ت) جدا

(Sarah: The city is situated in the south of Italy and it is very famous).

The three alternatives provided are selected among the same categories outlined above: the correct answer; an answer that seems correct but it is in the wrong variety; and a wrong answer. The students show their understanding of the use of language varieties in relation to the formality of the context, by selecting appropriate SA choices in the presentation section of the dialogue, and appropriate CA choices in the informal conversation section.

The written test comprises five exercises in total. The first two are written exercises based on listening activities and the third consists of two lists of words in SA. I explain the remaining two exercises in the following sections. The nature of the first three exercises is intentionally selected to reduce the amount of written text in CA that the students are requested to concentrate on. There are two reasons that justify my choice. First, this reduces the risk of bias in data collection. The experimental group has very little exposure to material written in CA during the language course. The language material employed for their language lessons is based on SA orthography, and CA vocabulary items are used only if they differ greatly from their SA equivalents. The experimental group is therefore less used to CA orthography than the control group, and I therefore consider it essential to minimise the use of written CA in the language test to reduce the chance of bias towards the control group. Second, both groups have limited experience of written CA in comparison with their familiarity with written SA, and neither is as accustomed to CA spelling as it is to SA spelling. Thus, in order to minimise the potential sources of confusion during the test, I make a limited use of CA in the form of written text. I have explained in detail the written part of the post-test and delayed post-test, and I proceed to outline the oral test.

Oral test

The explanation of the exercises that compose the oral test replicates that of the written test: I firstly focus on illustrating how the cognitive levels of Bloom's taxonomy are assessed with regard to CA vocabulary receptive-productive knowledge, and I focus on diglossic code-switching afterward. The ability of the students to recognise and recall newly learned CA vocabulary, i.e. the first

cognitive level, is tested by simulating various situations in which a variety of greetings can be used. Multiple conversations are simulated, and the students are expected to respond appropriately and to use expressions studied in the experimental language course. As per the second cognitive level, students are expected to be able to extrapolate and explain phonological and grammatical patterns that apply to CA. This is assessed during the oral test as follows. Students are asked to explain the CA phonological and form patterns of the sentence 'لېش ما بتحب الصيف؟' (Why don't you like summer?). Students are expected to (i) compare the CA word لېش with its equivalent in SA لماذا; (ii) illustrate how the negative particle ما is used differently in the two language varieties; (iii) contrast the conjugation of the present tense in SA with the same conjugation in CA; (iv) and explain the phonological characteristics of the pronunciation of الصيف. According to the third cognitive level of Bloom's taxonomy, students can form correct sentences in CA. This level is assessed by asking questions that replicate the questions of the pre-test, with the aim to giving the students the opportunity to use their language knowledge to form sentences based on their personal details. I ask them their name; their age; where they are from; their date of birth; what they study; and what the time is.

The fourth cognitive level is not assessed orally, as it is based on the ability of learners to infer meaning from a text. The fifth learning objective for CA vocabulary receptive-productive knowledge consists in the ability of the students to assess their language use based on their own judgement. They should also be aware of their progress over time. During the oral test, students are asked to assess their CA language use. Moreover, they have the opportunity to make comments on their mistakes and describe their progress over the weeks of instruction of the experimental language course.

I now illustrate how diglossic code-switching skills are assessed against the same taxonomy. The first cognitive level, according to which the students are able to distinguish groups of words as being CA or SA, is only assessed in the written test. The first element of the second level (which sees the students as being able to link linguistically and semantically newly learned CA vocabulary

against correspondent already known vocabulary in SA) is also assessed only in the written test. However, the remaining elements of the second cognitive level are measured orally: the ability to (i) retell SA texts in CA; (ii) extrapolate phonological and grammatical patterns that link SA and CA; (iii) and finally to use these patterns to predict how to form new vocabulary. In order to determine the students' ability to perform the afore mentioned cognitive knowledge, I specifically developed an oral exercise based on a written text. First, students read a passage in SA. The passage reads as follows:

فاطمة إنكليزية من أصل عربي. والد فاطمة أردني من عمان ووالدتها إنكليزية من لندن. تدرس فاطمة اللغات في الجامعة مع صديقتها سارة: اللغة الروسية واللغة الفرنسية. تسافر فاطمة إلى الأردن كل سنة في الصيف وأحيانا في عطلة عيد الميلاد أيضا. فاطمة تحب السفر إلى الأردن وزيارة عائلته.

(Fatima is English with Arabic origin. Her father is Jordanian from Amman and her mum is English from London. Fatima studies languages at University with her friend Sarah: she studies Russian and French. Fatima travels to Jordan every summer and sometimes during Christmas holidays as well. She loves travelling to Jordan and visiting her family).

Second, students are asked to retell the text in CA. Third, they are asked to extrapolate two recurrent patterns of switch, one phonological and one grammatical, from SA to CA. For example, a phonological pattern of change is the difference in pronunciation of the consonant ق /q/ and a grammatical change consists on using a different preposition after the verb سافر (to travel).

This applies to various verbs of movement, such as 'to go' and 'to return'. سافر is followed by إلى in SA and by the preposition على in CA. The preposition على can be used in its full spelling, or it can be abbreviated to ع /ʔ/ and attached to the word that follows it. Finally, the students are asked to use the two patterns of switch detected in the text, and make two original sentences in CA, each sentence respectively containing one pattern of change.

In the following cognitive level, learners apply their sociolinguistic knowledge to decide whether it is appropriate to use CA or SA in authentic contexts through role plays. This is assessed in the oral test by means of simulating two situations characterised by different degrees of formality. In each situation, the students are asked to perform a character and to use CA or SA appropriately. Finally, the fifth cognitive level is assessed by asking the students to evaluate their use of diglossic code-switching, and how their ability to code-switch between language varieties has developed throughout the period of instruction received in the experimental language course.

The oral test is scored on a scale of 0 to 25, whereas the written test is scored on a scale of 0 to 70. The total score of CA vocabulary retention knowledge across both exams is 65 (50 points in the written test and 15 in the oral test), and the total score of diglossic vocabulary skills is 30 (20 points in the written test and 10 in the oral test). The remaining 5 points are scored in the exercise based on the ability of the students to retell the SA text in CA, by performing diglossic code-switching accurately. More precisely, the points are attributed to the following changes: (i) والد in SA becomes أبو in CA (father); (ii) والدة in SA becomes أم in CA (mother); (iii) تدرس in SA becomes بتدرس in CA (she studies); (iv) تسافر in SA becomes بتسافر in CA (she travels); (v) تحب in SA becomes بتحب in CA (she loves). The overall test is scored on a scale of 0 to 100.

I have illustrated the design and content of the post-test and delayed post-test and I analyse in the following section the correlation between the results of the pre-test with those of the post-test and the delayed post-test among experimental and control groups.

5.4 Results on language tests of participant universities

This section reports the data on the development of diglossic vocabulary collected through the language tests by the experimental and control groups. I compare here the results of the experimental and control groups within the three participant Universities, with the aim to providing an overview of the impact that the two different FFI employed in this research have on the

students' development of diglossic vocabulary knowledge. The results of the post-test and delayed post-test are analysed in light of the results of the pre-test, which allow to understand the development of students' diglossic knowledge considering their beginner proficiency levels in SA.

A detailed analysis of the data is reported in the following section and it is divided into three parts. I divide the experimental and control groups in three subgroups, based on their scores on the pre-test: low-average proficiency (scores between 33 and 60), average proficiency (scores between 61 and 80); and high-average proficiency (scores between 81 and 100). This allows me to analyse the results on CA vocabulary development; diglossic code-switching; and on presence of language change of the post-test and delayed post-test, in light of the actual proficiency of the students in SA at the beginning of the course. I start by comparing the results of the control and experimental groups of the University of Milan.

Table 5 Language tests, University of Milan

CONTROL GROUP	Pre Test	Post Test	Delayed Post-Test	EXPERIM. GROUP	Pre Test	Post Test	Delayed Post-Test
Student				Student			
MICA1	90.0	84.0	75.0	MICASA1	51.7	56.0	49.0
MICA2	88.3	93.5	83.0	MICASA2	56.7	68.0	62.0
MICA3	93.3	88.0	82.0	MICASA3	58.3	63.0	58.0
MICA4	68.3	78.0	73.0	MICASA4	95.0	91.0	83.0
MICA5	90.0	80.0	74.0	MICASA5	95.0	92.0	87.0
MICA6	50.0	42.0	35.0	MICASA6	90.0	85.0	78.0
MICA7	53.3	40.0	33.0	MICASA7	46.7	69.0	60.0
MICA8	50.0	60.0	53.0	MICASA8	65.0	79.0	73.0
MICA9	88.3	70.0	62.0	MICASA9	58.3	81.0	75.0
MICA10	56.7	62.0	56.0	MICASA10	78.3	75.0	70.0
MICA11	71.7	74.0	68.0	MICASA11	71.7	63.0	56.0
MICA12	68.3	73.0	65.0	MICASA12	65.0	71.0	65.0
MICA13	98.3	97.0	90.0	MICASA13	50.0	56.0	51.0
MICA14	83.3	77.0	76.0	MICASA14	58.3	66.0	61.0
MICA15	61.7	68.0	68.0	MICASA15	50.0	63.0	56.0
				MICASA16	91.7	88.0	82.0
				MICASA17	75.0	71.0	64.0
				MICASA18	85.0	96.0	93.0
Average	74.1	72.4	66.2	Average	69.0	74.1	67.9
SD	17.1	16.5	16.4	SD	16.7	12.5	12.8
CoV	23.0	22.8	24.8	CoV	24.2	16.8	18.9

Source: author.

The data show that the control group performed better than the experimental group in the pre-test by 5.1 points. However, it is outperformed by the experimental group in the post-test and in the delayed post-test. The scores of the post-test are higher than those of the delayed post-test in both groups, and the average results of the delayed post-test lowers by the same amount of points for the two groups. It is interesting to note that not only are the average results scored by the experimental group in both tests higher than those of the control group, but also the results of the experimental group are more clustered around the average value and are therefore more homogeneous. The standard variation of the control group shows, instead, that the results are more spread out around the average. If we look at the individual results of the tests, we see that they spread out both below and above the score of the average. It is also

interesting to note that the level of internal variation in the control group remains almost constant across the three tests (the results are respectively 17.1, 16.5 and 16.4), whereas the internal variation of the experimental group lowers significantly from an initial score of 16.7 to the scores of the post-test (12.5) and the delayed post-test (12.8). The most striking feature of the comparison between the two data sets is the coefficient of variation. The two groups have a similar coefficient of variation in the pre-test and therefore they have a similar probability of distribution of the data. However, this changes significantly in the post-test and in the delayed post-test. As we have seen, the coefficient of standard deviation shows the extent of variability in relation to the average and it represents the ratio of the standard deviation to the average. It represents the probability of distribution. The figures show that the experimental group is less likely to present dispersed results in comparison with the control group.

The Wilcoxon-Mann-Whitney test elucidates that the difference of the results scored in the post-test by the two groups is not statistically significant, as the p-value is 0.492 and therefore it is higher than the value set at 0.05. The difference in the results scored in the delayed post-test is also not statistically significant, as the p-value is higher than 0.05 (p-value 0.928). As the difference between the results of the two group is not statistically significant, it is not possible to state that the experimental group outperformed the control group due to the intervention and the experimental form of FFI it is exposed to. However, it is worth noting that there is a pattern of scores: the experimental group outperformed the control group in both the post-test and in the delayed post-test, and in both tests the values are more homogeneously spread than the results of the control group. Meanwhile, the Wilcoxon 2-sample Rank Sum test shows that the differences between post-test and delayed post-test of both groups are statistically significant: the p-value for the control group is 0.00096 and the p-value of the experimental group is 0.0002 and thus they are both lower than 0.05. The significance between the two scores can be assessed by comparing the two p-values. The experimental group shows a lower value and therefore the difference is more significant than that of the control group. It is worth noticing that in both groups the performance declines between the post-test and the delayed post-test. The measurement of Wilcoxon 2-sample Rank Sum shows that the in the case of the experimental group the difference is more

significant than in the other group, which means that the results between the post-test and the delayed post-test are more significantly lower than those of the control group.

To sum up, the data show that the experimental group performed better than the control group in the post-test and in the delayed post-test; that the values of both tests are more homogeneously spread around the average in comparison with those of the control group; and finally, that in both tests the experimental group presents a lower probability of internal variation than the control group. This means that the results are likely to be closer to the average in comparison to the results of the control group. However, in the long term the performance of the experimental group declines more significantly than that of the control group.

Scores of students of the University of Genoa

I compare here the results of the control and experimental groups of the University of Genoa.

Table 6 Language tests, University of Genoa

CONTROL GROUP	Pre Test	Post Test	Delayed Post-Test	EXPERIM. GROUP	Pre Test	Post Test	Delayed Post-Test
Student				Student			
GECA1	46.7	46.0	49.0	GECASA1	93.3	92.0	86.8
GECA2	43.3	49.0	51.0	GECASA2	46.7	60.0	51.8
GECA3	33.3	36.0	37.0	GECASA3	66.7	72.0	64.9
GECA4	38.3	41.0	39.0	GECASA4	83.3	98.0	93.0
GECA5	76.7	77.0	77.0	GECASA5	80.0	78.0	71.1
GECA6	63.3	69.0	69.0	GECASA6	78.3	62.0	46.5
GECA7	75.0	65.0	67.0	GECASA7	40.0	48.0	46.5
				GECASA8	51.7	58.0	53.5
Average	53.8	54.7	55.6	Average	67.5	71.0	64.3
SD	17.7	15.6	15.6	SD	19.4	17.4	18.1
CoV	32.9	28.4	28.0	CoV	28.7	24.5	28.1

Source: author.

The data show that the experimental group outperformed the control group in every language test. The average of the experimental group is 13.7 points higher than the control group in the pre-test, it is 16.3 points higher in the post-test, and 8.7 points higher in the delayed post-test. However, it is worth noticing that the three-week period between the post-test and the delayed post-test had a positive effect on the control group, whereas it had a negative effect on the experimental group. The former improves its language performance by 0.9 points on average, whereas the latter decreases its performance by 6.7 points on average. The standard deviation of the two groups is also worth noticing. The control group reduces the level of internal dispersion from its value in the pre-test to that of the post-test, and it maintains a constant value between the post-test and the delayed post-test. The experimental group, instead, shows a different pattern: the internal dispersion of values decreases from the pre-test to the post-test by 2 points, but it increments by 0.7 in the delayed post-test. Therefore, the internal dispersion of values is less homogeneous than that of the control group and it is less consistent over a delayed period time. As per the

results collected at the University of Milan, the most striking feature of the comparison between the two data sets is the coefficient of variation. The likelihood of internal variation is higher for the control group in the pre-test, and it steadily decreases in the post-test and in the delayed post-test. In contrast, the experimental group does not show a steady pattern: the value of the coefficient of variation decreases between the pre-test and the post-test, but it increases in the delayed post-test to almost reach the same figure registered in the pre-test. The Wilcoxon-Mann-Whitney test illustrates that the difference of the results scored in the post-test by the two groups is not statistically significant, as the p-value is 0.059 and therefore it is higher than the value set at 0.05. The difference of the results scored in the delayed post-test is also not statistically significant, as the p-value is higher than 0.05 (p-value 0.384). As per the figures gathered at the University of Milan, it is not possible to claim that the results of the language tests are influenced by the experimental language intervention. However, it is worth noting that there is a pattern of scores: the control group performs steadily better over time and it reduces its level of internal dispersion. Thus, the language performances of the students in the group become increasingly more homogeneous over time. The experimental group, instead, shows an initial positive response to the language intervention by outperforming the control group in the post-test and by reducing its internal level of dispersion. However, its results in the delayed post-test are less homogeneous and the overall average language score is significantly lower. Finally, the size of the sample is not large enough for the distribution of the Wilcoxon 2-sample Rank Sum test statistic to form a normal distribution and it is therefore not possible to calculate an accurate p-value.

Scores of students of the University of Exeter

I compare here the results of the control and experimental groups of the University of Exeter.

Table 7 Language tests, University of Exeter

CONTROL GROUP	Pre Test	Post Test	Delayed Post-Test	EXPERIM. GROUP	Pre Test	Post Test	Delayed Post-Test
Student				Student			
EXCA1	71.7	72.0	60.0	EXCASA1	71.7	69.0	60.0
EXCA2	60.0	56.0	57.0	EXCASA2	56.7	39.0	34.0
EXCA3	81.7	63.0	72.0	EXCASA3	83.3	65.0	62.0
EXCA4	46.7	46.0	44.0	EXCASA4	75.0	71.0	71.0
EXCA5	93.3	58.0	57.0	EXCASA5	60.0	76.0	71.0
EXCA6	70.0	68.0	64.0	EXCASA6	100.0	95.0	90.0
EXCA7	88.3	87.0	82.0	EXCASA7	100.0	93.0	83.0
EXCA8	76.7	77.0	74.0				
Average	73.5	65.9	63.8	Average	78.1	72.6	67.3
SD	15.2	13.0	12.0	SD	17.4	18.8	18.2
CoV	20.6	19.7	18.8	CoV	22.3	26.0	27.0

Source: author.

The data show here that the experimental group performed better than the control group in all language tests, respectively by 4.6, 7.7 and 3.5 points on average. The scores of the language tests in both groups are progressively lower, with the highest results in both groups being those of the pre-test, and the lower results being those of the delayed post-test. It is interesting to note that the standard deviation and the coefficient of variation, instead, follow a different pattern. Both the standard deviation and the coefficient of variation show that the results of the control group are increasingly more homogenous over time and the likelihood of internal dispersion of values decreases steadily. As per the experimental group, instead, the level of internal dispersion is not stable and swings between the language tests. Interestingly, the likelihood of internal variation increases steadily across the three language tests. This situation could be partly due to a particularly low language performance in the post-test and delayed post-test of student EXCASA2, who scored 39 and 34 points, against the average figures of the experimental group that respectively scored 72.6 and 67.3. The results of the two test are widely dispersed around

the average in comparison to the remaining scores of the group and to the scores of the control group, and this is likely to play a significant role in the figures of the standard deviation and of the coefficient of variation. As is the case of the results collected at the Universities of Milan and Genoa, the Wilcoxon-Mann-Whitney test shows that the difference of the results scored in the post-test by the two groups is not statistically significant. This is because the p-value is 0.192 and therefore it is higher than the value set at 0.05. The difference of the results scored in the delayed post-test is also not statistically significant, as the p-value is higher than 0.05 (p-value 0.562). It is therefore not possible to state that the experimental group outperformed the control group due to the impact of the experimental language course and of the language treatment. However, it is worth noting that, as is the case of the two previous Universities, there exists a pattern among the scores. On the one hand the experimental group outperforms the control group in all the language tests, but on the other hand, its levels of internal distribution of the scores are increasingly less homogeneous and significantly more spread out from the average. On the contrary, the control group shows a steadier distribution of scores that progressively decreases over time. Similarly, the likelihood of internal variation also lowers in the long-term. Finally, as is the case of the groups from the University of Genoa, the size of the sample is not large enough for the distribution of the Wilcoxon 2-sample Rank Sum test statistic to form a normal distribution and it is therefore not possible to calculate an accurate p-value.

5.5 Results of subgroups on diglossic vocabulary development

In this section I divide the participants into subgroups based on their average scores in the pre-test. I identify three subgroups within each participant University: low-average proficiency; average proficiency; and high-average proficiency. The average is calculated here by dividing the sum of the average values of the experimental and control groups by the number of groups. The total sum of the average scores is 416. The sum of the average scores divided by the number of groups, which is 6, provides an average of 69. I divide the three groups as follows. The lowest result scored in the pre-test is 33 and the highest score is 100. Within the range 33-100, I created three groups of approximately the same range of scores and in which the value of 69 is at the

centre of the average group. This resulted in the creation of the following subgroups: low-average proficiency (scores from 33 to 60); average proficiency (scores from 61 to 80); and high-average proficiency (scores from 81 to 100). This allows me to analyse the results on CA vocabulary development; diglossic code-switching; and on presence of language change of the post-test and delayed post-test, in light of the actual proficiency of the students in SA at the beginning of the course.

The results are reported as follows. The scores are displayed in tables of nine columns. The first column reports the results of the pre-test. I then report the scores of the post-test on CA vocabulary retention; diglossic vocabulary skills; and ability to perform diglossic code-switching. The following column displays the total scores of the post-test. I then repeat the same sequence for the delayed post-test. CA vocabulary retention and diglossic vocabulary skills are collected by means of both written and oral tests. The ability to perform diglossic code-switching accurately is instead measured only through the oral test.

5.5.1 University of Milan

Low-average proficiency subgroups

Table 8 Diglossic vocabulary in low-average control group, Milan

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw	Total	CA Reten.	CS Skills	Diglos. Co-Sw	Total
MICA6	50.0	27.0	13.0	2.0	42.0	20.0	13.0	2.0	35.0
MICA7	53.3	26.0	11.0	3.0	40.0	20.0	10.0	3.0	33.0
MICA8	50.0	49.0	9.0	2.0	60.0	40.0	12.0	1.0	53.0
MICA10	56.7	45.0	15.0	2.0	62.0	38.0	16.0	2.0	56.0
Average	52.5	36.8	12.0	2.3	51.0	29.5	12.8	2	44.3
SD	3.2	12.0	2.6	0.5	11.6	11.0	2.5	0.8	11.9
CoV	6.1	32.5	21.5	17.8	22.8	37.3	19.6	32.7	27.0

Source: author.

Table 9 Diglossic vocabulary in low-average experimental group, Milan

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
MICASA1	51.7	37.0	16.0	3.0	56.0	32.0	14.0	3.0	49.0
MICASA2	56.7	48.0	18.0	2.0	68.0	44.0	16.0	2.0	62.0
MICASA3	58.3	46.0	14.0	3.0	63.0	42.0	13.0	3.0	58.0
MICASA7	46.7	52.0	14.0	3.0	69.0	45.0	13.0	2.0	60.0
MICASA9	58.3	56.0	21.0	4.0	81.0	51.0	20.0	4.0	75.0
MICASA13	50.0	37.0	15.0	4.0	56.0	32.0	16.0	3.0	51.0
MICASA14	58.3	42.0	19.0	5.0	66.0	40.0	16.0	5.0	61.0
MICASA15	50.0	45.0	15.0	3.0	63.0	39.0	14.0	3.0	56.0
Average	53.8	45.4	16.5	3.4	65.3	40.6	15.3	3.1	59.0
SD	4.7	6.7	2.6	0.9	8.0	6.5	2.3	1.0	8.0
CoV	8.7	14.8	15.5	24.1	12.3	15.9	15.2	28.2	13.5

Source: author.

The experimental group in the low-average proficiency subgroup in Milan shows a better performance both in the post-test and in the delayed post-test. The score is a sum of overall higher results in all the three language skills tested in both tests. It is worth noticing that although the control group shows a more homogeneous distribution of scores in the pre-test than the experimental group, the latter registers lower levels of dispersion in both the post-test and the delayed post-test. The control group has a noticeably great likelihood of high internal dispersion in the results of all language skills and in particular in the ability to perform diglossic code-switching. On the contrary, the experimental group shows a stable probability of internal dispersion between the post-test and delayed post-test in both CA vocabulary retention and diglossic vocabulary skills, and a slightly higher dispersion for the ability to perform diglossic code-switching.

Average proficiency subgroups

Table 10 Diglossic vocabulary in average control group, Milan

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
MICA4	68.3	56.0	19.0	3.0	78.0	52.0	18.0	3.0	73.0
MICA11	71.7	53.0	18.0	3.0	74.0	47.0	19.0	2.0	68.0
MICA12	68.3	53.0	17.0	3.0	73.0	45.0	17.0	3.0	65.0
MICA15	61.7	50.0	16.0	2.0	68.0	49.0	16.0	3.0	68.0
Average	67.5	53.0	17.5	2.8	73.3	48.3	17.5	2.8	68.5
SD	4.2	2.4	1.3	0.5	4.1	3.0	1.3	0.5	3.3
CoV	6.2	4.6	7.4	14.5	5.6	6.2	7.4	14.5	4.8

Source: author.

Table 11 Diglossic vocabulary in average experimental group, Milan

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
MICASA8	65.0	52.0	23.0	4.0	79.0	47.0	23.0	3.0	73.0
MICASA10	78.3	51.0	20.0	4.0	75.0	48.0	18.0	4.0	70.0
MICASA11	71.7	41.0	19.0	3.0	63.0	37.0	16.0	3.0	56.0
MICASA12	65.0	47.0	20.0	4.0	71.0	42.0	19.0	4.0	65.0
MICASA17	75.0	46.0	21.0	4.0	71.0	32.0	28.0	4.0	64.0
Average	71.0	47.4	20.6	3.8	71.8	41.2	20.8	3.6	65.6
SD	6.0	4.4	1.5	0.4	5.9	6.8	4.8	0.5	6.5
CoV	8.4	9.3	7.4	9.8	8.3	16.4	22.9	12.7	9.9

Source: author.

The experimental and control groups in the average proficiency subgroup of the University of Milan show a different pattern than those of the previous subgroup. There are remarkable figures both in the average scores of the language tests and in the internal dispersion of the scores. Firstly, the experimental group outperforms the control group in the pre-test, but its scores on the post-test and delayed post-test are lower than those of the control group. The overall higher scores of the control group are due to notably better results on CA vocabulary

retention in both tests. In fact, the scores on diglossic vocabulary skills and on the ability to perform diglossic code-switching are slightly lower, but their difference to the scores of the experimental group is not sufficient to have an impact on the overall score of the language tests. Secondly, the standard deviation of the control group is steadily lower than the standard deviation of the experimental group both in diglossic vocabulary skills and CA vocabulary retention, which means that the results are placed more homogeneously around the average value for both the post-test and the delayed post-test. It is almost equal in the ability to perform diglossic code-switching.

High-average proficiency subgroups

Table 12 Diglossic vocabulary in high-average control group, Milan

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
MICA1	90.0	57.0	24.0	3.0	84.0	50.0	22.0	3.0	75.0
MICA2	88.3	62.5	27.0	4.0	93.5	56.0	24.0	3.0	83.0
MICA3	93.3	53.0	30.0	5.0	88.0	49.0	28.0	5.0	82.0
MICA5	90.0	51.0	25.0	4.0	80.0	47.0	24.0	3.0	74.0
MICA9	88.3	47.0	20.0	3.0	70.0	41.0	18.0	3.0	62.0
MICA13	98.3	64.0	29.0	4.0	97.0	59.0	27.0	4.0	90.0
MICA14	83.3	51.0	23.0	3.0	77.0	50.0	23.0	3.0	76.0
Average	90.2	55.1	25.4	3.7	84.2	50.3	23.7	3.4	77.4
SD	4.7	6.3	3.5	0.8	9.5	5.9	3.3	0.8	8.8
CoV	5.2	11.5	13.8	20.4	11.2	11.7	13.9	22.9	11.4

Source: author.

Table 13 Diglossic vocabulary in high-average experimental group, Milan

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
MICASA4	95.0	61.0	26.0	4.0	91.0	57.0	23.0	3.0	83.0
MICASA5	95.0	62.0	26.0	4.0	92.0	57.0	26.0	4.0	87.0
MICASA6	90.0	56.0	26.0	3.0	85.0	50.0	25.0	3.0	78.0
MICASA16	91.7	58.0	26.0	4.0	88.0	52.0	26.0	4.0	82.0
MICASA18	85.0	63.0	28.0	5.0	96.0	46.0	42.0	5.0	93.0
Average	91.3	60.0	26.4	4.0	90.4	52.4	28.4	3.8	84.6
SD	4.1	2.9	0.9	0.7	4.2	4.7	7.7	0.8	5.7
CoV	4.5	4.9	3.4	14.7	4.6	9.0	27.1	18.3	6.7

Source: author.

The scores of the experimental group in the high-average proficiency subgroup of the University of Milan exceed the results of the control group in the pre-test, post-test and delayed post-test. The results follow the same pattern in the three language tests: the average score of the experimental group is higher and the standard deviation and the coefficient of variation are lower. A detailed analysis of the results of the specific skills included in the test shows a repetition of the same pattern. On the one hand the experimental group displays higher results in the average scores of CA vocabulary retention; diglossic vocabulary skills; and ability to perform diglossic code-switching. On the other hand, it has a lower standard deviation and coefficient of variation for each skill, which means that the performances of the students in this group are more uniform than those of the control group.

5.5.2 University of Genoa

Low-average proficiency subgroups

Table 14 Diglossic vocabulary in low-average control group, Genoa

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
GECA1	46.7	31.0	13.0	2.0	46.0	27.0	13.0	2.0	42.0
GECA2	43.3	33.0	15.0	1.0	49.0	30.0	14.0	1.0	45.0
GECA3	33.3	23.0	11.0	2.0	36.0	21.0	11.0	1.0	33.0
GECA4	38.3	25.0	14.0	2.0	41.0	20.0	11.0	2.0	33.0
Average	40.4	28.0	13.3	1.8	43.0	24.5	12.3	1.5	38.3
SD	5.8	4.8	1.7	0.5	5.7	4.8	1.5	0.6	6.2
CoV	14.4	17.0	12.9	28.6	13.3	19.6	12.2	38.5	16.2

Source: author.

Table 15 Diglossic vocabulary in low-average experimental group, Genoa

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
GECASA2	46.7	37.0	20.0	3.0	60.0	34.0	17.0	3.0	54.0
GECASA7	40.0	34.0	12.0	2.0	48.0	32.0	13.0	2.0	47.0
GECASA8	51.7	41.0	15.0	2.0	58.0	34.0	17.0	3.0	54.0
Average	46.1	37.3	15.7	2.3	55.3	33.3	15.7	2.7	51.7
SD	5.9	3.5	4.0	0.6	6.4	1.2	2.3	0.6	4.0
CoV	12.7	9.4	25.8	24.7	11.6	3.5	14.7	21.7	7.8

Source: author.

The scores of the experimental group in the low-average proficiency subgroup of the University of Genoa are higher than those of the control group in the pre-test, post-test and delayed post-test. The scores are higher also in each individual skill both in the post-test and in the delayed post-test. The standard deviation and coefficient of variation do not show a steady pattern and therefore it is not possible to provide an adequate interpretation of these figures.

Average proficiency subgroups

Table 16 Diglossic vocabulary in average control group, Genoa

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
GECA5	76.7	51.0	22.0	4.0	77.0	43.0	21.0	4.0	68.0
GECA6	63.3	42.0	23.0	4.0	69.0	42.0	17.0	2.0	61.0
GECA7	75.0	46.0	17.0	2.0	65.0	39.0	18.0	4.0	61.0
Average	71.7	46.3	20.7	3.3	70.3	41.3	18.7	3.3	63.3
SD	7.3	4.5	3.2	1.2	6.1	2.1	2.1	1.2	4.0
CoV	10.2	9.7	15.6	34.6	8.7	5.0	11.2	34.6	6.4

Source: author.

Table 17 Diglossic vocabulary in average experimental group, Genoa

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
GECASA3	66.7	48.0	21.0	3.0	72.0	41.0	21.0	4.0	66.0
GECASA5	80.0	51.0	23.0	4.0	78.0	47.0	23.0	3.0	73.0
GECASA6	78.3	34.0	24.0	4.0	62.0	25.0	19.0	3.0	47.0
Average	75.0	44.3	22.7	3.7	70.7	37.7	21.0	3.3	62.0
SD	7.3	9.1	1.5	0.6	8.1	11.4	2.0	0.6	13.5
CoV	9.7	20.5	6.7	15.7	11.4	30.2	9.5	17.3	21.7

Source: author.

The experimental and control groups in the average proficiency subgroup of the University of Genoa show very similar results between the post-test and the delayed post-test and there is no notably difference between them. Their scores on the pre-test show a better performance of the former instead. It is worth noticing that the control group outperforms the experimental group in CA vocabulary retention on both post-test and delayed post-test. The results for the two remaining skills, instead, are predominantly higher in the experimental group. The coefficient of variation indicates a high probability of internal

variation among the scores of the control group in the ability to perform language change diglossic code-switching, and among the scores of the experimental group in CA vocabulary retention. None of the students in the control group of the University of Genoa scored more than 80 in the pre-test and therefore it is not possible to create the above-average subgroup for this University.

5.5.3 University of Exeter

Low-average proficiency subgroups

Table 18 Diglossic vocabulary in low-average control group, Exeter

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
EXCA2	60.0	40.0	13.0	3.0	56.0	46.0	10.0	1.0	57.0
EXCA4	46.7	31.0	13.0	2.0	46.0	29.0	13.0	2.0	44.0
Average	53.3	35.5	13.0	2.5	51.0	37.5	11.5	1.5	50.5
SD	9.4	6.4	0.0	0.7	7.1	12.0	2.1	0.7	9.2
CoV	17.7	17.9	0.0	28.3	13.9	32.1	18.4	47.1	18.2

Source: author.

Table 19 Diglossic vocabulary in low-average experimental group, Exeter

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
EXCASA2	56.7	28.0	10.0	1.0	39.0	22.0	11.0	1.0	34.0
EXCASA5	60.0	53.0	19.0	4.0	76.0	46.0	22.0	3.0	71.0
Average	58.3	40.5	14.5	2.5	57.5	34.0	16.5	2.0	52.5
SD	2.4	17.7	6.4	2.1	26.2	17.0	7.8	1.4	26.2
CoV	4.0	43.6	43.9	84.9	45.5	49.9	47.1	70.7	49.8

Source: author.

Before proceeding with the analysis of the data, I would like to mention that there is a risk of inaccuracy in this subgroup. This is because there are only two students in each group and therefore the number of observations in the two data sets is extremely limited. As there is no fixed rule on the size of the number to analyse in education research, I consider these two groups valid for data analysis but I am aware that the estimates are likely to have larger variances than for larger sample sizes. I now continue with the analysis of the data. The experimental group in the low-average proficiency subgroup in Exeter shows a better performance in all language tests. However, its level of interval variation is extremely higher and less uniform than that of the control group. The latter shows a reduced degree of dispersion in the post-test and a slightly higher variation in the delayed post-test. In the experimental group, instead, there is an elevated heterogeneity in the data collected both on the post-test and the delayed post-test. It is worth noticing that the level of internal variation of the experimental group in the pre-test is actually low, which signifies that the two students performed similarly in the pre-test.

Average proficiency subgroups

Table 20 Diglossic vocabulary in average control group, Exeter

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
EXCA1	71.7	47.0	22.0	3.0	72.0	44.0	13.0	3.0	60.0
EXCA7	70.0	41.0	23.0	4.0	68.0	43.0	18.0	3.0	64.0
EXCA8	76.7	55.0	18.0	4.0	77.0	52.0	18.0	4.0	74.0
Average	72.8	47.7	21.0	3.7	72.3	46.3	16.3	3.3	66.0
SD	3.5	7.0	2.6	0.6	4.5	4.9	2.9	0.6	7.2
CoV	4.8	14.7	12.6	15.7	6.2	10.6	17.7	17.3	10.9

Source: author.

Table 21 Diglossic vocabulary in average experimental group, Exeter

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
EXCASA1	71.7	47.0	19.0	3.0	69.0	45.0	13.0	2.0	60.0
EXCASA4	75.0	51.0	17.0	3.0	71.0	49.0	18.0	4.0	71.0
Average	73.3	49.0	18.0	3.0	70.0	47.0	15.5	3.0	65.5
SD	2.4	2.8	1.4	0.0	1.4	2.8	3.5	1.4	7.8
CoV	3.2	5.8	7.9	0.0	2.0	6.0	22.8	47.1	11.9

Source: author.

The experimental group in the average proficiency subgroup of the University of Exeter registers a higher score in the pre-test but is outperformed by the control group in the post-test and delayed post-test. The control group reports higher results on diglossic vocabulary skills and on the ability to perform diglossic code-switching, but its scores in CA vocabulary retention are lower than those of the experimental group. The overall internal variation within the control group is on average higher than the experimental group, but it is steady on similar values. The internal variation of the experimental group, instead, is low and homogeneous in the post-test but it increases dramatically in the delayed post-test on both diglossic vocabulary skills and ability to perform diglossic code-switching.

High-average proficiency subgroups

Table 22 Diglossic vocabulary in high-average control group, Exeter

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
EXCA3	81.7	38.0	22.0	3.0	63.0	49.0	20.0	3.0	72.0
EXCA5	93.3	34.0	20.0	4.0	58.0	33.0	21.0	3.0	57.0
EXCA6	88.3	57.0	26.0	4.0	87.0	54.0	24.0	4.0	82.0
Average	87.8	43.0	22.7	3.7	69.3	45.3	21.7	3.3	70.3
SD	5.9	12.3	3.1	0.6	15.5	11.0	2.1	0.6	12.6
CoV	6.7	28.6	13.5	15.7	22.4	24.2	9.6	17.3	17.9

Source: author.

Table 23 Diglossic vocabulary in high-average experimental group, Exeter

Student	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
EXCASA3	83.3	43.0	20.0	2.0	65.0	47.0	13.0	2.0	62.0
EXCASA6	100.0	63.0	28.0	4.0	95.0	58.0	28.0	4.0	90.0
EXCASA7	100.0	60.0	28.0	5.0	93.0	51.0	27.0	5.0	83.0
Average	94.4	55.3	25.3	3.7	84.3	52.0	22.7	3.7	78.3
SD	9.6	10.8	4.6	1.5	16.8	5.6	8.4	1.5	14.6
CoV	10.2	19.5	18.2	41.7	19.9	10.7	37.0	41.7	18.6

Source: author.

The scores of the experimental group in the high-average proficiency subgroup of the University of Exeter, show a similar pattern to the average proficiency subgroup of the same university. The results of the control group in the pre-test, post-test and delayed post-test are lower than the results of the experimental group. However, the level of dispersion and the likelihood of internal variation of the experimental group are strikingly high and are higher than those of the control group.

5.6 Comparative analysis of results on diglossic vocabulary building

In this section I add the results of the pre-test, post-test and on the delayed post-test of the Universities of Milan, Genoa and Exeter together. I group the results of the control groups together and the scores of the experimental groups together. The groups are divided into the three proficiency groups I have outlined in the previous section. The sum of these scores is of fundamental importance because it allows me to compare the performance of control and experimental groups in a cross-university analysis. I analyse the average score and the standard deviation in each pair of control and experimental groups. I do not examine the coefficient of variation as the number of participant is too narrow to permit a reasonable estimation. We have seen in the previous section that a small sample has a considerable impact on the size of the coefficient of variation, which tends to be extraordinarily high. Since the cross-university analysis is carried out with the purpose of generalising the results and make broad assumptions on the values and distribution of the data, I believe that the calculation of the coefficient of variation in these circumstances is likely to lead to incorrect interpretations and conclusions. The data are presented in tables displaying the average scores and standard deviation for both control and experimental groups for each subgroup of language proficiency. I start with the low-average proficiency subgroup, I proceed to analysing the average proficiency subgroup and I finally evaluate the high-proficiency subgroup.

5.6.1 Low-average proficiency subgroups

Diglossic vocabulary building

Table 24 Diglossic vocabulary across low-average control groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	52.5	36.8	12.0	2.3	50.3	29.5	12.8	2.0	44.3
Genoa	40.4	28.0	13.3	1.8	43.1	24.5	12.3	1.5	38.3
Exeter	53.3	35.5	13.0	2.5	51.0	37.5	11.5	1.5	50.5
Average	48.8	33.4	12.8	2.2	48.4	30.5	12.2	1.7	44.4

Source: author.

Table 25 Diglossic vocabulary across low-average experimental groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	53.8	45.4	16.5	3.4	65.3	40.6	15.3	3.1	59.0
Genoa	46.1	37.3	15.7	2.3	55.3	33.3	15.7	2.7	51.7
Exeter	58.3	40.5	14.5	2.5	57.5	34.0	16.5	2.0	52.5
Average	52.7	41.1	15.6	2.7	59.4	36.0	15.8	2.6	54.4

Source: author.

The scores on the language tests of the low-average proficiency subgroups show a consistent and regular pattern of higher results scored by the experimental group in every language skills and in the three language tests. The ability to perform diglossic code-switching remains on approximately the same level only in the experimental language group, whereas the control group shows a higher fluctuation. Finally, results on diglossic vocabulary skills remain steady in both groups, whereas scores on CA vocabulary retention drop in both groups.

Standard deviation

Table 26 SD in diglossic vocabulary of low-average control groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	3.2	12.0	2.6	0.5	15.1	11.0	2.5	0.8	14.3
Genoa	5.8	4.8	1.7	0.5	7	4.8	1.5	0.6	6.9
Exeter	9.4	6.4	0.0	0.7	7.1	12.0	2.1	0.7	14.8
SD Average	6.2	7.7	1.4	0.6	9.7	9.3	2.0	0.7	12

Source: author.

Table 27 SD in diglossic vocabulary of low-average experimental groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	4.7	6.7	2.6	0.9	10.2	6.5	2.3	1.0	9.8
Genoa	5.9	3.5	4.0	0.6	8.1	1.2	2.3	0.6	4.1
Exeter	2.4	17.7	6.4	2.1	26.2	17.0	7.8	1.4	26.2
SD Average	4.3	9.3	4.3	1.2	14.8	8.2	4.1	1.0	13.3

Source: author.

Although the experimental group shows higher scores than the control group in the language test, it displays also a higher standard deviation. It is interesting to notice that its results are more homogeneous than those of the control group in the pre-test. However, they are noticeably higher in both the post-test and the delayed post-test. Therefore, although the group scores higher results, the students at the three universities have performed more heterogeneously than those of the control group.

5.6.2 Average proficiency subgroups

Diglossic vocabulary building

Table 28 Diglossic vocabulary across average control groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	67.5	53.0	17.5	2.8	73.3	48.3	17.5	2.8	68.5
Genoa	71.7	46.3	20.7	3.3	70.3	41.3	18.7	3.3	63.3
Exeter	72.8	47.7	21.0	3.7	72.3	46.3	16.3	3.3	66.0
Average	70.6	49.0	19.7	3.3	72.0	45.3	17.5	3.1	65.9

Source: author.

Table 29 Diglossic vocabulary across average experimental groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	71.0	47.4	20.6	3.8	71.8	41.2	20.8	3.6	65.6
Genoa	75.0	44.3	22.7	3.7	70.7	37.7	21.0	3.3	62.0
Exeter	73.3	49.0	18.0	3.0	70.0	47.0	15.5	3.0	65.5
Average	73.1	46.9	20.4	3.5	70.8	42.0	19.1	3.3	64.4

Source: author.

The scores on the pre-test of the average proficiency subgroups show that the level of proficiency of the experimental group at the beginning of the course is higher than the level of the control group. However, the latter outperforms the former in the post-test and in the delayed post-test, due to higher scores in CA vocabulary retention. The other two language skills are mastered slightly better by the experimental group instead. It is worth noticing that the difference in scores is so minimal that I do not feel confident in generalising the differences in scores between the two groups. On the contrary, I believe that in this it is

reasonable to argue that the results between the two groups are so close so as to be considered almost equivalent.

Standard deviation

Table 30 SD in diglossic vocabulary of average control groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	4.2	2.4	1.3	0.5	4.1	3.0	1.3	0.5	4.8
Genoa	7.3	4.5	3.2	1.2	8.9	2.1	2.1	1.2	5.4
Exeter	3.5	7.0	2.6	0.6	10.2	4.9	2.9	0.6	8.4
SD Average	5.0	4.7	2.4	0.7	7.8	3.3	2.1	0.7	6.1

Source: author.

Table 31 SD in diglossic vocabulary of average experimental groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	6.0	4.4	1.5	0.4	6.3	6.8	4.8	0.5	12.1
Genoa	7.3	9.1	1.5	0.6	11.2	11.4	2.0	0.6	14
Exeter	2.4	2.8	1.4	0.0	4.2	2.8	3.5	1.4	7.7
SD Average	5.2	5.4	1.5	0.3	7.2	7.0	3.4	0.8	11.2

Source: author.

The scores on standard deviation reproduce the same pattern of results than the pattern of the scores on language tests, in that there is not a substantial difference between the two groups.

5.6.3 High-average proficiency subgroups

Diglossic vocabulary building

Table 32 Diglossic vocabulary across high-average control groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	90.2	55.1	25.4	3.7	84.2	50.3	23.7	3.4	77.4
Genoa	-	-	-	-	-	-	-	-	-
Exeter	87.8	43.0	22.7	3.7	69.4	45.3	21.7	3.3	70.3
Average	87.8	32.7	24.1	3.7	76.7	47.8	22.7	3.3	73.8 5

Source: author.

Table 33 Diglossic vocabulary across high-average experimental groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	91.3	60.0	26.4	4.0	90.4	52.4	28.4	3.8	84.6
Genoa	88.3	62.0	28.0	5.0	95.0	58.5	26.5	4.5	89.5
Exeter	94.4	55.3	25.3	3.7	84.3	52.0	22.7	3.7	78.4
Average	91.4	59.1	26.6	4.2	89.9	54.3	25.9	4.0	84.1

Source: author.

The comparison between the scores in this subgroup is not completely accurate as there are no students in the control group of the University of Genoa within the high-average proficiency subgroup. However, on the basis of the figures of the other two Universities it is reasonable to claim that the experimental group is likely to outperform the control group in all the language skills but it is not possible to determine if the difference between the two groups is significant or whether it is relatively small.

Standard deviation

Table 34 SD in diglossic vocabulary of high-average control groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	4.7	6.3	3.5	0.8	10.6	5.9	3.3	0.8	10
Genoa	-	-	-	-	-	-	-	-	-
Exeter	5.9	12.3	3.1	0.6	16	11.0	2.1	0.6	13.7
SD Average	5.3	9.3	3.3	0.7	13.3	8.4	2.7	0.7	11.8

Source: author.

Table 35 SD in diglossic vocabulary of high-average experimental groups

University	Pre Test	Post Test	Post Test	Post Test	Post Test	Del. Post Test	Del. Post Test	Del. Post Test	Del. Post Test
	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total	CA Reten.	CS Skills	Diglos. Co-Sw.	Total
Milan	4.1	2.9	0.9	0.7	4.5	4.7	7.7	0.8	13.2
Genoa	7.1	2.8	1.4	0.0	4.2	2.1	0.7	0.7	3.5
Exeter	9.6	10.8	4.6	1.5	16.8	5.6	8.4	1.5	15.2
SD Average	6.9	5.5	2.3	0.7	8.5	4.1	5.6	1.0	10.7

Source: author.

The analysis of the scores in this subgroup faces the same challenge as the previous comparison, since there are no students in the control group of the University of Genoa within the high-average proficiency subgroup. In my view, the remaining data do not show patterns that are sufficiently stable so as to draw generalizable conclusions.

5.7 Conclusion

This chapter analysed the development of diglossic vocabulary within and across the three participant Universities. An analysis of the statistical formulae used to measure the results of the language tests was given, with a detailed

description of the features of each formula and the nature of the information gained through their application to the language scores. A detailed explanation of the language tests (pre-test, post-test, and delayed post-test) was also provided. More specifically, I identified the way in which each exercise is designed to prompt performances in the specific language skills measured: CA vocabulary retention; diglossic vocabulary skills; and ability to perform diglossic code-switching. I thereafter reported the results gathered within each participants University, and I finally compared the scores of the control and experimental groups across the three Universities.

Empirically, this chapter shows that the experimental groups in all three Universities outperform the control groups in the total scores of the language tests. However, a more detailed analysis that divides the two groups into three subgroups based on their proficiency levels in SA at the start of the course (low-average, average and high-average proficiency) shows that this pattern does not apply to every subgroup and that, especially among the average-proficiency groups, there is not a striking difference between the results of the two groups.

The chapter also illustrates that the standard deviation and the coefficient of variation show us that the internal dispersion of the scores within the three universities is not as homogeneous as their scores in the language tests. The values of the language tests of the experimental groups at the universities of Milan and Genoa are characterised by more homogeneity than the correspondent values scored by the control groups. At the University of Exeter, instead, the results of the experimental group are more dispersed than the results of the control group. A detailed analysis of the subgroup categories outlines that the experimental groups tend to show a higher degree of internal variation in comparison with the control group. It is reasonable to think that this is due to a more divergent response of the students to the nature of the instruction received and the cognitive engagement requested by the constant comparison of the two language varieties. This supposition is reinforced by the comparison of the results of language scores and the comparison of the standard variation values among the three participant Universities.

The scores from the language tests of the experimental groups across the three Universities are higher than the scores of the control groups. However, the standard deviation and coefficient of variation are also higher in most language

skills for the experimental groups. This applies always to the development of diglossic skills and to the ability to perform diglossic code-switching. It is not a regular pattern, in contrast, in CA vocabulary retention. This means that the performances of the experimental groups on skills completely based on diglossia and on the use of the two varieties are more heterogeneous and varied than those of the control groups. The Wilcoxon-Mann-Whitney test showed that this difference is not statistically significant, and therefore it is not possible to undoubtedly state that it is due to the differences among the language instruction. However, I cannot generalise the result of the test because the number of students participating from two universities is not sufficient to run the test, and I therefore only run it for the results of one university. Although the results are not statistically different, there is a pattern among the results that is almost constant: the control group does not perform as well as the experimental group in the overall language tests, but its results are more compact and homogeneous than those of the former. This has numerous implications on the integration of the two methods of instruction within the classroom. On the one hand, the FFI based on the comparison of CA and SA forms seems to better prepare the students to perform diglossic skills. However, it does not seem to have the same positive impact on all the students of the group, whose development of diglossic skills show different outcomes. On the other hand, the FFI based on the teaching of only one variety does not seem to lead to results as high as those of the other form of instruction used in this research, but they are more homogeneously spread among the students.

In order to better understand these results and the perspective of the students on their performances and the instruction received, I conduct an in-depth analysis in chapter eight, where I combine the data on language tests with the students' perceptions and motivation towards variation in Arabic and towards its study.

6 CHAPTER 6 – Language awareness

6.1 Introduction

This chapter presents the second empirical step of this thesis and answers the first sub-question: is focus-on-form instruction more effective in promoting diglossic language awareness when (a) it links forms in Standard and Colloquial Arabic and when (b) it focuses only on one variety? This thesis argues that language awareness is a combination of the ability to perform language change diglossic code-switching correctly; to verbalise the grammatical rules applied to realise such change; and the presence of meta-awareness. Meta-awareness, in turn, is the consciousness, on the part of the speakers, of the cognitive experience they undergo when they perform diglossic code-switching. The presence of meta-awareness changes on an individual basis and speakers can be aware or unaware of it, i.e. of their cognitive involvement during their performance of diglossic code-switching. The level of awareness is also subject to each speaker and it is defined on an individual basis. This thesis also argues that language awareness is tightly interwoven with knowledge of diglossic vocabulary and code-switching skills and that it adds to them an element of cognitive consciousness. As seen throughout this thesis, these three elements combined - language awareness, diglossic vocabulary knowledge and code-switching skills - allow the speakers to:

- i. use two varieties of Arabic simultaneously by means of their knowledge of diglossic vocabulary;
- ii. switch between language varieties in relation to pragmatics; subjective individual choices; and linguistic knowledge, by means of their diglossic code-switching skills;
- iii. realise correct changes between language varieties and be aware of their cognitive engagement during diglossic code-switching, by means of language awareness.

I have analysed the development of diglossic vocabulary knowledge and code-switching skills in chapter five. This chapter, instead, sets out to determine the

impact that the two different forms of instruction employed in this research have on the development of language awareness. The focus of instruction in the control group is only on CA, and the method of instruction emphasises the accuracy of CA language forms in communicative classrooms. Thus, the students are not trained to switch from SA forms and vocabulary to their equivalents in CA. They do not compare phonological and morphological change between the two varieties, and the exercise designed for their language course are based only on CA. On the contrary, the focus of instruction in the experimental group is on both language varieties, and its aim is to train the students to be cognitively engaged in the switching processes between the two varieties. FFI exposes the students to a regular and consistent comparison between correspondent vocabulary and forms in SA and CA, and it trains them to translate between one variety into the other; to use the two varieties simultaneously; and to identify patterns of switch between them.

This research intends to answer the question of whether one form of instruction is more effective for the development of language awareness than another. On the one hand, my assumption would be that the experimental form of instruction raises language awareness more effectively than the instruction used in the control groups. This is because it explicitly trains the participants to be cognitively engaged in diglossic code-switching. On the other, I am conscious that the experimental instruction could prove to be highly demanding for students with low and average proficiency levels of SA. The constant reference that this instruction makes to SA forms and vocabulary could create frustration and confusion, and this in turn could prove to be counterproductive to the development of language awareness. I therefore prefer not to have a defined assumption in favour of one of the two methods of instructions, and I answer the above-mentioned research question by means of the data analysis provided in this chapter.

The evaluation of language awareness takes place in three different stages. I identified the first level, i.e. the ability to perform language change diglossic code-switching correctly, as part of the development of diglossic vocabulary, as

explained in the previous chapter. I therefore already explained its measurement in section 5.3.1. The remaining two levels of language awareness are measured using information provided by the students and collected via retrospective think-aloud protocols. This is explained in section 6.1.

This chapter is structured as follows. The aim of the first section is to explain how the data are collected, and to identify how their assessment is carried out. I explain the language exercise and the questions I designed to gather data on language awareness. I also explain the features of the scoring system and how retrospective think-aloud protocols are converted into quantitative figures. The second section compares scores on language awareness across the universities studied.

This section is divided into three parts. First, I analyse the results within the three universities by adding the scores on language awareness to the results on the language tests analysed in the previous chapter. I subsequently focus on the results scored only on language awareness within the participant universities. I report the results of the groups on the pre-test for comparison as well. In the second part, I compare the results on the language tests and language awareness scored by the subgroups identified in the previous chapter (low-average, average and high-average proficiency groups). The third, and last, part focuses on comparing the average scores and standard deviation of language tests and language awareness across the participant universities

6.2 Data collection on language awareness

This section explains how data on the three levels of language awareness are collected and measured. First, I focus on the ability to perform language change diglossic code-switching. As explained in the previous chapter, this consists in linguistically changing vocabulary and form from one variety into another. It is therefore assessed as part of diglossic vocabulary-building knowledge. I assess it through one of the exercises included in the post-test and in the delayed post-test, which is designed specifically for this purpose. The exercise consists in providing the student with a text written in SA and ask them to retell it in CA.

There are five lexical items in the text that the students are expected to change: (i) والد in SA become أبو in CA (father); (ii) والدة in SA becomes أم in CA (mother); (iii) تدرس in SA becomes بتدرس in CA (she studies); (iv) تسافر in SA becomes بتسافر in CA (she travels); (v) تحب in SA becomes بتحب in CA (she loves). This exercise is scored on a scale of 0 to 5.

Second, the students are asked to verbalise the grammatical rules they have applied to realise the changes. The rules are as follows. First, والد and والدة are not used in CA and two different lexical items are used instead. Second, the three verbs تدرس - تسافر - تحب are subject to both linguistic and phonological changes. They all take the prefix ب in the CA present tense before the conjugated verb, but they undergo different phonological changes. بتدرس /btədros/ undergoes a change that entails placing the helping vowel /ə/ after the consonant ت; in بتحب /bətħebb/ and بتسافر /bət̪sa:fər/ the helping vowel is instead placed after the consonant ب. Seven points are given for the verbalisation of these rules. Two points are given for the vocabulary change in والد and والدة; three are given to the correct explanation of the use of prefixes for the present tense in CA. Finally, two are given to the correct use of the helping vowel /ə/.

Third, in retrospective moments of reflection, the participants describe the mental processes experienced in performing the code-switching tasks during the language test. They are also specifically asked to describe their experience of switching between language varieties in the exercise described above. The description provided by the students about their experience of performing code-switching is matched with their actual performance to establish whether they

have developed a correct awareness of their code-switching processes and abilities. This is discussed in the last chapter of this thesis, in which I combine quantitative and qualitative data. The description by the students of their experience of switching between language varieties and of realising the switching between varieties mentioned above is instead turned into quantitative figures for measurement. Seven points are attributed in total to the verbalisation of grammar rules, and the equivalent is attributed to the meta-awareness of the realisation of these changes. The scores for language awareness are in total five (ability to perform language change diglossic code-switching); seven (verbalisation of rules to realise the change); and seven (meta-awareness). The scores of the language tests and language awareness are set as follows: 65 points are attributed to CA vocabulary retention; 30 points are attributed to CA diglossic vocabulary skills; five points are attributed to the first level of diglossic language awareness; seven points are attributed to the awareness of morphosyntactical rule applied to perform language change diglossic code-switching; and finally seven points are attributed to the presence of meta-awareness. The scores of three language skills that compose the post-test and the delayed post-test, sum up to 100. The addition of the points attributed to the second and third level of language awareness raises the sum to 114. In order to provide a percentage value of the scores of the language tests together with diglossic language awareness, I calculate the percentage corresponding value of the sums of the scores for each student. The three levels of language awareness added to each other provide a maximum score of 19 (5+7+7). The percentage value of the scores on language awareness is shown by calculating the percentage corresponding value of numbers ranging from 0 to 19. For this reason, there are only 19 different values, although expressed in percentage, in the language awareness and delayed language awareness columns. These last two columns show scores on language awareness collected during and after the post-test and the delayed post-test.

Finally, data are measured using the same formulae employed in the previous chapter: average score; Wilcoxon-Mann-Whitney test; standard deviation; coefficient of standard deviation; and Wilcoxon 2-sample Rank Sum. It is

important to highlight that the p-values calculated through the Wilcoxon-Mann-Whitney test always fail to reject the null hypothesis in this research and therefore do not provide positive evidence that the difference between the scores of the control and experimental groups are due to the differences between the two methods of language instructions used. There are nonetheless recurrent and remarkable patterns among the scores collected that are worth paying attention to. I identify them within each section of data analysis.

6.3 Results on language awareness of participant universities

Scores of students of the University of Milan

Table 36 Language awareness in control group, Milan

Student	Pre-Test	Post-Test + Language Awareness	Delayed Post-Test + Language Awareness	Language Awareness	Delayed Language Awareness
MICA1	90.0	86.0	77.2	89.5	84.2
MICA2	88.3	89.9	79.8	68.4	57.9
MICA3	93.3	86.8	80.7	84.2	78.9
MICA4	68.3	77.2	72.8	68.4	68.4
MICA5	90.0	78.1	71.9	68.4	57.9
MICA6	50.0	43.0	37.7	47.4	52.6
MICA7	53.3	41.2	34.2	52.6	47.4
MICA8	50.0	56.1	50.0	31.6	26.3
MICA9	88.3	68.4	60.5	57.9	52.6
MICA10	56.7	58.8	52.6	36.8	31.6
MICA11	71.7	70.2	64.9	47.4	42.1
MICA12	68.3	70.2	62.3	52.6	47.4
MICA13	98.3	94.7	88.6	78.9	78.9
MICA14	83.3	72.8	70.2	47.4	36.8
MICA15	61.7	66.7	66.7	52.6	57.9
Average	74.1	70.7	64.7	58.9	54.7
SD	17.1	15.9	15.6	16.9	17.3
CoV	23.0	22.5	24.1	28.7	31.6

Source: author.

Table 37 Language awareness in experimental group, Milan

Student	Pre-Test	Post-Test + Language Awareness	Delayed Post- Test + Language Awareness	Language Awareness	Delayed Language Awareness
MICASA1	51.7	55.3	49.1	52.6	52.6
MICASA2	56.7	64.9	59.6	42.1	42.1
MICASA3	58.3	61.4	57.0	52.6	52.6
MICASA4	95.0	91.2	83.3	89.5	78.9
MICASA5	95.0	93.0	87.7	94.7	89.5
MICASA6	90.0	82.5	76.3	63.2	63.2
MICASA7	46.7	66.7	57.9	52.6	42.1
MICASA8	65.0	77.2	71.1	68.4	57.9
MICASA9	58.3	79.8	72.8	73.7	63.2
MICASA10	78.3	73.7	69.3	68.4	68.4
MICASA11	71.7	62.3	55.3	57.9	52.6
MICASA12	65.0	69.3	63.2	63.2	57.9
MICASA13	50.0	55.3	50.9	57.9	52.6
MICASA14	58.3	64.0	58.8	63.2	57.9
MICASA15	50.0	62.3	55.3	57.9	52.6
MICASA16	91.7	87.7	81.6	84.2	78.9
MICASA17	75.0	71.1	63.2	73.7	63.2
MICASA18	85.0	96.5	93.0	100.0	94.7
Average	69.0	73.0	67.0	67.5	62.3
SD	16.7	13.0	13.1	15.9	14.8
CoV	24.2	17.7	19.6	23.5	23.8

Source: author.

The control group of the University of Milan performed better than the experimental group in the pre-test by 5.1 points. However, the experimental group shows higher results in the post-test and in the delayed post-test. Moreover, the results of the experimental group are more homogeneous than those of the control group. Both groups show a constant level of dispersion between their results in the post-test and in the delayed post-test. It is interesting to notice that, as it was already the case for the results on diglossic vocabulary knowledge, the coefficient of variation in the control group remains almost constant across the three tests (the results are respectively 23.0, 22.5 and 24.1). The value of the coefficient for the experimental group, instead, lowers significantly in the post-test. Its value increases in the delayed post-test, but it remains lower than the value of the pre-test. The scores on language

awareness show a remarkable decrease of internal dispersion between the post-test and the delayed post-test for the experimental group, and an opposite trend for the control group. However, the experimental group outperforms the control group in both tests. I run the Wilcoxon-Mann-Whitney test to assess the difference in scores of the post-test and of the delayed post-test. The difference of the results scored in the post-test by the two groups is not statistically significant. The p-value is 0.90 and therefore it is higher than the value set at 0.05. The difference of the results scored in the delayed post-test is also not statistically significant, as the p-value is higher than 0.05 (the p-value is 0.97). The difference between the results of the two groups is therefore not attributable to the instruction received in the language course. However, as is the case with diglossic vocabulary-building analysed in the previous chapter, there is a pattern of scores: the experimental group scored higher than the control group in the post-test and in the delayed post-test, and in both tests the values of the former are more homogeneously spread. As per the Wilcoxon 2-sample Rank Sum, the p-value for the control group is 0.00096 and the p-value of the experimental group is 0.0002. They are both lower than 0.05 and therefore the differences between post-test and delayed post-test of both groups are statistically significant. The value of the experimental group is lower and this means that the difference is more significant than that of the control group. This replicates the findings gathered on diglossic vocabulary-building at the University of Milan. In both groups the results are lower between the post-test and the delayed post-test, and we can therefore claim that the results between the post-test and the delayed post-test are more significantly lower in the experimental group than in the control group.

Scores of students of the University of Genoa

Table 37 Language awareness in control group, Genoa

Student	Pre-Test	Post-Test + Language Awareness	Delayed Post-Test + Language Awareness	Language Awareness	Delayed Language Awareness
GECA1	46.7	45.6	43.0	42.1	47.4
GECA2	43.3	48.2	44.7	36.8	36.8
GECA3	33.3	36.0	32.5	36.8	26.3
GECA4	38.3	42.1	34.2	47.4	42.1
GECA5	76.7	76.3	67.5	73.7	68.4
GECA6	63.3	68.4	60.5	68.4	52.6
GECA7	75.0	64.0	58.8	52.6	52.6
Average	53.8	54.4	48.7	51.1	46.6
SD	17.7	15.1	13.7	14.8	13.4
CoV	32.9	27.8	28.0	28.9	28.7

Source: author.

Table 38 Language awareness in experimental group, Genoa

Student	Pre-Test	Post-Test + Language Awareness	Delayed Post-Test + Language Awareness	Language Awareness	Delayed Language Awareness
GECASA1	93.3	93.0	86.8	100.0	84.2
GECASA2	46.7	57.0	51.8	42.1	42.1
GECASA3	66.7	70.2	64.9	57.9	63.2
GECASA4	83.3	98.2	93.0	100.0	100.0
GECASA5	80.0	76.3	71.1	68.4	57.9
GECASA6	78.3	58.8	46.5	47.4	47.4
GECASA7	40.0	47.4	46.5	42.1	42.1
GECASA8	51.7	57.0	53.5	47.4	52.6
Average	67.5	69.7	64.3	63.2	61.2
SD	19.4	18.3	18.1	24.4	20.9
CoV	28.7	26.2	28.1	38.6	34.1

Source: author.

The data collected at the University of Genoa show that the experimental group outperformed the control group in every language test and in the development

of language awareness. The standard deviation of the experimental group, however, is consistently higher than that of the control group. Similarly, the dispersion of values of the experimental group is steadily higher as well. It is worth noticing that both groups decrease their internal dispersion over time.

The comparison between the coefficient of variation in the two data sets is remarkable: the likelihood of internal variation is higher for the control group in the pre-test and in the post-test but it equals that of the experimental group in the delayed post-test. As per language awareness, the likelihood of the control group to have disperse values is significantly lower than the likelihood of the experimental group both in the post-test and in the delayed post-test. The difference of the results scored in the post-test by the two groups is not statistically significant, as the p-value calculated through the Wilcoxon-Mann-Whitney test is 0.134, which is higher than the p-value set at 0.05. Likewise, the difference of the results scored in the delayed post-test is also not statistically significant: the p-value of the comparison between the two test is 0.93, which is higher than 0.05. It is therefore not possible to claim that the figures gathered on diglossic vocabulary-building and language awareness between the experimental and control groups are influenced by the instruction received in the experimental language course. However, in my view, it is important to highlight the significant patterns of results in the figures: the results of the experimental group are higher than the control group, but the individual scores are more spread out from the average than those of the control group. Finally, as we have already seen in the previous chapter, the size of the sample is not large enough to run the Wilcoxon 2-sample Rank Sum test and it is therefore not possible to compare the results of the post-test and of the delayed post-test.

Scores of students of the University of Exeter

Table 39 Language awareness in control group, Exeter

Student	Pre-Test	Post-Test + Language Awaren.	Delayed Post- Test + Language Awaren.	Language Awaren.	Delayed Language Awaren.
EXCA1	71.7	70.2	58.8	57.9	52.6
EXCA2	60.0	54.4	55.3	47.4	36.8
EXCA3	81.7	60.5	66.7	47.4	36.8
EXCA4	46.7	45.6	44.7	42.1	47.4
EXCA5	93.3	57.0	56.1	57.9	52.6
EXCA6	70.0	66.7	63.2	63.2	57.9
EXCA7	88.3	86.0	81.6	78.9	78.9
EXCA8	76.7	78.1	71.9	84.2	63.2
Average	73.5	64.8	62.3	59.9	53.3
SD	15.2	13.2	11.3	15.1	13.9
CoV	20.6	20.3	18.1	25.3	26.1

Source: author.

Table 40 Language awareness in experimental group, Exeter

Student	Pre-Test	Post-Test + Language Awaren.	Delayed Post- Test + Language Awaren.	Language Awaren.	Delayed Language Awaren.
EXCASA1	71.7	66.7	57.9	52.6	42.1
EXCASA2	56.7	37.7	33.3	26.3	26.3
EXCASA3	83.3	67.5	61.4	73.7	52.6
EXCASA4	75.0	71.9	70.2	73.7	68.4
EXCASA5	60.0	75.4	69.3	73.7	57.9
EXCASA6	100.0	95.6	91.2	94.7	94.7
EXCASA7	100.0	93.9	84.2	100.0	94.7
Average	78.1	72.7	66.8	70.7	62.4
SD	17.4	19.4	18.9	25.0	25.7
CoV	22.3	26.7	28.3	35.4	41.2

Source: author.

The data collected at the University of Exeter show a similarity with the data gathered in Genoa. The experimental group performed better than the control

group in all language tests and in the development of language awareness. In the latter, the scores on language awareness are remarkably higher. The scores of the language tests decrease progressively in both groups from the pre-test to the delayed post-test and scores on language awareness follow a similar pattern. However, the two groups show a contrasting pattern of distribution of values on internal dispersion: the control group decreases its dispersion steadily over-time, whereas the experimental group increases it. This translates into a pattern that sees the experimental group's performance becoming progressively less homogeneous, whereas the students of the control group score more homogeneous results in the long term. Similarly, the likelihood of internal dispersion of values shows a steady increase in the experimental group and a less regular behaviour: it decreases in the results of the language tests combined with language awareness, but it increases in language awareness' scores.

The Wilcoxon-Mann-Whitney test shows the same results collected for the other two universities, both for the post-test and for the delayed post-test: the difference of the results scored by the two groups is not due to the impact of the experimental language course and of the language treatment. The p-value of the post-tests is 0.291 and therefore it is higher than the value set at 0.05. The difference of the results scored in the delayed post-test, instead, shows a p-value higher than 0.05 (p-value 0.453). It is nonetheless possible to identify a pattern among the scores. As is often the case in the sets of data analysed so far, the experimental group outperforms the control group in all the language tests. The level of internal distribution of the experimental group's scores is, however, increasingly more heterogeneous and the values are more and more dispersed. The control group shows a contrasting pattern of steadier distribution values over time. Similarly, the likelihood of internal variation lowers in the control group in the long-term. Finally, and similarly to the results from the University of Genoa, the size of the sample is not large enough to calculate the Wilcoxon 2-sample Rank Sum test.

6.4 Results of subgroups on language awareness

This section replicates the division of the participants within the three universities into subgroups based on their average scores in the pre-test. The three subgroups are identical to those identified in the previous chapter: low-average proficiency (scores from 33 to 60); average proficiency (scores from 61 to 80); and high-average proficiency (scores from 81 to 100). This is to analyse the results of on the language tests combined with language awareness, and on language awareness alone, in light of the actual proficiency of the participants in SA at the beginning of the course.

The scores are displayed in tables as follows. The first column reports the results of the pre-test. The remaining columns are organised in three sections. They show the results of the post-test; the delayed post-test; and on language awareness. Every section is composed of two columns. On the left hand-side of the sections displaying the results on the post-test and delayed post-test there are the scores on the language tests. On the right hand-side I add the scores on language awareness to the language tests. The last section shows scores on language awareness and delayed-language awareness.

6.4.1 University of Milan

Low-average proficiency subgroups

Table 41 Language awareness in low-average control group, Milan

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
MICA6	50.0	42.0	43.0	35.0	37.7	47.4	52.6
MICA7	53.3	40.0	41.2	33.0	34.2	52.6	47.4
MICA8	50.0	60.0	56.1	53.0	50.0	31.6	26.3
MICA10	56.7	62.0	58.8	56.0	52.6	36.8	31.6
Average	52.5	51.0	49.8	44.3	43.6	42.1	39.5
SD	3.2	11.6	9.0	11.9	9.0	9.6	12.5
CoV	6.1	22.8	18.0	27.0	20.7	22.8	31.7

Source: author.

Table 42 Language awareness in low-average experimental group, Milan

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
MICASA1	51.7	56.0	55.3	49.0	49.1	52.6	52.6
MICASA2	56.7	68.0	64.9	62.0	59.6	42.1	42.1
MICASA3	58.3	63.0	61.4	58.0	57.0	52.6	52.6
MICASA7	46.7	69.0	66.7	60.0	57.9	52.6	42.1
MICASA9	58.3	81.0	79.8	75.0	72.8	73.7	63.2
MICASA13	50.0	56.0	55.3	51.0	50.9	57.9	52.6
MICASA14	58.3	66.0	64.0	61.0	58.8	63.2	57.9
MICASA15	50.0	63.0	62.3	56.0	55.3	57.9	52.6
Average	53.8	65.3	63.7	59.0	57.7	56.6	52.0
SD	4.7	8.0	7.7	8.0	7.2	9.2	7.1
CoV	8.7	12.3	12.2	13.5	12.4	16.3	13.7

Source: author.

The experimental low-average proficiency group shows a better performance in every data set. Although the results between the two groups are not greatly

different in the pre-test, and the control group is more homogeneous than the experimental group, the situation changes drastically in the other sets of data. The control group shows not only significantly lower scores, but its level of internal dispersion, and above all its likelihood of internal dispersion, are noticeably higher than those of the experimental group. Interestingly, the internal dispersion of the control group within language awareness scores increases in the long term whereas the experimental group becomes more homogeneous.

Average proficiency subgroups

Table 43 Language awareness in average control group, Milan

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
MICA4	68.3	78.0	77.2	73.0	72.8	68.4	68.4
MICA11	71.7	74.0	70.2	68.0	64.9	47.4	42.1
MICA12	68.3	73.0	70.2	65.0	62.3	52.6	47.4
MICA15	61.7	68.0	66.7	68.0	66.7	52.6	57.9
Average	67.5	73.3	71.1	68.5	66.7	55.3	53.9
SD	4.2	4.1	4.4	3.3	4.5	9.1	11.7
CoV	6.2	5.6	6.2	4.8	6.7	16.5	21.6

Source: author.

Table 44 Language awareness in average experimental group, Milan

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
MICASA8	65.0	79.0	77.2	73.0	71.1	68.4	57.9
MICASA10	78.3	75.0	73.7	70.0	69.3	68.4	68.4
MICASA11	71.7	63.0	62.3	56.0	55.3	57.9	52.6
MICASA12	65.0	71.0	69.3	65.0	63.2	63.2	57.9
MICASA17	75	71	71.1	64	63.2	73.7	63.2
Average	71.0	71.8	70.7	65.6	64.4	66.3	60.0
SD	6.0	5.9	5.6	6.5	6.2	6.0	6.0
CoV	8.4	8.3	7.9	9.9	9.7	9.0	10.0

Source: author.

The scores of the average groups in Milan show a better performance of the experimental group in the pre-test and in the scores on language awareness. The control group, instead, shows better results in the post-test and in the delayed post-test and it outperforms the experimental group in the scores on language tests combined with language awareness. Moreover, in these data sets the control group shows a lower degree of internal dispersion and a lower likelihood of internal variation. The scores on language awareness, instead, see not only a better performance of the experimental group in the average results, but also in standard deviation and coefficient of variation. If we compare this results with the analysis of the data on diglossic vocabulary development (Tables 10 and 11), it is reasonable to argue that the control group outperforms the experimental group greatly in CA vocabulary retention and the gap between the two groups' scores is not overcome by the experimental groups' better scores on language awareness and code-switching skills.

High-average proficiency subgroups

Table 45 Language awareness in high-average control group, Milan

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
MICA1	90.0	84.0	86.0	75.0	77.2	89.5	84.2
MICA2	88.3	93.5	89.9	83.0	79.8	68.4	57.9
MICA3	93.3	88.0	86.8	82.0	80.7	84.2	78.9
MICA5	90.0	80.0	78.1	74.0	71.9	68.4	57.9
MICA9	88.3	70.0	68.4	62.0	60.5	57.9	52.6
MICA13	98.3	97.0	94.7	90.0	88.6	78.9	78.9
MICA14	83.3	77.0	72.8	76.0	70.2	47.4	36.8
Average	90.2	84.2	82.4	77.4	75.6	70.7	63.9
SD	4.7	9.5	9.6	8.8	9.0	14.8	17.3
CoV	5.2	11.2	11.6	11.4	11.9	21.0	27.1

Source: author.

Table 46 Language awareness in high-average experimental group, Milan

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
MICASA4	95.0	91.0	91.2	83.0	83.3	89.5	78.9
MICASA5	95.0	92.0	93.0	87.0	87.7	94.7	89.5
MICASA6	90.0	85.0	82.5	78.0	76.3	63.2	63.2
MICASA16	91.7	88.0	87.7	82.0	81.6	84.2	78.9
MICASA18	85.0	96.0	96.5	93.0	93.0	100.0	94.7
Average	91.3	90.4	90.2	84.6	84.4	86.3	81.1
SD	4.1	4.2	5.4	5.7	6.3	14.2	12.1
CoV	4.5	4.6	5.9	6.7	7.5	16.5	14.9

Source: author.

The high-average proficiency experimental group outperforms the control group in every data set almost mirroring the situation of the low-average proficiency groups. The control group shows higher levels of internal dispersion, and a

strikingly higher likelihood of internal dispersion. Above all, the scores on the delayed language awareness drop significantly in comparison with the first test on language awareness.

6.4.2 University of Genoa

Low-average proficiency subgroups

Table 47 Language awareness in low-average control group, Genoa

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
GECA1	46.7	46.0	45.6	42.0	43.0	42.1	47.4
GECA2	43.3	49.0	48.2	45.0	44.7	36.8	36.8
GECA3	33.3	36.0	36.0	33.0	32.5	36.8	26.3
GECA4	38.3	41.0	42.1	33.0	34.2	47.4	42.1
Average	40.4	43.0	43.0	38.3	38.6	40.8	38.2
SD	5.8	5.7	5.3	6.2	6.2	5.0	9.0
CoV	14.4	13.3	12.4	16.2	16.0	12.4	23.6

Source: author.

Table 48 Language awareness in low-average experimental group, Genoa

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
GECASA2	46.7	60.0	57.0	54.0	51.8	42.1	42.1
GECASA7	40.0	48.0	47.4	47.0	46.5	42.1	42.1
GECASA8	51.7	58.0	57.0	54.0	53.5	47.4	52.6
Average	46.1	55.3	53.8	51.7	50.6	43.9	45.6
SD	5.9	6.4	5.6	4.0	3.7	3.0	6.1
CoV	12.7	11.6	10.4	7.8	7.2	6.9	13.3

Source: author.

As we have seen for the low-average proficiency group of the University of Milan, the experimental group shows higher results than the control groups. This is true for every language test and for language awareness as well. The control group seems to record a greater level of internal variation in the first tests (pre-test; post-test; and post-test combined with language awareness), but these results change in the long term (delayed post-test) and in language awareness scores. The likelihood of internal variation for the control group is instead always higher than the probability of variation within the experimental group, within all data sets. It and the control group is more homogeneous than the experimental group, the situation changes drastically in the other sets of data. The control group shows not only significantly lower scores, but its level of internal dispersion, and above all its likelihood of internal dispersion, are noticeably higher than those of the experimental group.

Average proficiency subgroups

Table 49 Language awareness in average control group, Genoa

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
GECA5	76.7	77.0	76.3	68.0	67.5	57.9	63.2
GECA6	63.3	69.0	68.4	61.0	60.5	68.4	57.9
GECA7	75.0	65.0	64.0	61.0	58.8	47.4	47.4
Average	71.7	70.3	69.6	63.3	62.3	57.9	56.1
SD	7.3	6.1	6.2	4.0	4.6	10.5	8.0
CoV	10.2	8.7	8.9	6.4	7.5	18.2	14.3

Source: author.

Table 50 Language awareness in average experimental group, Genoa

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
GECASA2	66.7	72.0	70.2	66.0	64.9	57.9	63.2
GECASA7	80.0	78.0	76.3	73.0	71.1	68.4	57.9
GECASA8	78.3	62.0	58.8	47.0	46.5	47.4	47.4
Average	75.0	70.7	68.4	62.0	60.8	57.9	56.1
SD	7.3	8.1	8.9	13.5	12.8	10.5	8.0
CoV	10.2	8.7	8.9	6.4	7.5	18.2	14.3

Source: author.

The comparison between the average proficiency groups in Genoa does not show a relevant and stable pattern. The experimental group performs better in the pre-test, and its scores are slightly higher in the post-test. The control group display higher scores in the delayed post-test, and in the scores that combine the delayed post-test with language awareness. The internal dispersion of the experimental group is higher in the experimental group. The coefficient of variation is identical in the two group. As per language awareness, instead, the scores between the two groups are identical, which is likely to be an error in data gathering, and I therefore cannot analyse the results.

6.4.3 University of Exeter

Low-average proficiency subgroups

Table 51 Language awareness in low-average control group, Exeter

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
EXCA2	60.0	56.0	54.4	57.0	55.3	47.4	36.8
EXCA4	46.7	46.0	45.6	44.0	44.7	42.1	47.4
Average	53.3	51.0	50.0	50.5	50.0	44.7	42.1
SD	9.4	7.1	6.2	9.2	7.4	3.7	7.4
CoV	17.7	13.9	12.4	18.2	14.9	8.3	17.7

Source: author.

Table 52 Language awareness in low-average experimental group, Exeter

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
EXCASA2	56.7	39.0	37.7	34.0	33.3	26.3	26.3
EXCASA5	60.0	76.0	75.4	71.0	69.3	73.7	57.9
Average	58.3	57.5	56.6	52.5	51.3	50.0	42.1
SD	2.4	26.2	26.7	26.2	25.4	33.5	22.3
CoV	4.0	45.5	47.1	49.8	49.6	67.0	53.0

Source: author.

Within the low-average proficiency group, the experimental group shows a better performance in every data set. However, it also shows a strikingly high level of internal dispersion and likelihood of internal variation throughout the data sets. As there are only two students in each group, the results of on the standard variation and coefficient of variation are likely to be extremely high (for example, the coefficient of variation is almost 50% in the post-test and delayed

post-test, as well as in the results of the two language tests combined with language awareness). However, it is worth noting that the control group is also composed of two members only and do not show extraordinary high levels of variation.

Average proficiency subgroups

Table 53 Language awareness in average control group, Exeter

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
EXCA1	71.7	72.0	70.2	60.0	58.8	57.9	52.6
EXCA7	88.3	87.0	86.0	82.0	81.6	78.9	78.9
EXCA8	76.7	77.0	78.1	74.0	71.9	84.2	63.2
Average	78.9	78.7	78.1	72.0	70.8	73.7	64.9
SD	8.6	7.6	7.9	11.1	11.4	13.9	13.2
CoV	10.8	9.7	10.1	15.5	16.2	18.9	20.4

Source: author.

Table 54 Language awareness in average experimental group, Exeter

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
EXCASA1	71.7	69.0	66.7	60.0	57.9	52.6	42.1
EXCASA4	75.0	71.0	71.9	71.0	70.2	73.7	68.4
Average	73.3	70.0	69.3	65.5	64.0	63.2	55.3
SD	2.4	1.4	3.7	7.8	8.7	14.9	18.6
CoV	3.2	2.0	5.4	11.9	13.6	23.6	33.7

Source: author.

The control group in the average proficiency subgroup outperforms the experimental group in all language test and in the development of language awareness. Both groups show an interesting pattern among language tests and

correspondent delayed language tests in that their scores lower very little from the former to the latter. The scores on language awareness, instead, show a high decrease for both groups in the data collected in after a three-week period. The control group is steadily less homogeneous than the experimental group and, and its coefficient of variation is higher, bar for language awareness. In this section, the data show, on the contrary, opposite figures.

High-average proficiency subgroups

Table 55 Language awareness in high-average control group, Exeter

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
EXCA3	81.7	63.0	60.5	72.0	66.7	47.4	36.8
EXCA5	93.3	58.0	57.0	57.0	56.1	57.9	52.6
EXCA6	70.0	68.0	66.7	64.0	63.2	63.2	57.9
Average	78.9	78.7	78.1	72.0	70.8	56.1	49.1
SD	8.6	7.6	7.9	11.1	11.4	8.0	11.0
CoV	10.8	9.7	10.1	15.5	16.2	14.3	22.3

Source: author.

Table 56 Language awareness in high-average experimental group, Exeter

Student	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post-Test + Lan. Awaren.	Lan. Awaren.	Delayed Lan. Awaren.
EXCASA3	83.3	65.0	67.5	62.0	61.4	73.7	52.6
EXCASA6	100.0	95.0	95.6	90.0	91.2	94.7	94.7
EXCASA7	100.0	93.0	93.9	83.0	84.2	100.0	94.7
Average	94.4	84.3	85.7	78.3	78.9	89.5	80.7
SD	9.6	16.8	15.7	14.6	15.6	13.9	24.3
CoV	10.2	19.9	18.4	18.6	19.8	15.6	30.1

Source: author.

Finally, within the high-average proficiency subgroup of the University of Exeter the experimental group outperforms the control group in all language test and in the development of language awareness. However, it also shows a steadily higher level of probability of internal variation, and the values are less homogeneous than the scores of the control group.

6.5 Comparative analysis of results on language awareness

This part of the data analysis mirrors the comparative analysis I carried out in the previous chapter section 5.6 by adding together the results from the pre-test, post-test and the delayed post-test of the Universities of Milan, Genoa and Exeter. The results of the subgroups within the control groups are analysed together, and so are the results of the subgroups within the experimental groups. In line with the explanation provided in section 5.6, I do not examine the coefficient of variation. The data analysis starts with the low-average proficiency subgroup, it subsequently proceeds to the average proficiency subgroup, and finally it assesses the high-proficiency subgroup.

6.5.1 Low-average proficiency subgroups

Language awareness

Table 57 Language awareness across low-average control groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan. Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	52.5	51.0	49.8	44.3	43.6	42.1	39.5
Genoa	40.4	43.0	43.0	38.3	38.6	40.8	38.2
Exeter	53.3	51.0	50.0	50.5	50.0	44.7	42.1
Average	48.8	48.3	47.6	44.3	44.1	42.5	39.9

Source: author.

Table 58 Language awareness across low-average experimental groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	53.8	65.3	63.7	59.0	57.7	56.6	52.0
Genoa	46.1	55.3	53.8	51.7	50.6	43.9	45.6
Exeter	58.3	57.5	56.6	52.5	51.3	50.0	42.1
Average	52.7	59.4	58.0	54.4	53.2	50.1	46.6

Source: author.

The development of language awareness across the low-average proficiency subgroups shows higher results in the experimental groups. This also applies to the language tests and to the scores that combine language tests with language awareness.

Standard deviation

Table 59 SD in language awareness of low-average control groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	3.2	11.6	9.0	11.9	9.0	9.6	12.5
Genoa	5.8	5.7	5.3	6.2	6.2	1.0	1.7
Exeter	9.4	7.1	6.2	9.2	7.4	3.7	7.4
SD Aver.	6.2	8.1	6.8	9.1	7.5	4.8	7.2

Source: author.

Table 60 SD in language awareness of low-average experimental groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	4.7	8.0	7.7	8.0	7.2	9.2	7.1
Genoa	5.9	6.4	5.6	4.0	3.7	3.0	6.1
Exeter	2.4	26.2	26.7	26.2	25.4	33.5	22.3
SD Aver.	4.3	13.5	13.3	12.7	12.1	15.3	11.8

Source: author.

Despite having more positive results in the average scores of all the data sets, the low-average proficiency experimental subgroups are also characterised by a consistently higher degree of internal variation. It is worth notice that, on the contrary, its results were more homogeneous than those of the control group in the pre-test.

6.5.2 Average proficiency subgroups

Language awareness

Table 61 Language awareness across average control groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	67.5	73.3	71.1	68.5	66.7	55.3	53.9
Genoa	71.7	70.3	69.6	63.3	62.3	57.9	56.1
Exeter	78.9	78.7	78.1	72.0	70.8	73.7	64.9
Average	72.7	74.1	72.9	67.9	66.6	62.3	58.3

Source: author.

Table 62 Language awareness across average experimental groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	71.0	71.8	70.7	65.6	64.4	66.3	60.0
Genoa	75.0	70.7	68.4	62.0	60.8	57.9	56.1
Exeter	73.3	70.0	69.3	65.5	64.0	63.2	55.3
Average	73.1	70.8	69.5	64.4	63.1	62.5	57.1

Source: author.

Language awareness across the average proficiency subgroups shows an opposite pattern in comparison with the low-average proficiency subgroups: the control groups scored higher results, bar in the pre-test. Its results are also slightly lower in the scores on language awareness that refer to the post-test data collection phase. The difference between the two scores is minimal (the control group scored 62.3 and the experimental group 62.5) and in my opinion this can be considered an equivalent level of language awareness.

Standard deviation

Table 63 SD in language awareness of average control groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	4.2	4.1	4.4	3.3	4.5	9.1	11.7
Genoa	7.3	6.1	6.2	4.0	4.6	10.5	8.0
Exeter	8.6	7.6	7.9	11.1	11.4	13.9	13.2
SD Aver.	6.7	6.0	6.2	6.2	6.9	11.2	11.0

Source: author.

Table 64 SD in language awareness of average experimental groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	6.0	5.9	5.6	6.5	6.2	6.0	6.0
Genoa	7.3	8.1	8.9	13.5	12.8	10.5	8.0
Exeter	2.4	1.4	3.7	7.8	8.7	14.9	18.6
SD Aver.	5.2	5.1	6.1	9.2	9.2	10.5	10.9

Source: author.

The pattern of correspondence of higher scores and higher internal variation, already noticed within the low-average proficiency subgroups, applies to the average proficiency subgroups. In this case, the control group shows higher results in the average scores of all the data sets, but it also displays a steadily higher degree of internal variation.

6.5.3 High-average proficiency subgroups

Language awareness

Table 65 Language awareness across high-average control groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	90.2	84.2	82.4	77.4	75.6	70.7	63.9
Genoa	-	-	-	-	-	-	-
Exeter	81.7	63.0	61.4	64.3	62.0	56.1	49.1
Average	86.0	73.6	71.9	70.9	68.8	63.4	56.5

Source: author.

Table 66 Language awareness across high-average experimental groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	91.3	90.4	90.2	84.6	84.4	86.3	81.1
Genoa	88.3	95.0	95.6	89.5	89.9	100.0	92.1
Exeter	94.4	84.3	85.7	78.3	78.9	89.5	80.7
Average	91.4	89.9	90.5	84.1	84.4	91.9	84.6

Source: author.

Finally, language awareness in the high-average proficiency subgroups is consistently higher in the experimental group. It is worth notice that the control group consists only in scores of the universities of Milan and Exeter, as there is no high-average proficiency control group at the University of Genoa.

Standard deviation

Table 67 SD in language awareness of high-average control groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	4.7	9.5	9.6	8.8	9.0	14.8	17.3
Genoa	-	-	-	-	-	-	-
Exeter	11.7	5.0	4.9	7.5	5.4	8.0	11.0
SD Aver.	8.2	7.2	7.2	8.2	7.2	11.4	14.1

Source: author.

Table 68 SD in language awareness of high-average experimental groups

University	Pre Test	Post Test	Post Test + Lan. Awaren.	Delayed Post-Test	Delayed Post- Test + Lan Awaren.	Lan Awaren .	Delayed Lan Awaren.
Milan	4.1	4.2	5.4	5.7	6.3	14.2	12.1
Genoa	7.1	4.2	3.7	3.5	4.3	0.0	11.2
Exeter	9.6	16.8	15.7	14.6	15.6	13.9	24.3
SD Aver.	6.9	8.4	8.3	7.9	8.7	9.4	15.9

Source: author.

The pattern of internal dispersion of the data in the high-average proficiency subgroups is not regular. The experimental group is less homogeneous in the post-test section, it increases its value of internal distribution in the delayed post-test section in the long-term period, and it similarly shows a higher internal variation in language awareness in the long-term period.

6.6 Conclusion

This chapter analysed the data collected in this study to answer the first research sub-question: whether focus-on-form instruction is more effective for diglossic language awareness when it links Standard and Colloquial Arabic forms, or when it focuses only on one variety. The chapter was divided into three sections to contextualise and discuss the data collected. The first section explained in detail the collection of the data through a language exercise and questions designed specifically to collect information on language awareness. The section also explained how their assessment is conducted and how retrospective think-aloud protocols are transformed into quantitative figures. The methodological tools detailed in this section have been developed specifically for this research as, to the best of my knowledge, the collection of data on language awareness on diglossic code-switching processes of higher-education students, has not been carried out before. The following two sections compared scores on language awareness within and across the participant universities.

The former compared the results of experimental and control groups within the three universities. It added the scores on language awareness to the results analysed in the previous chapter on the development of diglossic vocabulary and code-switching skills. It also compared the development of language awareness levels between the two groups. Finally, it carried out a comparison between the results in language awareness levels and the scores obtained by the combination of these results with the scores on diglossic vocabulary building, for the three subgroups outlined in the previous chapter (low-average, average and high-average proficiency groups). Finally, results on language awareness and the standard deviation of values collected across the participant universities were analysed.

Empirically, this chapter shows that the experimental groups outperform the control groups in language awareness development, when the two groups are analysed as two entireties, without making distinctions on the basis of the participants' actual proficiency levels in SA at the beginning of the language course. However, the division of the two groups into three SA proficiency-based subgroups at the start of the course, shows that this pattern does not apply to every subgroup. This is particularly the case of average proficiency groups, in which the results for the control group on language awareness tend to be higher than those of the experimental groups. This reflects the behaviour of the findings on diglossic vocabulary building. Another shared feature between the analysis of data on diglossic vocabulary development and data on language awareness, is that the experimental groups' values are often more heterogeneous than the control groups.

The scores on the language tests combined with values on language awareness across the participant universities show that the low-average and high-average proficiency experimental groups outperform the control group. As per the average proficiency subgroups, the control groups score higher. The link between scores on language tests and skills, and the behaviour of values on internal variation is discussed in the final chapter of this thesis.

7 CHAPTER 7 – Perceptions and attitudes towards variation

7.1 Introduction

This chapter is the third and last of three empirical chapters presenting the analysis of the threefold focus of investigation of this research. It aims to observe students' perception and attitudes towards variation in Arabic reality and learning, and it answers the second research sub-question: what impact does focus-on-form instruction have on students' perceptions of Arabic variation, when (a) it links forms in Standard and Colloquial Arabic and when (b) it focuses only on one variety? The viewpoints and motivations of the participants towards Arabic variation are collected through two open-ended questionnaires: the former is administered before the beginning of the experimental language course and the latter is administered at the end. As mentioned above, this thesis does not use attitude or perception scales and it relies on self-reporting questionnaires.

This chapter is divided into two sections. In the first I analyse the answers of the participants to the questionnaire conducted before the beginning of the language course, while the second section concentrates on the answers to the second questionnaire. In the first questionnaire, students provide information on their perceptions of SA, their learning experiences and the reasons that led them to study Arabic. In the second questionnaire, instead, the questions investigate the students' perceptions of integrating the teaching of CA within their SA programme through two different and experimental focus-on-form methods, and their experience and assessment of studying the two varieties of Arabic simultaneously.

As previously explained, the difference between the two methods of FFI employed in this research is based on the nature of the forms that are at the core of the two methods. On the one hand, FFI in the experimental group introduces CA vocabulary forms through comparisons and links with their correspondent forms in SA. On the other, FFI in the control group focuses only

on CA vocabulary forms without making explicit references to their correspondent forms in SA. Since the research took place during the academic year, the attendance at the CA lessons of the language course occurred simultaneously to the attendance at SA classes at the academic institutions of the participants. This means that the participants experienced studying CA and SA concurrently during this research. The questionnaire administered after the completion of the language course allows me to determine the experience of the students in terms of studying the two varieties simultaneously, and the impact that the two methods of FFI have on the perceptions and attitudes of the students towards Arabic variation. Moreover, I collected the participants' evaluations of the impact that CA instruction had on their knowledge of SA: if it interfered with it, either positively or negatively, or if it did not have a direct impact on it. This information was collected with the view to gathering an in-depth understanding of whether and how the two different methods of instruction used in this research influence the students' prior knowledge of SA and if they lead to a conscious development of diglossic knowledge. Students are also asked to provide information on their perception of the usefulness of the instructions received and, finally, their enjoyment of the language course.

7.2 Questionnaire administered before language course

7.2.1 Reasons for learning Arabic

I argued in section 3.5.3 that there is evidence that students are drawn to studying Arabic with the goal of mastering language skills in the way they are mastered by native speakers. They also share the goal of being able to interact with native speakers and engage in different levels of communication, together with the interest to deepen their knowledge and understanding of the Arab world. With a view to adding information on this aspect, I asked the participants to explain the reasons underlying their choice of learning Arabic. The answers I collected confirm a great interest in gaining deeper insights into the cultural, political and religious values of the Arab world, and especially of the Middle East.

One of the reasons for studying Arabic that emerged from the questionnaire is the desire to challenge the negative stereotypes about the Arab world, and “to challenge the negative images portrayed by the media, which can lead to mistrust and miscommunication” (GECA3, 2014). Many students are eager to develop a better cultural understanding of the region with a view to being able to act as cultural mediators and to support facilitation of communication and understanding between the Arab world and different cultures.

Another factor that led several participants to study Arabic, as emerged from the questionnaire, is their perception that the language is in high demand in the job market both within the government sector, and among humanitarian and development organisations. Most students hope to work, travel and live in Arab-speaking countries in the future, and the Middle East is often the preferred destination over other regions. In the case of the students of the University of Exeter, Jordan is the destination of choice of the entire cohort for their study year abroad, as mentioned in section 4.5.2. Student EXCA7 (2015) states: “our University offers to spend the year abroad in Jordan or Morocco. All the class has chosen Jordan mainly because of their interest in the Middle East.” The student also added that this choice “makes the studying of the Levantine dialect more relevant and more useful for the future” (EXCA7, 2015). Muslim students added the role of Arabic as the liturgical language of Islam and their desire to better understand the language of their religion among their motivations for studying Arabic.

Participants also reported to be fascinated by some features of Arabic, and that their decision to study Arabic at university was influenced by their interest in these features. For example, students are fascinated by the Arabic script and by its phonological system which is perceived to be highly different from their own mother tongues. Moreover, as student EXCASA6 describes, students are “eager to undertake the challenge of studying a language from the very beginning that not only is so difficult to learn, but is also spoken by very few foreigners” (EXCASA6, 2015). Several students stated that their positive attitude towards the challenge involved in studying Arabic wavered throughout

the course of their studies due to the complexity of the language and their inability to make continuous progress, especially in their speaking skills. In this regard, participant MICASA16 states that students find it difficult to maintain a steady positive attitude towards Arabic because of “the size of the challenge that Arabic entails, mainly due to the complexity and difficulty of the language.” It is worth noting that the variety of Arabic that the student refers to is SA, which was the only known variety when the questionnaire was administered.

The motivations outlined by the students show a deep interest in gaining an insight into the culture, politics, history and religion of the Arab world with a predilection for the Middle East. The students see their ability to master the language and understand the cultural values of the region as instrumental to promoting intercultural communication, and for improving their career opportunities. Finally, the participants showed an awareness of the need to be provided with effective language tools to perform in real-life situations, i.e. with both SA and at least one variety of CA.

7.2.2 Language skills development and teaching methodologies

The previous section shows that AFL students find Arabic, or, more specifically, SA, fascinating but also complex and difficult, and requiring arduous language learning. Among the most complex language skills, students pinpoint grammar and speaking. In the words of student MICA1, “The complex grammar rules, such as case endings and number agreement, are difficult and burdensome. Speaking is also very difficult because [students] rarely have the chance to practice and so [they] are not used to speaking in Arabic” (MICA1, 2014). Most students define grammar as intricate and too heavily “filled with rules” (GECASA3). In fact, some participants expressed strongly negative views and described Arabic grammar as “almost impossible” (EXCASA2, 2015); “very alien to speakers of Indo-European languages” (EXCA5, 2015); and as posing a large burden on the learners who “need to put a great effort into remembering each grammatical rule needed to make a sentence accurate at every level” (MICASA17, 2014). It is interesting to note the link that the students perceive as

existing between grammar and speaking. Speaking is described as a complicated skill to develop, as a result of two main factors: lack of speaking practice and development of communicative competence in the classroom, and the challenge of “remembering countless grammatical rules to form correct sentences while speaking” (GECA4, 2014). Other features of language learning mentioned by the students as difficult to develop are: memorisation of vocabulary and pronunciation.

Interestingly, student MICA6 mentioned diglossia among the main challenges that can hinder learners from improving the different language skills. The student voiced a fear of not being able to master both varieties, of being confused, and, above all, of being overwhelmed by studying two varieties that are highly different from each other (MICA6, 2014). Since the students had not been exposed to any variety of CA at the moment of the experimental language course, I assume that the impressions expressed were gained from surfing social media in Arabic, and from encountering Arabic language textbooks that use both varieties. In my opinion the answer of student MICA6 is of great importance because it reveals how a lack of structured exposure to Arabic variation in the classroom can result in demotivation, frustration and insecurity. The teaching approach and methodology that students are exposed to in the classroom is therefore crucial not only for the development of their language skills, but it is also crucial in influencing the students’ perception of Arabic variation.

This leads us to the following question of the questionnaire, in which participants are asked to describe the teaching methodology they have been exposed to in their course of study. Students from all the three participant universities describe the teaching methodology in the classroom to be teacher-centred, with limited opportunity to interact and develop communicative skills. The role of the teacher is that of the information provider and there is heavy dependence on the textbook, which has a leading role in the classroom instead of being only used as a facilitator of language learning. Grammar is taught through lessons specifically dedicated to it, followed by the application of

grammatical rules through written exercises. Translation, reading aloud for pronunciation purposes and dictation are among the most common exercises carried out in the classroom. In general, the students' main concerns are the lack of speaking practice, the imbalance in the amount of written and oral exercises, and that the lessons are mainly taught in their mother tongues instead of Arabic.

Participants were also asked to delineate the strategies they employ to memorise vocabulary. Their answers to this question are useful in order to understand the study exercises that students benefit from. Most participants studied lists of words and repeated them until memorisation. They would subsequently write the memorised words against a list of correspondent words in English, in order to verify their vocabulary and spelling knowledge. The newly-learned words were then used by the students within sentences to verify the ability to use them in context. In the words of student EXCA7, "putting new vocabulary into sentences helps me to memorise direct translations of sentences from English into Arabic and thanks to this exercise I am able to recall words in Arabic more easily." It is worth mentioning that many students make use of online learning tools, flash cards and vocabulary memorisation exercises and games. The combination of listening and speaking exercises is also mentioned as particularly useful in memorising new vocabulary, grammar rules through their application in real-life sentences. Finally, student GECA1 declared that they find it "very useful to have mock conversations with other students in Arabic. We ask each other questions and try to use the language in context. Our lessons in class are mainly lecture style lessons and there is not enough speaking practice, so I try to do it outside of the class."

This section has shown us the participants' perception of the difficulty of AFL language skills and the teaching methodologies they have been exposed to. I now proceed to analyse the participants' attitudes towards SA.

7.2.3 Students' attitudes towards SA

In the last question of the pre-test questionnaire, participants were asked to list and explain the most exciting and discouraging aspects related to studying Arabic. Most of the positive aspects recalled the answers to the first question, i.e. the reasons for learning Arabic. In my view this shows that the majority of the students maintain the same positive motivations throughout their first years of study. For example, students mentioned a “great sense of achievement and fulfilment that is gained by mastering a language as difficult as Arabic” (EXCA1, 2015). They also find it exciting to become familiar with a language system that is completely new to them, and although they perceive Arabic to be almost “impossible sometimes, this makes any progress even more rewarding” (GECASA7, 2014). Students find it exciting and motivating to have access to the liturgical language of Islam and therefore to the Muslim community. They also enjoy having access to Arabic media and literature, but they do not feel that their desire to be able to communicate with Arab native-speakers can be fulfilled by the variety of Arabic they study in their academic language courses. In this respect, the students do not feel that the expectations they had before enrolling onto their Arabic courses, of being able to communicate with native-speakers and to have direct access to in-group realities, were being achieved.

Among other discouraging aspects elucidated in the questionnaire, students recalled grammar and diglossia. As student MICA14 puts it, “the idea of diglossia and having to learn two different types of Arabic, formal and spoken, is very discouraging. This is because we do not study Colloquial Arabic at University and so we will have to study it after the completion of our University degree” (MICA14, 2014). Student GECASA5 adds: “I am now in my second year at university and not only do I still struggle with grammar, memorisation of vocabulary and speaking, but I also do not know any colloquial words that would allow me to interact with native speakers.” This shows that the students find it discouraging to dedicate all their energies to studying Standard Arabic and mastering its grammar rules without developing alongside it the skills to communicate in the colloquial variety. Student EXCASA1 claims: “The process

of learning, particularly at the beginning, seems to be more slow-paced than other languages. Also, Arabic does not seem to be taught with the goal of acquiring everyday-life communicative skills” (EXCASA1, 2015). Finally, very few students reported being familiar with Arab music and films.

In this section I have identified the answer provided by the students in the questionnaire administered before the language course and I now proceed to outline the answers provided to the questionnaire administered after the completion of the language course.

7.3 Questionnaire administered after language course

The questionnaire administered after the post-test comprises of four questions on the simultaneous study of both CA and SA. The questions are aimed at understanding the overall personal experience of the participants; their evaluation of what the positive outcomes and main challenges are in the simultaneous study of CA and SA; and finally, their assessment of the impact that the study of CA has had on their knowledge of SA. Within the answers to each question I firstly outline those of the control groups, and I subsequently focus on the experimental groups in order to be able to compare them and assess whether the two different teaching methodologies used have a different impact on the students’ attitudes and perceptions of language variation.

7.3.1 Overall experience of studying Arabic simultaneously

Control Group

The answers received in the control groups show a general agreement that the participants’ overall experience was positive and satisfying. The majority of the students described the language course as useful, helpful and engaging, albeit challenging and sometimes confusing. Some of the students were concerned about their ability to keep the two varieties separate before the beginning of the experimental course, but the majority reported feeling only occasionally confused during the course. For example, student EXCA8 stated that the

experimental language course was not as confusing as the student initially expected. In student EXCA3's words, "It was extremely helpful. I feel that it has helped me improve my Standard Arabic because it has enabled me to speak more confidently" (EXCA3, 2015). It is interesting to note that, although the teaching method employed in the control group does not create links between correspondent CA and SA vocabulary and forms, many students in the group reported having constantly compared CA learned in class with their previous knowledge of SA. GECA6 states "I compared Colloquial and Standard Arabic throughout the course to find the similarities between the two. It is useful to learn them at the same time so that the differences do not seem too huge, as opposed to learning one after the other where they would seem radically different. They have similarities as they are the same language and it is sensible to study them at the same time together." According to another student, "everything we covered in Colloquial Arabic was similar to what we had already studied in Standard Arabic. I am glad I knew it before learning Colloquial Arabic because I used it as a basis to learn Colloquial Arabic" (MICA9, 2014).

Experimental Group

The answers collected from the experimental groups are characterised by a deep dichotomy and are less homogeneous than those of the control groups. On the one hand, some students described the course as very useful, enjoyable, and slightly confusing but less than they had anticipated before the beginning of the empirical research. On the other, a few students found it too confusing, overwhelming and demanding. Many students appreciated its engaging activities and reported having enjoyed it greatly. As student MICASA2 puts it, "I find it a bit confusing, but also fun. I expected it to be even more difficult so I have found it easier than I expected. There should be a balance between hours of Standard Arabic and hours of Colloquial Arabic in the week" (MICASA2, 2014). Student EXCASA3 states: "I have enjoyed it greatly and I think that the learning of some Colloquial Arabic alongside Standard Arabic

gave me a source of enjoyment that I needed” (EXCASA, 2015). According to student GECASA1, the course “made me more confident in speaking Arabic so it had a good impact on my speaking skills both in Standard and Colloquial Arabic. Studying them both at the same time was not only motivating, because I was finally learning the variety that is actually spoken by native speakers, but also somehow crystallised some of my understanding of Standard Arabic” (GECASA1, 2014). In general, students found it useful to see the differences between the two varieties and to employ SA instead of English to understand the meaning of CA words. Moreover, they enjoyed learning “what needs to be changed when adapting Standard Arabic to Colloquial Arabic (for example pronunciation). The sheets provided, with different colours that differentiated Standard and Colloquial Arabic, and the clear distinction between prefixes and suffixes in the two varieties, were very helpful to see the differences clearly” (MICASA8, 2014). However, a few students described the course as overwhelming and confusing. EXCA2, for example, states that “it is too hard and confusing to learn Colloquial Arabic words by linking them to vocabulary in Standard Arabic that the students have only just learned” (EXCA2, 2015). It is interesting to notice that most of the students that found the course too demanding are placed in the low-average proficiency subgroups, whereas the higher the students scored in the pre-tests the more positive their comments towards the experimental teaching methodology are. This division is not so clear-cut in the control group and, as we have seen, the opinions of the students are more homogeneous.

In both groups many students shared the opinion that learning Standard and Colloquial Arabic simultaneously helped them to learn vocabulary and forms in CA and to consolidate their knowledge of correspondent SA vocabulary and forms that they were unsure of. Interestingly, the link between CA and SA was drawn by the majority of the students independently of the instruction received so independently of whether the relationship between the two varieties was explicitly highlighted in the classroom, i.e. in the experimental group, or whether the instruction only focused on CA, i.e. in the control group. Students in both groups agree that, although there are benefits from studying CA and SA

simultaneously, the pitfall of not teaching both varieties simultaneously from the very start of the academic year entails the risk that the students rely on one variety more than the other and that they do not have enough time to practise and develop code-switching skills.

7.3.2 Integration of CA into a SA programme

In this section I outline an overview of the main challenges and the positive aspects of integrating CA instruction within a SA programme.

Control Group

The main challenges mentioned by the control group are as follows. First, the students believe it is hard to memorise the conjugations of CA verbs in the present tense and their pronunciation; second, a few students mentioned that the differences in vocabulary can be difficult to memorise. The differences in grammar are also mentioned as potentially burdensome, but “since grammar is often easier in Colloquial Arabic, the two varieties are hardly confused” (MICA10, 2014). It is interesting to observe that, although some students found it difficult to study the two varieties simultaneously at the beginning of the language course, they soon started to identify the differences and commonalities between them and to treat them as “two different languages” (EXCA2, 2015) with a significant reduction in confusion between them. It is generally agreed that the confusion and mixing of the two language varieties decreases with time and in line with improvement in students’ CA language skills. Finally, numerous students mentioned that it would have been useful to study the two varieties side by side so as to compare them better.

Among the positive aspects of integrating CA into a SA programme, students mentioned that the vocabulary of the two varieties is similar and that the pronunciation in CA is simpler, and therefore that both are easier to remember. Students also added that their opinions “might not be accurate” (MICA3, 2014) and may simply be influenced by the enthusiasm generated by studying Colloquial Arabic. Colloquial Arabic is perceived by the students as a “relief”

(GECA2, 2014) not only because it is characterised by easier grammar than Standard Arabic, but also because it responds to a main communication need of the students. As student GECA5 puts it, “my motivation for studying Arabic is to be able to communicate with people in the Middle East, so I think it is more useful (and enjoyable) to learn the language that people actually speak. It is also more rewarding for the same reason.” The last positive aspect mentioned is the fact that the vocabulary covered in class was already known in SA and the language course played, among other things, an important role in giving the students the opportunity to refresh and revise vocabulary previously learned.

Experimental Group

The experimental group identified similar challenges to the ones mentioned by the control group. These challenges include differences in vocabulary and pronunciation. For example, they include “false friends, such as verbs that exist in both varieties but with different uses, diphthongs, and in general remembering the words that can be used in both varieties and the words that change” (EXCASA4, 2015). According to student MICASA10, “the language course raised two main challenges. First, it is difficult to speak clearly either Standard or Colloquial Arabic without mixing them. Second, [the students] at beginner-level are used to forming sentences in Standard Arabic by translating them from their mother tongue. The exposure to Colloquial Arabic through Standard Arabic implies that the students translate from their mother tongues to Standard Arabic and then from Standard to Colloquial Arabic. A course in which the comparison between Standard and Colloquial Arabic is fundamental to grasping the two varieties and making progress needs to consider the familiarity of the students with speaking in Arabic without relying on their mother tongues, and the real level of knowledge of the students in Standard Arabic” (MICASA10, 2014)

Among the positive aspects of integrating CA into a SA programme, the experimental group mentioned a general reinforcement of their knowledge in SA and a great sense of motivation generated by studying a colloquial variety. The

participants also appreciated a rise in self-confidence and a great improvement in their speaking skills in SA, influenced by their development of speaking skills in CA. Participants also mentioned the opportunity to learn everyday-life and informal words and expressions, as a positive aspect of learning the two varieties simultaneously. Interestingly, a sense of frustration was highlighted in this regard, due to the fact that the students had previously learned these words and expressions in SA, with the awareness that native speakers only use informal and everyday-life expressions in CA. In the words of student GECASA5, "I am glad to have finally learned greetings and colloquial expressions in the appropriate language variety. Although I found it useful to know them already in Standard Arabic and to translate my knowledge into Colloquial Arabic, I would have preferred not to have studied them in Standard Arabic because this variety is not appropriate for informal contexts and I believe that any unnecessary duplication of study is counterproductive for the students" (GECASA5, 2014). Finally, almost every participant stated that the most positive aspect of studying the two varieties at the same time consists in the ability to create parallels between the two varieties and gain a wide vocabulary in both varieties.

7.3.3 Influence of CA on SA passive and active skills

In this section I analyse the participants' perceptions of the influence that the CA language course had on their knowledge of SA. The answers analysed here refer to two different questions, focussing respectively on the potential influence of CA on the acquisition of SA vocabulary and forms, and on the prior SA knowledge of the participants.

Control Group

Most students in the control group feel that CA has a positive influence on SA because it helps to improve oral skills and it provides opportunities to practise speaking and listening skills that are transferrable from one variety to the other. It is interesting to note that almost all students mentioned a potential negative

influence among aspects of the two varieties that are complicated in one variety and simpler in the other, such as numbers and prefixes of verbs in the present tense. In these instances, many students switch from SA to CA, which is perceived as easier than SA, while they speak. Not only do participants describe this tendency as a negative, albeit useful, influence of CA on SA, but they also express the fear of not being able to communicate properly in writing, where the switching between SA and CA is not appropriate.

A widespread answer among the students of the control groups is that CA positively influence the acquisition and consolidation of vocabulary in SA. This is because, as we have seen previously, the students compare the vocabulary and forms of the two varieties to better remember newly learned words and revise familiar ones. Student MICA4 states “it is less likely to be confused when the differences and similarities among the two varieties are clear. However, the disproportion between the amount of instruction received in Standard Arabic and the instruction received in Colloquial Arabic creates an unbalanced situation in which one variety prevails over the other instead of complementing it” (MICA4, 2014).

Experimental group

Eight students of the experimental group do not believe that CA influences SA, either positively or negatively. It is worth noting that all the students that share this opinion are among the high-average proficiency students. In their opinion it does not positively influence the consolidation of SA vocabulary because the vocabulary covered in the CA lessons was already known in SA. It also does not facilitate the acquisition of SA vocabulary because the topics that the students study in the SA and CA lessons are different. Finally, it does not influence linguistic performances in SA. Some students assume that this is due to the fact that the constant comparison between SA and CA made in the language course has helped them to keep the two varieties clearly distinct. Other students suggest that the lack of influence on SA active and passive skills is attributable to the difference in their depth of knowledge of the two varieties:

students' knowledge of SA is deeper and stronger than their knowledge of CA and therefore the latter does not affect the language skills of the former.

The majority of the students, on the contrary, find that the development of knowledge in CA reinforces their SA language skills. This is because the methods of instruction received stimulate their ability to create constant links between the two varieties and therefore to consolidate their knowledge of SA while learning CA. It is worth mentioning that some students, despite appreciating the usefulness of understanding the links between the two varieties and being actively trained in developing code-switching skills, believe that the method of instruction is cognitively too demanding and burdensome if carried out while studying SA. This is because it adds rules to be learned and remembered to the study of SA which is already heavily based on grammar and rules.

7.4 Conclusion

This chapter analysed the participants' motivations for studying Arabic, the teaching and learning methodologies for SA that the participants are familiar with, and their perceptions of SA before being exposed to CA. It also analysed the participants' attitudes towards variation in Arabic reality and learning. The information gathered through two open-ended questionnaires conducted before and after the experimental language course was used to inform the chapter.

Empirically, this chapter has shown that the students hold a deep interest in gaining an insight into the culture, politics, history and religion of the Arab World with a preference for the Middle East. The students see their ability to master the language and understand the cultural values of the region as instrumental in promoting intercultural communication and for their career opportunities. They also have an awareness of their need to be provided with effective language tools to perform in real-life situations, i.e. with both SA and at least one variety of CA. The majority of the participants are interested in learning a variety of Colloquial Arabic that is spoken in the Middle East in order to be able to travel,

live and work in the region. The Middle East is the destination of choice for the greater number of participants, mainly for political reasons.

With regard to the difficulty of acquiring AFL language skills as perceived by participants, the students identified grammar and speaking as the most challenging abilities to develop. In my opinion, these results are partly intrinsic to studying a new language that (i) is based on a different grammatical system than that of the mother tongue; and (ii) and is studied in a classroom environment without any form of immersion in real-life situations and therefore with few opportunities for speaking practice. However, the results could also be a consequence of the teaching methodology the students are exposed to. As we have seen, SA lessons are described by the students to often be teacher-centred and lecture-based with great emphasis on grammar and little active engagement, especially oral engagement, on the part of the students. I will discuss this further in the following chapter. The learning strategies employed by the participants show their preference for activities that use newly-learned vocabulary in context, through original sentences created by the students to practise their vocabulary knowledge. This is also useful for reinforcing knowledge of grammar and pronunciation skills. In my opinion the answers of the students show the need to provide more task-based communicative activities in the classroom based on a combination of familiar and newly-learned vocabulary and topics. This not only helps the students to memorise vocabulary and use it correctly in context, but it also builds their confidence in speaking Arabic.

The absence of formal CA instruction is perceived by the students as a missed opportunity to learn a colloquial variety and therefore to communicate with native-speakers, to be engaged in everyday-life activities, and to develop a comprehensive understanding of the culture and values of the Arab world. It is in this light that both the students of the control group and those of the experimental group showed a great interest in participating in this research. The questionnaires conducted after the language course confirms great eagerness on the part of the students to study CA. They also share positive attitudes

towards studying CA and SA simultaneously. Although the majority of students in both groups find it challenging and sometimes confusing to integrate the study of CA within SA, the enthusiasm injected by developing communicative skills in a colloquial variety acts as a driving force for the development of language knowledge in both language varieties. I will discuss the differences and similarities between the two groups in detail in the following chapter. Chapter eight is the discussion and conclusion chapter of this thesis, in which I provide a critical discussion of the data analysed in the three empirical chapters.

8 CHAPTER 8 – Conclusion

8.1 Introduction

This thesis offers several main theoretical contributions to the literature of Arabic variation, TAFL, and second language acquisition. It contributes to the field of TAFL and specifically to the acquisition of diglossic vocabulary and the development of diglossic awareness. It elaborates on a theoretical framework based on the view that variation is an integral feature of Arabic and on its representation through the reference-packagings model outlined by Giolfo and Sinatora (2011). The reference-packagings model interprets Arabic variation diglossia as being the presence of two linguistic and metalinguistic models in the mind of native speakers, and native speakers' language attitudes and linguistic proficiency as being the main catalysts for variation in language use. This moves away from several theories criticised in chapter two, which see speakers in a diglossic context as having no freedom in choosing which language variety to use, and as passively abiding by superimposed boundaries. My view in this thesis is that, on the contrary, language variety is determined by speakers' subjective choices, sociolinguistic and pragmatic factors, and their command of language varieties.

Arabic is presented in this research as a multi-faceted language composed of multiple language varieties, which fulfil distinct linguistic and metalinguistic functions. Native-speakers switch between language varieties through code-switching. Code-switching is interpreted here as being determined by numerous linguistic, sociolinguistic and behavioural factors, as well as to be a function of language pragmatics. Not only do Arabic varieties constantly alternate and mix through code-switching and code-mixing, but also they are intrinsically heterogeneous. However, as previously said, this thesis claims that the varieties are not distinct units, but are part of a single linguistic and cultural unit which is Arabic.

The application of this theoretical framework to the field of TAFL implies that, in order to develop near-native language proficiency, students of AFL need to

learn at least two varieties of Arabic, SA and CA, and develop diglossic code-switching skills and diglossic language awareness to switch between language varieties. This is theoretically innovative, as it has not been experimented systematically before.

The empirical research in this thesis was based on an experimental language course that aimed at teaching CA using two different methods of FFI. I used the first method to teach CA vocabulary forms through a consistent comparison with their correspondent SA forms. The comparison was made through visual examples, through matching and translation exercises that linked the two varieties together, and through explicit explanations. I employed the second method of FFI to teach CA vocabulary forms without any explicit reference to their SA correspondents. Empirical findings from students of AFL enrolled in three different universities completed the thesis. The empirical data represented the development of diglossic vocabulary and diglossic language awareness; the three universities involved in this study were the University of Exeter, Genoa and Milan. This choice of cases contributes adequately to empirical findings. Firstly, results from students of AFL with different mother tongues on their development of diglossic vocabulary has never been studied jointly, despite the benefit of comparing results based on students with different first languages. Secondly, the development of diglossic vocabulary knowledge and of diglossic language awareness, have also never been studied jointly. In fact, to my knowledge, neither of the above has been empirically studied to date. Lastly, this thesis also analysed the perceptions of students towards Arabic and its variation in order to better understand how diglossic knowledge develops and also the attitudes of the students towards variation in Arabic reality and learning.

This thesis set out to answer one main research question: does focus-on-form instruction lead to diglossic vocabulary development more effectively, when (a) it links forms in Standard and Colloquial Arabic, or when (b) it focuses only on one variety? It also focused on three additional sub-questions. The first two sub-questions aim at understanding if focus-on-form instruction is more effective in promoting diglossic language awareness, and what impact it has on students'

perceptions of Arabic variation, when it links forms in Standard and Colloquial Arabic or when it focuses only on one variety. The last sub-question investigates the extent to which the development of diglossic vocabulary knowledge is a function of the method of focus-on-form instruction received. The main focus of this thesis was to contribute to vocabulary-building theories, particularly regarding what concerns the development of diglossic vocabulary. The thesis was built under the premise that the dependent variables are the three aspects of AFL acquisitions mentioned above: diglossic vocabulary development; diglossic language awareness, and students' perceptions and attitudes towards language variation. Thus, the phenomenon studied has a threefold nature. The independent variable is the creation of links and comparisons between SA and CA forms, which is used only in the experimental group.

Following the methodological chapter (chapter 4), in which I explained the structure and process of employing a mixed-methods study and how the theoretical framework is used to draw conclusions in relation to the system in focus, I presented three empirical chapters. In these chapters I analysed the quantitative and qualitative data collected in this research. Each chapter focuses on one aspect of the focus of investigation of this thesis and analyses the data that inform the answers to the research questions. Chapter five reported data on diglossic vocabulary building, chapter six focused on language awareness, and finally chapter seven explained the perceptions of the students towards language variation. The discussion and conclusion chapter reviews the contribution and findings of the above-mentioned chapters and it merges the analysis of the data collected with the view to answering the last sub-question and to shedding light on the impact that the two different methods of FFI used had on the three aspects of AFL acquisition that are at the core of this research.

This chapter is divided into three sections. The first summarises the theoretical standpoint of this research. The second discusses the analyses of all three empirical chapters. The first part of this section reviews and discusses the conclusions on diglossic vocabulary building; the second part focuses on

diglossic language awareness; and finally, the third part addresses students' perceptions of Arabic variation. The third section of this chapter presents the limitations to these findings and discusses the generalisability of the research. Limitations concerning the methodology and the empirical analysis framework are the key focus. Finally, the part on generalisability proposes areas of research for further study.

8.2 Position of the research towards TAFL

This thesis recognised that TAFL is mainly characterised by teaching methodologies that favour the teaching of SA and only rarely integrate it with the teaching of CA. I have reported in section 2.5 the reasons that have been brought forward in the literature for not incorporating CA within SA. First, SA is seen as a springboard for students to acquire any spoken dialect, and as a basis of reference upon which to develop knowledge of CA. Second, SA is claimed to be the variety used in formal and academic contexts, and it is therefore considered appropriate to be taught to students of AFL at higher-education level. I have explained that this interpretation of Arabic, mainly based on Ferguson's dichotomic distinction of H and L, is not accurate and it does not reflect the reality of Arabic. Arabic varieties are characterised by features that make them sociolinguistically more appropriate for certain situations than others, but native-speakers' linguistic performances are nonetheless constantly identified by code-switching, as we have seen previously. The position of this research is thus that Arabic is not composed of two distinct varieties, but it is characterised by variation and heterogeneity. This means that it is very rare to find situations in which a pure variety, being it either SA or CA, is used. For this reason, students of AFL need to be exposed to the nuances of Arabic language varieties and learn how to code-switch between them. Finally, this thesis also reports the position of numerous academic institutions that claim the need to train students of intermediate and advance levels to express high levels of

abstraction and therefore need to master the formal variety. I agree that the more the students advance in their knowledge of Arabic the more that need to master SA, but I do not believe that this is mutually exclusive with developing knowledge of CA as well.

Although there are TAFL teaching methodologies that include the teaching of CA and SA, or that start with CA and introduce SA at a later stage, these are currently less widespread than the methodologies based on SA. This thesis aims to contribute to the existing research in the field of Arabic Applied Linguistics by exploring experimental methods to incorporate the teaching of Colloquial Arabic vocabulary within a Standard Arabic-based programme. This is carried out with the view to incorporating innovative approaches to teach CA into mainstream practice. The experimental methods employed to teach CA in this thesis, both based on focus-on-form approaches, were explained in detail in section 4.3.

I have called in this thesis for the need to provide AFL students with the necessary tools to efficiently communicate in real-life circumstances and therefore to be able to master both language varieties. This position is supported by the guidelines of the American Council on the Teaching of Foreign Languages (ACTFL) (2012), which posit that proficiency in a foreign language is the ability to efficiently communicate sociolinguistically in real-life circumstances, and the Common European Framework of Reference for Languages (CEFR) (2001). The CEFR also focuses on using the language in authentic situations. This has led to the description of language use as follows: SA is used in reading and writing, whereas both CA and SA are used in listening and speaking. The second group of skills is not duplicate in the two language varieties, but it is used for different purposes: beginner's level students of AFL need to master speaking and listening in CA for ordinary conversations, while advanced students need to be able to use SA for delivering speeches or for formal interviews, and to adjust their CA to produce a form that is suitable for semi-formal conversations.

On these premises, my position towards proficiency sees it as the knowledge of language varieties and the ability to use them in the language skills that differentiate their use. Moreover, I interpret proficiency as the ability to code-switch between language varieties. As seen previously, the switch is a response to different sociolinguistic, pragmatics, metalinguistic and linguistic factors. On the basis of this conceptualisation of proficiency, I have delineated the focus of this research as (i) the development of diglossic vocabulary and diglossic code-switching skills; (ii) the development of diglossic language awareness; and (iii) lastly students' attitudes and motivations towards Arabic variation developed while focussing on building diglossic vocabulary. I discuss the analysis of the data of the threefold focus of investigation in the following section.

8.3 Main findings and contribution

8.3.1 Diglossic vocabulary development

The data collected on diglossic vocabulary development show that there is no statistical difference between the results of the control and the experimental groups. Thus, it is not possible to determine whether the differences in the performance of the two groups are due to the differences in the language instruction that they received in the experimental language course. Nonetheless, as I highlighted in the data analysis chapter, the scores of both groups tend to follow stable patterns that offer a deeper insight into the development of diglossic vocabulary. The analysis of the results of the language tests for the three participant universities shows that all the experimental groups outperformed the control groups both in the post-test and in the delayed post-test. At the University of Milan, the experimental group scored more homogenous results than the control group, and its results are more likely to remain close to the average in the long-term. However, in the long term, the overall performance of the experimental group lowers more significantly than the performance of the control group. At the University of Genoa, on the other hand, the control group becomes increasingly more homogeneous over time, and the overall average language score of the experimental group lowered

significantly in the long term. The findings at the University of Exeter confirmed those of the University of Genoa in that the standard deviation and the coefficient of variation of the control group are increasingly more homogenous over time and the likelihood of internal dispersion of values decreases steadily. On the other hand, the experimental group showed an unstable level of internal dispersion, and the likelihood of it affecting the language scores of the students increased steadily across the three language tests.

These results show that, although the experimental groups outperformed the control groups in the language tests, their results are less stable in the long-term and they fall at a faster pace than the results of the control group. Moreover, the results of the experimental groups are more heterogeneous than those of the control group. My assumption is that the experimental groups outperformed the control groups as a consequence of the FFI method of CA instruction they were exposed to in the language course. This method trained them to switch between SA and CA and to constantly compare the forms of the two varieties. As explained in section 5.3, the language tests were composed of exercises on CA vocabulary retention and code-switching. The latter exercises were introduced in the attempt to recreated real-life use of Arabic. The students that had been exposed to code-switching during the language course performed better in the tests. The findings show that having knowledge of SA and CA as two independent varieties, as is the case of the control group, does not provide the skills to perform code-switching as proficiently as the knowledge of how to code-switch between the two varieties and it also does not reinforce CA vocabulary retention. On the other hand, when FFI links Standard and Colloquial Arabic forms, it facilitates the ability to code switch and to retain newly learned CA vocabulary. We have seen in chapter seven that a great number of students from the control group autonomously compared CA vocabulary forms with their correspondents in SA. However, the results of the language tests were not equivalent to those of the experimental group. This could be caused by two factors: on the one hand, not all the students of the control group trained themselves to create links between Standard and Colloquial Arabic forms, whereas all the students of the experimental groups

were explicitly trained to do so. Thus, there exists a disproportion among the number of students within the two groups that were exposed to diglossic code-switching and this disproportion affected the results of the language tests. On the other, it could be argued that self-training in creating links between language varieties is not as effective as receiving explicit diglossic code-switching instruction. Further research is needed to inform these results and better understand the behaviour of the figures collected.

These are crucial findings in that they contradict the approach to teaching Arabic that claims that knowledge of SA is sufficient to learn CA since it acts as springboard to CA, and dismisses the need for CA instruction.

However, as we have seen earlier, the results of the experimental groups are more heterogeneous and less stable in the long-term, and decrease at a faster pace than the results of the control group. I suggest an interpretation of these results as follows. First, the fact that the results of the experimental groups are more heterogeneous than those of the control groups is likely to mean that the students of the former group respond in a more internally uneven manner than that of the students of the control group. Chapter seven has shown that the method of FFI employed in the experimental group can be perceived by the student as extremely useful and enjoyable, but it can also be seen as cognitively demanding and overwhelming. This could in turn mean that the method of FFI used with the experimental group can have either exceptionally positive results for students that benefit from a high-level of cognitive engagement, it also triggers affirmative reactions in those students, but it is also likely to have negative results for the students that do not benefit from a highly cognitive-based instruction. In addition to this, the analysis of the subgroups show that this dichotomy is particularly strong among high-average and low-average proficiency subgroups.

The impact of the experimental FFI is positive and steady for the high-average subgroup both with regard with results on diglossic vocabulary development and with internal variation. As per the low-average proficiency subgroup, although the experimental group showed higher scores than the control group in

the language test (probably due to the element of code-switching training as we have seen above), its results were more heterogeneous than those of the control group and less stable. This is likely to be caused by the fact that students in the experimental group draw on their background knowledge of SA, which implies a challenge for students who do not have a solid knowledge of SA. Moreover, as we have seen in chapter seven, most students do not feel comfortable in speaking in Arabic because they do not practice speaking in their SA lessons, and this adds a second challenge for the low-average proficiency subgroup.

In the long-term, the results of the students with high-proficiency levels and those students who respond positively to the instruction, remain steadily positive, but the results of the students that do not benefit from the instruction, or that are included in the low-proficiency subgroups, drop considerably.

It is interesting to note that not all experimental subgroups outperformed the control subgroups. Two experimental subgroups out of three (low-average and high-average proficiency subgroups) performed better than the correspondent control subgroups in all universities. As the majority of the experimental subgroups outperformed the control groups, the overall results of the experimental groups were higher than those of the control groups. The results of the average subgroup, instead, do not show a significant difference between the two groups. The results between the two groups are not identical but they are very close and can be considered almost equivalent. This means that the instruction received does not influence the performance of the average proficiency subgroups.

8.3.2 Diglossic language awareness

The second part of the threefold focus of investigation of this thesis was diglossic language awareness. This research set out to answer the question of whether the form of FFI instruction used in the experimental group is more effective for the development of language awareness than that used in the control group. My assumption before collecting the data was that the

experimental form of instruction would raise language awareness more effectively than the other form of FFI. This assumption was based on the fact that the experimental group explicitly trains the participants to be cognitively engaged in diglossic code-switching and therefore creates an environment in which the students are aware of the reasons and modalities in which diglossic code-switching takes place. However, I was also conscious that the experimental instruction could be highly cognitively demanding for students and that the constant references that it creates between SA and CA vocabulary forms could create confusion and be counterproductive to the development of language awareness. This thesis therefore did not have a defined assumption in favour of one of the two methods of instructions, and I answered the research question by means of the data analysis provided in chapter six.

The data collected on diglossic language awareness confirmed the results obtained for diglossic vocabulary development in that there is no statistical difference between the results of the control and the experimental groups. It is therefore not possible to determine statistically whether the differences between the two groups are a consequence of the language instruction that the two groups received in the experimental language course. However, the results of both groups tend to follow stable patterns that provided an insight into diglossic language awareness raising. The data analysed in chapter six showed that the experimental groups scored higher results than the control group, when the two groups were analysed as two units without distinguishing the subgroups on the basis of the participants' actual proficiency levels in SA at the beginning of the course. This was also the case for diglossic vocabulary development and, in my opinion, this happened for the same reason: the experimental group is more trained in code-switching. As a consequence, it is more aware of it and its overall performance is better than that of the control group. Although lower than those of the experimental group, the results of the control group are more homogeneously spread around the average. Those of the experimental group, on the contrary, are significantly heterogeneous in nature and the scores are either significantly higher or lower than the average of the group. This reflects the impact that the method of FFI used with the experimental group had on the

students: it was received extremely positively by some of the participants, who performed very well and significantly above the average in the language tests, but it also had a more negative impact on students who in turn performed significantly below the average. In the control group, instead, the method of FFI had a more uniform and consistent impact. In the long term, the dichotomy of results and attitudes of the participants is reflected in the results of the group being less stable and lowering at a faster pace than those of the control group. It is worth noting that, as per the development of diglossic vocabulary, the average proficiency subgroups do not show substantial differences between their results and are considered in this thesis almost identical.

The impact of the instruction on the participants' development of language awareness is reflected in the answers on language awareness collected orally through retrospective think-aloud protocols and answers collected through the post-test questionnaires. The students of the experimental group that responded positively to the instruction received reported to be able to link the forms of the two varieties but maintain them separate in their minds (as native speakers do). For example, MICASA7 stated: "What I found fascinating is that, thanks to the comparison and exercises to switch between the two varieties, I now feel that switching it is becoming a natural language practice: it is natural to use greetings in Colloquial Arabic, read in Standard Arabic and tell something about my day in Colloquial Arabic" (MICASA7, 2014). Similarly, EXCASA1 reported: "I do get confused between the two because with time and practice I associate one variety with informal speaking and listening, and the other with formal speaking and listening, while reading and writing remain in Standard Arabic. These lessons have helped me to start definitely differentiating between the two varieties in a very clear and natural way" (EXCASA1, 2015). The same students also responded positively to the questions of the post-test questionnaire on their experience of mixing the two language varieties. They described the constant comparison between SA and CA as being useful in keeping the two varieties clearly distinct without being burdensome or confusing.

The students of the experimental group who responded less positively to the instruction received described the instruction and the course as difficult, albeit useful. They shared a sense of frustration due to the impractical duplication of efforts required to study the same vocabulary and forms twice, and due to a perceived burden of being exposed to lessons in SA and CA heavily characterised by explanations of rules and attention to forms. Various answers to the post-test questionnaire reported puzzlement on the part of the students over learning vocabulary in SA related to familiar and informal topics, such as greetings, that is in fact only used in CA. The answers of these students on language awareness show a significant difference between the post-test and the delayed post-test. Although the students show awareness of diglossic switching in the post-test, they do not retain substantial awareness over time.

The results of the control groups are lower but more homogeneous than those of the experimental group. Lower results are likely to be a consequence of lack of training in code-switching, as we have previously seen. However, the form of instruction received triggered a more uniform response from the students. This helped them in maintaining their scores consistently more stable in the long term, in comparison with the experimental group. The students showed in their retrospective think-aloud protocols a lack of fully developed awareness of their ability to perform code-switching, and they seemed to overlook inaccuracies in their code-switching performances. However, we have seen in chapter seven that most students of the control group autonomously created links between CA vocabulary forms learned in the experimental language course and their correspondent forms in SA. Since the students of the control group maintained their scores more stable than the experimental group, it is reasonable to assume that they benefit in the long term from the freedom to create a comparison between the two varieties autonomously, without any explicit instruction in the classroom. Students of the control group seem to agree that instruction in CA provides them with the “necessary amount of fun that learning a language needs, and that was not a part of Standard Arabic lessons” (EXCA7, 2015). In the words of students GECA4, “I feel that Colloquial Arabic has improved my performance in speaking and has motivated me to speak more

Arabic. What I enjoy is also that Colloquial Arabic compensates for the hard parts of Standard Arabic. Grammar for example is extremely complicated and hard to learn. Learning Colloquial Arabic has showed me that speaking doesn't necessarily have to be precise and it revived my interest in Standard Arabic. In fact, I feel that my Standard Arabic has improved because I am interested in the links between this and Colloquial Arabic.”

I now proceed to discuss the information analysed in the last empirical chapter, on students' perception towards Arabic variation.

8.3.3 Students' perceptions of variation

The third, and last, empirical chapter focussed on the impact that the methods of FFI employed in this research have on the students' perceptions of Arabic variation. I also assessed their evaluation of the effectiveness of the instruction received, and I gathered in-depth information on their experience of studying CA and SA simultaneously. I compared this information with information that I collected before the beginning of the experimental language course on students' attitudes and motivations towards SA and its study as a foreign language. Students also explain effective learning strategies that they implement to remember vocabulary; the teaching methodologies they were exposed to in their SA lessons; their assessment of Arabic complexity; and the challenges entailed by the development of Arabic language skills. The comparison between the data collected before and after the language course allowed me to better understand the impact that the methods of instruction used in this research had on how the perceptions of the students towards Arabic developed, as well as on their views and motivations towards Arabic language and learning.

This chapter has shown that the students are aware of the need to master both CA and SA to perform in real-life situations and they expressed a sense of frustration caused by not receiving instruction in both varieties. An interesting aspect that emerged from the chapter analysis was that the majority of the participants identified grammar and speaking as the two most challenging and

complicated language skills to develop. Students often identified their lack of speaking practice as one of the main factors that prevented them from feeling comfortable with speaking Arabic. They also attributed their lack of ability to speak to having to pay attention to complicated grammar rules. In my opinion, this is likely to be a consequence of receiving teacher-centred SA lessons that place great emphasis on grammar and have little active engagement. This study showed that, in contrast, students gained motivation and enthusiasm by receiving CA instruction. In student GECA5' words, "I didn't encounter big challenges during the course. I did confuse the vocabulary between the two varieties every now and then, and I struggled at times, but studying Colloquial Arabic is so motivating and useful that it motivated me not to give up and carry on studying." (GECA5, 2014). Student MICASA2 added: "I found it extremely useful and I think it should be introduced from the beginning of the first year." (MICASA2, 2014).

Students also share positive attitudes towards studying CA and SA simultaneously. The majority of the participants stated that their acquired knowledge of CA had helped them in consolidating their previous knowledge of SA and also improved their language performances in SA, because developing knowledge in CA had triggered their self-confidence. There were two main differences between the experimental and control groups towards the method of instruction received. The control group showed a general agreement that the experience was positive and useful and many students reported having constantly compared CA learned in class with the corresponding SA vocabulary. The answers collected from the experimental groups, in contrast, had a dichotomous nature and were less homogeneous than the control groups: some students found the course useful and enjoyable, albeit confusing, while other students found it too rule-based and demanding. As the results of experimental groups were higher than those of the control groups, I believe that the experimental FFI method has proven to prepare the students more efficiently for using Arabic in real-life language situations. However, this instructional method needs to be structured and developed with the goal of including plenty of communicative activities, role-plays, vocabulary games, etc.

in order to offer a differentiated programme that does not place its focus solely on cognitive engagement.

8.4 Limitations

I have sought to gather the data for this thesis in an objective and ethical manner. The research design, theoretical and conceptual frameworks and operationalisation of variables were developed based on the literature of Arabic variation, second language acquisition and TAFL. However, a research project measuring diglossic vocabulary building and language awareness simultaneously has never been carried out before, to the best of my knowledge. For this reason, several challenges and limitations have been present all through the process, from decisions regarding the research design to the data gathering and empirical analysis.

The main theoretical limitation was the absence of internationally agreed-upon guidelines to measure learning outcomes for diglossic vocabulary building. I developed the learning outcome standards used in this thesis by applying Bloom's taxonomy of educational objective as explained in chapter 4. Although I carried out this application in a very thorough and informed manner, Bloom's taxonomy was not developed specifically to measure diglossia, and mine is therefore an attempt at a useful adaptation.

Another limitation is that the absence of guidelines to be followed to measure the main focus of research of this thesis meant that I decided how to define, collect, measure and assess the learning outcomes of the students. An additional limitation is the size of participant cohort. In order to avoid bias in data collection, I selected participants that had never been exposed to CA. This had an impact on the number of students that could participate in the research, and therefore decreased the size of the cohort. For example, second-generation students and students enrolled in the second year who had spent a period abroad during summer in an Arab country could not participate. As I highlighted in empirical chapters 5 and 6, the reduced size of participant cohort of the

research implied that the quantitative analysis of the data was not always possible.

Finally, it is possible that the influence of other, unaccounted-for variables was at play during the fieldwork. This is the case, for example, when looking at the amount of self-study that each student carried out during the course and prior to the language tests. To avoid overwhelming the students with homework during their already busy academic terms, I did not ask them to study and prepare themselves for the language tests. I also decided not to measure self-study as a variable because this study focusses on another variable, which is the proficiency level of the students in SA. However, some of the students might in any case have studied or revised in preparation for the tests, and this could have had an impact on their results

To compensate for this sort of limitations, a systematic approach to analysing the data was taken and explained in detail in the methodological section. It is nonetheless not impossible, that other interpretations could add explanatory value to changes and continuities.

8.5 Conclusion

This thesis outlined several views on Arabic variation and elaborated a theoretical framework to observe the development of diglossic vocabulary knowledge; of diglossic language awareness; and of students' perceptions and attitudes towards variation in Arabic language and learning. The theoretical framework was built on two concepts at the core of this research. The first core concept was that Arabic reality is characterised by variation. With regard to language use, this means that Arabic varieties are intrinsically heterogeneous and are characterised by diglossic variation. As for the theoretical representation of the language in the mind of native speakers, the varieties of Arabic are well-defined and complementary and their alternation is subject to individual choices of the speakers. These choices are influenced by linguistic, sociolinguistic and pragmatic factors, but they ultimately depend on the

speakers. The central role of the speaker in determining and triggering language variation was the second core concept that supported the theoretical framework of this thesis. This conceptualisation of Arabic variation and of the crucial role of the speakers in determining diglossic code-switching has a decisive impact on TAFL, in that it requires us to define proficiency in Arabic in light of diglossic variation. Proficiency is therefore understood in the theoretical framework of this thesis as the ability to master SA and at least one variety of CA, and to code-switch accurately and efficiently between them through diglossic code-switching.

The conceptualisation of proficiency offered by the theoretical framework of this thesis required delineating language skills and expected language outcomes for students of AFL against which I could measure the development of diglossic language skills. I identified the language skills to be measured as follows: CA vocabulary retention of vocabulary items already known in SA; development of diglossic code-switching skills; and presence of diglossic language awareness. With regard to learning outcomes I developed original suggestions on the basis of Bloom's taxonomy of educational objectives and Leow's (1997) levels of language awareness. The learning outcomes for the first two skills, i.e. CA vocabulary retention of vocabulary items already known in SA and development of diglossic code-switching skills, were developed using Bloom's taxonomy because of its versatility and adaptability to different learning and cognitive environments. Leow's levels of presence of diglossic language awareness were instead applied without amendments.

The focus of the suggested expected learning outcomes was on beginner-level Arabic students in higher-education. The level of the students was chosen to meet one of the objectives of this research. The objective was to experimentally analyse two methods of FFI to teach CA within widespread Arabic academic programmes. Since widespread and common academic syllabi focus on teaching SA, this research was placed within SA-based Arabic programmes. Within these academic programmes, beginner-level students are rarely exposed to any variety of CA, whereas students at higher levels are likely to have been

exposed to CA at least once. The two experimental methods of FFI were based on the traditional approach to isolated FFI as outlined by Ellis (2001), but they were originally modified in this thesis to focus on diglossic forms.

The discussion in this chapter helped to answer the last research question: to what extent is the development of diglossic vocabulary knowledge a function of the method of focus-on-form instruction received? The answer in my opinion is that it is greatly influenced by the instruction received. The quantitative analysis of data on diglossic vocabulary development showed that when FFI links Standard and Colloquial Arabic forms, it is more effective in developing diglossic vocabulary knowledge and language awareness than when it does not link the two varieties. However, the results of the experimental groups lower more rapidly than those of the control group and the results of the groups are more heterogeneous. We have also seen that, although only the students of the experimental groups were explicitly exposed to a constant comparison between varieties, most of the students of the control groups autonomously created links between the two varieties. However, the results showed that the students of the experimental groups perform code-switching more effectively than the control group, and that the experimental method of FFI reinforces vocabulary retention of newly-learned CA vocabulary items. It could therefore be argued that self-training is not sufficient to develop the necessary skills to perform code-switching effectively. Finally, it is worth highlighting that the two methods of instructions were received slightly differently by the students. On one side, students of both groups enthusiastically welcomed the language course and CA instruction. On the other, the response of the control group was homogenous while that of the experimental group was less consistent. In this case the students showed two distinct attitudes in which they either enjoyed it greatly or they felt overwhelmed and confused.

In my opinion, these findings suggest that a method of TAFL based on the combination of the two forms of instruction would be beneficial for Arabic language students. For example, the instruction in the classroom could focus prevalently on the forms of CA and on communicative exercises, while

consolidation exercises, intended to be carried out at home and checked in class, could instead be based on the comparison between the two varieties. Moreover, the exercises could be developed in a way that makes the correction in class communicative, interactive and enjoyable. The findings also seem to suggest that CA instruction should be introduced at an earlier stage in the Arabic programme. This would allow the students to have more opportunities to practise their speaking skills, and it would allow for greater opportunities for linkage between the two varieties and a more consistent approach in general.

To conclude, several suggestions for subsequent research would be especially useful, to further develop research in this area. Firstly, additional research would be needed to test the framework developed in other cases. Second, this thesis used non-parametric quantitative analysis. It would be fascinating to test the variables on a greater scale and analyse large-n data. This would provide greater potential to draw inferences between variables and analyse these frameworks from a different perspective. This would be particularly interesting for diglossic language skills, as no such study has yet been done on a large scale. However, due to the constraints of sample sizes in education research, it is likely that joint research is needed to conduct such large-scale quantitative research on diglossic language skills. Furthermore, this thesis has focused on students that had received between 100 and 150 hours of SA instruction before the beginning of the language course. It would be interesting also to analyse the impact of the two methods of FFI on students that have received less instruction in SA, in order to evaluate the role that it plays in determining students' results and motivations towards Arabic variation.

Appendices

I) Questionnaires

APPENDIX A Questionnaire administered before language test

- 1) Name and Surname:
- 2) Age:
- 3) E-mail address:
- 4) Nationality:
- 5) Mother tongue:
- 6) Where will you spend your year abroad?
- 7) Why did you choose that destination?
- 8) Why do you study Arabic?
- 9) Have you ever studied any Spoken Arabic? If yes, which variety and for how long?
- 10) What do you think is the hardest aspect of Arabic?
- 11) Rate the following language skills according to their difficulty level
(1= EASY; 2= RELATIVELY EASY; 3= FAIR; 4= CHALLENGING; 5= QUITE HARD, 6= ALMOST IMPOSSIBLE)

Listening Comprehension:

Speaking:

Pronunciation:

Reading Comprehension:

Writing:

Grammar:

Remember Written Word Forms:

Remember Vocabulary:

- 12) What are, in your opinion, the most exciting and/or discouraging aspects of studying Arabic? List them and explain.
- 13) Describe the Arabic teaching methodology/methodologies you have been exposed so far.
- 14) Which strategy do you use to remember Arabic vocabulary?
- 15) Which of the above listed strategies do you find effective? Why?
- 16) Do you listen to Arab music? If yes, what is your favourite? If not, why?
- 17) Do you watch Arab movies? What do you like/dislike about them?

APPENDIX B Questionnaire administered after language test

- 1) What are the elements of this experimental course that you have enjoyed?
- 2) Please describe any negative aspects of this course and the challenges you have faced.
- 3) Please describe the impact that this experimental course had on your previous knowledge of Standard Arabic (both receptive and productive knowledge).
- 4) If you could make any changes to this course, what would you change?

II) Language Material for Experimental Group

APPENDIX C Experimental Group - Lesson 1

جُمَلٌ وَأَسْئَلَةٌ مَفِيدَةٌ



عَفْوًا، مَا فِهِمْتِ

عِنْدِي سَوَّالٌ



شو يَعْنِي...؟

عَفْوَاً، يُمْكِنُ تَعِيدُ/تَعِيدِي؟



كَيْفَ بِنُقُولُ... بِالْعَامِيَّةِ؟



كَيْفَ بِنُقُولِي... بِالْعَامِيَّةِ؟

تحیات

Ex. 1 – Match the following images with the correct greetings



صباح الخير



مساء الخير

السَّلام عليكم

مَعَ السَّلامَة



تَشَرَّفْنَا!

أهلاً وسهلاً

مَرْحَبًا!

Ex. 2 – Write the correct response to the previous greetings.

_____ أهلين _____

_____ مرحبتين _____

معلومات عنك/عنك



تَمْرِين ١ - من وين أنتِ

Ex. 1 – Work in pairs and get to know your classmates

Student A	Student B
شو اسمك؟ من وين أنتِ؟ من أي مدينة؟	جاين المملكة المتحدة مدينة لندن

Identities for Students B

كورين كانادا مدينة فانكوفر	محمد قطر مدينة الدوحة
توفيق المغرب مدينة فاس	سارة أستراليا مدينة سيدني

تَمْرِين ٢-

Read the following words. Guess the meaning of new terms

then match the words in عامية with the correct terms in فصحى

بين، بيت، عين، مين، يوم، هون،

ليش، لون، صوت، هيك، فوق

<u>عامية</u>	<u>فصحى</u>
مين	لماذا
هون	مَنْ
ليش	هكذا
هيك	هنا

ل+ماذا

= ل+ايش

Diphthongs in Standard Arabic are special category of vowels: they are monosyllabic sounds that begin with one short vowel and glide into a long vowel that has a different sound (ex. خَيْرٌ، يَوْمٌ). A common example in English is the sound at the end of the word “toy”. **Diphthongs in Spoken Arabic** tend to eliminate the short vowel and the sound of the long vowel prevails (ex (يوم، لون

البيت بيتك!

تَمْرِين ٣ - Listen and repeat the numbers in عامية from 1 to 10.

Which differences can you hear from the same numbers in فصحي?

١ واحد

٢ اثنان

٣ ثلاثة

٤ أربعة

٥ خمسة

٦ ستة

٧ سبعة

٨ ثَمَانِيَة

٩ تِسْعَة

١٠ عَشْرَة

فصحى ≠ عامية

● ت ⇒ ث

● :ثَلَاثَة , ثَمَانَة

(long vowel - consonant - َconsonant)

تَمْرِين ٤ -

Work in pairs with three different classmates. Ask them their phone number and write it below here

سَوِ رَقْم تَلْفُونِكِ؟

تَمْرِين ٥ - Listen and repeat the numbers in عامية from 11 to 19.

ايد عش / احد عش	أَحَدَ عَشَرَ	١١
تتَّعش	اِثْنَا عَشَرَ	١٢
تأطَّعش	ثَلَاثَةَ عَشَرَ	١٣
أرْبَعَطَّعش	أَرْبَعَةَ عَشَرَ	١٤
خَمْسَطَّعش	خَمْسَةَ عَشَرَ	١٥
سِتَّعش	سِتَّةَ عَشَرَ	١٦
سَبَّعش	سَبْعَةَ عَشَرَ	١٧
تَمْنَطَّعش	ثَمَانِيَةَ عَشَرَ	١٨
تِسْعَطَّعش	تِسْعَةَ عَشَرَ	١٩

Can you spot three main differences between عامية and فصحي in the numbers above?

- _____
- _____
- _____

تَمْرِين ٦ -

Work in pairs. Listen to the numbers in mathematical sequence that your classmate reads out loud and add three numbers that follow the same sequence.

Ex.

- ١-٣-٥...
- ٧-٩-١١

١) ١٦، ١٢...	١) ١، ٢، ٣...	١) ١٩، ١٨، ١٧...
٢)	٢)	٢)

١) ٢، ٤، ٦...	١) ١٨، ١٥، ١٢...	١) ٣، ٦، ٩...
٢)	٢)	٢)

تَمْرِي ١ - Work in pairs. Ask your classmate the information below-



الاسم:

الجنسية:

العمر:

بلد:

تَمْرِين ٢ - Listen and repeat the numbers in عامية from 20 to 30.

قَدَيْش عُمْرَك؟

قَدَيْش عُمْرَك؟ عُمْرِي ٢١ / ١٩

! قَدَيْش = قَد + ايش = قَد (to exist) + ايش (شَو synonym of)

امْتِي عيد ميلادَك؟

امْتِي عيد ميلادَك؟ عيد ميلادي ب... ٧ شهر ٣

MONTHS : There are two options you can use

شهر ١	كانون الثاني	شهر ٧	تموز
شهر ٢	شباط	شهر ٨	آب
شهر ٣	آذار	شهر ٩	أيلول
شهر ٤	نيسان	شهر ١٠	تشرين الأول
شهر ٥	أيار	شهر ١١	تشرين الثاني
شهر ٦	حزيران	شهر ١٢	كانون الأول

تَمْرِين ٣ -

Ask your classmate information about themselves and introduce them to the rest of the class

اسمه / اسمها _____

عمره / عمرها _____

عيد ميلاده ب / عيد ميلادها ب _____

تَمْرِين ٤ - Fill in the table with the missing question words.

فصحى		عامية
ما/ماذا		_____
من		_____
أين		_____
كيف		_____
متى		_____
من		_____
لماذا		_____
كم		_____ _____
كم		_____

تَمْرِين ١ - الحياة اليومية

أنا بِالْعَادَةِ بَعْدَ الْجَامِعَةِ بَرَجَعْتُ، بَدَرْتُ، بَعْدِينَ بَطَّلَعْتُ (أَخْرَجْتُ)
مَعَ رِفْقَاتِي عَلَى الْقَهْوَةِ، بَحَكِي (أَتَحَدَّثُ) مَعَهُنَّ، مَنَّمَشَيْتُ، هَيْكَ
الْحَيَاةِ.

بِأَمْرِيكَ النَّاسَ طَوَّلَ الْوَقْتَ مَسْتَعْجِلِينَ، بِيَرْكُضُوا كَثِيرًا.
أَمَّا النَّاسُ بِسُورِيَا عَلَى مَهْلَهُنَّ.

List the verbs according to their subject

هِيَ

نَحْنُ

أَنَا

تَمْرِين ٢ - حَيَاتِي اليَوْمِيَّة

بِالْعَادَةِ يَبْلُشُ شُغْلِي السَّاعَةَ تِسْعَةَ فَبِطَّلِعُ مِنَ الْبَيْتِ السَّاعَةَ ثَمَانَةَ أَوْ

السَّاعَةَ سَبْعَةَ وَالذَّوَامَ بِيَسْتَمِرُّ تَقْرِيْبًا لِّلْسَاعَةِ تِنْتَيْنِ أَوْ السَّاعَةَ ثَلَاثَةَ.

وَالْمَسَاءَ أحيانًا بِطَّلِعُ عَلَى الْقَهْوَةِ أَوْ عَلَى الْكَافِيْتَرِيَا، بِرُوحِ

عَلَى السَّيْنِمَا أَوْ بِطَّلِعُ مِشْوَارًا.

عَامِي ≠ فَصْحِي

بِالْعَادَةِ = فِي الْعَادَةِ

بَلَّشْ، يَبْلُشُ = بَدَأَ، يَبْدَأُ

لَ = إِلَى

السَّاعَةَ تِنْتَيْنِ = السَّاعَةَ الثَّانِيَةَ

ت = لِ

رَاحَ، بِرُوحِ عَلَى = ذَهَبَ، يَذْهَبُ إِلَى

تَمْرِين ٣ - Describe your typical day:

تَمْرِين ٤ - Read the following nouns out loud :

صَدِيق، وَقْت، فَوْق، تَقْرِيْباً شَرْق، وَرَق، أَوْراق، رَقْم، أَرْقَام،
رَفِيق، رِفَاق (رِفَقَات)

تَمْرِين ٢ - Work in pairs and play battleship:

Student A

٣٠									
٢٠									
١٠									
	١	٢	٣	٤	٥	٦	٧	٨	٩

Student B

٣٠									
٢٠									
١٠									
	١	٢	٣	٤	٥	٦	٧	٨	٩

تَمْرِين ٣ - Telling the time

قَدِّيش الساعة؟



- الساعة واحدة
- الساعة تَنتَين
- الساعة ثلاثة
- الساعة أربعة
- الساعة خَمسة
- الساعة سِتَّة
- الساعة سَبعة
- الساعة ثمانية
- الساعة تِسعة
- الساعة عَشرة
- الساعة اِحدَ عشر

قَدِّيش الساعة؟



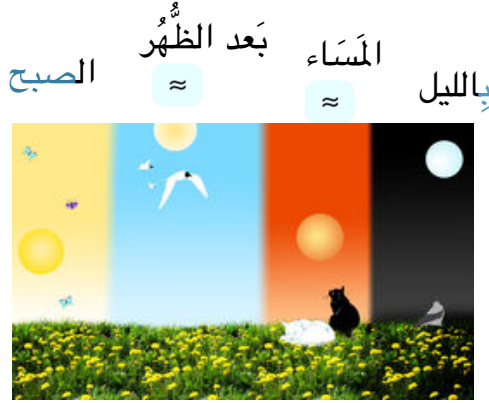
أي ساعة؟

- ٢:٢٠ (١)
- ٧:٥٠ (٢)
- ٣:٣٥ (٣)
- ٨:٥٥ (٤)
- ١:٤٥ (٥)

Work with a classmate and
in turns ask each other the time

- ١:٠٠ (١)
- ٨:٣٠ (٢)
- ٥:٤٥ (٣)
- ٤:١٥ (٤)
- ٩:٥٥ (٥)

امتی؟ ... امتی الدرس / دَوَامك؟



This week's schedule

دَرْس العَرَبِي (٩:٠٠)

(٧:٠٠)

الرَّحْلة



دَرْس الإنكليزي

(٤:٤٥)

العطلة

(كُلّ اليوم)

الاجتماع

(١١:٣٠)

الحفلة (١٠:٠٠)

الحفلة الموسيقية (٨:٠٠)

شمال

غَرْب



شَرْق

جنوب

بِ + شمال / جنوب / غَرْب / شَرْق + (اسم)

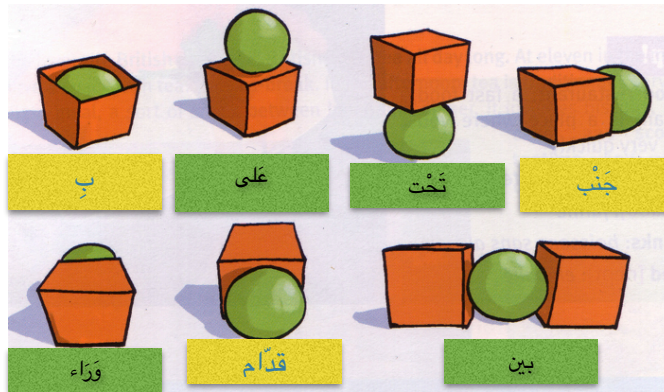
بِ + جنوب غَرْب + (اسم)

بِ + جنوب شَرْق + (اسم)



لندن بِجنوب إنجلترا
إكزتر بِجنوب غَرْب إنجلترا

بِ + الشمال / الجنوب / الغَرْب / الشَرْق



Summary of diglossic code-switching rules covered:

- 1) ق often pronounced as ء

رَقَم، أَرْقَام
قَدَّيْش
قَهْوَة
وَقْت
تَقْرِيْباً
رَفِيْق، رَفَقَات
قَدَّام
قَلَم، أَقْلَام
قَصِيْر، ة

- 2) ث often pronounced as ت (يوم) التتتين: ت

- 3) Diphthongs: 3 letter words (consonant + vowel + consonant) ex. بَيْت، بَيْن، يَوْم...
the ة of the first letter is dropped and the sound of the long vowel prevails:

(ī) كَيْف، مِيْن
(ē) وِيْن، خَيْر، بِيْن، بِيْت، هِيْك، لِيْش
(ou) يَوْم، هُوْن،
(o) فُوْق

- 4) V C° C

Long vowel preceded by two consonants: the first consonant from the right is vocalised with sukun.

شَمَال، جَنُوب، مُنِيْح-ة، رَفِيْق - رَفَقَات، كَثِيْر، كُبِيْر-ة، صُغِيْر-ة، مَرْحَبَتِيْن، قَنِّيْنَتِيْن
check conjugation tables 1c and 1e!(NB) +
check conjugation table 1a when verb suffix is used! (NB) +

Exceptions: the consonant will carry a short vowel on top the first consonant from the right. Ex. قَدِيْم، قَدِيْم

- 5) أَرْبَعِيْن، خَمْسِيْن، تِسْعِيْن: /īn/ tents ending= always pronounced ين

- 6) أَسْبُوْعِيْن، يَوْمِيْن: /ēn/ dual ending= always pronounced ين
Feminine word= ة + ين form the sound /tēn/ مَرْحَبَتِيْن، قَنِّيْنَتِيْن

III) Language Material for Control Group

APPENDIX I Control Group - Lesson 1

جُمَلٌ وَأَسْئَلَةٌ مَفِيدَةٌ



عَفْوًا، مَا فِهِمْتُ

عِنْدِي سَوَّالٌ



شو يَعْنِي...؟

عَفْوَاً، يُمْكِنُ تَعِيدُ/تَعِيدِي؟



كَيْفَ بِنُقُولُ... بِالْعَامِيَّةِ؟



كَيْفَ بِنُقُولِي... بِالْعَامِيَّةِ؟

تحیات

Ex. 1 – Match the following images with the correct greetings



صباح الخير



مساء الخير

السَّلَام عَلَيْكُمْ

مَعَ السَّلَامَةِ



تَشْرَفْنَا!

أَهْلًا وَسَهْلًا

مَرْحَبًا!

Ex. 2 – Write the correct response to the previous greetings.

_____ أهلين _____

_____ مرحبتين _____

معلومات عنك/عنك

شو اسمك؟

من وين انت؟

من أي مدينة؟

تَمْرِين ١ - من وين أنتِ

Ex. 1 – Work in pairs and get to know your classmates

Student A	Student B
شو اسمك؟ من وين أنتِ؟ من أي مدينة؟	جاين المملكة المتحدة مدينة لندن

Identities for Students B

كورين كانادا مدينة فانكوفر	محمد قطر مدينة الدوحة
توفيق المغرب مدينة فاس	سارة أستراليا مدينة سيدني

تَمْرِين ٢ -

Read the following words – Can you guess the meaning of new terms?

بين، بيت، عين، مين، يوم، هون،

ليش، لون، صوت، هيك، فوق

Diphthongs are special category of monosyllabic sounds that begin with one short vowel and glide into a long vowel that has a different sound: a common example in English is the sound at the end of the word “toy”. **Spoken Arabic** tends to eliminate diphthongs and the sound of the long vowel prevails over the

sound of the short vowel (ex لُون، يَوْم).

البيت بيتك!

تَمْرِين ٣ - Listen and repeat the numbers in عامية from 1 to 10.

١ واحد

٢ تْنين

٣ تَلاتة

٤ أَرْبعة

٥ خَمْسَة

٦ سِتَّة

٧ سَبْعَة

٨ ثَمَانَة

٩ تِسْعَة

١٠ عَشْرَة

❖ ثَمَانَة , ثَلَاثَة : (long vowel - consonant - °consonant)

تَمْرِين ٤ -

Work in pairs with three different classmates. Ask them their phone number and write it below here

شو رقم تِلْفُونِك؟

تَمْرِين ٥ - Listen and repeat the numbers in **عامية** from 11 to 19.

١١	ايد عَش / اَحْدَ عَش
١٢	تَعَش
١٣	تَلْطَعَش
١٤	أَرْبَعَطَعَش
١٥	خَمْسَطَعَش
١٦	سِتَّ عَش
١٧	سَبْعَطَش
١٨	تَمْنَطَعَش
١٩	تِسْعَطَعَش

تَمْرِين ٦ -

Work in pairs. Listen to the numbers in mathematical sequence that your classmate reads out loud and add three numbers that follow the same sequence.

Ex.

- ١-٣-٥...
- ٧-٩-١١

١) ١٦، ١٢...	١) ١، ٢، ٣...	١) ١٩، ١٨، ١٧...
٢)	٢)	٢)

١) ٢، ٤، ٦...	١) ١٨، ١٥، ١٢...	١) ٣، ٦، ٩...
٢)	٢)	٢)

APPENDIX K Control Group - Lesson 3

تَمْرِي ١ - Work in pairs. Ask your classmate the information below-



الاسم:

الجنسية:

العمر:

بلد:

تَمْرِين ٢ - Listen and repeat the numbers in عامية from 20 to 30.

قَدِّيشْ عُمْرَكَ؟

قَدِّيشْ عُمْرَكَ؟ عُمْرِي ٢١ / ١٩

امْتِي عِيدِ مِيلَادِكَ؟

امْتِي عِيدِ مِيلَادِكَ؟ عِيدِ مِيلَادِي ب... ٧ شَهْر

MONTHS : There are two options you can use

شهر ١	كانون الثاني	شهر ٧	تموز
شهر ٢	شباط	شهر ٨	آب
شهر ٣	آذار	شهر ٩	أيلول
شهر ٤	نيسان	شهر ١٠	تشرين الأول
شهر ٥	أيار	شهر ١١	تشرين الثاني
شهر ٦	حزيران	شهر ١٢	كانون الأول

تَمْرِين ٣ -










Ask your classmate information about themselves and introduce them to the rest of the class

اسمه / اسمها _____

عمره / عمرها _____

عيد ميلاده ب / عيد ميلادها ب _____

تَمْرِين ٤ - Fill in the table with the missing question words.

English		عامية
What		_____
From		_____
Where		_____
How		_____
When		_____
Who		_____
Why		_____
How many		_____
How much		_____

تَمْرِين ١ - الحَيَاة اليَوْمِيَّة

أنا بِالْعَادَةِ بَعْدَ الْجَامِعَةِ بَرُجَعُ، بَدْرُسُ، بَعْدِينَ بَطَّلَعُ (I go out)
مَعَ رِفْقَاتِي عَلَى الْقَهْوَةِ، بَحْكِي (I speak) مَعَهُنَّ، مَنْتَمَشِّي، هَيْكَ
الْحَيَاةِ.

بِأَمْرِيكَ النَّاسَ طُولَ الْوَقْتِ مَسْتَعْجِلِينَ، بِيَرْكُضُوا كَثِيرًا.
أَمَّا النَّاسُ بِسُورِيَا عَلَى مَهْلَهُنَّ.

List the verbs according to their subject

هِنَّ

نَحْنُ

أَنَا

تَمْرِين ٢ - حَيَاتِي اليَوْمِيَّة

بِالْعَادَةِ يَبْلُشُ شُغْلِي السَّاعَةَ تِسْعَةَ فَبَطَّلِعُ مِنَ الْبَيْتِ السَّاعَةَ ثَمَانَةَ أَوْ
السَّاعَةَ سَبْعَةَ وَالذَّوَامَ بِيَسْتَمِرُّ تَقْرِيْبًا لِّلْسَاعَةِ تَنْتِيْنِ أَوْ السَّاعَةَ ثَلَاثَةَ.
وَالْمَسَاءِ أْحْيَانًا بَطَّلِعُ عَلَى الْقَهْوَةِ أَوْ عَلَى الْكَافِيْتَرِيَا، بِرُوحِ
عَلَى السِّيْنِمَا أَوْ بِطَّلِعُ مِشْوَارِ.

New vocabulary

بِالْعَادِ = Usually

بَلَّشْ، يَبْلُشُّ = to start

لَ = until

السَّاعَةَ تَنْتِي = 2 o'clock

بِيْرُوْحِ = to go

Describe your typical day: - تَمْرِين ٣

Read the following nouns out loud : - تَمْرِين ٤

صَدِيق، وَقْت، فَوْق، تَقْرِيْباً شَرْق، وَرَق، أَوْرَاق، رَقْم، أَرْقَام،
رَفِيق، رِفَاق (رِفَقَات)

Work in pairs and play battleship: - تَمْرِين ٢

Student A

٣٠									
٢٠									
١٠									
	١	٢	٣	٤	٥	٦	٧	٨	٩

Student B

٣٠									
٢٠									
١٠									
	١	٢	٣	٤	٥	٦	٧	٨	٩

قَدِّيش الساعة؟



- الساعة واحدة
- الساعة تَنْتَيْنِ
- الساعة تَلَاثَة
- الساعة أَرْبَعَة
- الساعة خَمْسَة
- الساعة سِتَّة
- الساعة سَبْعَة
- الساعة ثَمَانَة
- الساعة تِسْعَة
- الساعة عَشْرَة
- الساعة أَحَد عَشْر
- الساعة ثَنْعَش

- ٢:٢٠ (١)
- ٧:٥٠ (٢)
- ٣:٣٥ (٣)
- ٨:٥٥ (٤)
- ١:٤٥ (٥)

Work with a classmate and
in turns ask each other the time

- ١:٠٠ (١)
- ٨:٣٠ (٢)
- ٥:٤٥ (٣)
- ٤:١٥ (٤)
- ٩:٥٥ (٥)

تَمْرِين ٤ - Days of the week

امتی؟ ...امتی الدرس/ دَوامک؟



This week's schedule

دَرْسُ العَرَبِي (٩:٠٠)

(٧:٠٠)

الرَّحْلَةُ



دَرْسُ الإنكليزي

(٤:٤٥)

العَطْلَةُ

(كُلِّ اليَوْمِ)

الاجْتِمَاع

(١١:٣٠)

(١٠:٠٠)

الحَفْلَةُ

الحَفْلَةُ الموسيقية (٨:٠٠)

شُمال

غَرْب



شَرْق

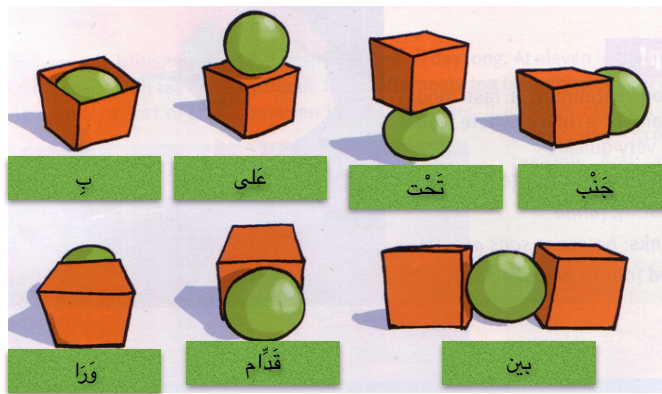
جُنُوب

ب + شُمال / جُنُوب / غَرْب / شَرْق + (اسم)
ب + جُنُوب غَرْب + (اسم)
ب + جُنُوب شَرْق + (اسم)



لندن بجنوب إنجلترا
إكزتر بجنوب غرب إنجلترا

ب + الشُمال / الجُنُوب / الغَرْب / الشَرْق



IV) Language Tests

APPENDIX 0 Pre-Test

Written part

Ex. 1

A new student has joined your Arabic module and introduces himself to the class.
Read the presentation and fill in the form with the information provided.

إِسْمِي سَامِي حَسَنٌ وَأَنَا طَالِبٌ جَدِيدٌ فِي هَذِهِ الْجَامِعَةِ. أَنَا مِنَ الْمَغْرِبِ، وَلِدْتُ وَكَبِرْتُ فِي مَدِينَةِ
الرَّبَاطِ وَلَكِنِّي أَسْكُنُ فِي مَدِينَةِ إِكْزِيْتِرِ الْآنَ. عُمُرِي تِسْعَ عَشْرَةَ سَنَةً. فِي الْحَقِيقَةِ، سَابُلُغُ عَشْرِينَ
سَنَةً فِي الشَّهْرِ الْقَادِمِ: عِيدَ مِيلَادِي فِي يَوْمِ ٢٠ مَارْسِ/ آذَارِ. لِي أَخٌ وَأُخْتٌ:
أَخِي يَسْكُنُ هُنَا فِي أَنْكَلْتِرَا بَيْنَمَا أُخْتِي تَسْكُنُ فِي الْمَغْرِبِ مَعَ أَهْلِهَا، زَوْجُهَا وَبِنْتُهَا. ذَهَبْتُ إِلَى
الرَّبَاطِ السَّنَةَ الْمَاضِيَةَ وَزُرْتُ أَبْنَاءَ أَعْمَامِي وَكُلَّ أَقَارِبِي.

Name:

Surname:

Nationality:

Age:

Birthday:

Where is the student from?

Write as much information as you can about the members of his family:

Ex. 2

Match the following nouns in English with their corresponding terms in Arabic

English	Arabic
May	تسعة
Twenty-eight	يوم الاثنين
Thirty-six	يناير/ كانون الثاني
Saturday	خمسة عشرة
Nine	يوم الثلاثاء
Tuesday	أيار/ مايو
Monday	ستة وثلاثون
Fifteen	يوم السبت
January	ثمانية وعشرون


Ex. 3

Write the opposite of the following words

- 1) غَدًا ≠
- 2) قَرِيب ≠
- 3) طَوِيلَة ≠
- 4) سَهْل ≠
- 5) كَبِيرَة ≠
- 6) جَدِيد ≠
- 7) يَمِين ≠

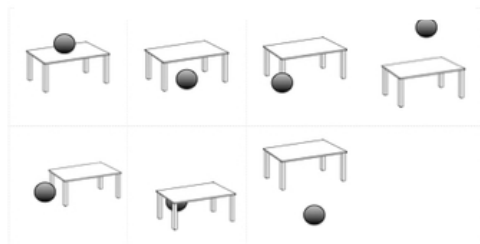
Ex. 4

Fill in the gaps by adding the missing letters

	شال	
غر—		شر—
	نبـ	

Ex. 5

Translate the following prepositions



- 1)
- 2)
- 3)
- 4)
- 5)
- 6)
- 7)

Ex. 6

Conjugate the following verbs into the present tense (المضارع المرفوع) and translate the conjugated verb in English.

Subject	Verb	Present Tense	Translation
أنا	درَّسَ		
أنتَ	لعبَ		
أنتِ	عملَ		
هو	شربَ		
هي	قالَ		
نحن	سافرَ		
أنتم	أكلَ		
هم	قرأَ		

Oral part

Greetings



صباح الخير - صباح النور

مساء الخير - مساء النور

Introductions



مرحبا

كيف حالك؟

ما اسمك؟

كم عمرك؟

من أين أنت؟

متى عيد ميلادك؟

كم الساعة؟

Classroom – The students describe the illustrations below





Wh-words - students listen to recorded questions, individuate the Wh-words they hear and translate them in their mother tongues.

ما اسمك؟

كم عمرك؟

من أين أنت؟

كم الساعة؟

كم لغة تتكلم؟

متى عيد ميلادك؟

أي يوم اليوم؟

لماذا؟

Written part

Ex. 1

Listen to the recording in Standard Arabic and write the verbs you recognise in Colloquial Arabic.

Keys: يياكل; ييقراً; ييشغل; ييلعب; بيدرس.

Ex. 2

Listen to twelve sentences in Colloquial Arabic and translate them into English:

- 1) How old are you?
- 2) It's 10:45.
- 3) The pen is on the table.
- 4) I am 25
- 5) Today is Tuesday
- 6) I usually study until 6 pm
- 7) How are you?
- 8) The bottle is behind the door
- 9) I have a meeting at 4:00
- 10) My phone number is 03 763 550
- 11) I read in the evenings
- 12) The chair is between the table and the door

Ex. 3

Read the text below on the right-hand side of the page and fill in the gaps using the words provided in the left-hand side.

			سارة: مَرَحَبًا
			محمد:
مَرَحَبَتَيْنِ	سلام	صباح الخير	_____
			كيف حالك؟
			سارة: ماشي الحال. كيف حالك أنت؟
			محمد:

			أخبارك؟
			سارة: كل _____ تمام
			محمد:
			_____ بتعملي _____؟
			سارة:
			_____ البيت _____
			الساعة _____
			عندي كثير _____ اليوم
			عشان تَقْدِمي (because of my presentation)
			عن مدينة نابولي وهو تقريبا هيك:
			تقع نابولي في _____ إيطاليا
			والمدينة على _____ الأبيض
			المتوسط. يبلغ سكان المدينة حوالي مليون نسمة
			ولذلك هي أكبر مدن جنوب إيطاليا.
			نابولي مشهورة _____ لأنها مدينة
			تاريخية و _____ جدا... و أصل
			البيتزا "
			سارة: _____ تَقْدِمي؟
			محمد: ماشي، يمكن قصير شوي
			سارة: شكرا! _____ هلق
			وبرجع على المكتبة، بشوفك!
تمام	مش	حلو	
شو	شي	ما	
شي	شو	شيئًا	
شي	شو	شيئًا	
اليوم	يوم	البكرة	
برجع	بترجع	بيرجع	
على	إلى	ل	
التاسعة	تسعة	التسعة	
عمل	شغل	اشتغال	
جَنُوبِ	جَنُوبِ	جَنُوبِ	
البحر	بحر	الساحل	
دولياً	بالعالم	بالدنية	
جميلة	حلوة	حلوة	
كَيْفِ	كيف	شو	

بذهب	بروح	أذهب	محمد: بشوقك !
------	------	------	---------------

Ex. 4

Which Arabic variety (عامية or فصحي) combines the words in the following two lists? Translate accordingly:

A	B
شمال	على
أمام	يمين
في	سهل
كم	عمل
قنينة	ما (?)
حلو	مشغول

List A:

Variety:

Translation:

List B:

Variety:

Translation:

Ex. 5

Provide Colloquial Arabic vocabulary that corresponds to the following in Standard Arabic:

جميلة

عمل

لماذا

يسار

هم

Oral part

- 1) A teacher and a student meet in the school canteen. The student greets the teacher and have a quick conversation: how does the student greet the teacher? Which fixed expressions is the student going to use? How is the teacher going to respond? Why?
- 2) Two students meet in the school's canteen. They greet each other and have a quick conversation. Which fixed expressions are they going to use? Why?
- 3) Explain the phonological and form patterns (CA) of the sentence
'لئش ما بتحب الصيف?' (Why don't you like summer?).
- 4) Answer the following questions:

شو اسمك؟

قديش عمرك؟

من وين أنت؟

امتى عيد ميلادك؟

شو بتدرس/ي

قديش الساعة

- 5) How would you self-assess your use of Colloquial Arabic? What are your most common mistakes? Can you describe your progress over the weeks of instruction of the experimental language course?
- 6) Read the following text and when you are ready answer the questions below:

فاطمة إنكليزية من أصل عربي. والد فاطمة أردني من عمان ووالدتها إنكليزية من

لندن. تدرس فاطمة اللغات في الجامعة مع صديقتها سارة: اللغة الروسية واللغة الفرنسية. تسافر فاطمة إلى الأردن كل سنة في الصيف وأحيانا في عطلة عيد الميلاد أيضا. فاطمة تحب السفر إلى الأردن وزيارة عائلته.

6a) Please retell the text in Colloquial Arabic;

6b) Extrapolate two recurring patterns of code-switching, one phonological and one grammatical (For example, the difference in pronunciation of the consonant ق /q/ and the different prepositions in the two varieties that follow the verb سافر (to travel);

6c) use the two patterns of change detected in the text, and make two original sentences in CA, each sentence respectively containing one pattern of change.

Bibliography

- ABDEL-JAWAD, H. R. 1986. The Emergence of an Urban Dialect in the Jordanian Urban Centres. *International Journal of the Sociology of Language*, 61, 53-63.
- ABDEL-MALEK, Z. 1972. The Influence of Diglossia on the Novels of Yuusif al-Sibaa'l. *Journal of Arabic Literature*, 3, 132-141.
- ABU-ABSI, S. 1986. The 'Simplified Arabic' of Iftah Ya Simsim: Pedagogical and Sociolinguistic Implications. *Al-'Arabiyya*, 24, 111-121.
- ABU-ABSI, S. 1990. A Characterization of the Language of Iftah Ya Simsim: Sociolinguistics and Educational Implications for Arabic. *Language Problems and Language Planning*, 14, 33-46.
- ABU-MELHIM, A. R. 1991. Code-Switching and Linguistic Accommodation in Arabic. *In: COMRIE, B. & EID, M. (eds.) Perspectives on Arabic Linguistics : Papers from the Annual Symposium on Arabic Linguistics. Volume III: Salt Lake City, Utah 1989.* Amsterdam/ Philadelphia: John Benjamins Publishing.
- ACTFL 2012a. Performance Descriptors for Language Learners. Alexandria, VA.
- ACTFL 2012b. Proficiency Guidelines
- AL-BATAL, M. 1992. Diglossia Proficiency: The Need for an Alternative Approach to teaching. *In: ROUCHDY, A. (ed.) The Arabic Language in America.* Detroit, Michigan: Wayne State University Press.
- AL-BATAL, M. 1995. *The Teaching of Arabic as a Foreign Language: Issues and Directions.* Provo, Utah: American Association of Teachers of Arabic (AATA).
- AL-BATAL, M. 2002. Identity and Language Tension in Lebanon: the Arabic of Local News at LCBI. *In: ROUCHDY, A (ed) Language Contact and Language Conflict in Arabic,* London: Routledge Curzon, 91-115.
- AL-BATAL, M. B., KIRK R. 2006. The Teaching and learning of Arabic in the United States: Realities, Needs, and Future Directions. *In: WAHBA, K., TAHA Z.A. & ENGLAND, LIZ (ed.) Handbook for Arabic Language Teaching Professionals in the 21st Century* Mahwah, N.J. : Lawrence Erlbaum Associates.
- ALBIRINI, A. 2011. The Sociolinguistic Functions of Code-Switching Between Standard Arabic and Dialectal Arabic. *Language in Society*, 40, 537-562.
- ALBIRINI, A. 2016. *Modern Arabic Sociolinguistics.* Oxford: Routledge.

- ALLEN, R. 1992. Teaching Arabic in the United States: Past Present and Future. In: ROUCHDY, A. *The Arabic Language in America*. Detroit: Wayne State University Press, 222-249.
- ALLPORT, A. 1988. What Concept of Awareness? In: MARCEL, A.J. & BISIACH, E. (eds.) *Consciousness in Contemporary Science*. London: Clarendon Press, 159-186.
- ALOSH, M. 1992. Designing a Proficiency-oriented Syllabus for Modern Standard Arabic as a Foreign Language. In: Rouchdy, A. *The Arabic Language in America*. Detroit: Wayne State University Press, 222-249.
- ALOSH, M. 1997. *Learner, Text, and Context in Foreign Language Acquisition*. Columbus, Ohio, National Foreign Language Resource Center (Ohio State University).
- ALOSH, M. 2009. The Arabic Language Continuum: at Which Point to Start? NECTL Review, 64, 53-58.
- AL-SHARKAWI, M. 2010. *The Ecology of Arabic. A Study of Arabicization*. Leiden: Brill.
- AL-WER, E. 2002. Education as a Speaker Variable. In: ROUCHDY, A. (ed.) *Language Contact and Language Conflict in Arabic: Variations on a Sociolinguistic Theme*. London: Routledge Curzon.
- AL-WER, E. 2007. The Formation of the Dialect of Amman: from Chaos to Order. In: MILLER, C., AL-WER, E., CAUBET, D. & WATSON, J. C. E. (eds.) *Arabic in the City: Issues in Dialect Contact and Language Variation*. London: Routledge.
- AL-WER, E. 2014. Language and Gender in the Middle East and North Africa In: EHRLICH, S., MEYERHOFF, M. & HOLMES, J. (eds.) *The handbook of language, gender and sexuality*. Oxford Wiley-Blackwell.
- ALBIRINI, A. 2015. *Modern Arabic sociolinguistics : diglossia, codeswitching, attitudes and identity*, New York, NY Routledge.
- ALTOMA, S. 1969. *The Problem of Diglossia in Arabic: a Comparative Study of Classical and Iraqi Arabic* Cambridge, MA, Harvard University Press.
- AMARA, M. 2006. Teaching Arabic in Israel. In: WAHBA, K., TAHA Z.A. & ENGLAND, LIZ (ed.) *Handbook for Arabic language teaching professionals in the 21st century* Mahwah: N.J: Lawrence Erlbaum Associates.

- AMER, F.H., ADAILEH, B.A. & RAKHIEH, B.A. 2011. Arabic Diglossia: a Phonological Study. *Argumentum* 7, 19-36.
- ANDERSON, L. W. and KRATHWOHL, D. R. 2001. *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Allyn & Bacon: Boston, MA. Pearson.
- ANGHELESCU, N. 1974. Arabic Diglossia and its Methodological Implications, *Romano-Arabica*, Bucarest, pp. 82-83.
- BADAWI, S. A. 1973. *Mustawayat al-Lugha al-'Arabiyya al-Mu'asira fi Misr*, Cairo, Dar al-Ma'arif.
- BASSIOUNEY, R. 2006. *Functions of Code Switching in Egypt: Evidence from Monologues*, Leiden, BRILL.
- BASSIOUNEY, R. 2009. *Arabic Sociolinguistics*, Edinburgh, Edinburgh University Press.
- BELNAP, R.K. & BISHOP, B. 2003. Arabic Personal Correspondence: a Window on Change in Progress? *International Journal of the Sociology of Language*, 163, 9–25.
- BENTAHILA, A. & DAVIES, E. E. 1991. Constraints on Code-switching: a Look Beyond Grammar. *In: Papers for the symposium on code-switching in bilingual studies: theory, significance and perspectives*. Strasbourg: European Science Foundation, 369~403.
- BLANC, H. 1960. Style Variations in Arabic: A Sample of Interdialectal Conversation. *In* FERGUSON, C. (ed.), *Contributions to Arabic Linguistics*. Cambridge, MA: Harvard University Press. 81-156.
- BLAU, J. 1977. *The Beginning of the Arabic Diglossia: A Study on the Origins of Neoarabic*. Malibu, CA: Undena Publications.
- BLOOM, B.S. and KRATHWOHL, D. R. 1956 *Taxonomy of Educational Objectives: The Classification of Educational Goals, by a committee of college and university examiners*. Handbook I: Cognitive Domain. NY, NY: Longmans, Green.
- BOUSSOFARA-OMAR, N. 2003. Revisiting Arabic Diglossic Switching in Light of the MLF Model and its Sub-Models: the 4-M Model and the Abstract Level Model. *Bilingualism: Language and Cognition*, 6, 33-46.
- BOUSSOFARA-OMAR, N. 2005. Neither Third Language Nor Middle Variety but Diglossic Switching. *Journal of Arabic Linguistics* 45, 55-80.

- BRUSTAD, K., AL-BATAL, M., & AL-TONSI, A. 2004. *Al-kitaab fii ta^callum al-^carabiyya. A Textbook for Beginning Arabic, Volume I* (Third Edition). Washington, D.C.: Georgetown University Press, 2011.
- CANALE, M. & SWAIN, M. 1980. Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing. *Applied Linguistics*, 1, 1980, 1–47.
- CORRIENTE, F. 1976. *T From Old Arabic to Classical Arabic Through the Pre-Islamic Koine: Some Notes on the Native Grammarians' Sources, Attitudes and Goals*. Manchester: Manchester University Press.
- COWELL, M. W. 1964. *A Reference Grammar of Syrian Arabic, Based on the Dialect of Damascus* Washington, D.C., Georgetown University Press.
- CRESWELL, J. W., PLANO, C, GUTMANN, M. L. & HANSON, W. E. 2003. Advanced mixed methods research designs. In TASHAKKORI A. & TEDDLIE C. (Eds), *Handbook on mixed methods in the behavioral and social sciences*. Thousand Oaks, CA: Sage Publications. 209-240.
- EID, M. 1982. The non-randomness of diglossic variation in Arabic, *Glossa*, 16, 54-84.
- EID, M. 1988. Principles of Code-Switching Between Standard and Egyptian Arabic. *Al-'Arabiyya*, 21. 51-79.
- EISELE, J. C. 2002. Approaching Diglossia: Authorities, Values and Representations. In: Rouchdy A. (ed) *Language Contact and Language Conflict in Arabic*. London: Curzon, 3-23.
- EISELE, J. C. 2006. Developing Frames of Reference for Assessment and Curricular Design in a Diglossic L2: From Skills to Tasks (and Back Again). In: WAHBA, K. M., TAHA, Z.A. & ENGLAND, L. (eds.) *Handbook for Arabic language teaching professionals in the 21st century*. Mahwah: N.J.: Lawrence Erlbaum Associates.
- ELGIBALI, A. 1993, Stability and Language Variation in Arabic: Cairene and Kuwaiti Dialects In (eds.) Eid, M. and Holes, C. *Perspectives on Arabic Linguistics* Philadelphia & Amsterdam: John Benjamins. 75-96.
- ELLIS, N.C. 1997, Sequencing in SLA: phonological memory, chunking and points of order, *Studies in Second Language Acquisition*, 18, 91-126.
- ELLIS, N.C. 2006. Cognitive perspectives on SLA: The associative-cognitive CREED. *AILA Review*, 19, 100-121.

- ELLIS, R. 2001. Introduction: Investigating Form-Focused Instruction. *A Journal of Research in Language Studies*, 51, 1-46.
- EI-HASSAN, S.A. 1977. Educated Spoken Arabic in Egypt and the Levant: A Critical Review of Diglossia and Related Concepts, *Archivum Linguisticum*, 8/2, 112-132.
- ENNAJI, M. 2007. Arabic Sociolinguistics and Cultural Diversity in Morocco. In: BENMAMOUN, E. (ed.) *Perspectives of Arabic Linguistics XIX: Papers from the Nineteenth Annual Symposium on Arabic Linguistics*. Amsterdam: John Benjamins.
- EUROPE, C.O. 2001. Common European Framework of Reference for Languages.
- EVIATAR, Z. I., IBRAHIM, R. 2000. Bilingual is as Bilingual Does: Metalinguistic Abilities of Arabic-speaking Children. *Applied Psycholinguistics* 21. 451–471.
- FAKHRI, A. 1995. Arabic as a Foreign Language: Bringing Diglossia into the Classroom. In: HAGGSTROM, M. A. M., LESLIE ZARKER; WIECZOREK, J. A. (ed.) *The Foreign Language Classroom: Bridging Theory and Practice*. New York: Garland.
- FARGHALY, A. 2005. A case for an inter-Arabic grammar. In: ELGIBALI, A. (ed.) *Investigating Arabic: Current Parameters in Analysis and Learning*. Boston: Brill.
- FASOLD, R. 1984 *The Sociolinguistics of Society*. Oxford: Blackwell.
- FERGUSON, C. 1959. Diglossia. *Word*, 15, 325-340.
- FERGUSON, C. 1959b. The Arabic Koine. *Language*, 35 (4), 616-630.
- FERGUSON, C. 1971. *Language structure and language use: Essays by Charles Ferguson*. Stanford: Stanford University Press.
- FERGUSON, C. 1996. *Sociolinguistic Perspectives*. Oxford: Oxford University Press.
- FISHMAN, J. 1967 Bilingualism with and without diglossia: diglossia with and without bilingualism. *Journal of Social Issues*, 23, 29-38.
- GARDNER-CHLOROS, P. 2009. *Code-switching*, Cambridge, UK Cambridge University Press.

- GIOLFO, M. & SINATORA, F. 2011. Rethinking Arabic Diglossia. Language Representations and Ideological Intents. In: VALORE, P. (ed.) *Multilingualism. Language, Power and Knowledge*. Pisa: Edistudio.
- GIOLFO, M. & SALVAGGIO, F. (In print). A Digitally Assisted Model of Integration of Standard and Colloquial Arabic Based on the Common European Framework. In: AL-BATAL, M. (ed.) *Arabic as One: Integrating the Dialect within the Arabic Curriculum*. Washington, D.C.: Georgetown University Press.
- GIOLFO, M., SALVAGGIO, F. (in print). *Mastering Arabic Variation. A Common European Framework of Reference integrated model*. ROMA: Aracne Editrice.
- GODFROID, A., HOUSEN, A., & BOERS F. 2010. A procedure for testing the Noticing Hypothesis in the Context of Vocabulary Acquisition. In PUTZ M. & SICOLA L. (Eds.), *Cognitive Processing in Second Language Acquisition*. Amsterdam: Benjamins. 169–197.
- GRANDE, 2012. Arabic Language Teaching and Valorization of Roots: the Italian Experience. Mother tongue and intercultural valorization: Europe and its migrant youth. Franco Angeli, 123-147.
- HADDAD, Y.Y. 2004. *Not Quite American? The Shaping of Arab and Muslim identity in the US*. Baylor University Press. Waco, Texas.
- HAERI, N. 1991. *Sociolinguistic Variation in Cairene Arabic: Palatalization and the Qaf in the Speech of Men and Women*. Philadelphia: Univeristy of Pennsylvania.
- HAERI, N. 1991. Sociolinguistic Variation in Cairene Arabic: Palatalization and the Qaf in the Speech of Men and Women. Philadelphia: Univeristy of Pennsylvania.
- HAERI, N. 2003. *Sacred Language, Ordinary People: Dilemmas of Cultures and Politics in Egypt.*, New York: Palgrave Macmillan.
- HARY, B. 1996. The Importance of Language Continuum in Arabic Multiglossia. In: ELGIBALI, A., (ed.) *Understanding in Arabic Essays in Contemporary Arabic Linguistics in honor of El-Said Badawi*. Cairo: American University in Cairo Press. 69-90.
- HAUGEN, E. I. 1972. *The Ecology of Language*. Redwood City: Stanford University Press.
- HAYWOOD, J. A. & NAHMAD, H. M. 1998. *A new Arabic grammar of the written language* London Lund Humphries.

- HEE-MAN, S. E.-K., F. 2006. Teaching Arabic in Korea. *In: WAHBA, K., TAHA Z.A. & ENGLAND, LIZ (ed.) Handbook for Arabic language teaching professionals in the 21st century* Mahwah: N.J.: Lawrence Erlbaum Associates.
- HOLES, C, 1986. The Social Motivation for Phonological Convergence in Three Arabic Dialects. *International Journal of the Socioligy of Language*,61, 33-51.
- HOLES, C. 1993. The Uses of Variation: a Study of the Political Speeches of Gamal Abd al-Nasir. *In: Eid M. & Holes C. Perspectives on Arabic Linguistics V*, Amsterdam: Benjamins, 13-45.
- HOLES, C, 1995. Community, Dialect and Urbanisation in the Arabic-Speaking Middle East. *Bulletin of the School of Oriental and African Studies*,58(2), 270-287.
- HOLES, C. 2004. *Modern Arabic: Structures, Functions, and Varieties*. Washington, D.C., Georgetown University Press.
- HUDSON, A. 1980. *Sociolinguistics*. Cambridge: Cambridge University Press.
- HUDSON, A. 2002. Outline of a Theory of Diglossia. *International Journey of the Sociology of Language*, 157, 1-48.
- HUSSEIN, R. & EL-ALI, N. 1989. Subjective Reactions of Rural University Students Toward Different Varieties of Arabic. *Al-'Arabiyya*, 22 (1-2), 37-54.
- IBRAHIM, M. H. 1986. Standard and Prestige Language: A Problem in Arabic Sociolinguistics. *Anthropological Linguistics*, 28, 115-126.
- IBRAHIM, R. & AHARON-PERETZ, J. 2005. Is Literary Arabic a Second Language for Native Arab Speakers?: Evidence from Semantic Priming Study. *Journal of Psycholinguistic Research*, 34, 51-70.
- KAYE, A. S. 1970. Modern Standard Arabic and the colloquials, *Lingua*, XXIV, 374-91.
- KAYE, A. 1994. Formal vs. Informal in Arabic: Diglossia, Triglossia, Tetraglossia, etc., Polyglossia – Multiglossia Viewed as a Continuum. *Zeitschrift für arabische Linguistik*, 47-66.
- KHAMIS-DAKWAR, R. & FROUD, K. 2007. Lexical Processing in Two Language Varieties: An event-related brain potential study of Arabic native speakers. *In: MUGHAZY, M. (ed.) Perspectives on Arabic linguistics XX* Amsterdam: John Benjamins.
- KHAMIS-DAKWAR, R. & FROUD, K. 2014. Neurocognitive Modeling of the Two Language Varieties in Arabic Diglossia. *In: KHAMIS-DAKWAR, R. & FROUD,*

- K. (eds.) *Perspectives on Arabic linguistics XXXI*. Amsterdam: John Benjamins.
- KRATHWOHL, D. R. 2002. A Revision of Bloom's Taxxonomy. *Theory into Practice*, 41. Ohio State University. Retrieved at <http://www.depauw.edu/files/resources/krathwohl.pdf>
- LABOV, W. 1972. *Language in the Inner City: Studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press.
- LABOV, W. 1973. The boundaries of words and their meanings. In BAILEY, J. & SHUY, R. (eds.), *Ways of analysing variation in English*. Washington, DC: Georgetown University.
- LEOW, R. P. 1997. Attention, Awareness, and Foreign Language Behavior. *A Journal of Research in Language Studies*, 47(3), 467–505.
- LEOW, R. 2000. A Study of the Role of Awareness in Foreign Language Behavior: Aware versus Unaware Learners. *Studies in Second Language Acquisition*, 22(4), 557-584.
- LIDDICOAT, M., LENANNE, R., ABDUL RAHIM, I. 2012. *Syrian Colloquial Arabic: A Functional Course*. Mary-Jane Liddicoat Publisher. Available online at: www.syrianarabic.com (Last accessed on December 20th 2017).
- LIGHTBOWN, P.M. & SPADA, N. 2006. *How Languages are Learned*. Oxford: Oxford University Press.
- LYSTER, R. 1994. The Effect of Functional-Analytic Teaching on Aspects of French Immersion Students' Sociolinguistic Competence. *Applied Linguistics*, 15 (3),263–287.
- OWENS J. 2000. *Arabic as a Minority Language*. Berlin: Mouton De Gruyter.
- OWENS J. 2001. Arabic Sociolinguistics. *Arabica*, XLVIII, 419-469.
- MAAMOURI, M. (1998). *Language Education and Human Development: Arabic diglossia and its Impact on the Quality of Education in the Arab region*. Discussion Paper Paper prepared for The World Bank Mediterranean Development Forum, Marrakesh, Philadelphia: University of Philadelphia Literacy Institute.
- MACKEY, A. & PHILIP, J. 1998, Conversational interaction and second language development: recasts, responses, and red herrings?. *Modern Language Journal*, 82, (3), 338-356.

- MAGIDOW, A. 2013. Competing Pressures in Diglossia: Avoiding Colloquial Elements in Writing Modern Standard Arabic. *Journal of Arabic and Islamic Studies*, 13, 146-168.
- MANFREDI, S. 2012. Dialect Mixing and Dialect Levelling in Kordofanian Baggara Arabic. In BARONTINI, A., PEREIRA, C., VINCENTE, A., & ZIAMARI, K. (eds.), *Dynamiques langagières en Arabophonies : variations, contacts, migrations et créations artistiques. Hommage offert à Dominique Caubet par ses élèves et collègues*. Universidad de Zaragoza, Colección Estudios de Dialectología Árabe, 7, 141-162.
- MANSOURI, F. 2000. Grammatical Markedness and Information Processing in the Acquisition of Arabic [as] a Second Language. Munich: Lincom Europa.
- MARÇAIS, W. 1930. "La Diglossie Arabe", *L'Enseignement Public*, 97, 401-409.
- MAZRAANI, N. 1997. Aspects of Language Variation in Arabic Political Speech-Making, Richmond, Curzon.
- MEJDELL, G. 2006. *Mixed Styles in Spoken Arabic in Egypt*. Leiden: Brill.
- MEJDELL, G. 2012. Diglossia, Code Switching, Style Variation, and Congruence - Notions for Analyzing Mixed Arabic. *Al-'Arabiyya*, 44-45, 29- 39.
- MEISELES, G. 1980. Educated Spoken Arabic and the Arabic Language Continuum. *Archivum Linguisticum*, 11, 118-143.
- MILLER, C. 2004. Variation and Change in Arab Urban Vernaculars. In: Haak et al. (eds) *Approaches to Arabic Dialects*, 177-206.
- MILLER, C. 2006. Juba Arabic. In: VERSTEEGH, K., EID, M., ELGIBALI, A., WOIDICH, M. & ZABORSKI, A. (eds.) *Encyclopedia of Arabic Language and Linguistics* Leiden: Brill.
- MILLER, C. 2007. Arabic Urban Vernaculars: Development and Change. In: MILLER, C., AL-WER, E., CAUBET, D. & WATSON, J. C. E. (eds.) *Arabic in the City: Issues in Dialect Contact and Language Variation*. London: Routledge.
- MILROY, J. 2001. Language Ideologies and the Consequences of Standardization. *Journal of Sociolinguistics*, 5, 530-555.
- MITCHELL, T. 1975. Some Preliminary Observations on the Arabic Koine. *Bulletin of the British Society for Middle Eastern Studies* 2(2), 70-86.

- MITCHELL, T. 1986. What is Educated Spoken Arabic?. *International Journal of the Sociology of Language*, 61, 7-32.
- MELKA, F. 1997. Receptive vs. Productive Aspects of Vocabulary, *In: SCHMITT, N. & MCCARTHY, M. (eds.) Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- MYERS-SCOTTON, C. 1993. *Social motivations for codeswitching: Evidence from Africa*, Oxford: Oxford University Press.
- NATION, I.S.P. & LAUFER, B. 2012. Vocabulary. *In: GASS, S. & MACKEY, A. (eds) The Routledge Handbook of Second Language Acquisition*. New York: Routledge. 163-176.
- NATION, I.S.P. 2012. Teaching language in use: Vocabulary. *In EISENMANNE M. & SUMMER, T. (Eds.) Basic Issues in EFL Teaching and Learning*. University Winter Verlag, 103-116.
- NATION, I.S.P. 2013. Vocabulary acquisition in second language acquisition. *The Encyclopedia of Applied Linguistics*, ed. Chapelle, C.A. Oxford: Wiley-Blackwell.
- NORRIS, J.M. & ORTEGA, L. 2000. Effectiveness of L2 Instruction: A Research Synthesis and Quantitative Meta-analysis. *A Journal of Research in Language Studies*, 50(3), 417–528.
- O'LEARY, D. L. 1923. *Arabic History and Culture*. London: Routledge.
- OWENS, J. 2006. *A Linguistic History of Arabic*. Oxford: Oxford University Press.
- PALMER, J. 2007. Arabic Diglossia: Teaching Only the Standard Variety Is a Diservice to Students. *Arizona Working Papers in SLA & Teaching*, 14, 111-122.
- PALMER, J. 2008. Arabic Diglossia: Student Perceptions of Spoken Arabic After Living in the Arabic-Speaking World. *Arizona Working Papers in SLA & Teaching*, 15, 81-95.
- PALVA, H. 2006. Dialects: Classification. *In: VERSTEEGH, K., EID, M., ELGIBALI, A., WOIDICH, M. & ZABORSKI, A. (eds.) Encyclopedia of Arabic Language and Linguistics*. Leiden: Brill.
- PARKINSON, D. 1991. Searching for Modern Fusha: Real-Life Formal Arabic. *Al-'Arabiyya*, 24, 31-64.

- PARKINSON, D. B. 1992. Good Arabic: Ability and Ideology in the Egyptian Arabic Speech Community. *Language Research*, 28, 225-253.
- PARKINSON, D. B. 1996. Variability in Standard Arabic Grammar Skills. In: ELGIBALI, A., (ed.) *Understanding in Arabic Essays in Contemporary Arabic Linguistics in honor of El-Said Badawi*. Cairo: American University in Cairo Press, 91-101.
- RYDING, K. C. 1991. Proficiency Despite Diglossia: A New Approach for Arabic. *Modern Language Journal* 75, 212-218.
- RYDING, K. C. 2005. *A Reference Grammar of Modern Standard Arabic*, Cambridge, Cambridge University Press.
- RYDING, K. C. 2006. Teaching Arabic in the United States. In: WAHBA, K., TAHA Z.A.; ENGLAND, L. (ed.) *Handbook for Arabic Language Teaching Professionals in the 21st Century*. Mahwah: N.J: Lawrence Erlbaum Associates.
- RYDING, K. C. 2013. *Teaching and Learning Arabic as a Foreign Language : A Guide For Teachers* Washington, DC Georgetown University Press.
- SAEED, A. 1997. *The Pragmatics of Code-Switching. Fusha Arabic to Aammiyyah Arabic in Religious Oriented Discourse*. (Unpublished Doctoral Dissertation). Muncie, IN: Ball State University.
- SAIDAT, A. 2010. Language Attitude: The Case of Jordan. *International Journal of Academic Research*, 2, 235-243.
- SCHMITT, N. & MCCARTHY, M. 1997 *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- SAVILLE-TROIKE, M. 1989. The Ethnographic Analysis of Communicative Events. In SAVILLE-TROIKE, M (ed.) *The Ethnography of Communication: an Introduction*. Oxford: Blackwell. 107-139.
- SHIRI, S. 2002. Speak Arabic Please! Tunisian Arabic Speakers' Linguistic Accommodation to Middle Easterners. In: ROUCHDY, A. (ed.) *Language Contact and Language Conflict in Arabic*. London: Routledge Curzon.
- SNOW, D. 2013. Revisiting Ferguson's Defining Cases of Diglossia. *Journal of Multilingual and Multicultural Development* 34, 61-76.
- SOLIMAN, A. 2008. *The Changing Role of Arabic in Religious Discourse: A Sociolinguistic Study of Egyptian Arabic*. (Unpublished Doctoral Dissertation). College Park: University of Maryland.

- SOLIMAN, R. 2014. Arabic Cross-dialectal Conversations: A missing Element in the Teaching of Arabic as a Second Language. *In: AGUILAR, V., MANZANO, M. Á., PÉREZ CAÑADA, L. M., SALEH, W. & SANTILLÁN GRIMM, P. (eds.) ARABELE 2012.* Murcia: Universidad de Murcia.
- SULEIMAN, Y. 1999. *Language and Society in the Middle East and North Africa: Studies in Variation and Identity.* Surrey: Curzon.
- SULEIMAN, Y. 2003. *The Arabic Language and National Identity: a Study in Ideology.* Edinburgh: Edinburgh University Press.
- SULEIMAN, Y. 2004. *A War of Words: Language in Conflict in the Middle East.* Cambridge: Cambridge University Press.
- SULEIMAN, Y. 2011 *Arabic, self and identity: A study in conflict and displacement.* Oxford: Oxford University Press.
- TAHA, Z. 2006. Toward Pragmatic Competency in Arabic. *In: KASSEM M. WAHBA, TAHA Z.A. & ENGLAND., L. (eds.) Handbook for Arabic language teaching professionals in the 21st century.* Mahwah, N.J.: Lawrence Erlbaum Associates, 353-361.
- TAHA, Z. 1995. The Grammar Controversy: What to Teach and Why?. *IN (ed.) The Teaching of Arabic as a Foreign Language: Issues and Directions.* Al-'Arabiyya Monograph Series No: 2, Salt Lake City: The American Association of Teachers of Arabic, University of Utah 175-183.
- THEODOROPOULOU, I. & TYLER, J. 2014. Perceptual Dialectology of the Arab World. *Al-'Arabiyya*, 47, 21-39.
- TOMLIN, R., & VILLA, V. 1994. Attention in Cognitive Science and Second Language Acquisition. *Studies in Second Language Acquisition*, 16(2), 183-20.
- TRENTMAN, E. 2011. L2 Arabic Dialect Comprehension: Empirical Evidence for the Transfer of Familiar Dialect Knowledge to Unfamiliar Dialects. *L2 Journal* 3, 22-49.
- VAN MOL, M. 2003. *Variation in Modern Standard Arabic in Radio News Broadcasts: A Synchronic Descriptive Investigation Into the Use of Complementary Particles.* LEUVEN: PEETERS PUBLISHERS
- VAN PATTEN, B. & CADIerno, T. 1993. Explicit Instruction and Input Processing. *Studies in Second Language Acquisition*, 15(2), 225-243

- VAN PATTEN, B. & CADIerno, T. 1993b. Input Processing and Second Language Acquisition: A Role for Instruction. *The Modern Language Journal*, 77(1).45-57.
- VAN PATTEN, B., WILLIAMS J. & ROTT, S. 2004 Form-Meaning Connections in Second Language Acquisition. *Lawrence Erlbaum Associates*, 1-27.
- VAN PATTEN, B. & FERNANDEZ, C. 2003 Processing Instruction: Theory, Research, and Commentary. *Lawrence Erlbaum Associates*, 277-293.
- VERSTEEGH, K. 2001. *The Arabic Language*. Edinburgh: Edinburgh University Press.
- WAHBA, K. 1996. Linguistic Variation in Alexandrian Arabic: the Feature of Stress. In: ELGIBALI, A. (ed.) *Understanding Arabic: Essays in Contemporary Arabic Linguistics in Honor of El-Said Badawi*. Cairo: American University in Cairo Press.
- WAHBA, K. 2006. Arabic Language Use and the Educated Language User. In: KASSEM W., TAHA Z. A. & ENGLAND, L. (eds.) *Volume 36 of Handbooks of Linguistics and Communication Science*. Mahwah: N.J.: Lawrence Erlbaum Associates.
- WALTERS, K. 1996. Diglossia, linguistic variation, and language change in Arabic. In: *Perspectives on Arabic Linguistics. Papers from the Annual Symposium on Arabic Linguistics. Volume VIII*. Amherst: Massachussets.
- WALTERS, K. 2003. Fergies Prescience: the Changing Nature of Diglossia in Tunisia. *International Journal of the Sociology of Language*, 163, 77-109.
- WEHR, H. 2012. *A Dictionary of Modern Written Arabic*, New York, Snowball Publishing.
- WENINGER, S. 2011. *The Semitic Languages: An International Handbook*. Berlin; Boston: De Gruyter Mouton 2011.
- WEBB, S. & Nation, I.S.P. 2013. Teaching vocabulary. In *The Encyclopedia of Applied Linguistics*, Oxford, UK: Wiley-Blackwell.
- WESELY, P. 2013. Describing and Illustrating Data in Mixed Methods Research. *The Encyclopedia of Applied Linguistics*. 1–7.
- WILMSEN, D. 2010. Dialects of Written Arabic: Syntactic Differences in the Treatment of Object Pronouns in Egyptian and Levantine Newspapers. *Arabica*, 57, 99-128.

YOUNES, M. 2006. Integrating the Colloquial with Fusha in the Arabic-as-a-Foreign-Language Classroom. *In: KASSEM M. WAHBA, TAHA Z.A. & ENGLAND., L. (eds.) Handbook for Arabic language teaching professionals in the 21st century.* Mahwah, N.J.: Lawrence Erlbaum Associates.

YOUNES, M. 2015. *The Integrated Approach to Arabic Instruction*, Oxon, Routledge.

ZUGHOUL, M. R. 1980. Diglossia in Arabic: Investigating Solutions. *The Integrated Anthropological Linguistics* 22, 201-217.