

**Digital Vernaculars: An Investigation of Najdi Arabic in  
Multilingual Synchronous Computer-Mediated  
Communication**

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

**2012**

**Ebtesam S Alothman**

School of Languages, Linguistics and Cultures

Linguistics and English Language

## TABLE OF CONTENTS

LIST OF FIGURES.....	10
LIST OF TABLES.....	12
TRANSLITERATION SYSTEMS.....	14
LIST OF ABBREVIATIONS I.....	15
LIST OF ABBREVIATIONS II.....	16
ABSTRACT.....	17
DECLARATION.....	18
COPYRIGHT STATEMENT.....	19
ACKNOWLEDGMENT.....	20
DEDICATION.....	21
THE AUTHOR.....	22
<b>CHAPTER ONE.....</b>	<b>23</b>
<b>INTRODUCTION.....</b>	<b>23</b>
1.1 Preliminaries.....	23
1.2 Multilingual CMC: A selective history.....	24
1.3 Statement of the problem.....	30
1.4 Aims of the study.....	32
1.5 Research questions.....	33
1.6 Overview of methodology.....	34
1.7 Significance of the study.....	35
1.8 Limitations.....	36
1.8 Organization of the thesis.....	37
<b>CHAPTER TWO.....</b>	<b>41</b>
<b>COMPUTER-MEDIATED COMMUNICATION.....</b>	<b>41</b>
2.1 Computer-mediated communication.....	41
2.1.1 Definition of computer-mediated communication.....	41
2.1.2 Internet Relay Chat (IRC).....	43
2.1.3 Virtual community.....	47

<b>CHAPTER THREE.....</b>	<b>51</b>
<b>METHODOLOGY.....</b>	<b>51</b>
3.1 Najdi Arabic documentation.....	52
3.1.1 Roman-Morpho Syntactic (RMS) Dialect Survey.....	52
3.1.2 Informants.....	55
3.2 Computer-Mediated Discourse Analysis Approach.....	57
3.2.1The “IRC Language Questionnaire”.....	61
3.2.1.1 Internet-Mediated Research.....	62
3.2.1.2 The “IRC Language Questionnaire”.....	65
3.2.1.3 Pilot study.....	68
3.2.1.4 Strategy of sampling.....	69
3.2.1.5 Validity and reliability.....	71
3.2.1.6 Data analysis.....	72
3.2.2 Discourse-analytic method.....	73
3.2.2.1 Setting.....	73
3.2.2.2 Data collection.....	77
3.2.2.3 Data presentation.....	78
3.2.2.4 Data analysis.....	79
3.3 Ethical considerations.....	80
3.5 Summary.....	82
<b>CHAPTER FOUR.....</b>	<b>84</b>
<b>NAJDI ARABIC.....</b>	<b>84</b>
4.1 The Arabian Peninsula: Historical background.....	84
4.2 Najd.....	86
4.3 Najdi Arabic and Modern Standard Arabic.....	89
4.4 Phonology.....	92
4.4.1 Vowels.....	92
4.4.2 Consonants.....	93
4.5 Morphology.....	97
4.5.1 Nominal morphology.....	98
4.5.1.1 Noun class types.....	98
4.5.1.2 Number.....	100

4.5.1.3 Gender.....	101
4.5.1.4 Definiteness.....	101
4.5.1.5 Indefinites pronouns.....	102
4.5.1.6 Personal pronouns.....	104
4.5.1.7 Object pronouns.....	106
4.5.1.8 Demonstratives.....	107
4.5.1.9 Numerals.....	108
4.5.1.10 Quantifies.....	110
4.5.1.11 Adjectives.....	110
4.5.1.12 Adverbs.....	112
4.5.1.13 Local relations.....	113
4.5.1.14 Possessions.....	114
4.5.1.15 Interrogatives.....	115
4.5.2 Verbal morphology.....	116
4.5.2.1 Verb derivation.....	116
4.5.2.2 Verb inflection classes.....	117
4.5.2.3 Tense.....	119
4.5.2.4 Modality.....	120
4.5.2.5 Aspect.....	120
4.5.2.6 Negation.....	121
4.5.2.7 Modal verbs.....	123
4.6 Syntax.....	126
4.6.1 Simple sentence.....	127
4.6.1.1 Declarative clauses.....	127
4.6.1.2 Interrogative clauses.....	128
4.6.1.3 Construct.....	129
4.6.1.4 Modification.....	130
4.6.1.5 Conditional clauses.....	131
4.6.2 Complex sentence.....	132
4.7 Conclusion.....	134

<b>CHAPTER FIVE.....</b>	<b>139</b>
<b>IRC PARTICIPANTS: SOCIAL DIMENSIONS AND ATTITUDES TOWARDS USING NAJDI ARABIC ONLINE.....</b>	<b>139</b>
5.2 Background CMC research.....	140
5.3 Results of the “IRC Language Questionnaire” .....	143
5.3.1 Participants’ social profiles.....	144
5.3.2 Participants’ internet usage.....	145
5.3.3 Participants’ choice of script and language.....	147
5.3.4 Social context of Romanized NA in CMC.....	155
5.3.5 Participants’ attitudes towards the use of NA online.....	159
5.4 Conclusion.....	168
<b>CHAPTER SIX.....</b>	<b>171</b>
<b>ORTHOGRAPHIC REPRESENTATION OF NAJDI ARABIC ONLINE.....</b>	<b>171</b>
6.1 Orthography in CMC.....	172
6.1.1 Conceptual basis of orthography: A terminological framework.....	172
6.1.2 The sociolinguistics of writing systems.....	175
6.1.2.1 The development of new writing systems.....	175
6.1.2.2 Digraphia.....	177
6.1.2.3 Sociolinguistic issues of writing.....	182
6.1.3 Adaptation of writing systems in CMC.....	183
6.1.4 Romanized Arabic systems.....	186
6.2 Methodology.....	188
6.3 Results: Romanized NA in IRC.....	189
6.3.1 Romanized NA: Consonants.....	189
6.3.1.1 One phoneme-to-one grapheme correspondence.....	189
6.3.1.1.1 Digraphs.....	193
6.3.1.1.2 The pharyngealized voiced interdental fricative / d <sup>ʕ</sup> /.....	193
6.3.1.1.3 The dental-alveolar affricates /ts/and /dz/.....	194
6.3.1.1.4 Geminates.....	194
6.3.1.2 One phoneme –to- two or three graphemes correspondence.....	195

6.3.1.2.1 The voiced uvular fricative /ɣ/.....	196
6.3.1.2.2 The voiced velar stop /g/.....	196
6.3.1.2.3 The voiced dental fricative /ð/.....	198
6.3.1.2.4 The voiceless uvular fricative /x/.....	198
6.3.1.3 Two phonemes –to- one grapheme correspondence.....	200
6.3.1.4 No phoneme –to- one grapheme correspondence.....	201
6.3.2 Romanized NA: Vowels.....	202
6.3.2.1 Short vowels.....	203
6.3.2.2 Long vowels.....	204
6.3.2.3 Trends of vowel representation in Romanized NA.....	206
6.4 Conclusion.....	207
<b>CHAPTER SEVEN.....</b>	<b>212</b>
<b>CODE-SWITCHING AND SCRIPT-SWITCHING IN NAJDI IRC.....</b>	<b>212</b>
7.1 Code-switching and CMC.....	215
7.1.1 Code-Switching: A terminological framework.....	216
7.1.1.1 Code.....	216
7.1.1.2 Code-switching.....	217
7.1.2 Code-switching vs. borrowing.....	218
7.1.3 Analytical approaches to code-switching.....	222
7.1.3.1 Linguistic approaches to code-switching.....	222
7.1.3.2 Sociolinguistic approaches to code-switching.....	224
7.1.3.2.1 Gumperz: Situation-Metaphor—We code and they code.....	224
7.1.3.2.2 Myers-Scotton’s Markedness Model .....	226
7.1.3.2.3 Auer’s Conversationalist Approach.....	227
7.1.4 Code-switching in CMC.....	233
7.2 Methodology.....	240
7.3 Morphological integration of borrowings.....	242
7.3.1 Morphological integration of borrowings: Roman script.....	243
7.3.1.1 The definite article.....	243
7.3.1.2 Prepositions.....	245
7.3.1.3 Pluralisation.....	245
7.3.1.4 Possessive constructions.....	246

7.3.1.5 Verb conjugations.....	247
7.3.2 Morphological integration of borrowings: Arabic script.....	248
7.4 The communicative functions of Najdi IRC code-switching.....	252
7.4.1 Discourse-related code-switching.....	252
7.4.1.1 Code-switching for qualification.....	253
7.4.1.2 Code-switching for specifying an addressee.....	255
7.4.1.3 Code-switching of tags.....	255
7.4.1.4 Code-switching for reiteration.....	257
7.4.1.5 Code-switching for conversational coherence.....	258
7.4.1.6 Code-switching for personalization vs. objectiveness.....	259
7.4.1.7 Code-switching for requesting attention.....	261
7.4.1.8 Code-switching for quotation.....	263
7.4.1.9 Code-switching of institutional terms.....	264
7.4.1.10 Code-switching for signalling topic shift.....	266
7.4.1.11 Code-switching for self-repairs.....	268
7.4.1.12 Code-switching for indicating dispreference.....	269
7.4.2 Participant-related code-switching.....	271
7.4.2.1 Preference-related code-switching.....	271
7.4.2.2 Competence-related code-switching.....	272
7.5 The communicative functions of script-switching in IRC.....	273
7.5.1 Discourse-related functions of script switching.....	276
7.5.1.1 Script-switching for signaling change of mode.....	276
7.5.1.2 Script-switching for reiteration.....	280
7.5.1.3 Script-switching for topic shifting.....	282
7.5.1.4 Script-switching for requesting attention.....	284
7.5.2 Participant-related functions of script switching.....	286
7.5.2.1 Script-switching for inclusion and exclusion of IRC participants.....	286
7.5.2.2 Competence-related script-switching.....	288
7.6 Conclusion.....	289

<b>CHAPTER EIGHT.....</b>	<b>294</b>
<b>THE SPECIFIC CHARACTERSTICS OF NAJDI IRC LANGUAGE.....</b>	<b>294</b>
8.1 The relationship between speech and writing: A theoretical framework.....	297
8.2 The language of CMC: A mixed medium of communication.....	303
8.2.1 The language of CMC between speech and writing.....	305
8.2.2 The specific characteristics of CMC language.....	307
8.3 Methodology.....	311
8.4 Results.....	313
8.4.1 E- abbreviation.....	313
8.4.1.1 Acronyms of sentence.....	315
8.4.1.2 Phonetic acronyms.....	316
8.4.1.3 Abbreviations.....	317
8.4.1.4 Letter homophones.....	318
8.4.1.5 Numerals.....	318
8.4.1.6 Combination of initial letters and numeral homophones.....	320
8.4.2 Addressitivity.....	321
8.4.3 Paralinguistic cues.....	322
8.4.3.1 Emoticons.....	322
8.4.3.2 Capitalization.....	325
8.4.3.3 Multiple punctuation.....	326
8.4.3.4 Reduplication.....	327
8.4.3.5 Vocalization.....	328
8.5 Conclusion.....	329
<b>CHAPTER NINE.....</b>	<b>333</b>
<b>CONCLUDING REMARKS AND IMPLICATION.....</b>	<b>333</b>
9.1 Summary.....	334
9.2 Implications.....	342
9.3 Suggestions for further study.....	346

<b>REFERENCES</b> .....	349
<b>APPENDICES</b> .....	365
Appendix A: The Arabic Version of the Roman1-Morph-Syntax Dialect Survey.....	365
Appendix B: Informed Consent Form & Voluntary Consent Form.....	404
Appendix C: The “IRC Language Questionnaire” .....	407
Appendix D: The most common abbreviations in English CMC.....	420
Appendix E: The most common English emoticons.....	421

Word count 78,308

## LIST OF FIGURES

Figure 2.1: Screenshot of IRC windows.....	45
Figure 3.1 Screenshot of Yahoo messenger.....	74
Figure 3.2 Screenshot of Yahoo messenger. The actual room conversation.....	75
Figure 3.3 Screenshot of homepage of www.ksavip.com.....	76
Figure 5.1 IRC participants' internet usage.....	146
Figure 5.2 IRC participants' frequency of chatting.....	147
Figure 5.3 Participants' choice of script.....	148
Figure 5.4 IRC participants' use of MSA in IRC.....	149
Figure 5.5 Participants' use of NA in IRC.....	150
Figure 5.6 Participants' use of both MSA and NA.....	151
Figure 5.7 Participants use of English only in IRC.....	152
Figure 5.8 Participants' use of both English and NA.....	153
Figure 5.9 Participants' code-switching to English.....	154
Figure 5.10 Participants' reasons for using Romanized NA.....	156
Figure 5.11 Where did Romanized NA come from?.....	158
Figure 5.12 Participants' use of Romanised NA in offline settings.....	159
Figure 5.13 Histogram for Belief 1.....	161
Figure 5.14 Histogram for Belief 2.....	162
Figure 5.15 Histogram for Belief 3.....	163
Figure 5.16 Histogram for Belief 4.....	164
Figure 5.17 Histogram for Belief 5.....	165
Figure 5.18 Histogram for Belief 6.....	166
Figure 5.19 Histogram for Belief 7.....	167
Figure 5.20 Histogram for Belief 8.....	168

Figure 6.1 : Representation of short and long vowels in Arabic..... 203

## LIST OF TABLES

Table 3.1 Five discourse analysis paradigms (Herring 2004: 17).....	59
Table 3.2 Profile of IRC data.....	78
Table 3.3 The categories of IRC data.....	78
Table 4.1 Short vowels.....	92
Table 4.2 Long vowels.....	93
Table 4.3 Najdi Arabic consonants.....	94
Table 4.4 Personal free pronouns in NA.....	98
Table 4.5 Bound forms of personal pronouns.....	105
Table 4.6 Demonstratives.....	108
Table 4.7 Cardinal numerals.....	108
Table 4.8 Ordinal numerals.....	109
Table 4.9 Adverbs.....	112
Table 4.10 Days of the week, seasons of the year.....	113
Table 4.11 Local prepositions.....	113
Table 4.12 Interrogatives.....	115
Table 4.13 The internal passive (perfective).....	116
Table 4.14 The internal passive (imperfective).....	116
Table 4.15 Perfective (strong verbs).....	117
Table 4.16 Perfective (weak verbs) .....	118
Table 4.17 Imperfective (strong verbs).....	118
Table 4.18 Imperfective (weak verbs) .....	118
Table 4.19 Imperative .....	120
Table 4.20 The Modal elements.....	123
Table 5.1 IRC participants' social profile.....	144

Table 6.1 Representation of Arabic sounds by a single Roman character.....	190
Table 6.2 Representation of Arabic sounds that do not have English equivalents by a single Roman characters.....	192
Table 6.3 Representation of Arabic sounds by more than one Roman representation..	195
Table 6.4 Representation of short vowels in Najdi IRC.....	204
Table 6.5 Representation of long vowels in Najdi IRC.....	205
Table 8.1 Common features of digital writing (Danet 2001:17).....	309
Table 8.2 The most common abbreviations and acronyms in Najdi IRC.....	320
Table 8.3 The most frequent emoticons in IRC (Alothman 2010).....	325
Table 8.4 The most frequent multiple punctuation marks in IRC.....	327
Table 8.5 Taxonomy of Najdi IRC features .....	332

## TRANSLITERATION SYSTEM

I followed the transliteration system used in Arabic dialectology. Below is an illustration of the symbols used in this thesis.

<b>Consonants</b>								
	labial	dental	dental-pharyngealized	palate-alveolar	velar	uvular	pharyngeal	glottal
VI stop	p	t	ṭ	č	k	q		
Vd stop	b	d	ḍ	dž	g	G		ʿ
Nasal	m	n						
Lateral		l			ɭ			
Trill		r						
VI fricative	f	θ			x		ħ	h
Vd fricative	v	ð	ḏ		ġ		ç	
VI sibilant		s	š	š				
Vd sibilant		z	ẓ	ž				
Semi-vowel	w			y				

## Vowels

ī		ū
i		u
	(ə)	ō
ē		o
e	a	ā

## LIST OF ABBREVIATIONS I

- 1:** first person
- 2:** second person
- 3:** third person
- ACC:** accusative
- COMP:** complementizer
- CONS:** construct state marker
- COP:** copula
- DEF:** definite article
- DM:** discourse marker
- ED:** Arabic Eḏāfe attribute marker
- F:** feminine
- INDEF:** indefinite article
- M:** masculine
- NEG:** negation
- PAST:** simple past (perfective)
- PL:** plural
- POSS:** possessive
- REL:** relative particle
- SG:** singular

## **LIST OF ABBREVIATIONS II**

**CMC:** Computer-Mediated Communication

**IRC:** Internet-Relay Chat

**MUDs:** Multi-User Dungeons

**MOOs:** Mud-Object Oriented

**WELL:** Whole Earth Lectronic Link

**BBS:** Bulletin Board System

**ASCII:** American Standard Code for Information Interchange

**IMR:** Internet-Mediated Research

**CA:** Classical Arabic

**MSA:** Modern Standard Arabic

**RASs:** Romanized Arabic Systems

**CMDA:** Computer-Mediated Discourse Analysis

**NLS:** New Literacy Studies

## **ABSTRACT**

The University of Manchester

Ebtesam S. Alothman

PhD Linguistics

### **Digital Vernaculars: An Investigation of Najdi Arabic in Multilingual Synchronous Computer-Mediated Communication**

**2012**

The present study is conducted within the borders of multilingual Computer-Mediated Communication (CMC) . It aims at investigating the orthographic representation of Najdi Arabic in Internet Relay Chat (IRC). Along with this basic purpose, the study examines the communicative functions of multilingual practices in representing Najdi Arabic online. These practices include code-switching between languages (English-Arabic) and script-switching between scripts (Arabic-Roman). In addition, it is within the intentions of this study to analyze the specific characteristics of Najdi Arabic online which make it a hybrid between written and spoken languages. The investigation starts with the hypothesis that Najdi Arabic online is a language form that shows interrelations of languages, scripts and features of CMC.

In order to realize these intentions, an online questionnaire was administered among internet users and a large corpus of IRC data was collected. Within Herring's (2004) computer-mediated discourse analysis approach, qualitative and quantitative approaches to data analysis were conducted. Responses to the online questionnaire were coded and analyzed in order to establish social profiles of who use Najdi Arabic online. IRC data was subjected to textual and interactional analysis to investigate the orthographic and linguistic features of this language form as well as the mechanism underlying code-switching and script-switching.

The analysis of data yielded a number of important findings. First, youth language creatively uses non-standard spellings in constructing anti-standard orthography that constitutes a cyberspace code of communication. Regarding the orthographic representation of Najdi Arabic, the analysis reveals interrelations among languages (Arabic/English) and ASCII-characters. The process of Romanizing Najdi Arabic involves a combination of both transliteration of the Arabic orthography and transcription of spoken Najdi Arabic. Second, this writing system functions as a code of communication that carries social and cultural meanings and it is argued that switching between scripts is not entirely arbitrary. Applying the methodological tools of Auer's (1995) Conversational Analysis to script-switching online, it is found that alternating between scripts online achieves different communicative functions that define and enhance the IRC discourse. Third, code-switching has been found to serve communicative functions similar to those of face-to face conversation and stimulated by the synchronicity of online discourse. Finally, investigating the specific features of Najdi Arabic online reveals another level of language alternation between common English CMC features identified in the literature and innovative CMC features based on the Arabic orthography.

## **DECLARATION**

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

## COPYRIGHT STATEMENT

- i.** The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and she has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.
- ii.** Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made **only** in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.
- iii.** The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproduction”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.
- iv.** Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see <http://www.campus.manchester.ac.uk/medialibrary/policies/intellectual-property.pdf>), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see <http://www.manchester.ac.uk/library/aboutus/regulations>) and in The University’s policy on presentation of Theses.

## ACKNOWLEDGMENT

Praise be to God the Exalted (Lord of the worlds) for his blessings that he bestowed upon me and favors that he granted unto me. After all due praise to God, the Almighty, thanks go to all those who have helped me in the various stages of writing this thesis. I am greatly indebted to numerous people who have contributed to the making of this work by assistance, sincere love, and support.

I owe an immeasurable debt of gratitude to my supervisor Professor Yaron Matras for his insightful guidance and scholarly contributions to this thesis. His sound advices, insightful assistance and valuable remarks contributed immensely toward the accomplishment of this thesis in its present form. Above all and the most needed, he provided me with constant encouragement and support in various ways. His past years of supervision and guidance have exceptionally inspired and enriched my growth as a student, a researcher and an academic. I am indebted to him more than he knows.

I am also privileged to have had the opportunity, over the last three years, to discuss issues of CMC at the professional level with many pioneers in the field. Among linguists who I have referred in the thesis are scholars who deserve special thanks, including Jannis Androutsopoulos and Susan Herrings for encouraging and providing with an initial opportunity to research Arabic CMC. For many hours of thoughtful-provoking discussion, I wish to thank Mark Sebba, Yukiko Nishimura, Tereza Spilioti, Simon Yates and John C. Paolillo.

My mother deserves special mention for her inseparable support and prayers. She is the person who sincerely raised me with her caring and love, showing me the joy of intellectual pursuit ever since I was a child. My deceased father showed me the example of a hard-working man. I am grateful to them both as well as my brothers and sisters.

Words fail me to express my appreciation to my husband, Ali, whose dedication, love and persistent confidence in me, has taken the load off my shoulder. I really appreciate his forbearance during the long periods of obsession and obliviousness that seemed to be a part of my writing this thesis. To him, I shall always be indebted. Heartfelt thanks are also due to my bilingual daughters, Almahat and Alhanouf, who shared with me countless hours of reading and writing. They were a constant source of happiness and inspiration which cheered my writing efforts

I extend my acknowledgement and sincere appreciation to all those who, at various times and in different places, have paved the way for me during the course of my study. My deep thanks go to Dr. Muhammad Alahmade, my administrative supervisor in the Royal Embassy of Saudi Arabia in London for his cooperation and encouragement. My gratitude is extended to Princess Nora Bint AbdulRuhman University for funding my PhD study.

Ebtesam AlOthman  
September 2012, Manchester UK

## **DEDICATION**

This thesis is lovingly dedicated to my mother, my dearest husband Ali, and to my bilingual children Almahat and Alhanouf.

The researcher

## THE AUTHOR

### ACADEMIC QUALIFICATIONS:

- First class Degree BA (Hons) in English Language, Princess Nora bint AbdulRuhman University, Saudi Arabia, Riyadh, 2000.
- M.A. in Applied Linguistics, Princess Nora bint AbdulRuhman University, Saudi Arabia, Riyadh, 2008.
- Diploma in Project Management and Quality Control, East End Computing & Business College UK, 2011

### RELEVANT EXPERIENCE

- Teaching Assistant at the Department of English Language. Princess Nora Bint AbdulRuhman University 2002-2007
- Lecturer at the Department of English. Princess Norah Bint Abdulrahman University 2008-now
- A member in Arabic Online Writing Project Organization
- A member in Saudi TESOL

### ACADEMIC PUBLICATIONS

Papers that have been accepted or still under review for publication include:

- Alothman, E. 2010. Computer mediated communication: New varieties of language. *Proceedings of 4th Saudi International Conference - SIC 2010*. University of Manchester , UK.
- Alothman, E. forthcoming. The impact of global communication on dialect use in Saudi Arabia: A linguistic investigation of Najdi Arabic in Internet Relay Chat. Paper presented at the Third Gulf Research Meeting (GRM), Cambridge, 11 - 14 July 2012, organized by the Gulf Research Center Cambridge (GRCC) at the University of Cambridge, UK.
- Alothman, E. forthcoming. Romanized Arabic: Past and present. *Proceedings of the 43th Seminar for Arabian Studies*. The British Museum, UK

# CHAPTER ONE

## INTRODUCTION

### 1.1. Preliminaries

With its rapid popularization, the internet has served various purposes since it was officially launched in the U.S.A in 1960s. The hundreds of computers connected to the internet in the 1960s increased astronomically over the following decades. Modern internet users not only search for information on the internet but bring new information to it as well. More pertinent to this study is the fact that people are increasingly using the internet as a means of communication. Troest (1998: 19) comments, “humankind has a tendency to seize every conceivable opportunity for communicating, no matter how obscure a medium”. This inclination towards using the internet for interaction lies at the heart of the emergence of Computer-Mediated Communication (henceforth, CMC). Most of the issues that have provoked heated debate regarding CMC involve the question of how the internet is rapidly changing the way we use language (Crystal 2001).

The language used on the internet has been regarded as a new form of language that is different from spoken and written languages and examined extensively for its idiosyncratic features (e.g., Crystal 2001, Herring 2001, 2002, Danet 2001, Segerstated 2002). Crystal defined internet language as “displaying features that are unique to the internet, and encountered in all the above [internet] situations, arising out of its character as a medium which is electronic, global, and interactive” (2001: 18). Numerous studies focused on the differences between CMC and real-life face-to-face interaction (December 1993, Collot & Belmore 1996, Yates 1996).

Scholarly research on CMC published in English has only recently begun to take into consideration the linguistic diversity of CMC (Danet & Herring 2007). In contrast to the huge body of research in linguistic, sociolinguistic, and pragmatic aspects of English CMC, there has been little research into multilingual CMC. Androutsopoulos (2006: 431) maintains that “similar publications” to English CMC “in other languages no doubt exist, and bringing their findings will be a major task of future scholarship in this area”. The following section is devoted to present a selective history on multilingual CMC that significantly supports this endeavour.

## **1.2 Multilingual CMC: A selective history**

Stressing the need for publications on linguistic diversity of CMC, Androutsopoulos (2006: 428) pointed out that:

“English dominated the internet landscape of the 1990s in terms of both the native language of estimated users and the language of available websites, but more recent years have witnessed a rapid increase in linguistic diversity, with the majority of users and websites today using a language other than English.”

The term “multilingualism”, then, came into existence to designate the linguistic diversity that Androutsopoulos indicated at in the above lines. With the remarkable increase of internet use among populations communicating online in other languages, the term “multilingual internet” has been coined by CMC scholars (Danet

& Herring 2007: 1) to refer to innovative practices in languages other than English. From the time the multilingual internet came into existence, three broad issues have been addressed in the literature: adaptation of writing systems online in non-English speaking environments, language choice in multilingual internet content, and the specific characteristics of text-based CMC compared to spoken and written communication (Danet & Herring 2007).

The most prominent issue tackled by CMC researchers is the use and adaptation of the writing systems to different modes of CMC. CMC researchers have investigated online transcription of languages with different writing systems when adapted to the restrictions of ASCII (American Standard Code for Information Interchange) (Danet & Herring 2007). In the early days of the internet, non-English speakers were suffering from limitations of digital encoding possibilities that hindered both word processing and communication. Danet and Herring (2007: 8) maintain that “because early planners of the internet were North Americans and sought only to facilitate communication in English, they did not anticipate problems encountered by speakers of other languages trying to communicate online”.

The impact of ASCII protocol on the representation of non-Roman languages is manifested in the Romanized transliteration of native scripts such as Greek, Arabic and Persian online (e.g., Warschauer *et al.* 2002, Palfreyman and al Khalil 2003, Tseliga 2007) despite the development of Unicode text-transmission protocol. Such representations are characterized by innovative correspondences between non-Roman languages and ASCII Roman characters by the use of both Roman alphabets and numerals. In order to find out more about these practices, studies have been carried out on Japanese (Nishimura 2008), Taiwanese (Su 2007), Arabic (Palfreyman

& Al Khalil 2007), Greek (Androutsopoulos 1998, 1999, 2000, 2001, Tseliga 2007) and Chinese (Yang 2007).

Not only standard languages are used in CMC but also regional and non-standard varieties are on the increase. It has been argued that the online languages share both spoken and written linguistic features, which means that although communication in CMC is achieved through writing (i.e., typing on the keyboard) the language produced, especially in informal online contexts, resembles that of speech (Collot & Belmore 1996, Crystal 2001, Werry 1996). Investigation into the transcription of non-Roman based languages online went further to include the use of regional and non-standard varieties. The innovative practices of transliteration persist when using regional and non-standard varieties that lack standard writing systems online. In this context, Warschauer *et al.* (2007) declared that text-based CMC “fosters written communication” in non-standard dialects that are used mainly for oral communication (2007: 304). In investigating Egyptian Arabic as used by Egyptian university students, Warschauer found that Romanized Egyptian Arabic (along with English) is widely used in text-based CMC.

Similarly, a number of regional linguistic features of other non-standard dialects were found to be used in text-based CMC (Androutsopoulos & Ziegler, 2004). Androutsopoulos (2010) maintains that the traditional variationist focus on phonology is replaced by the analysis of written representations of phonological variation between standard and dialect or formal and casual style. Themistocleous (2008) investigated the orthography in online text-based communication among young Greek Cypriots using Cypriot Greek, a variety of Greek spoken in the Island of Cyprus. The use of Mauritian Creole online was investigated by Rajah-carrim (2008) who found that the use of the internet has an impact in shaping the official

orthography of Mauritian Creole. Commenting on the impact of electronic communication upon the standardization of Mauritian Creole, Rajah-Carrim (2008: 223) argues that “it is possible that increased usage of Creole in new technologies will lead to the emergence of orthographic conventions among users of the language without the intervention of language policy-makers”.

The aforementioned cases demonstrate the fact that when engaging in CMC, participants have a variety of codes that can be used. These codes include the official standard language, English (being the lingua franca of the internet), and the regional and non-standard variety. The choice between these codes has been mapped out in an extensive investigation among numerous CMC researchers in different CMC modes (e.g., Paolillo 1996, Androutsopoulos 2007, Warschauer et al. 2007). The main thrust of these studies was to investigate how the process of alternation between different codes occurs, and the functional reasons of why bilingual speakers engage in code-switching online.

As far as Modern Standard Arabic (henceforth, MSA) and Arabic dialects are concerned, very few studies have investigated the phenomenon of code-switching in CMC. Through examining online communication among Egyptian Arabs, Warschauer and his colleagues found a tendency among Egyptian speakers to mix Romanized Egyptian Arabic with MSA (Warschauer *et al.* 2002). Their analysis reveals that English is used predominantly in formal Bulletin Board Systems (BBS) postings and formal e-mails, whereas Egyptian Arabic written in Roman characters is used extensively in informal interaction (e.g., informal e-mails and IRC chat). In 2007, Palfreyman and al-Khalil carried out an empirical study among female university students in the United Arab Emirates. Their data was based on a corpus of instant messenger conversations supplemented by e-mail surveys of participants’

experience with this structure of writing. Their analysis reveals that “there was a fair amount of code-switching (changing mid-utterance or mid-sentence from one language to another) and code-mixing (using words or phrases from one language within sentences in the other language)” (2007: 53).

A different issue that is addressed by CMC researchers is identifying the specific characteristics of text-based CMC as a hybrid between spoken and written communication. At the onset of English CMC research, the focus of investigation was identifying idiosyncratic characteristics of text-based CMC in different modes in terms of addressivity, abbreviations, prosody, and gestures (e.g., Werry 1996, Crystal 2001, Herring 2001, 2002, Danet 2001, Segerstedt 2002). The main premise that underlies this body of research is the fact that these specific characteristics of text-based CMC “are emerging as a direct reflection of the physical constraints on the medium combined with a desire to create language that is as ‘speech-like’ as possible” (Werry 1996: 48).

In order to find out more about the features of internet languages other than English, studies have been conducted on Chinese internet language (Yu *et al.* 2001, Gao 2006). Other investigations have been done within a Japanese online context. For example, Nishimura (2007) investigated linguistic innovations and interactional features of casual online communication in online Japanese context. Using (BBS) messages as the primary sources of data, Nishimura investigated the uses of *kanji* and examined the incorporation of informal spoken features as the use of final particles. Young Japanese BBS users are found to employ colloquial language online and use such spoken features as final particles in order to interact with other users online as if they were in face-to-face conversation. She also found that the varying degrees of style mixture in Japanese text-based CMC are related to the extent to

which the users feel involvement and closeness to each other. Segerstedt (2002) investigated the characteristics of “e-style” in different modes of CMC among Swedish internet users. He proposed a taxonomy of these characteristics that includes unconventional abbreviations, non-alphabetical graphical means (such as emoticons), non-standard spelling and punctuations, as well as colloquial lexicon and dialect expressions.

Investigation into CMC among bilingual users of the internet demonstrates that the internet constitutes an increasingly multilingual communicative environment. Contrary to the early history of the internet, English is no longer the lingua franca of the internet. The dominance of the English language in the professional scene and the lack of software support for non-Roman languages at the early days of the internet promoted speakers of the Arabic language to develop a new form of the Arabic language and new forms of writing that adapt to their language using the Roman script (Warschauer *et al.* 2002).

The present study aims at presenting an empirical investigation of a regional variety online: Najdi Arabic (henceforth, NA) online. Specifically, this study aims at an investigation of synchronous CMC as used by Najdi users engaged in IRC. When investigating the myriad of issues within multilingual CMC, the use of NA online offers unique insights. Not only does NA require a non-Roman writing system, but it also lacks an official standard writing system and represents a diglossic linguistic situation (Ferguson 1959). NA online will be investigated in terms of the three issues that have marked the history of multilingual CMC: the orthography, code-switching, and the specific features of Najdi text-based CMC.

### 1.3 Statement of the problem

The present study presents an account of how NA is used online. The short extract below which is taken from Najdi IRC interactions shows a sample of the type of the language investigated in this study<sup>1</sup>.

A: hi everybody, w r finally in riyadh. s there any meeting 2morrow?

B: 7emd ullah 3ala salamatkum

C: r u going to meet el director?

A: I honestly do not know who s.... I will go 2morrow check.

B:Allah y3een. dr khaled replied my email saying they r still working on deciding a new director.

C: ahlaan :o :o :o :o

B: I really don't know :

A: >>> C how s ur son?

C: 7emd ullah...he is doing great; he turned 1 two days ago.

C: How are the girls?

A:walla b5eer they went back to school

A: esta3adeeeti ??

B ??????????

A: استعديتي؟

B: marrrrra :OO

A: do not worry.. el interview just to know you

B: ان شالله

B: الله يسهل

The above extract, which includes a mixture of English and NA written in both Roman and Arabic scripts, represents the main features of NA online. This language, whose distinctive features are the use of Roman script and excessive use of code-switching, is a language Najdi internet users prefer using online compared to other

---

<sup>1</sup> IRC participants' nicknames have been replaced with capital alphabets (A, B, C, etc.) for the sake of maintaining participants' confidentiality.

language choices available to them (for example, English only, MSA only, or NA in Arabic script). The language investigated in this study (NA online) is composed of three main resources: languages (English, Arabic), writing systems (Arabic script, Roman script) and specific features of the medium (e.g., emoticons and punctuations). These three resources highlight the nature of text-based CMC as being multilingual, multiscriptual, and multimodal.

It is important to investigate who uses this form of NA and why. It is equally important to investigate the attitudes and beliefs towards the use of this language form. An important aim of this study is making social profiles of IRC participants in terms of their social background, their linguistic and literacy choices and their attitudes towards these choices.

Using the Roman script to represent NA is the first and clearest sign of NA online. This emergent form of Romanized Arabic is different from conventional Romanized Arabic Systems (RASs) in making use of ASCII characters and numerals. Therefore, it is important to investigate how consonants and vowels of NA are represented in Roman script. It is also significant to investigate patterns of script-switching in an effort to understand motivations and functions underlying the mechanism of script-switching in the IRC discourse.

Beside the use of Roman script, an important feature that marks this kind of online language form is the tendency to freely code switch between English and NA. It seems that the ideology of 'correct' writing (Milroy & Milroy 1991: 66) that inhibits language alternation in written languages is less powerful online. What facilitates this written code switching in this particular kind of writing is that boundaries which separate speech from writing are eliminated in what Tagliamonte

and Denis (2008: 8) consider as “the emerging tendency for written genres to be more like speech, a process referred to as colloquialization”.

This tendency to be more like written speech is revealed by the use of a variety of specific features of text-based CMC that make it a hybrid medium which shares features of both spoken and written communication. Some of these features (e.g., abbreviations) are developed to facilitate writing as quickly as possible and, thus, to approximate speech. Other features (e.g., emoticons and multiple punctuations) are designed to express paralinguistic features of spoken communication that are expressed by facial expressions, tone of the voice, and body movements in spoken discourse. It is, therefore, important to investigate the specific characteristics of Najdi IRC discourse that make it a hybrid between spoken and written language.

#### **1.4 Aims of the study**

Given the fact that NA online is a form of language different from spoken NA, and that it lacks a standardized writing system, a need arose to investigate how this non-standard dialect is represented in synchronous CMC. Based on the need for further research in this area, and the rationale given in the preceding section, this study aims to:

- (1) Introduce NA, which is the subject of the current investigation in terms of phonology, nominal morphology, verbal morphology, and the syntax of this dialect;
- (2) Establish the social profile of individuals who use NA online in terms of their social background, linguistic choices, and attitudes towards these choices;

- (3) Investigate how consonants and vowels of NA are represented by Roman script;
- (4) Trace patterns of code-switching and script-switching online used by Najdi internet users, and investigate the motivations and functions underlying the mechanism of code switching and script-switching in IRC interactions;
- (5) Analyze the specific characteristics of NA discourse online compared to spoken and written communication, and propose a systematic taxonomy of these characteristics in Najdi CMC; and
- (6) Give an interpretation of the results and provide implications on the use of NA in text-based CMC.

### **1.5 Research questions**

In particular, this study will attempt to answer the following questions:

- (1) What are the characteristics of spoken NA in terms of phonology, morphology, and syntax?
- (2) Who uses NA online ?
  - What are the general patterns of their linguistic and literacy choices online?
  - What are the general patterns of attitudes the IRC participants hold concerning the use of NA online?
- (3) What are the orthographic features of NA representation online?
  - What are the orthographic features of Romanized NA online? Specifically, how are consonants and vowels represented in Roman script?

(4) What are the communicative functions of switching between different codes?

- How are borrowed English items integrated into both Roman and Arabic script?
- What are the communicative functions of code-switching to English/NA?
- What are the communicative functions of script-switching to Arabic/Roman script?

(5) What are the specific features of Najdi IRC that make it a hybrid between spoken and written language?

## **1.6 Overview of methodology**

In order to address the aims posed in this study, three main procedures are followed for the collection and analysis of data. Below is a brief outline of the methodological framework within which the present study is conducted, while more details on the methodological tools of data collection and analysis are given in the relevant chapters.

- An exploratory survey is used with five Najdi speakers to present a systematic description of NA, the subject of the current investigation, in terms of phonology, morphology, and syntax.
- In order to present a social profile of IRC participants, an online questionnaire is designed and utilized to investigate the IRC participants'

social profiles and their linguistic and literacy practices on line, as well as their attitudes towards these practices.

- IRC data is collected in an electronic format. The target of investigating this CMC data is the observable linguistic and orthographic features, as well as patterns of code switching. Analytical tools of language-focused content analysis of Herring's (2004) *Computer-Mediated Discourse Analysis Approach* (henceforth, CMDA) are utilized throughout the investigation.

### **1.7 Significance of the study**

This study aims at presenting an empirical systematic account of the representation of NA online. A study of the representation of a non-standard Arabic dialect online has significant contributions to fill gaps in different areas of research. Despite the fact that scholarly investigation into multilingual CMC is growing in different directions, there still remains a justified need for more academic research describing the representation of different languages in CMC. Observing this phenomenon, Danet and Herring (2007: 3) made the following comment that relates to the purpose of the present study:

“In today’s multilingual, global world, people are communicating on the internet not only in its established lingua franca, English, but also in a multitude of other languages...To date, however, the research literature in English on computer-mediated communication has focused almost exclusively on emergent practices in English, neglecting developments within populations communicating online in other languages.”

Research tackling the linguistics of text-based CMC, which is the focus of the present study, has been limited in comparison to other sociological and psychological areas of the CMC domain. With the exception of a few studies (e.g., Nishimura 2007, Su 2007, Yang 2007), most of CMC research analyzes the use of English online. Thus, there remains much to be done in regard to examining the use of languages other than English online.

There is also a shortage of research addressing the use of non-standard and regional varieties online (Warschauer *et. al.* 2002, Androutsopoulos & Ziegler 2004, Palfreyman & Al Khalil 2007). This study, therefore, is an attempt to investigate the use of an Arabic dialect on line—an area that has never been investigated before. Except for Palfreyman & Al Khalil's (2007) study of instant messaging among Arabic speakers in Dubai, and Warschauer and his colleagues' (2007) investigation of identity among Egyptian users of internet, there is no published study on CMC within the Arabic context. This study is the first to address issues of examining the linguistic features and orthographical representation of a Saudi dialect. Consequently, the results obtained from this study will enrich cross-cultural research on CMC as it is conducted within a novel context.

### **1.8 Limitations**

Though the present study contributes to the study of Arabic CMC, there are some limitations that need to be taken into consideration before making generalizations from the findings.

First of all, this study is concerned with investigating the use of NA in synchronous CMC (IRC). It aims at examining orthographic and linguistic practices as well as code switching practices in Najdi IRC. Since Najdi users are likely to use the same practices elsewhere in cyberspace (such as discussion forums, emails, Facebook interactions, and Bulletin Boards), it is likely that the research findings will be applicable elsewhere. However, the present study will make no such claim, and will instead assert that a wider research be conducted, perhaps even comparing the use of NA in different CMC modes.

In addition to limitations of space, there exists a more pressing limitation of time. Indeed, technology is progressing and developing at an alarmingly rapid speed. Crystal (2004: 224) claims that “any attempt to characterize the language of the internet, whether as a whole or with reference to one of its constituent situations, immediately runs up against the transience of technology”. Thus, any results of this research must be seen in the immediate framework of the time within which the data was collected.

## **1.9 Organization of thesis**

Chapter One gives an introduction to the study. This chapter presents a selective history of multilingual internet research and, thus, locates the present study in an ongoing process of researching text-based CMC in a non-English context. The issues addressed in this section underpin the scope and aims of the present study. This chapter also presents the statement of the research problem and the motivations underlying the present investigation. The statement of the problem identifies the

three key issues of this study: orthographic representation of NA online, code switching in Najdi IRC, and specific characteristics of Najdi IRC discourse.

Chapter Two presents an account of relevant facets of CMC in order to provide the reader with initial understanding of the phenomenon under investigation. The chapter begins with an introduction of the notion of CMC followed by an overview of its basic properties with emphasis being given to the context of IRC. Then, the theoretical background and foundation of the concept of “virtual community” will be reviewed.

Chapter Three outlines the methodological design of the study and the tools used for collecting and analyzing the data. It provides a description of the three tools used in conducting the study and procedures followed in using them. A description of the adapted Arabic version of the Roman- Morpho Syntax (RMS) dialect survey will be given. An account of Herring’s (2004) CMDA adopted in the present study will be presented. Then, the “IRC Language Questionnaire” that was designed to make social profiles of IRC participants will be described in terms of data collection and data analysis. Third, a description of the procedures of collecting authentic IRC data, which serves the main tool in the present study, will be detailed. Ethical issues will be addressed in this chapter.

Chapter Four presents a systematic description of NA, which is the first completed written documentation of this dialect and will be relevant to understanding how this dialect is represented online. The chapter starts with an introduction of the dialect followed by a detailed description of its phonology, nominal morphology, verbal morphology, and syntax.

Chapter Five presents IRC' participants' social profile in terms of age, gender, levels of education, frequency of using the internet and chatting, and their levels of English proficiency. Their general patterns of linguistic and literacy practices online will be traced and summarized in numerical terms. Their attitudes regarding the use of NA and the innovative literacy practices in representing it will be identified.

Chapters Six to Eight include the analysis of CMC data and are organized on the basis of the research questions. Each chapter begins with a review of related literature that serves two functions. First, it sets the theoretical foundations of the key issue examined in that chapter. Second, it synthesizes related text-based CMC studies, and summarizes their results. Details of the theoretical frameworks are represented first, followed by a review of CMC studies related to the key issue of the chapter. Reviews will be detailed in terms of how specific studies contribute to the undertaking of the present study and how this study is justified in relation to existing literature. Following the review of literature, related data are analyzed, and the results are discussed.

Chapter Six gives an analysis of the orthographic features of NA in synchronous online communication. It summarizes the different practices by which Najdi IRC participants represent their non-Roman based writing system online via Roman script.

Chapter Seven presents an examination of the communicative functions of code-switching to English/ Arabic and script-switching to Roman/Arabic script. It consists of three sections. The first section is concerned with an investigation of how borrowed items are integrated into the IRC data. Second, occurrences of code-switching will be examined in terms of their communicative functions of IRC

discourse. The last section addresses the communicative functions of switching between Roman and Arabic scripts.

Chapter Eight presents an analysis of the specific features that characterize NA in synchronous IRC discourse as a hybrid between spoken and written communication. The chapter concludes with a proposed taxonomy (of these characteristics) that best describes the language of Arabic CMC.

Chapter Nine provides a general conclusion of the study. It restates the study's major objectives and questions. Then, it summarizes the study's main findings. Practical implications of the study's conclusions are presented. The last section of Chapter Nine includes recommendations for further research.

## **CHAPTER TWO**

### **COMPUTER-MEDIATED COMMUNICATION**

This study is concerned with an investigation of the representation of NA in CMC. The study addresses three key issues: namely orthography, code-switching, and the language of IRC compared to speech and writing. An investigation of the bilingual and biscriptual case of NA in CMC should be firmly presented through an understanding of the nature of CMC on the one hand and the theoretical assumptions underlying these three key issues on the other hand. This chapter presents an account of relevant aspects of CMC. A number of issues that relate to this novel medium of communication are described to provide the reader with an initial and profound understanding of the online context in which this study is taking place. First, the definition of CMC is introduced, followed by an overview of its basic characteristics. Second, a description of the context of IRC is given and an account of the notion of virtual community and speech community is detailed.

#### **2.1 Computer-Mediated Communication**

##### **2.1.1 Definition of computer-mediated communication**

CMC is a relatively new field of study. However, as the internet has become an integral part of human life, the field is growing rapidly. The variety of purposes that the internet has served has increased the number of users and, hence, fuelled scholarly study of CMC. With the numerous efforts being exerted in CMC research, there has emerged a great interest to define CMC. Most of these definitions relate

CMC to both the medium being used and interpersonal communication conducted via computer.

The term 'computer-mediated communication' is roughly used to describe any form of communication via the medium of computer including CD-Resource, online database, word processing, and emails (December 1996). Herring (2001: 612) defines Computer-Mediated Discourse (CMD) as "the communication produced when human beings interact with one another by transmitting messages via networked computers". December (1996: 1) gives a very comprehensive definition of CMC. In his words,

"Internet-based, computer-mediated communication involves information exchange that takes place on the global, cooperative collection of networks using the TCP/IP protocol suite and the client-server model for data communication. Messages may undergo a range of time and distribution manipulations and encode a variety of media types. The resulting information content exchanged can involve a wide range of symbols people use for communication."

There are four basic types of text-based CMC system: E-mails, Mailing Lists, Newsgroups, and real-time communication. While the first three are asynchronous CMC system types, real time communication is synchronous. In a synchronous communication system, the interaction takes place in real time, with messages appearing as soon as they are composed. That is, a user joins an ongoing conversation in real time and sends messages that are displayed into a permanently scrolling screen with messages from other users. IRC serves as a means for this

electronic immediate communication, providing thousands of “rooms” dealing with different topics. The IRC system allows its users to engage in more than one chat window and participate in two or more conversations simultaneously. In an asynchronous mode of CMC communication, interaction takes place in forums with delayed posting. BBS is an example of this postponed interaction in which interaction are organized into conferences, identified by certain titles that gather all messages of a common theme (Collot & Belmore 1996).

The linguistic features of text-based CMC vary according to the messaging systems used and the social and cultural contexts. Nevertheless, IRC is the best example of what the internet has to offer with respect to immediate, real-time communication. Chat programs do not only store sent messages but also transmit a person’s typed utterance immediately to the screen of another person. Because IRC will be the context in which the data of the present project will be analyzed, the following section presents an overview of IRC system, summarising its features.

### **2.1.2 Internet Relay Chat (IRC)**

IRC is a synchronous form of communication that is gaining popularity among internet users throughout the globe. IRC rooms are described by Rheingold (1993: 42) as “a cross-cultural grab bag of conversations” as they are hosted by a series of internet servers through which people using different languages, cultures and social backgrounds, can interact with each other. Interaction in IRC is a written or, rather, typed form of communication that is composed, transmitted, received, and responded to in a way that is analogous to spoken communication. Investigating

communication via the IRC system has attracted the attention of scholars in the CMC field for its unique technical features that have an effect on language usage.

One of the essential ways in which IRC differs from other forms of CMC is the concept of creating a ‘channel’. These ‘channels’ are commonly referred to by internet users as “rooms”. The structure of IRC channels was developed to manage dealing with the relatively large number of internet users at one time. When first entering an IRC room, the chatter cannot participate in any conversation with other users. To do so, he/she must log in a room where other users are communicating. These channels are created by users with a name and, sometimes, a topic of some description. All names of channels appear prefixed with the symbol (#) with some orientation depending on their topic—for example, sports, computers, novels, teen chat, a technology, and so on.

IRC channels allow users to communicate with each other by sending messages to be read by all users currently in the channel. Only participants in the channel can read the posted messages and respond to them via a scrolling screen that is divided into two parts; a large, scrollable window showing the conversation and a smaller bottom section. In the bottom section, the participant can type a message and click the Enter key in order to be sent as a separate line, which serves to distinguish individual messages. When lines appear in the top section, other users in the same channel can respond to them by typing again in the bottom section. The flow of conversation inside IRC channels is typed as separate lines, where each turn is prefaced by the participant’s nickname, as seen in Figure 2.1, which illustrates the structure of IRC channel (#T2W.Forex 19 people, TOPIC: Welcome to trade2Win’s New Forex Chat room).

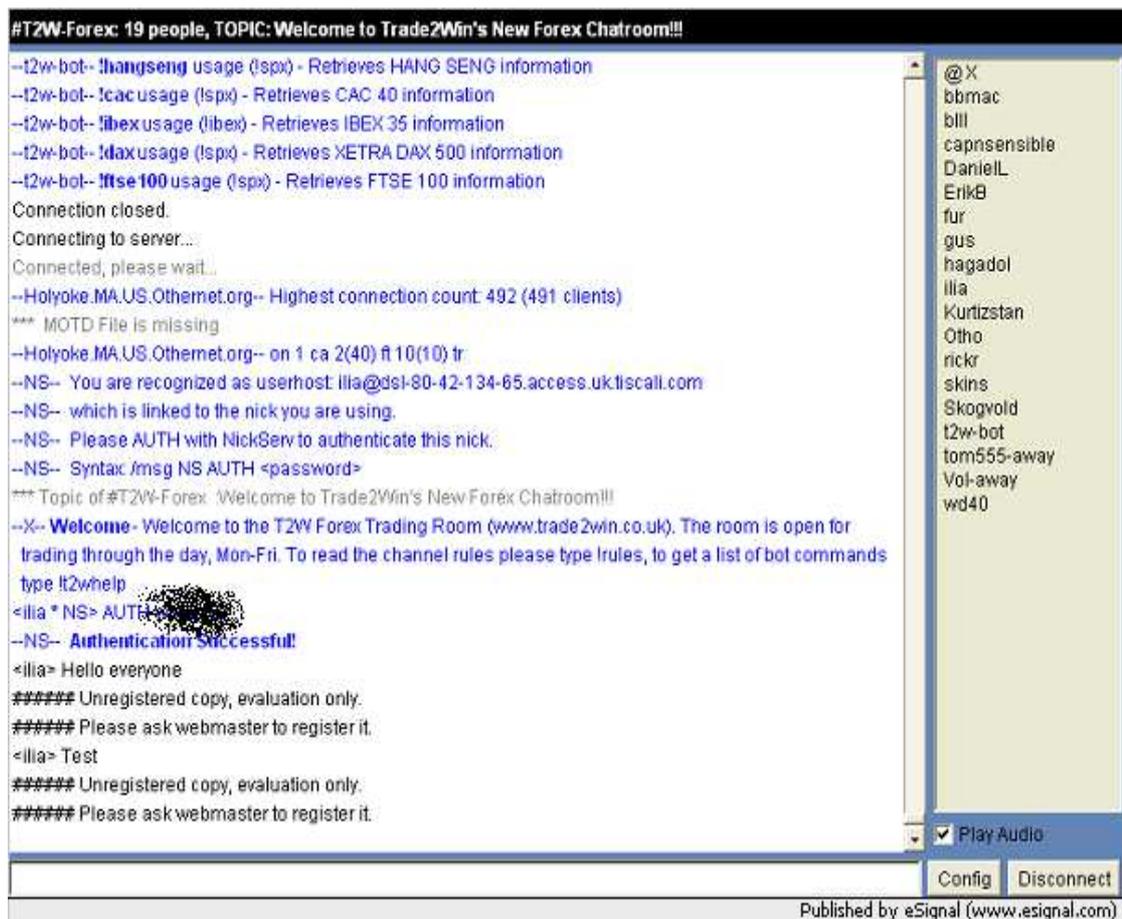


Figure 2.1: Screen shot of IRC windows

Sharing many features of MUDs<sup>2</sup> (Multi-User Dungeons) and MOOs<sup>3</sup> (Mud-Object Oriented), IRC features organize conversations that take place within its channels. The way IRC channels are constructed allows a limited and controlled number of participants to take part in chatting within these channel and, thus, make the flow of conversation controllable (Reid 1991).The column to the right of the channel image displays the list participants in the channel. To keep track of which

<sup>2</sup> Multi-User Dungeon (MUD) is a text-based CMC environment in which participants communicate with each other via synchronous chat messaging (Richard 2003: 9-10).

<sup>3</sup> Mud-Object oriented (MOOs) are text-based online virtual reality systems to which multiple users are connected at the same time. MOOs started out with text based adventure games. With the advent of the internet, MUD was formed as a networked version of one of those games(Richard 2003: 9-10).

users have joined which channels, and ensure that each user can only see the typed conversation of users within his/her channel, this list is continually updating as users join channels regularly.

The IRC program, accordingly, is a multi-user synchronous communication form. It allows its users to choose which person or group of people they wish to interact with. Thus, features of IRC form a *virtual community*, where participants are engaged in social network relations. That is, “users of IRC invent new concepts of culture and interaction and challenge the conventions of both” (Reid 1991: 32).

The widespread use of IRC has captured scholarly interest from different perspectives. Werry (1996) examined IRC and detailed its linguistic and interactional features. Analyzing IRC conversations, he found that participants employ a number of innovative linguistic strategies that function to compensate for contextual, spatial, temporal, and social constraints. Coherence in IRC rooms was intensively addressed by Cornelius and Boos (2003), who developed a coherence measure based on the topics of discussions. Siebenhaar (2006) investigated code choice and code-switching in Swiss-German internet relay chat rooms, and concluded that the choice of a particular variety depends on both individual preference and on the predominant variety used within a specific room chat. Paolillo (1999) was mainly concerned with analyzing structured social relations between participants and the linguistic variants they use. He maintained that sociolinguistic relations in IRC channels are more complex than predicted by sociolinguistic theory for offline interaction. Nevertheless, his study demonstrates that current social network theories are applicable to social interaction and linguistic variation on IRC. Keeping the research questions of the present study in mind, and the fact that social

relations within virtual communities are significant in understanding language use (Paolillo 1999), it is necessary to shed light upon the concept of virtual community in the following subsection.

### **2.1.3 Virtual community**

It has been argued that CMC is an alluring field of study not only because of its data channels, but also because of the virtual spaces of meeting and interaction that it offers (Dutton 1998). In these spaces, people can build and maintain social and cultural realities without even being physically present in CMC spaces. In these spaces, people construct social network relations within communities in which they share common interests.

Experiencing various online forums as basic grounds for his analysis of the effect of internet on the way people communicate, Rheingold (2000) explored the new communities that have emerged in CMC. Examining different online communities, including bulletin worlds (such as Whole Earth Lectronic Link (WELL), MUDs, and IRC), he analysed how these communities affect perception of the self, interpersonal relations, and cultural identity. This experience resulted in his work, *The Virtual Community: Homesteading on the Electronic Frontier*, where he defines virtual communities as ‘social aggregations that emerge from the Net when enough people carry on those public discussions, to form webs of personal relationships in cyberspace’ (2000: ii). In these virtual communities, people establish complex networks of relationships, through the use and the development of a specific language, which maintain the identity of the participants. The conception of an

identity, according to Rheingold, shared by community members is essential for maintaining the cohesion and sense of shared cultural identity in virtual communities.

The boundaries of virtual communities have been considered to be controversial. To Rheingold, these communities have definite boundaries that are defined by the nature of the medium. People in these virtual communities, in Rheingold's words, 'do just about everything people do in real life, but we leave our bodies behind' (2000: vii). Accordingly, Rheingold views CMC as having the power to dissolve physical boundaries of identity and shape communities, the members of which belong to different parts of the world. Rheingold's conception of virtual communities that are not restricted to physical frontiers contradicts the traditional definition of what a community is. In traditional terms, a community consists of a group of people who belong to the same geographical area and who are considered an entity because they share cultural and social identities (Wardhaugh 2006).

Investigating virtual communities online has resulted in a substantial body of research that asserts the fact that community should not be viewed as only physical existence of groupings of people in a digital world that sustains remote relations (Rheingold 1994). According to these studies, a community is a network of interpersonal relations that provide a sense of sociability, shared identity and belonging. Moreover, they maintain that online communities demonstrate the presence of the same behaviours that define offline communities. For example, membership is restricted to those who belong to the communities, and a leader is elected. Herring (2008) attempted to set minimal criteria for online groups

constituting virtual community. To Herring, a virtual community is one which includes:

- a shared reason for communication;
- the existence of norms or protocols;
- regular interaction; and
- some users experiencing emotional attachment to others in the group.

(Herring 2008: 5)

An important concept that should be also considered in the investigation of the present study is that of ‘speech community’. The term ‘speech community’ is an important concept in sociolinguistic studies. Tracing the most important shifts in meanings of ‘speech community’ since 1960s, Rampton (1998) claims that from the start of sociolinguistic discussion of speech community, the aim has been to show that social organization and language use are profoundly interwoven, and when one’s sense of speech community alters, there are often consequence for the kinds of language practice that speakers attend to.

Investigation into the impact of new information and communication technologies is complicated by various factors; by the range and interconnectedness of the media involved, by the complexity of the ways in which they are integrated in our human life, and by the diversity of interactional and textual genres they are bound up with. Research on CMC and on ‘virtual communities’ emphasizes that communication in electronic networks will continue to rely on written texts, unsupported by the semiotics of spoken face-to-face communication (Herring 1996). According to Rampton (1998: 6), “CMC permits the revitalisation of the public sphere, an arena of one-(and many-) to – many dialogue”. Compared with face-

toface interaction and writing, communication in CMC, according to Rampton, provides individuals with far greater opportunities to design different identities for themselves, and to elaborate these identities in ongoing interactive social lives.

IRC rooms, by applying Herring's criteria, are considered virtual speech communities, where participants share common interests and behave according to norms set by the community. In this regard, Paolillo (1999) maintains that the very nature of IRC allows its participants to construct ties in complex social patterns that are similar to patterns identifying the presence of offline communities. For example, IRC channels have moderators that have the power to include or exclude members. Norms and protocols of IRC virtual communities are set by moderators. Moreover, IRC channels include regular members who are engaged in regular contact with each other and, hence, develop emotional ties with each other.

## **CHAPTER THREE**

### **METHODOLOGY**

To address the questions posed in the present study, a combination of tools has been selected and applied. The research design conducted throughout this study is chosen to best answer the research questions by triangulating both numeric results of quantitative methods and the details of qualitative research. The research questions related to online practices in representing NA in IRC are addressed in four separate chapters (Chapter 5, 6, 7 and 8), contributing to the investigation of NA online.

Below is an account of the overall methodology design for data collection and analysis. Section 3.1 describes methods of NA documentation followed in this study. The following two main sections are concerned with presenting the main methodology of investigating online linguistic and literacy practices. Section 3.2 is concerned with locating the methodologies chosen in the present study in the field of CMC by presenting an overview of Herring's (2004) CMDA approach. Then a detailed description of the two main analytical procedures followed in this study to collect and analyze data is given. Section 3.2.1 provides a detailed description of the "IRC Language Questionnaire" utilized in this study. Section 3.2.2 describes how the particular data for analysis is gathered from Yahoo Messenger Chat rooms. A brief account of the ethical guidelines of collecting and handling data is given in the last section. The chapters in the thesis that deal with each key issue separately give a more detailed account of material and methods for each study individually.

### **3.1 Najdi Arabic documentation.**

Before collecting the corpus on NA, it was necessary to establish a well-founded research design. Attention was paid to issues of sampling, fieldwork orientation, and the instrument utilized to elicit aspects of NA. This section is devoted to addressing the methods of documentation developed in the present study to present a description of NA. Section 3.1 presents the Romani Morpho-Syntax (RMS) Dialect Survey, which is used to collect a corpus on NA. A description of its Arabic adapted translated version will be given. Then, an account of informants and the processes followed in applying data collection methods will be discussed.

#### **3.1.1 Romani-Morpho Syntactic (RMS) Dialect Survey**

The decision to use the RMS dialect survey as a data gathering method in this study was influenced by the fact that it is a powerful tool in conducting a language investigation. The RMS survey used in this study to present NA was mainly compiled by Matras (2001) to be the first Romani dialectological survey.<sup>4</sup> The RMS survey was originally developed with the aim of investigating Romani dialects that have not been thoroughly described so far. The RMS survey has been translated to numerous languages to elicit translations from native speakers of Romani. In 2008, the survey was translated into Arabic by Matras and Davey to be used in investigating the Coastal Dhofārī Arabic with the aim of presenting a sketch

---

<sup>4</sup> Details on the development and usage of RMS survey can be found at <http://romani.humanities.manchester.ac.uk/atmanchester/projects/rmsdatabase.shtml>

grammar of the dialect. The Arabic version of the survey, which is adapted from the original RMS survey, is the one which is used in the present study to present a complete description of NA in terms of phonology, morphology, and syntax.

The survey consists of 1065 items (see Appendix A). MSA was used in giving instructions and reading items out to informants. The organization of the survey is as follows:

- Items 1-311 are designed for eliciting singular and plural forms of a number of lexical items. The informants are given an item in MSA form and asked to provide its singular and plural forms.
- The rest of the survey is a mixture of sentences and verb conjugations. The verb items (312-376) are divided into three categories:

(a) The first part (items 312-363) is concerned with verb conjugation. The informants are given a verb in a simple form and asked to conjugate it in the past and present tense.

(b) The second part of verb items (items 364-371) is mainly to discover which lexical verbs are used in the dialect. The informants are given a verb in MSA and are asked to provide the verb they would use in their dialect. Then they are asked about its present and past forms for 3MS.

(c) The third part of verbs (372-376) is concerned with eliciting the imperative form of verbs. The informants are given a verb in MSA and are asked to provide its imperative form in their dialect.

Each informant was given a brief, verbal explanation of the purpose and nature of the study, and asked if s/he would be interested in taking part. The researcher pointed out that the research attempts to present a contemporary documentation of NA. The informants were also informed that the study focuses on the representation of NA in CMC and that their participation would facilitate the procedures of the study and, hence, enhance the results. Consequently, the participants would try to be as helpful and cooperative as possible. Moreover, they were assured that the personal data would be treated as confidential, and would not be made available to any third party. It would not even be used for research purposes. If interest was shown, each participant was sent written information about the recording session, including the purpose of the study, instructions, a consent letter, and contact numbers of the researcher. This material is shown in Appendix B.

All the recording sessions generally followed a similar arrangement. Each began with a warm-up period of informal discussion, eliciting biographical and demographic information, which also served to acclimatize the informant to the microphone and recording equipment. The first five minutes of warm up period of each recording session were not analyzed. Depending on which setting was most suitable for both the informants and the researcher, recordings were conducted in three main settings: the informant's home, the researcher's home and an empty and quiet conference room in SLLC in the University of Manchester. The average length of each interview was between three and four hours.

Survey items were read by the researcher in MSA to assure that the informants were using their own expressions. Informants were asked to provide the equivalent of each item in their own dialect. Survey sessions were recorded by utilizing a Sony

Digital Audio Tape (DAT) stereo recorder on a 90-minute stereo DAT micro cassette and Audio-Technical 831b lavalier microphones.

Responses to the survey were recorded and saved in the WAVE audio file format. They were then transcribed onto a pre-formatted spreadsheet, where each numbered sentence is pre-tagged for the relevant grammatical-semantic categories that appear in it.

### **3.1.2 Informants**

For the purpose of collecting reliable data, the researcher employed the strategy of judgment sampling. In the judgment sampling technique, criteria of selecting informants are predetermined (Milroy & Gosdon 2003). In employing this sampling technique, it was necessary to decide the criteria, bases, and standards necessary for informants to be included in the investigation and find a sample that matches these criteria (Merriam 2001).

The use of judgment sampling to select informants from the Najdi community in the present study is quite justifiable for two reasons. First, according to Milroy (1987:27) in communities with a well-established population whose features are well identified, judgment sampling is considered quite appropriate for a linguistic study. Second, as a native speaker of NA, the researcher relied on personal knowledge to select informants representing the Najdi population. Informants were selected initially via the researcher's network of friends, colleagues, and acquaintances. "Networking sampling" technique was used as some subjects were requested to ask their relatives and friends who might be interested in participating in the study (Milroy & Gordon 2003: 32). Eligible informants were those who were either born

in Najd or arrived before the age of five years. Najdi speakers who had spent long periods of their life outside the Najd area during the period of socialization were considered ineligible and, thus, excluded from the study (Labov 1994). The dialect study will include only informants who were born and raised in Najdi community and had not moved around much. Five informants were asked to participate in the present research by responding to the RMS survey. The informants were all aged between 20 and 35 and, thus, represent the young generation of the community. This sample size of participants would facilitate the analysis of the coded data obtained from the survey.

The researcher is a member of the Najdi community and this facilitated the access to the community and the elicitation of data. According to Milroy (1987: 80), “the closer the field worker is matched to subjects in terms of various social attributes, the more successful he or she is likely to be.” The researcher belongs to the Najdi community in Riyadh, and therefore matches the informants linguistically, socially, culturally and ethnically. The researcher descended from a well-known family from Sudair in Central Najd, and lived all her life between Sudair (*Almajmaçah*) and Riyadh. This helped her in getting access to the community and obtaining reliable data. Informants welcomed her in their homes to tape-record them and introduced other members of their family. Most importantly, the researcher speaking the levelled NA of Sudair and being familiar with the dialect of Qasim helped in understanding all of the informants’ responses. The ability of the researcher to comprehend the rapid speech of informants is a key issue for the success of dialect study (Labov 1972: 215).

### 3.2 Computer-Mediated Discourse Analysis Approach

The term “Computer- Mediated Discourse ” (CMD) came into existence only in 1995 as a consequence of the rapid increase in CMC research, and refers to all types of communication carried out on the internet (Herring 2001). Scholars from different linguistic disciplines, such as pragmatics, sociolinguistics, and discourse analysis, extended their interest in the CMC field, and text-based CMC is investigated from different linguistic perspectives. What facilitates this investigation, according to Herring (2004: 1), is the fact that internet users engage in “socially meaningful activities online in a way that typically leaves a textual trace, making the interactions more accessible to scrutiny and reflection than is the case in ephemeral spoken communication”.

Investigating the linguistics of CMC draws on subfields of linguistics such as language variation, corpus linguistics and dialectal variation. The adaptation and/or reconceptualising of existing traditional methodology depends on the purpose of each CMC research. Commenting on data and methodology of language-focused research in CMD, Androutsopoulos and BeiBwenger (2008: 2) argue that what is lacking is:

“critical reflection on the problems and challenges that arises when these research traditions are applied to new settings and environments of CMD. For example, does one-to-one transfer of research frameworks lead to contextually rich understandings of language use and interactional processes in CMD, or does it rather conceal some of its essential new aspects?”

Herring's (2001, 2004) CMDA is a response to this lack in the CMD methodological framework. CMDA considers online interactions through the means of language, and its results are grounded in speculation about language and language use. Herring (2004) maintains that CMDA applies tools adapted from a variety of language-focused disciplines such as linguistics, communication, and rhetoric, to the analysis of online interactions. CMDA may utilize qualitative or quantitative analysis, but "what basically identifies CMDA is the analysis of logs of verbal languages (characters, words, utterances, messages, exchanges, threads, archives, etc.)" (Herring 2004: 4). Analytical tools in CMDA are drawn from any language-related method, but the methodological orientation of CMDA is a language-focused content analysis

This language-focused approach is based on the same theoretical assumptions underlying traditional linguistic discourse analysis (Herring 2004). For example, the initial theoretical assumption underlying CMDA is that discourse exhibits recurrent patterns that are produced either consciously or unconsciously (Goffman 1959; cited in Herring 2004). Patterns in discourse may be produced consciously or unconsciously (Goffman, 1959); in the latter case, a speaker is not necessarily aware of what she is doing, and thus direct observation may produce more reliable generalizations than a self-report of her behavior. A basic goal of discourse analysis is to identify patterns in discourse that are demonstrably present, but that may not be immediately obvious to the casual observer or to the discourse participants themselves. Another theoretical assumption is that discourse involves individuals' choices that are not only governed by purely linguistic considerations, but also reveal cognitive and social dimensions (Herring 2004). It follows from this assumption that discourse analysis can provide insight into non-linguistic, as well as linguistic,

phenomena. However, CMDA adds to these two traditional assumptions underlying discourse analysis the assumption that CMD is shaped by the technological features of CMC. According to Herring (2004), it is a matter for empirical investigation in what ways, to what extent, and under what circumstances CMC technologies shape the communication that takes place through them.

CMDA is a set of methodological and analytical procedures from which a CMC researcher can choose those that best suit the purpose and the nature of the research. It is, in Herring’s terms, best viewed as a set of methodological lenses through which researchers can make observations and interpret the findings of empirical and experimental studies (Herring 2004: 4). The crucial methodological orientation of CMDA is “language-focused content analysis”, which might be qualitative, quantitative, or both. Herring (2004) identifies five main discourse analysis paradigms that might be adapted to suit the properties of CMC. These paradigms are summarized in Table 3.1.

Table 3.1 Five discourse analysis paradigms (Herring 2004: 17)

<b>Paradigms</b>	<b>Issues</b>	<b>Phenomena</b>	<b>Procedures</b>
Text Analysis	classification, description, “texture” of texts	genres, schematic organization, reference, salience, cohesion, etc.	identification of structural regularities within and across texts
Conversation Analysis	interaction as a jointly negotiated accomplishment	turn-taking, sequences, topic development, etc.	close analysis of the mechanics of interaction; unit is the turn
Pragmatics	language as an activity—“doing things” with words	speech acts, relevance, politeness, etc.	interpretation of speakers’ intentions from discourse evidence
Interactional	role of culture in	verbal genres,	analysis of the

Sociolinguistics	shaping and interpreting interaction	discourse styles, (mis)communication, framing, etc.	socio-cultural meanings indexed through interaction
Critical Discourse Analysis	discourse as a site in which power and meaning are contested and negotiated	verbal genres, discourse styles, (mis)communication, framing, etc.	interpretation of meaning and structure in relation to ideology, power dynamics

Being the most precise and fully articulated approach in CMD research (Androutsopoulos & Beißwenger 2008), CMDA has been applied, explored, and extended by many CMC researchers. Yates' (1996) study of the similarities and differences between a corpus of CMC data and spoken and written corpora is an example of a quantitative application of CMDA. Kendall's (2002) study of gender behaviours in MUDS is purely a qualitative example of CMDA as he applies methods of ethnography and participant observation.

The present study will adopt the CMDA ethnographic qualitative approach for observing, categorizing, coding, and interpreting data, supplemented by a quantitative analysis. The primary focus of this research design is to explore phenomena and to use the quantitative data to assist the qualitative interpretations. At the most basic level, it is conducted in two phases: qualitative and quantitative, with the priority given the qualitative phase. The findings of these two phases are then integrated during the interpretation of findings. Based on the research questions posed, the two main methodological procedures utilized in this phase of the study are:

- An online questionnaire that is designed and utilized to investigate the IRC participants' social dimensions that influence language use. Another aim of

the online questionnaire is to elicit general patterns of IRC participants' choices of language and script, as well as their attitudes towards using English, MSA, and NA online.

- A large corpus of IRC authentic data is collected through accessing Saudi chat rooms where Najdi users engage in online interaction. The target of investigating this IRC data is the linguistic and orthographic features, as well as patterns of code switching and script-switching.

It is worth mentioning that the selected methods of data collection and analysis in this study are integrated with each other. The qualitative CMDA adopted in this study comprises a quantitative phase. Codes of certain categories are identified and, when necessary, counted and summarized in terms of their relative frequencies. The quantitative method comprises a qualitative phase in deciding what the basic instances of codes to be counted are and in utilizing open-ended questions. The following sections discuss each method in detail.

### **3.2.1 The “IRC Language Questionnaire”**

The general aim of this phase is to provide an account of the social profiles of IRC participants. A survey of the IRC participants' different social dimensions provides the initial empirical ground on which to build a more comprehensive investigation of how NA is represented online. First, an overview of internet-research methods is presented. A detailed description of the “IRC Language questionnaire” follows, and then an account of the sampling issues will be given. The processes followed in developing and applying the main instruments are

reviewed. Finally, the issues of reliability and validity of data collection procedures, as well as measures of data analysis, are summarised.

### **3.2.1.1 Internet-Mediated Research**

Interest in using Internet-Mediated Research (IMR) as a primary research tool clearly has grown during the past decade, as indicated by an increasing body of literature on the topic (e.g., Couper & Traugott 2000, Mann & Stewart 2000, Hewson *et al.* 2003, Wright 2005, Fielding *et al.* 2008). This body of literature explores the advantages and disadvantages of IMR, pilot software procedures and programming codes, and sets guidelines and recommendation for designing and utilizing these tools to conduct research involving the internet-based population.

Sampling bias is one of the issues addressed extensively in the IMR literature. Researchers often portray the internet-based population as an inherently biased and non-random sample. These authors claim that the internet-based population consists primarily of well-educated, high-earning, and technologically-proficient males and females who work in academic fields. For this reason, critics fault IMR for sampling problems and question the validity of data acquired through IMR. Bordia warns that “in spite of the growing number of people who use computer communication, computers are still available to only a certain segment of the population” (1996; quoted in Hewson 2003: 27).

On the other hand, a number of studies that explicitly address the issue of internet-sample representation reported a great deal of similarity between the

internet-based population and the non internet-based population. For example, Smith and Leigh (1997) conducted a study that compares an internet and non-internet sample on several demographic variables; they found a great deal of comparability in five out of seven demographic variables. Kendall (1999; quoted in Mann 2000: 36) provides further support for the diversity of internet users; he reports that the use of the internet is increasing among children and the older generation. Recently, the view that the internet-based population consists mainly of educated upper-middle class males has been challenged by the continued explosive growth of the internet. According to a late-2011 statistics from Internet World States, two billion people—one third of the world's population—are using the internet<sup>5</sup>.

The IMR literature has also tackled the important issue of the response rate for surveys. A number of studies have studied the rate of response to email surveys (e.g., Parker 1992, Schuldt & Totten 1994, Tse *et al.* 1998), but the findings of these studies have varied greatly. According to Hewson (2003), the lack of consistency is due to the wide variation across these studies in the factors they investigated. Among studies that report low e-mail response rate is that conducted by Schuldt and Totten (1994) who report a 56.5% response rate in a postal survey and only a 19.3% response rate in internet surveys. Similarly, Tse *et al.* (1998) report a response rate to an internet survey of only 6% and a 27% response rate for postal surveys. Other studies found that internet surveys experienced a higher response rate than postal mail. Parker (1992) reported internet surveys with a response rate over 65%.

---

<sup>5</sup> <http://www.internetworldstats.com/>

Taking these considerations into account, researchers and scholars in the IMR field began suggesting ways to improve internet research methods. Smith (1997; quoted in Hewson 2003: 30) noted in her study that the response rate peaked immediately after a newsgroup posting. Coomber (1997) suggests that reposting the internet survey to newsgroups once a week is appropriate. Hewson (2003: 82) maintains that sending an initial participation requests is a good practice; doing so increases the response rate and maintains compliance with ethical codes. To maintain the highest possible response rate, all of these strategies were adopted when administrating the “IRC Language Questionnaire”.

After taking into account these considerations, an online questionnaire was designed and used to collect the data in this study as the most convenient way to contact the participants who use the internet. Mann and Steward (2000) maintain that online research eliminated the cost barriers of travel and saved both time and money. An online survey would facilitate access to IRC participants who would be difficult to reach using traditional data-collection methods. The online questionnaire, however, allows the researcher to reach thousands of participants in many different locations in a short period of time. Moreover, an online questionnaire is user-friendly: the participants can start and complete the questionnaire when and where it is convenient for them to do so. The online questionnaire also maintains a great deal of confidentiality and anonymity for the participants, allowing participants to “speak their mind” without worrying about being observed or criticized by the interlocutor (Mann 2000: 25).

From a methodological perspective, the online survey software package facilitates managing and analyzing data. The “SurveyMonkey ®” package has been utilized to conduct “The IRC Language Questionnaire” in the present study. The following subsection is a detailed description of the questionnaire design followed by procedures of developing and applying it among IRC participants.

### **3.2.1.2 The “IRC Language Questionnaire”**

The aim of this questionnaire is to provide an account of IRC participants’ profiles in terms of different social dimensions. Another aim is to explore the different literacy practices employed by IRC participants. It is equally important to explore IRC participants’ attitudes towards the use of unconventional innovative writing in IRC, and to investigate their metalinguistic awareness of the existence of the new language form emerging in the IRC setting.

As IRC participants’ ages, gender, and proficiency in English and their attitudes towards these choices are for the most part unobservable, verbal report data is used to identify these extralinguistic factors that cannot be captured by observation. Although self-reports may be inaccurate if the participant does not report truthfully, it is the only way to identify the participant’s choice of literacy practices, as well as his/her attitudes and beliefs towards these choices. As Grenfell and Harris (1999: 54) state: “... it is not easy to get inside the ‘black box’ of the human brain and find out what is going on there. We work with what we can get, which, despite the limitations, provides food for thought.” Therefore, the “IRC Language Questionnaire” (see Appendix C) was used in this study to investigate

participants' choices of language and script, and to explore the different attitudes and beliefs they hold about using NA online.

The researcher developed the questionnaire, which is based on preliminary investigation of IRC data and existing research (Androutsopoulos 2004; Warschauer *et al.* 2002; Androutsopoulos 2004; Palfreyman and Alkhalil 2003; Themistocleous 2008). It includes five sections.

Section 1 is an introduction, which briefly explains the purpose of the study. Participants were told their input would assist in completing a PhD research on the topic of CMC. They were assured that all information would be treated as confidential, and not made available to a third person. Subjects' anonymity could be maintained as their identities are not an essential part in the questionnaire. Time taken to fill the questionnaire (no more than 20 min.) is indicated at the end of the introduction. Participants were informed that their participation was voluntary, and that they could withdraw at any point without specifying a reason and without consequence. They were thanked for taking part in the experiment and were told not to hesitate in contacting the researcher via e-mail if they had any questions or comments concerning the study. The researcher's contact details and affiliation were provided. Hewson (2003) suggests online surveys having details of affiliation with an official institution increases response rate.

Section 2 concerns consent, and includes three compulsory questions participants must read and click "Yes" to in order to continue with the questionnaire.

The questions are:

1. Have you fully understood the nature of the study?

2. Do you understand that your participation in this study is on a voluntary basis and you have the right to withdraw at any point of the questionnaire without giving any reason and without any consequence?

3. Do you agree to take part in this questionnaire?

Section 3 collects personal and demographic information from each participant, such as age, gender, nationality, academic level, proficiency in English, and frequency of internet use and chatting. This data is required to determine the extent to which each subject is relevant to the study. For example, non-Saudi participants should be excluded from the sample.

Section 4— Participants' choices of languages and scripts—includes seven statements representing different choices. Participants were required to rate each statement on a five-point scale from “Never” (1) to “Always” (5). Three additional multiple-choice questions were also included in this part:

- (1) What are the reasons of using Romanized NA online?
- (2) Where did Romanized NA come from?
- (3) Do you use Romanized NA in offline settings such as writing down notes or writing letters?

Section 5 consists of eight items that evaluate different beliefs towards representing NA online by a variety of different literacy practices. The participant was asked to read each statement carefully and decide if he/she agreed or disagreed with it by making a rating on a five-point scale from “Absolutely Disagree” (1) to “Absolutely Agree” (5).

In order to maximise the response rate and reduce sampling bias, the questionnaire was designed with a clear and simple layout that closely resembles a paper-based questionnaire. A progression bar was also included so participants knew how much they had completed and how much was left. Couper & Traugott (2001: 232) maintains: “the benefit of progress indicators is that they inform respondents of their progress through the instrument and should motivate them to complete the survey”. Completing all items of the questionnaires is obligatory to avoid insufficient data, and a red warning message is displayed if a participant fails to complete an item. A variety of question types are included, such as Likert Scale and open-ended questions which ask for more clarification or comment. Displaying a few numbers of items that are brief, concise, and simple per screen is an essential factor to increasing response rate and truthfulness (Schonlau *et. al.* 2002). As SurveyMonkey® online software packages allow the use of different scripts, including Arabic, the questionnaire was designed in MSA. This helped to ensure that the language of the questionnaire would not affect the participants’ responses.

### **3.2.1.3 Pilot study**

As a validity measure to ensure the questionnaire was well-planned and understandable, a pilot study was carried out. Five participants (three female and two male) were involved in a think-aloud pre-test session. Each participant responded to a preliminary version of the questionnaire and was asked to verbalise their responses aloud, item by item. Some misleading items and wording confusions were identified during this process. These items were either deleted or reworded. Two of the

participants commented that the 7-point Likert scale used in the questionnaire was not well-designed enough to capture their responses. Hence, the researcher decided to use a 5-point scale in the questionnaire. Other comments included the need for clarifying academic terminology (such as “code-switching”, “script”, and “orthography”). Simplified Arabic translations of these items with illuminating examples were included in the final version of the questionnaire. One 18-year-old respondent commented that he was unable to identify his English proficiency level, so more clarification was added to responses related to that question item. These are: “Low = I cannot speak English”, “Pre-Intermediate = I can hardly use English”, “Intermediate = I can use English”, “Upper-Intermediate = I speak English in a good way”, and “Advanced = I speak English fluently”. These measurements are based on assessments of the most standardised English proficiency test.<sup>6</sup> Piloting the questionnaire proved very useful for identifying areas of ambiguity and misunderstanding. It also enabled the researcher to estimate the time needed to complete the questionnaire.

#### **3.2.1.4 Strategy of sampling**

The strategy of sampling is dependent on the type of the research being conducted (Creswell 2003). In quantitative research, the purpose of sampling is to define a group that is representative of a population. The ultimate objective of this sampling

---

<sup>6</sup> In Saudi Arabia, English is learned and used as a foreign language, and little exposure to the language is available to people. English is assessed among academics and educated people by standardized proficiency tests, such as the Test of English as a Foreign Language (TOEFL) and International English Language Testing System (IELTS).

strategy is to facilitate generalisation from the research sample to a larger, unmeasured population. The more representative the sample is, the greater the probability the research findings will have “population external validity” (Tashakkori & Teddlie 1998: 65).

In order to make sure that results can be reliably generalised, a sample needs to be large and randomly-selected. Probabilistic sampling strategy allows individuals to volunteer to take part in a study. This strategy was used in the present study. Creswell (2003: 156) states that such random sampling in which each individual in the population “has an equal probability of being selected (a systematic or probabilistic sample)” generates a sample that is large and diverse. With randomisation, according to Creswell, a representative sample from a population provides the ability to generalise results to the greater population.

Once the “IRC Language Questionnaire” was designed, a Web Link collector was created (<https://www.surveymonkey.com/s/EbtesamAlothman>). The link to the questionnaire was posted once a week in different IRC rooms. This technique of distribution proved extremely helpful in increasing response rate (Coomber 1997). An e-mail containing the Web Link collector and an invitation to take part in the study was sent to the researcher’s family and friends. They were kindly requested to forward the Web Link collector to their own contact lists. Worth mentioning is that SurveyMonkey® software is effective at making distribution of an online questionnaire very controlled. The Internet Protocol (IP) addresses of all participants were recorded so that each participant could complete the questionnaire only once. The Web Link collector was active for three months (August to October 2011).

### **3.2.1.5 Validity and reliability**

An important step when generalising the results of a questionnaire to a larger, unmeasured sample is to first confirm the questionnaire's validity. Joppe (2000: 1) defines validity in quantitative research: "validity determines whether the research truly measures that which it was intended to measure or how truthful the research results are". Hatch and Farhady (1983: 252) maintain that it "... is not a formal validity type... It is related to content validity". As "there is no statistical measurement of content validity", Hatch and Lazaraton (1991: 540) indicate that a practical procedure for achieving this kind of validity is that "... panels of experts may be asked to rate the representativeness and comprehensiveness of each part of a test".

To achieve this level of validity, three copies of the questionnaire were emailed to three experts in the field. The experts were requested to assess the structure, clarity of language, directness of instructions, and content of each item in the light of the research questions and theoretical framework of the research. They were also asked to provide any suggestions, changes, additions, omission, or comments that might improve the quality of the questionnaire. The experts' comments resulted in a few minor modifications concerning the wording of some items. Generally, there was an agreement that the content of all items explores the different choices of languages and scripts employed by IRC participants.

The second necessary step in the analysis of quantitative data analysis is checking for reliability. The traditional quantitative view of reliability is based on the assumption of replicability or repeatability: essentially, that the same results will be

obtained if we observe the same thing twice (though we cannot actually measure the same thing twice). Joppe (2000: 1) defines reliability as:

“the extent to which results are consistent over time and an accurate representation of the total population under study is referred to as reliability and if the results of a study can be reproduced under a similar methodology, then the research instrument is considered to be reliable.”

In choosing between the coefficient of stability (i.e., test-retest) and the coefficient of equivalence (i.e., internal consistency), the latter is preferred here for its practicality, as an “internal consistency” procedure involves a single administration of data (Alothman 2008). Among different equations of internal consistency reliability, Cronbach’s alpha coefficient model was chosen. The model is an estimate of full-test reliability based on the average inter-item correlation. A high alpha value means questionnaire items are related and that they all measure related contents. The Cronbach’s alpha coefficient ( $\alpha$ ) obtained for the subtests ranged from 0.63 to 0.78.

### **3.2.1.6 Data analysis**

Responses to the questionnaire were subjected to descriptive statistics to uncover general patterns of IRC participants’ choice of language and script, as well as their beliefs about using these innovative orthographic practices (Research Question 2). Descriptive analysis was conducted to describe, organise, and summarise the general characteristics of the data. Frequencies, percentages, and mean measures were

tallied collectively and cross-tabulated according to complexes. The data were used to create a series of tables, graphs, and charts to provide a concise summary and visual comparison of collected data.

All items in the questionnaire were first downloaded in an Excel summary spreadsheet, in which each question item was presented in its own tapped sheet. The spreadsheet shows the response count and percent for each item. Means and percentages were also calculated. All items in the questionnaire were given numeric codes in Excel format to make the process of exporting and analysing the data easier. Histograms were created for each belief statement to reveal the median and mean of each, and to display accurate levels of agreements and disagreements.

### **3.2.2 Discourse-analytic method**

The main methodological tool in this study is to collect data in an electronic readable format. A large corpus of data was collected through accessing Saudi chat rooms where Najdi speakers engage in online interaction. A brief outline of the setting, data collection, presentation, and analysis, and copyright regulations is given in the following subsections.

#### **3.2.2.1 Setting**

The main source of collecting IRC data is [www.yahoo.chat.sites](http://www.yahoo.chat.sites), which is part of [www.yahoo.site](http://www.yahoo.site). Yahoo is an American multinational internet corporation that was launched on 1 March, 1995, in Sunnyvale, California, United States. In order

to log into Yahoo chat groups, IRC participants have to log into Yahoo messenger (Figure 3.1).

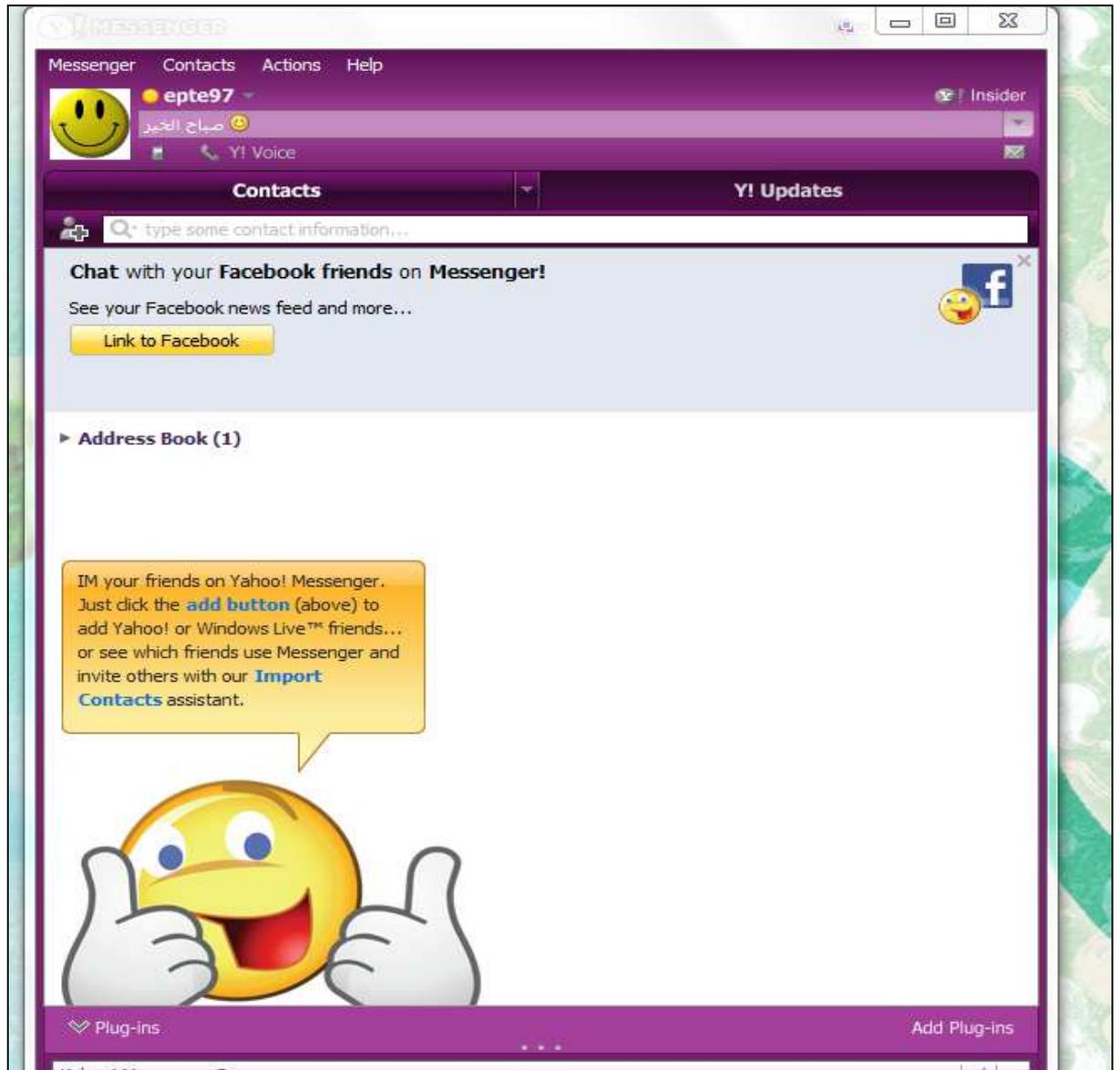


Figure 3.1 Screenshot of Yahoo messenger

The layout of Yahoo messenger window clearly indicates that English is the main language. This indicates that IRC participants logging into my Yahoo messenger are bilingual. The masthead at the top of the window incorporates a number of available actions that IRC participants can undertake or modify (Messenger, contacts, actions,

and help). A link to [www.facebook.com](http://www.facebook.com) is provided and instructions are given in English. The only available place that can display Arabic is IRC participants' personal comments or the description by which they present themselves. Figure 3.2 below illustrates the actual room of conversation, which also presents English as the main language. The chat window incorporates a slogan which changes every second, displaying English-language advertisements.

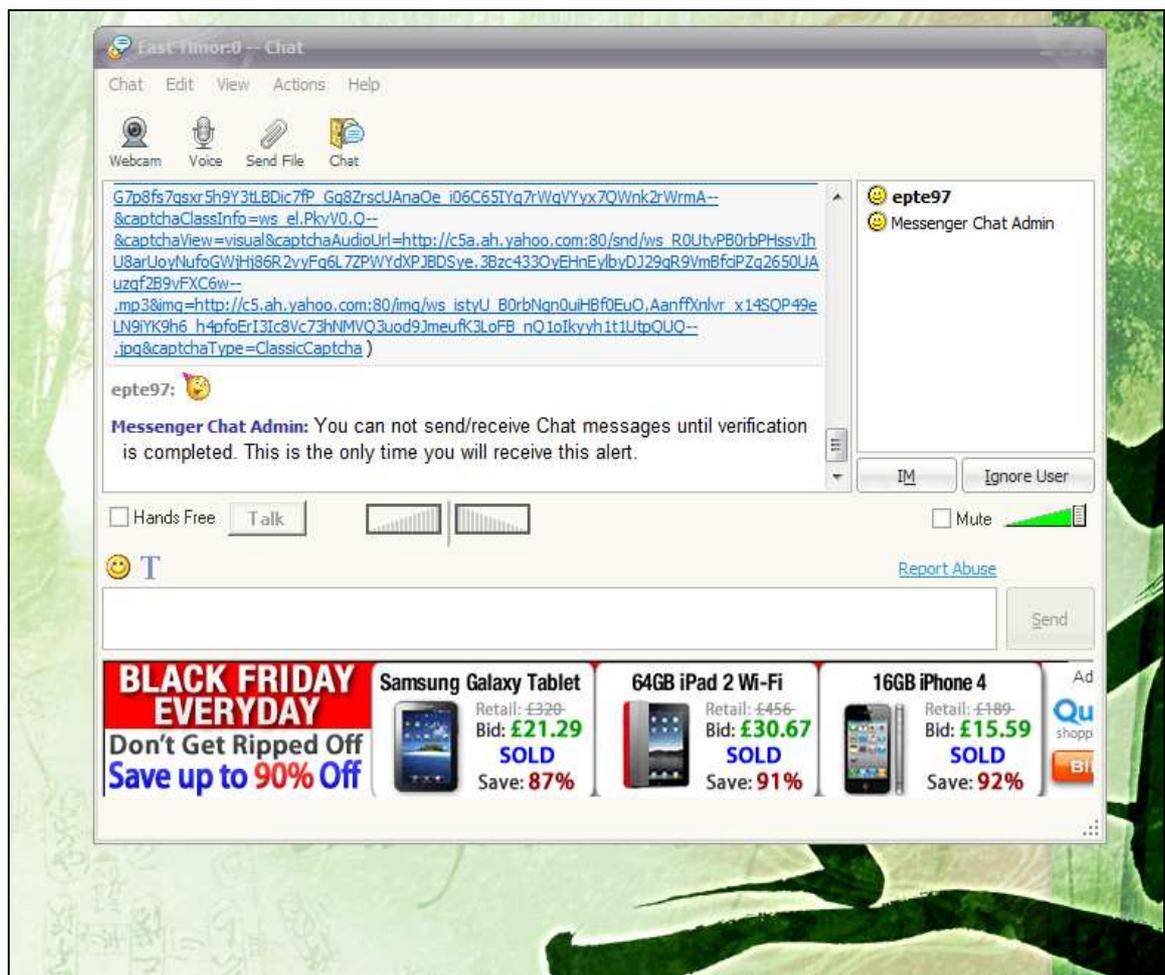


Figure 3.2 Screenshot of Yahoo messenger. The actual room conversation

The other main setting of IRC is [www.KSAVIP.com](http://www.KSAVIP.com) (Figure 3.3), which clearly reflects Arabic as the prominent language. The layout of the site is a clear reflection of the Saudi culture. It is in green, which is the colour of the Saudi flag. Up to the left hand side is a picture of a male in a traditional clothes with the Saudi flag behind

him. Up to the right hand is the name of the site, which is written in both English and Arabic, with the html name given in English ([www.KSAVIP.com](http://www.KSAVIP.com)). Towards the left hand side is a slogan of advertisements which changes every second between an English version and Arabic version. Though the instructions are given in Arabic, the layout of this site presents an environment where both Arabic and English can be used.



Figure 3.3 Screenshot of homepage of www.ksavip.com

In terms of data collection, those main sites are chosen for the flexibility of saving conversations that are taking place in chat windows. Internet Explorer does not facilitate this feature. However, with these two chat sites, it is possible to save and copy the whole room conversations.

### 3.2.2.2 Data collection

A large corpus of IRC authentic interactions was collected. The researcher logged in, accessing Najdi Chat rooms via Yahoo! Messenger, and [www.ksavip.com](http://www.ksavip.com). Once logged on into these channels, the researcher surfed the different rooms choosing the best for analysis among them. Once a room was selected, the chat box was left open for a period of about 20-60 minutes. The interactions between chatters were saved by clicking “select all”, “copy” and then “paste” to a word document. The saved data were carefully read to decide whether they were appropriate for analysis or not. If the sample was appropriate for analysis, it would be saved in an electronic retrievable system. Later, it would be printed for the coding and analysis process. Almost 56,077 words of different log files comprise the corpus of this study.

The data was collected between September 2009 and November 2010 by joining different Yahoo Chat rooms like (#Sahat AlRiyadh) and (# Riyadh Chat) where Najdi users usually communicate with each other. Data collection took place on different times of the day. After collecting the required data, some modifications were made to the original chat texts. These modifications include the following:

- Advertisements by robots or chatters were excluded from the analysis
- Lines announcing the entrance and exists of the different chatters were deleted.
- Since only textual data was to be analysed, lines commenting on or replying to something on voice chat were deleted.
- Nicknames were excluded from data analysis and were replaced by capital letters.
- Lines posting links, pictures and musical files were deleted.

### 3.2.2.3 Data presentation

A total of 56,077 words were collected. Table 3.2 gives a profile of IRC data used in this study.

Table 3.2 Profile of IRC data

Total number of words	56,077
Total number of IRC messages	41,725
Average number of words per IRC message	1.34

IRC data was further organized into two main categories: NA and English. NA was divided into two subcategories: NA in Arabic script and Romanized NA. English words that are rendered into Arabic script such as (welcome و لکم), (thanks ٹانکس), and (copy کوبی) are regarded as English and, therefore, included in the English category.

Table 3.3 The categories of IRC data

Category		Number of words
<b>NA</b>	NA in Arabic script	23,319
	Romanized NA	24,667
<b>English</b>	English	7,893
	English in Arabic script	198
		56,077

#### 3.2.2.4 Data analysis

The present study adopted a CMDA ethnographic approach for observing, categorizing, coding, and interpretation of data. Qualitative researchers tend to use inductive analysis of data, meaning that the critical themes emerge out of the data. Bogdan and Biklen (1982: 145) define qualitative data analysis as “working with data, organizing it, breaking it into manageable units, synthesizing it, searching for patterns, discovering what is important and what is to be learned, and deciding what you will tell others”. The qualitative analysis will be categorized into several categories for the purpose of exhibition and explication. The researcher adopted Sinclair & Coulthard’s (1975) criteria for categorizing the data. These can be summarized as:

- The descriptive categories should be finite in number.
- The descriptive categories should be precisely relevant to the data. Problems in interpreting marginal cases are one of the features of all practical classification.
- The descriptive system should be explicable and comprehensible through the data examined.
- Constraints on possible combinations of symbols should be placed in the descriptive system.

(Sinclair & Coulthard 1975:15)

The aim of investigation in this phase of the study is the textual properties of language (in the form of IRC messages posted) and interactions (in the form of IRC messages exchanged). The focus is on the observable linguistic and communicative

features of IRC messages in order to first explore the nature of linguistic and orthographic features, and, second, examine the communicative features of switching between different codes.

Through lenses of linguistic signs of CMDA, the researcher analysed online interactions in regard to the research key purposes (orthographic features, code-switching and specific features of NA online discourse). For the convenience of descriptive analysis and systematic categorization, categories are identified on the inductive basis of analysis. Specifically speaking, the textual analysis paradigm of CMDA (see Table 3.1) is followed for the analysis of linguistic and orthographic features of NA online. The aim of this analysis is the identification of structural regularities across IRC data. For the analysis of patterns of code-switching and script-switching, the interactional sociolinguistic paradigm of CMDA (see Table 3.1) is followed for the sake of interpretation of meanings indexed by patterns of code switching and script switching. In both cases, quantitative analysis is conducted whenever needed to confirm or cross-validate qualitative analysis. Details on analytical procedures are given in greater detail in each chapter dealing with each specific issue.

### **3.3 Ethical considerations**

Before collecting data for the present study, it had to be considered which kind of data to investigate and how data material should be handled ethically. Public websites in which a large number of participants communicate with each other in a real-time mode of communication via writing messages are a rather recent phenomenon. Studies in the CMC field adopt a number of different approaches to ethical issues, depending on a variety of issues such as the research questions, nature

and context of data, and level of researcher's participation in the target activities (e.g., Yates 1993, Mann & Stewart 2000). These previous studies researching online communication have employed a variety of practices for protecting research participants' rights and privacy. Mann and Stewart (2000: 15) endeavoured to outline an ethical framework for CMC research, and they state a number of established guidelines of fair information processing online:

- Personal data should be collected for one specific, legitimate purpose;
- People should have access to the data collected about themselves;
- Existence of data banks should be publicly known;
- Personal data should be reasonably guarded against risks such as loss, unauthorized access, modification, or disclosure; and
- Data should be collected in a context of free speech.

The present study fits within the framework suggested by Mann and Stewart. All of the guidelines suggested above have been considered while collecting and handling the data of the present study. Information that identifies participants' identities or locations is not exposed, so the data were made anonymous. Place names were changed and participants' nicknames were replaced by letters.

Collecting the data of the present study was carried out in accordance with the copyright regulations related to Yahoo! Chat site and ([www.KSAVIP.com](http://www.KSAVIP.com)). The researcher accessed the Yahoo! "Terms of service" site (<http://yahoo.com/info/terms>) and ([www.KSAVIP.com](http://www.KSAVIP.com)). Although, the terms of service prohibit copying or using any portion of the service for commercial purposes, there seemed to be no restrictions against using the site data for scientific purposes.

Moreover, an e-mail explaining the purposes of the study and requesting permission to use the data of the chat sessions was sent to the Yahoo! Copyright agent.

Permission to save and analyze the material was given to the researcher. The data collected in the present study are not used by a third person or for any purposes other than those stated in the e-mail which was sent to the Yahoo! Copyright agent.

Extracts of online public conversations are collected in a digital format that reflects the speech of people in virtual environments. Obtaining consent from each participant in IRC rooms is nearly impossible. Some researchers in CMC argue that the researcher has to participate actively in a chat room, as well as inform the other participants that there is an ongoing data collection in order to justify the use of the material exposed in the chat rooms (Fouser *et al.* 2000). On the other hand, Segerstad (2002) suggests that spontaneous and informal speech in chat rooms will be influenced if participants are informed that their contributions will be analyzed. Because of the intention of not interfering with the flow of discussion, the stance suggested by Segerstad (2002) was taken in the present study. When logging in chat rooms, all participants were informed by a system message that they are logging in public places and that their conversations are posted on publicly accessible websites. Participants are aware that anyone can get access to their conversation and may read their contributions. For these reasons, permission from those responsible for Yahoo! Chat site to use the material provided the necessary and sufficient consent.

### **3.5 Summary**

This chapter outlines the overall methodological approach to data collecting and data analysis of the present study. It discusses in detail the research design followed in the study covering five fundamental elements. First, it gives a description of NA

documentation procedures followed in this study, including an account of the survey used as well as the procedures in applying it. Then, it clarifies the study's methodological position in CMC research in following CMDA and justifies why combination of methods were chosen. Third, this chapter presents an in depth elaboration on designing and utilizing the "IRC Language Questionnaire". Other important issues related to the tools as their validity, reliability, and scoring systems are included. Fourth, the description of the qualitative approach of the study is followed by an account of the discourse-analytical approach and its basic paradigms. The last section of Chapter Three encompassed the ethical considerations followed in the study.

It is evident that this chapter paves the way for presenting the following five chapters that deal with the analysis of the data. Beginning with chapter Four, findings are presented, analysed, and discussed.

## CHAPTER FOUR

### NAJDI ARABIC

The present study is designed to explore and describe the linguistic practices employed by Najdi speakers when using their non-standard dialect online. Before such an investigation is carried out, it is necessary to introduce the dialect that is the subject of this investigation. The purpose of this chapter is to provide the first complete documentation of NA and present a systematic description of its phonology, morphology, and syntax. The analysis and presentation of data are based on two sources; the existing literature that dealt with specific features of NA and the data collected by the researcher utilizing the Arabic version of RMS survey.

The chapter begins with a brief historical background about the Arabian Peninsula followed by a review on the linguistic situation in Najd. Section 4.3 describes the diglossic situation between MSA and the Arabic dialects that diverged from it. The following sections address the issue of data presentation. The results described in this chapter are based on the analysis of data obtained by the Romani-Morpho Syntactic dialect survey (See section 3.1). Section 4.4 describes the phonology of NA, illuminating areas of differences between NA and MSA. Section 4.5 presents both nominal morphology and verbal morphology. Section 4.6 describes the syntax of NA. Section 4.7 concludes the chapter.

#### **4.1 The Arabian Peninsula: Historical background**

The original homeland of speakers of Arabic is the central and northern regions of the Arabian Peninsula (Watson 2002). The Early Arab grammarians have referred to the different dialects spoken in the Arabian Peninsula as *luḡāt* 'languages' and

went as far as to classify these dialects into two main groups based on the geographical distribution of these dialectal variations: the language of Hijaz in the western parts of the Arabian Peninsula and the language of Najd.

This linguistic variability in the Arabian Peninsula is denoted by the revelation of the Qur'an in variant readings. This fact is indicated by the saying of the Prophet Muhammad (Peace be upon Him): "This Qur'an was revealed in seven *‘ahruf* ‘letters’". Muslim scholars of Sunna have indicated that "seven *‘ahruf* stands for the dialects of the seven tribes of Arab in the Arabian Peninsula: Quraish, Huthayl, Thagif, Kinana, Tamim, and Yemen (Anis 1952). The revelation of the Qur'an in the dialects of these tribes made it easier to read and memorize by Muslims in the Arabian Peninsula.

Until the time of Prophet Muhammad (Peace be Upon Him), in the 7<sup>th</sup> century AD, the Arabic language was mainly spoken and not written. With the expansion of Islam in territories outside the Arabian Peninsula, the early Arab grammarians felt the need to preserve Classical Arabic (CA). Therefore, they undertook the task of preserving the purity of the literal language of Qur'an by codifying Arabic grammar. In their codification of Arabic grammar, they relied on two main sources of data: the Qur'an, and the speech of the Bedouins of the Arabian Peninsula. Then, many people began to write in Arabic. Among the first documents to be written was the Holy Qur'an; however, soon many scientific manuscripts and medical books were written in Arabic, and also narratives like *the Arabian Nights* or *The story of Aladdin*.

In the 14<sup>th</sup> century, Ibn Khaldun in his *Prolegomena* indicated the linguistic differences between the speech of Bedouins in the Arabian Peninsula and that of the urban inhabitants. One of the main differences was the realization of Qaf, voiceless

uvular stop /q/. The issue of Qaf realization was extensively investigated among different Arabic dialects to the extent that it had been considered a key criterion to characterize and classify Arabic dialects. The [g] realization of Qaf is found only in Bedouin speech. Studies were conducted to examine the linguistic change in the following centuries to reveal the existence of a gap between the spoken Arabic used by the Bedouin tribes in the Arabian Peninsula and that used by people who lived in the urban centres in the Arab world (Versteegh 1997). Moreover, such studies revealed that the dialects spoken in central Arabia remained more conservative in preserving many classical attributes of CA compared to other dialects outside the Arabian Peninsula (Versteegh 1997).

European interest in modern colloquial dialects of the Arabic language began in the 19<sup>th</sup> century. Based on geographical distribution, Western linguists (e.g., Johnstone 1967, Prochazka 1988, Ingham 1996, Versteegh 1997) characterized features of modern colloquial varieties of Arabic language. Within each area, dialects were further classified into Bedouin and sedentary dialects. The Najdi dialect is an example of the linguistic conservatism of an isolated dialect compared with other dialects in the Arabian Peninsula (e.g., Hijaz dialect).

## **4.2 Najd**

Occupying the central area of the Arabian Peninsula and spoken by more than 10,000,000 speakers, NA has attracted the attention of linguists, dialectologists and historians over the years. Due to the inaccessible position within the boundaries of *Nafud*, *Dahna*, and *Rubç alkali* deserts, NA displays very little non-Arabic influence and, thus, has a special status of preserving many features of the classical form of Arabic (Ingham 1996). When compared with other Arabic dialects, NA retains a

number of characteristics of the archaic type and, thus, has classical image. Early western travellers among Bedouins who were able to observe their speech over long periods often commented on this. Philby (1922, cited in Ingham, 1996: 6) remarked enthusiastically “their speech over long periods is not only unquestionably Arabic, but a particularly beautiful, almost Classical Arabic at that”. Doughty also remarked that “here first in Najd I heard the *nun* in the ending of nouns pronounced indefinitely, it is like an Attic sweetness in the Arabian tongue, and savours at the first hearing of self pleasing, but is with them a natural erudition” (1924, cited in Ingham 1996: 22). Accordingly, NA has acquired a prestigious position among the numerous dialects spoken in the Arabian Peninsula.

The linguistic area of NA is located between Central Najd which is a vast rocky plateau and *Jabal Shammar*. It is spoken by a large number of ethnically homogenous people who are spread over area within and outside Central Najd. In his study of NA, Ingham (1994: 4) provided a detailed classification of different groups who speak varieties that can be labelled as NA. These varieties include:

- The speech of the sedentary people of Central Najd (i.e the regions of Al-Riyadh, Al-Washim and Sudair), Qasim (i.e Buriadah, Unizah and AlRass) and Jabal Shammar.
- The speech of the Bedouin tribes of Central Najd, Qasim and Jabal Shammar. The main Bedouin tribes of these regions are Anizah, Utaibah, Subai’e, Suhuul, Buguum, Dawasir, Harb, Muṭair, Awazim and Rasheedi in the centre, Shammar and Ḍafir in the north and Gahtan , AlMurrah and ‘Ajman in the south and east.
- The speech of Bedouin tribes of the Syrian Desert (i.e ‘Aniza and Shammar extraction).

Based on geographical and shared linguistic features, Ingham (1994: 5) divided these three groups into sub-groups:

- Dialects of Central Najd which include dialects of central Najd, both sedentary people and Bedouin tribes in addition to ‘Anizah of the Syrian desert.
- Dialects of Northern Najd which include the dialect of Jabal Shammar and the Shammar tribes of Northern Najd.
- Dialects of mixed Central and Northern Najd which include the dialects of Qasim and the Dhafir tribes
- Dialects of Southern Najd which include the dialects of Najran and Gahtan tribe of the south and of the AlMurrah and ‘Ajman tribes of the east.

Najdi people speak a distinct variety of Arabic that is originally the dialect of Nomadic Bedouin mixed with the dialect of sedentary townspeople. NA is used informally in daily life conversations by high, middle, and low social classes in Saudi Arabia. Though NA is not used in a written form, it is the most common form used in poetry among nomads (AlSweel 1992).



Map 4.1: The Arabian Peninsula<sup>7</sup>

### 4.3 Najdi Arabic and Modern Standard Arabic

Arabic language has been cited as an example of diglossia (Ferguson 1959). Arabic exists in two forms; MSA and a number of regional and social vernaculars. Literary Arabic began to develop a standard form through the development of grammatical norms in the 8<sup>th</sup> century and came to be known as *al-fuṣḥā* (Watson 2002: 188). This standard form of Arabic can be referred to as Classical Arabic (CA)

<sup>7</sup> From <http://najd2.wordpress.com/> accessed 20/12/2009

and MSA being used to describe its medieval and modern forms respectively. CA is the language of the Qur'an, the holy book of Islam, and the language of classical literature of the pre Islamic era. Although MSA developed different lexical and stylistic features from CA, the morphology and syntax have remained basically the same (Watson 2002: 188). The Arabic vernaculars, by contrast, have developed different phonological, morphological, syntactic and stylistic features over the centuries. According to Lipinski (1997; cited in Watson 2002:75), these vernaculars began to emerge in the 6<sup>th</sup> century when poets of Arab tribes began to recite their poetry in a proto-classical Arabic based on archaic dialects which are quite different from CA.

Diglossia in the Najd area is an example of Ferguson's definition of diglossic situation as a quite stable language situation. MSA is a written form that is used only in formal sittings (e.g. parliamentary discussions, religious sermons, academic lectures, etc) whereas NA is used in everyday conversation by all people. Ferguson claims that in the situation of Arabic diglossia, MSA, which is based on CA, is the highly codified and superposed variety. It is viewed as superior to non-standard vernaculars, such as NA, due to widespread prejudices against vernaculars within the language community. The presence of superiority of MSA led to clear avoidance of writing Arabic vernacular since doing so is often viewed as undermining the prestigious status of MSA and corrupting its image (Palfreyman and Alkhalil 2003). Habash (2010) went further to demonstrate two facts about the Arabic linguistic situation: (a) the high degree of difference between MSA and Arabic dialects and (b) the fact that MSA is not any Arab's native language (2010: 1).

Preserving many attributes of CA, NA has been investigated by many researchers. A number of works had addressed particular areas of Najd. Abboud (1964) wrote the first published description on NA grammar, titled 'The Syntax of Najdi Arabic'. In this study, Abboud analyzed various syntactic properties of NA such as types of phrases, parts of speech and derivations. Badawi (1965) analysed intonation in the dialect of Riyadh. Lehn (1967) compared the vowel systems of the dialects of 'Unizah, Shaqra and Riyadh. Johnston (1967) gave a detailed description of the syllabic phonology of these dialects, based on the dialect of 'Unizah. At a later date, Abboud (1978) was the first who presented a phonological analysis of the dialect of Sudair based on corpus collected in Almajma'ah. Later, Ingham filled gaps by focusing on dialects of Bedouins of Mutair (1979) and the çajman, çawazim and Harb (1982). Al Sweel (1981) analyzed morphological and phonological features of NA. Moreover, he investigated weak and strong verbs with respect to their derivations in CA. Prochazka (1988) in his survey of the Saudi Arabian dialects divided his material into: (1) the dialects of Southern Hijāz and Tihama and (2) the Najdi and Eastern Arabian dialects. AlSweel (1981, 1992) has treated the phonology of 'Uniza and Kurpershoek (1993) studied the variety of 'Utibah based on an extensive collection of their literature.

It can be noticed that the above studies comprises a comprehensive coverage of different varieties spoken in Najd. However, the existing literature lacks documentation from central Najd (Riyadh) with which the researcher of the present study is concerned in this chapter. The main concern of this chapter is to present a comprehensive research detailing NA in terms of phonology, morphology and syntax.

## 4.4 Phonology

This brief section treats the phonology of NA dialect in terms of inventory of vowels and consonants.

### 4.4.1 Vowels

MSA has three short vowels, three long vowels and two diphthongs (composed by a combination of short /a/ and the semivowels /j/ and /w/). NA displays a little variation from the MSA vowels. Table 4.1 displays short vowels in NA.

Table 4.1 Short vowels

/i/	short high front	<i>ya-ktib</i> ‘M.3SG-write’, <i>çilim</i> ‘knowelge’
/u/	short high back	<i>kum</i> ‘sleeve’, <i>umm</i> ‘mother’, <i>gumaz</i> ‘jumped’
/a/	short mid central	<i>şabr</i> ‘patience’, <i>galb</i> ‘heart’, <i>çam</i> ‘auncle’
/o/	short mid back	<i>moçālləm</i> ‘teacher’,
/ə/	short mid central	<i>nijər</i> ‘shovel’, <i>əsm</i> ‘name’

NA has five long vowels /ī/, /ē/, /ū/, /ā/ and /ō/. The first two are closed and the last three are open. /ē/ and /ō/ have clear realization in most cases. However, when they precede a plosive, they have a gliding pronunciation as in *beit* ‘house’, *leit* ‘would’, *foug* ‘above’ as against *hēl* ‘cardamom’, *kbēr* ‘big’ and *şor* ‘advice’ (Johnstone 1964, Ingham 1996).

Table 4.2 Long vowels

/ī/	long high front	<i>kibīr</i> ‘big’, <i>θigīl</i> ‘heavy’
/ū/	long high back	<i>ḡūd</i> ‘lute’, <i>nūr</i> ‘light’
/ā/	long low back	<i>gām</i> ‘stood up.3SG.M’
/ō/	long mid back	<i>zōl</i> ‘figure’, <i>zōd</i> ‘extra’
/ē/	long mid front	<i>ḡiēl</i> ‘very much’, <i>hiēl</i> ‘cardamon’

The analysis of the data reveals that the short vowels /o/ and its long form /ō/ appear in some loanwords, e.g., كوكاكولا *kōkakōla* ‘coca.colā’, شوكلاتة *šōkolatah* ‘chocolate’ and دكتور *dk̄tōr* ‘doctor’. The short mid back vowel /o/ appears also in foreign names such as جون *jon* and توم *tom*. /i/ and /ī/ were also found to appear in stable loanwords like تشيك *čēk* ‘Czech’ and سكرتير *sekretēr* ‘secretere’. NA has only one diphthong; namely /ai/ as in خايأت *xaiyāt* ‘tailor’.

#### 4.4.2 Consonants

NA, like many Arabic dialects, is particularly rich in uvular, pharyngeal, and pharyngealized consonants. There are 33 consonants in NA, five consonants more than MSA which are /p, v, tz, dz, g/. The first two consonants are borrowed from English and can be found only in loan words and foreign names. The three remaining consonants are productive. Below is an analysis of consonants found in the collected data of NA

Table 4.3 Najdi Arabic consonants

	bilabial	Labio-dental	Inter-dental	Dental	alveolar	Palato-alveolar	palatal	velar	uvular	pharyngeal	glottal
stops											
vls.	p			t ṭ				k	q		ʔ
vd.	b		ðˤ	d ḍ				g			
fricative											
vls.		f	θ		s š		ʃ		x	ħ	
vd.		v	ð		z				χ	ʕ	h
affricate											
vls.				ts							
vd.				dz		dʒ					
nasals											
vd.	m				n						
lateral											
vd.					l						
trill											
vd.					r						
approximate		w									
vd.							y				

As revealed by Table 3.3 , there are eleven stops in NA. The stop positions are bilabial /p, b/, inter-dental / ðˤ / , dental alveolar/ t/, /d/ and their emphatic forms /ṭ/, /ḍ/ velar /g/, /k /, uvular /q/ and glottal /ʔ/. Examples include *bent* ‘girl’, *pakistan* ‘Pakistan’, *damm* ‘blood’, *tamer* ‘dates’, *qalam* ‘pen’, *kalb* ‘dog’, *kill* ‘all’, *gumər* ‘moon’, *amal* ‘hope’, *amer* ‘demand’, and *ṭalab* ‘order’.

Except the glottal stop /ʔ/ (hamza), all stops in occur in initial, medial and final position. The analysis of data reveals the absence of the medial glottal stop (hamza) and its replacement by long vowels (e.g., *rās* < *ra’s* ‘head’ and *ðib* < *ði’b* ‘wolf’). Like most Arabic dialects, the glottal stop /ʔ/ (hamza) occurs in initial positions in a

handful of words which might be considered as CA. It occurs in ‘*allah* (الله) ‘the name of God’. However, this initial *hamza* disappeared when it follows long or short /i/ (e.g. *bismillah* بسم الله ‘in the name of God’).

There are thirteen fricatives in NA and they occur in different positions. The fricative positions are labio-dental / f, v/, inter-dental / θ, ð/, alveolar /s, z/ and the emphatic form /s/, palatal /ʃ/, uvular /x, ɣ/, pharyngeal /ħ, ʕ/, and glottal /h/. Examples that include these fricatives in initial positions are *surʕa* ‘speed’, *ziēn* ‘good’, *zibda* ‘butter’, *villa* ‘villa’, *fikra* ‘idea’, *faham* ‘understood.3SG.M’, *ðīl* ‘tail’, *ðīb* ‘wolf’, *θalāθ* ‘three’, *šarr* ‘evil’, *xiēr* ‘well’, *ḥadr* ‘south’, *ʕilūm* ‘news’, *hayyin* ‘easy’, *ḏuma* ‘thirst’, and *ḏahar* ‘went out.3SG.M’.

There are three affricates in NA. These are the dental affricate /ts, dz/ and the palato-alveolar /dʒ/. Examples that include these affricates as initial sounds are *tsīf* ‘how’, *tsam* ‘how much’, *dzirīb* ‘close’, *dzdām* ‘in front of’, *džrīda* ‘newspaper’, and *džizā* ‘reward’

There are only two voiced nasal: /n/ a and its position is dental-alveolar and /m/ which is bilabial. Examples that include /n/ are *nās* ‘people’, *nāḥya* ‘direction’. There are one voiced lateral: /l/. /m/ is bilabial while /l/ is dental- alveolar. They are found in words like *malāhi* ‘fair’, *munāsaba* ‘occasion’, *līlan* ‘night’, and *li* ‘to’. There is one voiced trill which is also dental- alveolar /r/. Examples from NA that include this consonant include *rādžal* ‘man’, *ramil* ‘sand’. Two voiced approximates are found in NA; /w/ which is labio-dental and /y/ which is palatal.

In regard to differences between NA consonants and MSA consonants, NA differs from MSA in the following features:

- The merger of the pharyngealised voiced interdental fricative / d<sup>ʕ</sup> / and the pharyngealised voiced dental fricative / ð<sup>ʕ</sup>/ to a pharyngealized voiced

interdental fricative / ð<sup>s</sup>/. Thus, the MSA words *bayḍa* ‘an egg’ and *ḍel* ‘shadow’ both pronounced with the interdental fricative associated with *ḍel* ‘shadow’ in MSA.

- The introduction of two new units [ts] and [dz] which result from the affrication of velar stops /k/ and /g/. The affrication of velar stops /k/ and /g/ to voiceless alveolar affricate [ts] and voiced alveolar affricate [dz] in some Arabic vernaculars in the Arabian Peninsula has been traced in a number of previous studies. Johnstone (1963), Wallin (1868) and Socin (1901) studied the phenomenon of affrication in the different dialects of the Arabian Peninsula. Studying the dialect of `Aniza, Johnstone found that [ts] and [dz] occur consistently in the contiguity of front vowels (e.g. *tsib̄r* ‘big’, *iḥts* ‘speak!’, *tsabda* ‘liver’). Ingham (1996) went further to use the affricated units [ts] and [dz] as one of the features separating the central Najdi dialects from the minor ones.
- Borrowed consonants /p/ and /v/ are used only in loan words and foreign names (e.g. *pakistan* ‘Pakistan’, *kapris* ‘caprice’, *vayrus* ‘virus’, *nu:vamber* ‘November’ ). Such words are pronounced as in the original language/p/ and /v/ and they are usually transcribed as <ب> /p/ and <ف> /v/. However, the use of both sounds is considered to be marginal and Najdi speakers may pronounce them as /b/ and /f/.

#### 4.5 Morphology

The morphology of NA is similar to that of MSA in showing a morphological differentiation of different word categories. This morphological differentiation was indicated by early Arab grammarians (Ibn Aqeel 1980) who divided words into three

categories. These three classes are: (i) *ism* ‘nouns’ which literally means ‘name’; (ii) *fiʕl*, ‘verb’ which literally means ‘action’ and (iii) *ḥarf* ‘particle’ which literally means ‘letter’. The class of *ism* covers all nouns and adjectives while the class of *fiʕl* corresponds to the category of ‘verbs’ in English. The remaining class *ḥarf* covers everything else including, pronouns, demonstratives, prepositions, conjunctions and articles.

Displaying the most elaborate morphological differentiation, the class of verbs is inflected for number, gender, person, tense and aspect and transitive type. The class of nouns distinguishes number showing singular and plural and gender. The two classes show a consonantal root which is a central feature of Arabic morphology in general. This root, symbolized C1C2C3, is usually triconsonantal as in *kesar* ‘to break’ (root *ksr*). The lexical root makes the lexical component of the word and occurs in both the verb class as in *kitab* ‘to write’ and the noun class *kitabah* ‘writing’. The class of *ḥarf* ‘particle’ is the least morphologically complex of the three main classes. In fact, particles are non lexical items in nature and do not involve a triconsonantal root<sup>8</sup>.

The particle class will not be explained separately since most particles analyzed in this chapter are referred to under their respective grammatical functions. The nominal and verbal morphology will be discussed in the following sections.

---

<sup>8</sup> There are some exceptional particles that involve a triconsonantal root as *fōg* ‘above’ (root *fwg*) and *ʕala* ‘on’ (root *ʕly*).

### 4.5.1 Nominal morphology

Najdi nouns have three numbers (singular, dual and plural); two genders (masculine and feminine) and two states (indefinite, definite).

Nominal morphology of NA, like many Arabic dialects, shows both derivational and inflectional morphology. Derivational suffixes are used to produce verbal nouns from verbs ( *gara* ‘read.3SG.M’, *graya* ‘reading’), and nouns from adjectives (*samen* ‘fat’, *simna* ‘obesity’), adjectives from nouns (*zayn* ‘good’, *zayna* ‘goodness’) and specific categories of nouns such as nouns of instruments and places.

The inflectional morphology involves the formation of plurals and duals from the singular (*sāḩāt* ‘hours’, *sāḩatiēn* ‘two hours’, *sāḩa* ‘an hour’) and the addition of affixes to form feminine and masculine singular and the affixation for number, gender and indefiniteness.

The following sections will treat some of these derivational and inflectional morphological features of NA. It will be observed that nominal morphology also shows strong and weak stems which behave in the same way as those of the verb.

#### 4.5.1.1 Noun class types

Noun class types can be categorized in specific categories of nouns such as nouns of instruments and places, occupational nouns and active and particle nouns. The description of these nouns will be in the form of list of typical type classes of nouns. Ingham (1996: 32) presented his comprehensive categorization of noun class types that is based on a selection of large number of nouns. They are categorized according to their homogeneity as groups and their semantic characteristics as a class. Basic noun forms are:

- Nouns of instrument that has the underlying structure *mi-CCāC*. Examples include *milgaṭ* ‘tonge’, *migārḏ* ‘tweezers’, *miftāḥ* ‘key’.
- Nouns of places that are structured as *ma-CCaC* or *ma-CCi*. The recorded data includes many nouns of places like *maktab* ‘desk’, *masdžid* ‘mosque’, and *madžlis* ‘place of sitting’.
- Abstract nouns that have the underlying structures *CaCaC*, *CiCaC*, *CuCaC*, *CaCiC* and *CiCC*. Examples of abstract nouns are *baḥr* ‘sea’, *džibal* ‘mountain’, *gumar* ‘moon’, *walad* ‘boy’.
- Nouns that indicate occupations have the structure *CaCCāC*. Examples include *rassām* ‘painter’, *ḥammāl* ‘a porter’. Intensive nouns have also the same structure *CaCCāC* for example *caḏḏāb* ‘a liar’ and *naṣṣāb* ‘a liar’.
- Active participle nouns with the underlying structures *CāCiC* such as *rāseb* ‘a loser’, *rāyih* ‘walking’.
- Passive particle nouns which have the structure *ma-CCūC* such as *masmūṣ* ‘heard.3SG.M’, *maktūb* ‘written. 3SG.M’, *masmūḥ* ‘permitted.3SG.M’.
- Comparative adjectives that have the structure *aCCaC* such as ‘*azyan* ‘better’, ‘*aṭwal* ‘longer’, ‘*akbar* ‘bigger’.
- Nouns of colours that have the structure *aCCaC* such as *aswad* ‘black’, *azrag* ‘blue’, *axḏar* ‘green’.

To Ingham’s categories of noun types, I add two further categories that are based on the recorded data of the spoken dialect. They are:

- Nouns that indicate physical defects and have the structure *CaCaC* and *CaCa* for example *ṣama* ‘blindness’, *ḥawal* ‘strabismus’.

- Abstract nouns of movements and actions that have the structure *CaCC* such as *rakḏ* ‘running’, *ḏarb* ‘hitting’.

#### 4.5 1.2 Number

NA distinguishes between nouns based on quantity. All nouns are either singular *múfrad* when there is one, dual *muḥanna* when there are two, and plural *jamḥ* if there are three or more. The dual is formed by adding the suffix- *-ayn* to the noun stem (*sāḥṭayn* ‘two hours’, *ktābayn* ‘two books’). While the categories of singular and plural are shown in the noun, adjective, verb and personal pronouns, the category of dual is shown only in the nouns and adjectives.

The plurals are formed from the singular in two ways. First, the regular productive plural formation that are undergone through prefixes or suffixes and do not affect the form of the individual morphemes. Productive plurals in NA are formed by the addition of the suffixes *-īn* (masculine) and *-āt* (feminine) such as *swāg* ‘a driver’, *swagīn* ‘drivers’, *xabbāz* ‘a bread maker’, *xabbāzīn* ‘bread makers’. On the other hand, the feminine suffix *-āt* is used to form plurals from feminine singular nouns ending in *-a* (*sayara* ‘a car’, *sayarāt* ‘cars’, *modaresa* ‘a teacher.F’, *modaresat* ‘teachers.F’).

The other way of plural formation is what is known as the “broken plural” formation or “introflexion”. This is the process of alternating the singular form by vowel lengthening, vowel addition or vowel substitution. Other singular inanimate nouns are pluralized by alteration of vowels (*beīt* ‘a house’, *byūt* ‘houses’, *kaff* ‘a hand’, *kifūf* ‘hands’, *seīf* ‘a sword’, *siyūf* ‘swords’). A large number of animate and inanimate nouns are pluralized through the modification of the singular forms into *aCCāC* (*yōm* ‘a day’, *iyyām* ‘days’, *tirki* ‘a Turkish man’, *atrāk* ‘Turks’,

*adžnabi* ‘a foreigner’, *adžnāb* ‘foreigners’). Nouns that indicate origins are regularly structured *CaCātiC* and are pluralized *CaCāyiC* (*ḥamūla* ‘a clan’, *ḥamāyil* ‘clans’, *gibīla* ‘a tribe’, *gibāyil* ‘tribes’).

#### 4.5. 1.3 Gender

NA has two genders, expressed by nominal, verbal and adjectival agreement. The genders are usually referred to as masculine (*moḏākkar*) and feminine (*mo‘ānnaθ*). Masculine nouns are unmarked (*radžāl* ‘a man’, *šiēx* ‘an old man’, *ibn* ‘a son’). Feminine nouns are attached to the feminine gender suffixes *-a* (always follows a fatha, short vowel /a/) (*mara* ‘woman’, *modaresa* ‘a teacher.F’, *madina* ‘a city.F’).

A number of feminine nouns and adjectives are not overtly feminine in form but bear feminine features. That is, they are not attached to the feminine gender suffix but show feminine attributes (*bint* ‘a girl’, *čadžūz* ‘an old woman’, *ḥāmil* ‘a pregnant woman’, *qāšir* ‘a young girl’).

Inanimate nouns are classified according to their morphological form i.e. they will be considered masculine unless they have the feminine gender suffixes *-a*. Therefore, *sayārra* ‘a car’ and *deriša* ‘a window’ are feminine but *šubbak* ‘a window’ is masculine.

#### 4.5. 1.4 Definiteness

In NA, nouns and modifying adjectives are either definite or indefinite. Basically, a noun is definite if it has the definite article prefix *il-*, (*l-walad* ‘DEF-boy’, *l-kitab* ‘DEF-book’). The article (*‘adātu t-tačrīf*) *il-* expresses definite state of a noun of any gender and number. It is also prefixed to each of that noun's modifying adjectives (*l-džedīd* ‘DEF-new, *l-kibīr* ‘DEF-big’). The definite article *il-* is

assimilated to a number of consonants (alveolar, dental and interdental.), so that in these cases, the article in pronunciation is expressed only by geminating the initial consonant of the noun. The consonants that cause assimilation of the definite article are: /t/, /t̪/, /d/, /d̪/, /r/, /z/, /s/, /š/, /s̪/, /d̪/, /t̪/, /z̪/, /l/, /n/. (*it-tyra* ‘DEF -plane’, *iz-zaraç* ‘DEF-plant’). These consonants are called 'solar letters' in CA (*al-ḥurūf aš-šamsiyya*), while the remaining 14 are called 'lunar letters' or 'moon letters' (*al-ḥurūf al-qamariyya*).

A noun is also definite if it has a suffixed pronoun (*kitāb-ī* ‘my book’, *bīet-ik* ‘your house’), if it is inherently definite by being a proper noun (*najd* ‘Najd’, *šāmmar* ‘Shammer’, *maṣer* ‘Eygpt’) or if it is in a genitive state with a definite noun or nouns (*bint l-jirān* ‘daughter- ED neighbours’). Abstract, mass and collective nouns are all definite nouns and marked with definite article. Below are examples of definite nouns that appeared in the recorded data:

(1) *ən-nās ma-t-ḥibb l-kiḏeb*  
 DEF-people NEG-3SG-like DEF-lying  
 ‘people do not like lying.’

(2) *l-bared ṭawīl*  
 DEF-winter long  
 ‘the winter is long.’

(3) *bānat-ha ye-štaḡl-ūn fī māktab l-barīd*  
 daughters-POSS.3SG.F 3PL.work-3PL in office DEF-post  
 ‘her daughters work in the post office.’

#### 4.5. 1.5 Indefinite pronouns

For analyzing patterns of indefinites in NA data, I followed Haspelmath’s (1997) approach that has identified various types of indefinites based on their information status, and for each semantic domain: Person, Object, Location and Time. The

related types of the Object domain are Direct Negation, Indirect Negation, Questions, Free-Choice, Irrealis, and Specific. Examples illustrating the use of indefinite pronouns in NA are:

- Direct Negation:

(4) *maḥad* *yi-kalim-a*  
 nobody 3SG-talk-3SG.M  
 ‘nobody talks to him’

- Indirect Negation:

(5) *mā* *yi-fham* *ay-šay w haww muddares*  
 NEG 3SG.M-understand anything and he is teacher  
 ‘he is a teacher although he does not understand anything.’

(6) *mā-ḥaṣal-aw* *ayy-aḥḥad*  
 NEG-found-3PL anybody  
 ‘they did not find anybody’

- Question:

(7) *aḥad* *rāyih* *ar-riyāḏ*  
 anybody going-3SG.M DEF-Riyadh  
 ‘will anybody go to Riyadh?’

(8) *aštra-īt* *šay min əs-sūg?*  
 bought-2SG.M thing from DEF-market  
 ‘did you buy anything from the market?’

- Free-Choice:

(10) *ayyaḥad* *yi-gdar* *yi-rūḥ* *bukra*  
 anyone 3SG.M-can 3SG.M-go tomorrow  
 ‘anyone can go tomorrow.’

Irrealis:

(11) *a-bḡa* *a-štārī* *ayy-šay* *l-bint-ha*  
 1SG-want ISG-buy anything for-girl-POSS-3SG.F  
 ‘I want to buy anything for her daughter.’

- Specific:

(12) *a-dawwer ayy-šay ġarīb*  
 ISG-search anything exclusive  
 ‘I am searching for something exclusive.’

The prefix *-in* which is derived from the CA *Tanwīn* is used to indicate indefinite nouns in both singular and plural forms *radžālin* ‘a man’, *mritin* ‘a woman’, *radžādžīlin* ‘men, and *byūtin* ‘houses’.

#### 4.5. 1.6 Personal pronouns

The following table illustrates a pronominal paradigm of personal pronouns in NA. It consists of 9 forms in singular and plural. While the 1st person does not distinguish gender, the 2nd and 3rd persons do in the case of free pronouns but not in the bound forms of these pronouns. Unlike MSA, there are no dual personal pronouns. Duals are regularly referred to by plural forms.

Table 4.4 Personal Free pronouns in NA

		<b>Pronouns</b>
<b>1 SG</b>		<i>‘ana</i>
<b>2<sup>nd</sup></b>	masculine	<i>ənta</i>
	feminine	<i>ənti</i>
<b>3<sup>rd</sup></b>	masculine	<i>huwwa</i>
	feminine	<i>hiyya</i>
<b>1 PL</b>		<i>əħna</i>
<b>2 PL</b>	masculine	<i>əntum</i>
	feminine	<i>əntum</i>
<b>3 PL</b>	masculine	<i>hummm</i>
	feminine	<i>hənn</i>

In the manner of MSA, NA shows enclitic forms of pronouns (*aḍ-ḍamā'iru l-muttaṣila(tu)*). They may be affixed to nouns to indicate possession or to verbs indicating direct object. The 2<sup>nd</sup> and 3<sup>rd</sup> feminine bound pronouns in NA are different from their counterparts in MSA in having two variants (*-kum* vs. *-kin*; and *hum* vs. *-hin*). The forms (*-kin* and *-hin*) are the archaic forms of these pronouns and are seldom used by the participants in the present study. The following table is an illustration of these pronouns.

Table 4.5 Bound forms of personal pronouns

Person		Pronouns
1 <sup>st</sup>		<i>-i</i>
2 <sup>nd</sup>	masculine	<i>-k</i>
	feminine	<i>-itz -ik</i>
3 <sup>rd</sup>	masculine	<i>-h</i>
	feminine	<i>-ha</i>
1PL		<i>-nā</i>
2PL	masculine	<i>-kum</i>
	feminine	<i>-kum -kin</i>
3PL	masculine	<i>-hum</i>
	feminine	<i>-hum -hin</i>

Examples of these pronouns indicating possession:

(13) *saiyārt-i gedīma*  
 car-POSS-1SG old  
 'my car is old.'

(14) *walad-hum sāfār l-dubay*  
 son-POSS-3PL travelled.3SG.M to-Dubai  
 'their son travelled to Dubai.'

Example of these pronouns when affixed to verb to indicate direct objects:

(15) *kalamt-hum*    *ʕan*    *l-bīet*    *l-jedīd*  
talked. ISG-3PL about DEF-house DEF-new  
'I talked to them about the new house.'

#### 4.5. 1.7 Object pronouns

NA, like many Arabic dialects, is characterized by the ability of suffixing the direct object pronoun to the main verb. This object pronoun suffix is usually accompanied by some stress modifications and lengthening of vowel. Examples of verbs containing these suffixes are given below with dashes illustrating vowel lengthening and object pronouns.

*nes-i-ni*  
forgot-3SG.M- me  
'he forgot me.'

*nes-ik*  
forgot.3SG.M- you.2SG.M  
'he forgot you.'

*nes-a-ha*  
forgot-3SG.M- her  
'he forgot her.'

*nes-i-kum*  
forgot-3SG. M-you.2PL  
'he forgot you.'

*nes-i-hum*  
forgot- 3SG.M- them  
'he forgot them.'

*nes-ak*  
forgot- 3SG.M- you.2SG.M  
'he forgot you.'

*nes-ih*  
forgot-3SG.M- him  
'he forgot him.'

*nes-i-na*  
forgot-3SG.M-us

‘he forgot us.’

*nes-i-kin*  
forgot-3SG.M- you.2PL.F  
‘he forgot you.’

*nes-i-hin*  
forgot-3SG.M- them.3PL  
‘he forget them.’

These forms are applied to all varieties of central Najd. They can be added to certain prepositions such as *bi-* ‘by’ and *li* ‘for’ or “ as in:

(16) *ktb-ū l-i*  
wrote-3PL to-me  
‘they wrote to me.’

(17) *gilt-l-hum*  
said.ISG- to-3PL  
‘I said it to them.’

#### 4.5. 1.8 Demonstratives

All demonstratives in NA require that the nouns they precede should be marked as definite. Demonstratives agree in number and gender with their reference nouns. The basic element of all demonstrative is the syllable *-ḏa* which is used as a neutral demonstrative in regard to proximity. When *-ḏa* is preceded by *hā* forms, the demonstrative signifies nearness of an object. On the other hand, when *-ḏa* is followed by *-k* forms, the resulting demonstrative signifies farness of an object or something near the hearer rather than the speaker. Thus, two categories of demonstratives exist in NA: near- deictic ‘this’ and far-deictic ‘that’. On these bases, demonstratives in NA can be illustrated as follow:

Table 4.6 Demonstratives

<b>This, these</b>		masculine	feminine
<b>Near- deictic</b>	singular	<i>ða, haða</i>	<i>ði, hāði</i>
	plural	<i>ðōl, haðol</i>	<i>ðōl, haðol</i>
<b>That, those</b>	singular	<i>ðāk, haðāk</i>	<i>ðīk, haðīk</i>
	plural	<i>ðōlāk, haðōlāk</i>	<i>ðolīk, hðōlīk</i>
<b>Far-deictic</b>			

Demonstratives agree in number and gender with their reference nouns (*haða ir-radžāl* ‘this man’, *hāði l-bint* ‘this girl’, *ðolāk ar-radžadžīl* ‘those men’). Plural forms of inanimate nouns are treated as feminine plural and therefore modified by feminine plural demonstrative (*hðōlīk l-ktub* ‘those books’, *ðolīk aš- šwar* ‘those pictures’).

#### 4.5. 1.9 Numerals

Like the demonstratives, the numerals items in NA precede nouns. The numeral item two is preceded by dual noun, while those from three to ten are followed by plural nouns.

Table 4.7 Cardinal numerals

one	<i>wāḥid</i>	eleven	<i>eḥdaçaš</i>
two	<i>iḥnīn</i>	twelve	<i>ḥnaçaš</i>
three	<i>ḥalaḥ</i>	thirteen	<i>ḥaleḥḥaçaš</i>
four	<i>‘arbaç</i>	fourteen	<i>‘arbaçḥaçaš</i>
five	<i>xams</i>	fifteen	<i>xaməḥḥaçaš</i>

six	<i>sitt</i>	sixteen	<i>siṭaṭaṣṣ</i>
seven	<i>sabṭ</i>	seventeen	<i>sabṭaṭaṣṣ</i>
eight	<i>ḥiṣan</i>	eighteen	<i>ḥiṣanaṭaṣṣ</i>
nine	<i>tisṭ</i>	nineteen	<i>tasṭaṭaṣṣ</i>
ten	<i>ṭaṣar</i>	twenty	<i>ṭa ṣrīn</i>
hundred	<i>māyah</i>	thousand	<i>‘alaf</i>

Cardinal numbers, like demonstratives and most other bound elements, precede the head with the exception of one and two which follow the head noun. Numerals from two to ten are followed by a noun in the plural form except for two where the noun is in the dual, while those from ten upwards are followed by a noun in the singular form.

<i>radžāl wāḥid</i>	‘one man’	<i>radžālīn iṭnīn</i>	‘two men’
<i>ḥalaḥ ridžādžīl</i>	‘three men’	<i>arbaṭ ridžādžīl</i>	‘four men’
<i>xams ridžādžīl</i>	‘five men’	<i>sitt ridžādžīl</i>	‘six men’
<i>ḥnaṭaṣṣ radžāl</i>	‘twelve men’	<i>myat radžāl</i>	‘a hundred men’

Ordinal numerals *al-‘aṭdād at-tartiyabiyya* are adjectives, hence, there is agreement in gender with the nouns they modify. However, there is no plurality as with the cardinal numbers.

Table 4.8 Ordinal numerals

	<b>masculine</b>	<b>feminine</b>
first	<i>‘awwal</i>	<i>‘ulā</i>
second	<i>ḥāni</i>	<i>ḥāniya</i>
third	<i>ḥālāḥ</i>	<i>ḥālāḥa</i>
fourth	<i>rābiṭ</i>	<i>rābṭa</i>
fifth	<i>xāmis</i>	<i>xāmisa</i>
sixth	<i>sādis</i>	<i>sādasa</i>
seventh	<i>sābiṭ</i>	<i>sābiṭa</i>

eighth	<i>θāmin</i>	<i>θāmāna</i>
ninth	<i>tāsiç</i>	<i>tāsəça</i>
tenth	<i>çāšir</i>	<i>çāšəra</i>

#### 4.5. 1.10 Quantifiers

There are some quantitative elements which semantically function as modifiers although some of them may occur in isolation. Examples are *kill* ‘all, every’, *baçd* ‘some’, *wājid* ‘many, very’, *šwayy* ‘a little’ *marrāh and bilmarrah* ‘very much’. The quantitatives *kill* and *baçd* usually precede the nouns as in the following examples:

(18) *kill əl-ğānam həg ar-rajal*  
all DEF-sheep POSS DEF-man  
‘all he sheep belong to this man.’

(19) *šwayyat banat fī l-madersa*  
few girls at DEF-school  
‘there are few girls at school.’

(20) *banat šwayy fī l-madersa*  
girls few at DEF-school  
‘there are little girls at school.’

#### 4.5. 1.11 Adjectives

Adjectives in NA follow the nouns and are inflected for gender, number and definiteness. Examples are:

(21) *il-madersa l-ğādīma*  
DEF-school DEF-old  
‘the old school.’

(22) *moçalləma mašğūla*  
teacher busy  
‘a busy teacher.’

In predicative statements, there is no definiteness agreement between the nouns and adjectives.

(23) *il-moçallamīn mašgūlīn*  
 DEF-teachers busy  
 ‘the teachers are busy.’

In case of comparison, NA does not have separate forms for comparative and superlative adjectives and one form is used to indicate both comparative and superlative forms. The comparative/superlative forms of adjectives are formed through an internal vowel change in the adjective (*kəbīr* ‘big’ > ‘*akbar* ‘bigger’, ‘*gədīm* ‘old’>*agdam* ‘older’).

- Comparative:

(24) *çali ‘akbar min xawya-a*  
 Ali older than friend-POSS.3SGM.  
 ‘Ali is older than his friend.’

- Superlative:

(25) ‘*akbar walad fi l-faşal çali*  
 oldest boy in DEF-class Ali  
 ‘the oldest boy in the class is Ali.’

The position of an adjective in the sentence determines its function (superlative or comparative). Comparative adjectives follow the nouns and inflect for definiteness.

(26) *il-wālad l-‘akbar min çali ‘ahmad*  
 DEF-boy DEF-older than Ali Ahmad  
 ‘the boy who is older than Ali is Ahmad.’

In attribute forms, the superlative adjectives precede the nouns and do not inflect for definiteness, number, or gender.

(27) ‘*agdam bāt fī l-ḥara haww bāt-i*  
 oldest house in DEF-area is house-POSS-1SG  
 ‘the oldest house in the area is mine.’

#### 4.5. 1.12 Adverbs

Adverbs in NA are identical to their adjective forms. They can either precede or follow the verbs in sentences.

(28) *il-xabar haða galag-ni kiθīr*  
 DEF-new this worried-3SG-ISG really  
 ‘this news really worried me.’

Some adverbs do not have an adjective forms. The adverb *tagriban* ‘almost’ is an example.

(29) *tagriban xiliṣ ramaḏān*  
 almost finished- 3SG.M Ramadan  
 ‘Ramadan is almost over.’

Tables 4.9 lists below the most common adverbs in NA.

Table 4.9 Adverbs

<b>Adverbs of Time</b>	today	<i>əl-yūm</i>	early	<i>badrī</i>
	yesterday	<i>aməs</i>	finally	<i>‘xyran</i>
	tomorrow	<i>bukra-batsər</i>	late	<i>mət‘axər</i>
	now	<i>alḥīn</i>	usually	<i>dāyyəm</i>
	later	<i>baçdaīn</i>	usually, for the most part	<i>ğāliban</i>
	a long time ago	<i>zāman, zəmān</i>	sometimes	<i>aḥyanan</i>
	recently/soon	<i>girīb</i>	never	<i>abād</i>
	always	<i>dāyəm</i>		
<b>Adverbs indicating degree</b>	very	<i>džədān. kəθīr</i>	a lot; often	<i>dāyyəm</i>
	a little	<i>šwayya</i>	nearly	<i>tagr īban</i> <i>ḥwāli</i>
<b>Adverbs of Manner</b>	in this way	<i>kiða</i>	quickly	<i>bisurça</i>
	slowly	<i>bišweiš</i>	immediately	<i>alḥyn</i>
	on purpose	<i>gaşd</i>	presently	<i>ḥalyān</i>

Table 4.10 Days of the week, seasons of the year

Saturday	<i>is-sabət</i>	Spring	<i>ir-ribīç</i>
Sunday	<i>il-aḥad</i>	Summer	<i>iş-şyaf</i>
Monday	<i>il'-θnīən</i>	Autumn	<i>il-xarīf</i>
Tuesday	<i>iθ-θalāθa, əθ-θulūθ</i>	Winter	<i>əş-şəta</i>
Wednesday	<i>il-ərbəça, əlrubūç</i>		
Thursday	<i>il-xamīs</i>		
Friday	<i>il-jumça</i>		

#### 4.5. 1.13 Local relations

Table 4.11 summarizes local prepositions that express spatial relations.

Table 4.11 Local prepositions

<i>b-</i>	in	<i>barra</i>	out
<i>dāxəl, dżwaa</i>	inside	<i>mən</i>	out of
<i>b-</i>	into	<i>l-</i>	towards
<i>çala</i>	at	<i>çala tül</i>	along
<i>çənd</i>	next to	<i>layīn</i>	until
<i>l-</i>	to	<i>maç</i>	with
<i>gədām</i>	in front of	<i>wara</i>	behind
<i>fog</i>	above	<i>çala</i>	on
<i>fog</i>	over	<i>çaks</i>	opposite
<i>bīən</i>	between	<i>ḥol</i>	around

As table 4.12 reveals, *l-*, *çala* and *b-* are used to cover more than one meaning.

Only one meaning of these expresses location. Both *l-* and *b-* are prefixed to nouns and pronouns. Examples of the use of some local prepositions are given below.

(30) *tarāk*      *l-kitāb*      *çala*      *əḫ-ṭawla*  
left.3SG.M    DEF-book    on    DEF-table  
‘he left the book on the table.’

(31) *çalī çənd l-bāb*  
Ali at DEF-door  
'Ali is at the door.'

(32) *wigāf gadūm l-bīət*  
stood.3SG.M in front of DEF-house  
'he stood in front of the house.'

#### 4.5. 1.14 Possession

The possessive particle *ħag* which is derived from the classical word *ħaq* 'right', is regularly used to indicate possession. In such structures, either both nouns will be definite, both indefinite or the first is indefinite and the second is definite.

(33) *il-masdzəd ħag l-ħāra gedīm*  
DEF-mosque POSS DEF-area old  
'the mosque of the area is old.'

Possession is also expressed through the possessive pronouns suffixed to the noun.

(34) *kitāb-a*  
book.POSS-3MSG  
'his book.'

(35) *kitāb-i*  
book-POSS.1SG  
'my book'

Another way to express possession is the use of construct state where a possessor follow a possessed noun.

(36) *bāb l-beīt dzedīd*  
door-CONS DEF-house new  
'the door of the house is new.'

#### 4.5. 1. 15 Interrogatives

NA shows much variation in the use of interrogatives. The following table illustrates interrogatives in NA.

Table 4.12 Interrogatives

what	<i>wišu, wiš</i>
where	<i>wayn</i>
when	<i>meta</i>
why	<i>warā, līš, wišula</i>
how	<i>tsīf, šlūn</i>
how much	<i>kām, tsām</i>
which	<i>‘ay</i>
who	<i>mən, mīn</i>
whose	<i>l-mīn</i>

Some of these items are recent development in Arabic as CA does not have them like *wara, wišulah, wišu, wiš*. These interrogative elements occur only in the dialects of Najd. Some of these elements are derived from classical words. The interrogative *wīn* is derived from *ayyn* ‘where’ and *mən, mīn* are both derived from the classical interrogative *man* ‘who’. ‘*ay* and *kām* exist also in MSA with the same pronunciation and use. Most of these question words are inflected for number and gender as in the following:

(37) *wayn-ha?*  
where -2SG.F  
‘where she is?’

(38) *wayn -hum?*  
where-3PL  
‘where are they?’

### 3.5. 2 Verbal Morphology

#### 3.5. 2.1 Verb derivation

All verb forms are derived from triconsonantal stems (*k.t.b* ‘to write’, *š.r.b* ‘to drink’) which are then inflected for person, tense, and mood. Verb derivation patterns will be discussed in terms of voice, tense and aspect.

The formulation of the passive in NA, namely ‘internal passive’, reflects quite closely the *majhūl* form in MSA. The internal passive is generally formed by the modification of the structure of the verb stem. In perfective passive forms, *-in* is added to the verb stem (*šarag* ‘to rob’ > *in-šaragt* ‘I was robbed’. In cases of imperfective passive forms, *in-* or *yi-* is added to the verb stem (*ḡarab* ‘to hit’ > *in-ḡerb-t* ‘I am being hit’). The recorded material and my observation of the spoken language in Najdi community reveal that the internal passive is widely used as it can be formed for all verb forms. The formation of internal passive for the strong verbs in perfective and imperfective forms is illustrated below.

Table 4.13 The internal passive (perfective)

<i>in-šaragt</i> ‘I was robbed’	1SG	<i>in-šarag-na</i> ‘we were robbed’	1PL
<i>in-šaragt</i> ‘you were robbed’	2SG.M	<i>in-šarag-tu</i> ‘you were robbed’	2PL.M
<i>in-šaragt-i</i> ‘you were robbed’	2SG.F	<i>in-šarag-tu</i> ‘you were robbed’	2PL.F
<i>in-šarag</i> ‘he was robbed’	3SG.M	<i>in-šarag-aw</i> ‘They were robbed’	3PL.M
<i>in-šarag-at</i> ‘she was robbed’	3SG-F	<i>in-šarag-an</i> ‘They were robbed’	3PL.F

Table 4.14 The internal passive (imperfective)

<i>an- ḡarib</i> ‘I am being hit’	1SG	<i>tin- ḡerib-ōn</i> ‘you are being hit’	1PL
<i>tn- ḡerib</i> ‘you are being hit’	2SG.M	<i>tin- ḡerib-ōn</i> ‘you are being hit’	2PL.M
<i>tn- ḡeri- ayīn</i> ‘you are being hit’	2SG.F	<i>tin- ḡerib-ōn</i> ‘you are being hit’	2PL.F
<i>yin- ḡerib</i> ‘he is being hit’	3SG.M	<i>yin- ḡerib-ōn</i> ‘They are being hit’	3PL.M
<i>tin- ḡerib</i> ‘she is being hit’	3SG-F	<i>yin- ḡerib-ōn</i> ‘They are being hit’	3PL.F

#### 4.5. 2.2 Verb inflection classes

Like any other Arabic dialect, NA makes a distinction between strong and weak verbs. According to early Arab grammarians, a verb is identified as weak when one of the consonants of its triconsonantal stem is / ‘, w, or y/. On the contrary, a strong verb is the one that does not have one of these three consonants in its triconsonantal stem. On the basis of this definition, verbs like *kitab* ‘to write’ and *samç* ‘to listen’ are both categorized as strong verbs while *wagaf* ‘to stand’ and *šaf* ‘saw’ are weak ones. Strong verbs have a sub-category of doubled stems in which C2 and C3 are identical. In such verbs, the stem sometimes shows merging of C2 and C3 into a single consonant. Of the three preceding categories of verb stems strong, doubled and weak, the strong category constitutes the majority of verbs.

Verbs in NA, like other dialects of Arabic, show two verb forms: Perfective (past), and Imperfective (present). The perfective form indicates completed aspect of event and actions while the imperfective form is used to indicate actions in progress or actions that will happen in the future. These two verb aspects involve prefixing and suffixing and alteration in the syllabic structure of the stem. Below are examples of both strong and weak verbs in different conjugations in the three verb forms followed by an explanation of the formation of verb conjugation.

Table 4.15 Perfective (strong verbs)

<i>kitab-t</i>	‘I wrote’	1SG	<i>kitab-na</i>	‘we wrote’	1PL
<i>kitab-t</i>	‘you wrote’	2SG.M	<i>kitab-tw</i>	‘you wrote’	2PL.M
<i>kitab-ti</i>	‘you wrote’	2SG.F	<i>kitab-tw</i>	‘you wrote’	2PL.F
<i>kitab</i>	‘he wrote’	3SG.M	<i>ktib-aw</i>	‘they wrote’	3PL.M
<i>ktib-at</i>	‘she wrote’	3SG.F	<i>ktib-an ktib-aw</i>	‘they wrote’	3PL.F

Table 4.16 Perfective (weak verbs)

<i>šif-t</i>	‘I saw’	1SG	<i>šif-na</i>	‘we saw’	1PL
<i>šif-t</i>	‘you saw’	2SG.M	<i>šif-t-w</i>	‘you saw’	2PL.M
<i>šif-ti</i>	‘you saw’	2SG.F	<i>šif-t-w</i>	‘you saw’	2PL.F
<i>šāf</i>	‘he saw’	3SG.M	<i>šāf-aw</i>	‘I saw’	3PL.M
<i>šāf-t</i>	‘she saw’	3SG.F	<i>šāf-aw šāf-an</i>	‘I saw’	3PL.F

Table 4.17 Imperfective (strong verbs)

<i>a-ktib</i>	‘I write’	1SG	<i>na-ktib</i>	‘we write’	1PL
<i>ta-ktib</i>	‘you write’	2SG.M	<i>ta-ktib-ūn</i>	‘you write’	2PL.M
<i>ta-ktib-īn</i>	‘you write’	2SG.F	<i>ta-ktib-in</i>	‘you write’	2PL.F
<i>ya-ktib</i>	‘he writes’	3SG.M	<i>yi-ktib-ūn</i>	‘they write’	3PL.M
<i>ta-ktib</i>	‘she writes’	3SG.F	<i>yi-ktib-in</i>	‘they write’	3PL.F

Table 4.18 Imperfective (weak verbs)

<i>a-šūf</i>	‘I see’	1SG	<i>n- šūf</i>	‘we see’	1PL
<i>t- šūf</i>	‘you see’	2SG.M	<i>t- šūf-ūn</i>	‘you see’	2PL.M
<i>t- šūf-īn</i>	‘you see’	2SG.F	<i>t- šūf-ūn</i>	‘you see’	2PL.F
<i>y- šūf</i>	‘he sees’	3SG.M	<i>yi- šūf-ūn</i>	‘you see’	3PL.M
<i>t- šūf</i>	‘she sees’	3SG.F	<i>yi- šūf-ūn</i>	‘you see’	3PL.F

Both weak and strong triconsonantal stems form their perfective and imperfective in the same manner. A verb in its perfective conjugation is formed by the verb stem and a personal suffix (e.g. *ktib-aw* ‘wrote-3PL.M’). A verb of the imperfective conjugation consists of the verb stem and prefixes in singular cases (e.g. *ta-ktib* ‘3SG.F-writes’). In second person and plural cases, imperfective conjugation consists of the verb stem, prefixes and a personal suffix (e.g. *yi-ktib-ūn* ‘write-3PL’). As revealed by the preceding conjugations of the strong and weak verbs, they are attached to the same suffixes and prefixes to form their perfective and imperfective forms. However, some of the above suffixes show some modifications when followed by an object pronoun suffix. These modifications include doubling of the

–n of –in and n- and inserting long vowels in the place of –i, –u, –aw, –ti, –tu. Recorded verbs of such cases include *ṭalabnā-a* ‘ordered.1PL.3SG’ and *yi- ṭlbin-a* ‘3PL.ordered-3SG’.

The imperative conjugations of a verb constitute the verb stem with same set of suffixes attached to its imperfective form excluding the prefixes. However, it is only the second person form of a verb that can form an imperative verb form (e.g *i-ktib* ‘write! -2SG. M’ and *i-ktib-u* ‘write!-2PL’).

#### 4.5. 2.3 Tense

There are three tenses in NA: the present, past and perfect. The present form is formed by the stem of the verb plus personal suffixes and prefixes<sup>9</sup>, *ta-ktib* ‘2SG.M-write’, *ta-ktib-ūn* ‘write-2PL), *yi-ktib* ‘3SG.M-write’. Examples of present indicative are:

(39) *aštrī kill yom min l-bāgala*  
 buy.ISG every day from DEF-grocery  
 ‘I buy every day from the grocery.’

(40) *ḥamad yi-ktib əd-darəs*  
 Hamad 3MSG-writes DEF-lesson  
 ‘Hamad writes the lesson.’

The future tense is expressed by the use of the verb stem preceded by *rāḥ* and *ba-* which have the meaning of ‘will’.

(41) *rāḥ aštrī sayyāra bukra*  
 will buy.ISG car tomorrow  
 ‘I will buy a car tomorrow.’

---

<sup>9</sup> See tables 4-17 for the present tense conjugation

The past tense is formed by the stem of the verb in conjugation with personal suffixes<sup>10</sup>, *kitab-t* ‘wrote-1SG’, *kitab-ti* ‘wrote-3SG.F’. Examples are:

(42) *yom wagf-at*      *çind bāb*      *beīt-hum*  
 when stood-3SG.F    at    door-CONS    house-POSS.3PL  
 ‘when she stood at the door of their house?’

#### 4.5. 2.4 Modality

The imperative form is formed on the basis of the present stem with the same set of suffixes but without prefixes. It involves also alteration in the syllabic structure of the stem. Only the second person form of a verb can appear in the imperative mood. Below is an illustration of the imperative conjugation of the verb *kitab* ‘to write’.

Table 4.19 Imperative

<i>i-ktib</i>	2SG.M
<i>ikitb-i</i>	2SG.F
<i>ikitb-u</i>	2PL.M
<i>ikitb-u</i> , <i>ikitb-in</i>	2PL.F

#### 4.5. 2.5 Aspect

The progressive is the only category of aspect that can apply to the present tense to show the continuity of the action of the verb at the time of the event. The progressive aspect is manifested by the use of particles like *gāçad*, *jālās* literally ‘sitting’ preceding the present form of the verb. These particles are inflected for person, number and gender.

(43) *gāçad-a*      *ta-ktib*      *əd-dares*  
 sitting-3SG.F    3SG.F-writes    DEF-lesson  
 ‘she is writing the lesson.’

<sup>10</sup> See table 4-15 for the past tense conjugation

(44) *əl-çyāl gāçad-īn yi-taçaş-ōn*  
 DEF-kids sitting-3PL 3PL.eating-3PL.M  
 ‘the kids are eating their dinner.’

The particle *alḥīn* literally means ‘now’ is used to express the present continuous tense. It occurs at any position of the sentence and it does not inflect for person, number and gender.

(45) *alḥīn a-ktib rəsāla l-ixū-ī*  
 now 1SG-write letter to-brother-POSS-1SG  
 ‘I am writing a letter to my brother.’

(46) *əl-çyāl yi-tçaş-ōn alḥīn*  
 DEF-kids 3PL.eat.3PL now  
 ‘the kids are eating dinner now.’

The expressions *li-sāça*, *li-sana* are used to give the temporal duration. The entire construction (TEMP *w* S) is equivalent to the English present perfect continuous ‘have been -ing’.

(47) *li-sāçah w ana a-ktib rəsāla*  
 to-me hour and I 1SG-write letter  
 ‘I have been writing a letter for an hour.’

(48) *li-sanah w-ana a-ktib kitab*  
 to-me year and-I 1SG-write book  
 ‘I have been writing a book for a year.’

#### 4.5. 2.6 Negation

In NA, the two particles *ma* and *la* are basically used in the process of negation of verbal sentences. The item *la* is usually used with imperative verb forms while *ma* is used in all verbal statements. Examples of the types of these structures are given below:

(49) *la t-rūḥ-ēn l-l-masdžed*  
 NEG 2SG-go-2SG.F to-DEF-mosque  
 ‘don’t go to the mosque.’

(50) *la t-xaf-īn*  
 NEG 2SG.F.scare-2SG.F  
 ‘do not be afraid’

The particle *ma* is used with imperfective verbs indicating present and futures<sup>11</sup>.

Examples of such verbal statements are:

(51) *ma ken-na fī il-bar*  
 NEG was-1PL in DEF-desert  
 ‘we were not in the desert.’

(52) *ma t-ruḥīn maṣa-hum*  
 NEG 2SG.F-go with-them  
 ‘you will not go with them.’

A feature that is unique to speech of Central Najdi is the production of emphatic negation form by the use of the structure *wa* ( and)-*la* as in the following example:

(53) *wa-la šef-na-hum*  
 and-NEG saw-IPL-them  
 ‘and we did not see them.’

In other cases, both particles *ma* and *la* are used with *wala* (and) in coordinated sentences as in:

(54) *ma rāh il-madresa wa-la nām*  
 NEG went.3SG.M DEF-school and-NEG selpt.3SG.M  
 ‘he neither studied nor slept.’

The particle *ma* is used to negate the expletive *fē* (there). For example:

(55) *ma fē aḥad farḥān*  
 NEG there body happy  
 ‘nobody is satisfied.’

Some people of Riyadh use the particle *mū* and *mush* to negate verbal plain statements. For example:

---

<sup>11</sup> The particle *ma* is also used in imperative verbal statement as in *ma trūḥēn* ‘do not go’

(56) *mū mašgūl-ēn*  
 NEG busy-3PL  
 ‘they are not busy.’

#### 4.5. 2.7 Modal verbs

Modality in NA is achieved by a category of modal items that illustrate semantically the attitude of the speaker to the situation referred to in the context. These modal elements cannot stand independently. Structurally, some of these modal elements are prefixes that are morphologically bound to other elements and others constitute independent words. Modal elements in NA are reduced forms of verbs, adjectives but constitute extended semantic functions of the basic meanings of the items from which they are derived. However these reduced forms differ from their origins in their lack of concord as they have identical forms for all persons. They cover a range of different meanings including necessity, intent, approval, and judgment.

The recorded data contains many of these modal elements that have different semantic functions. Below is an illustration of the most common modals that were found in the data. Modals that end with a hyphen are those which may be followed by the object pronoun suffixes.

Table 4.20 The Modal elements

The Modal	Semantic meaning
<i>Jiçil-</i>	‘may’
<i>çasa</i>	‘hope’
<i>tsinn-a</i>	‘it seems’
<i>lēt-</i>	‘would’
<i>bağa</i>	‘will’
<i>y-abī</i>	‘will’
<i>xal-</i>	‘let’
<i>yalla</i>	‘hardly’
<i>kūd</i>	‘hope’
<i>illa</i>	‘must’
<i>widd-ik</i>	‘you want’

The above mentioned elements differ from each other in regard to their etymology, morphology, and their positions in sentences. Below is an account of some of these elements and examples of these forms.

The modal element *jiçil* implies a hope that something might happen. It is derived from the classical causative verb *jaçal* ‘to make’ has undergone some internal vowel alteration. This model is commonly used in Central Najd in conversation, poetry and conventional sayings. Examples are below:

- (57) *jiçil* *şadr-ik* *ma* *yi- ð̄yig*  
 may chest-POSS-2FSG NEG 3SG.M-narrow  
 ‘I hope your mood is not getting upset.’

*çasa* has the same semantic function as *jiçil* and it is usually used as an optative marker that means ‘I hope’. This particle exists also in CA with the same meaning. In NA, it is frequently accompanied by the negative particle *ma* as in the following frequent expression:

- (58) *çasa* *ma* *şar*  
 hope NEG wrong  
 ‘I hope nothing is wrong.’

*çasa* and *jiçil* may also be followed by a personal pronoun suffix:

- (59) *çasa-k* *xallē-t-ih*  
 hope-2SG.F left-2SG.F-3SG.M  
 ‘I hope you left it.’

- (60) *jiçil –ik* *t-waşal* *bi-s-salamah*  
 hope-2SG.M 2SG.M-arrive in-DEF-safety  
 ‘I hope you arrive safely.’

The modal *tsinn-* which means ‘it seems’ is used only in NA and it is derived from the classical particle ‘*ka’anna*’ which also means ‘like’. The particle ‘*ka’anna*’

is a combination of *ka* which means ‘like’ and ‘*anna*’ which means ‘that’. It is usually followed by object pronoun suffix as in the following statement:

(61) *tsinn-ik ma wed-ik t-rūh*  
 it seem-2SG.M NEG want-2SG.M 2SG.M-go  
 ‘it seems that you do not want to go.’

*Liēt* ‘would that’, please’ is basically the classical particle *liata* which means ‘would that’. This modal has another form *ya liēt* which also derived from the classical *ya liata*. It occurs with both verbal and non-verbal structures. This modal has the semantic function of expressing wish, hopes and unfulfilled desires. In some situation, it may imply a polite request as in :

(62) *liēt-ik t-kallam-hum*  
 would-2SG.M 2SG.M-speak-them  
 ‘I hope you speak to them.’

*Kūd* also has the meaning of ‘I hope’ but it has somewhat archaic flavour. It is regularly derived from the classical verb *kaad* which means ‘almost’. However, it has undergone a modification and result in a new item that has a different meaning.

(63) *kūd a-rūh maça-hum li-l-masdžed*  
 hope 1SG-go with-them to-DEF-mosque  
 ‘I hope I go with them to the mosque.’

*Yallah* ‘hardly’ has a different meaning than the above mentioned modals. This is a special use of the phrase *ya allah* ‘my God’ and carries a special function that implies the difficulty of the situation. It is regularly followed by a verbal structure as in the following:

(64) *yallah ti-gdar t-tašil fē-hum*  
 hardly 2SG.M-can 2SG.M-phone them  
 ‘you can hardly phone them.’

## 4.6 Syntax

This section is a treatment of syntactic structures in NA. Two basic sequence types of sentences exist in MSA: the verbal sentence and the nominal sentence. The verbal sentence, which is referred to as ‘uninodal’, is the one in which the verb precedes and is regarded as the basic element of the sentence. The information is presented as a unit and regarded as wholly new. In such structure, the verb is regarded as the root which carries much of the action of the sentence reflecting the verb subject and sometimes also its object. The actual word order in NA is both subject-initial and verb-initial. Examples of verb-initial structures:

(65) *zer-na eḥna ahāl-na*  
visited-1PL we family-POSS.1PL  
‘we visited our families.’

(66) *nis-a hawwa kill az-za‘al*  
forgot.3SG.M he all DEF-anger  
‘he is not angry anymore.’

The information in uninodal sentences may be all new as in examples (65) and (66). It might be that given information is added to the new information as in the following examples:

(67) *yi- ḡrib-ih walad l-jīrān*  
2SG.M-hit- him son-CONS DEF-neighbours  
‘he is hit by the neighbours’ son.’

The nominal sentence, which is referred to as ‘binodal’, has two parts—the *mubtada*‘ (the subject), and the predicate *xabr* (information about the subject). The subject-initial structure in NA is similar to to the nominal sentence in MSA. The subject is usually a definite noun, either a proper noun like *Fatma*, *Ali*, and *Maha*, a

noun with a definite article like *al-walad* ‘the boy’, *al-mādzlās* ‘the sitting’ or a pronoun like *haw* ‘he’, and *hum* ‘they’. The predicate is usually indefinite and agree in gender and number with the subject.

(68) *il-walad safer dubai*  
 DEF-boy travelled-3SG. Dubai  
 ‘the boy travelled to Dubai’

(69) *axt-i t-dres fī il-džamça*  
 sister-POSS.ISG studys-3SG.F in DEF-university  
 ‘my sister is studying in the university.’

An explanation of NA syntax in terms of simple and complex sentence will be given in the following sections.

#### 4.6.1 Simple sentence

##### 4.6.1.1 Declarative clauses

In the typical order of the constituents in a declarative sentence the subject has the initial position followed by the verb. Both subject and verb agree in gender and number. The object is suited after the verb.

(70) *hīyya t-zūr ahl-hā kill yom*  
 she 3SG.F-visits family-POSS.3SG.F every day  
 ‘she visits her family every day.’

(71) *ahl-ī raḥ-aw l-bār*  
 family-POSS.ISG went-3PL to-DEF-desert  
 ‘my family go to the desert.’

In few cases, the subject of a declarative clause can be located at the final position of the clause.

(72) *rāḥ*                    *maḥā-hum*    *ux-ī*  
 went.3SG.M with-them    brother-POSS.1SG  
 ‘my brother went with them.’

(73) *hum*   *ḥazm-aw*    *kill*   *ən-nās*  
 they invited-3PL all DEF-people  
 ‘they invited all the people.’

Declarative clauses are negated by the use of the particle *ma*.

(74) *ma*   *ḥazm-aw*    *kill*   *ən-nās*  
 NEG invited-3PL all DEF-people  
 ‘they did invite all the people.’

#### 4.6.1.2 Interrogative clauses

Both verbal and nominal sentences can be transformed into yes/no question by raising the voice pitch on the part of the sentence to be questioned.

(75) *kallam-t*            *ḥeṣṣah ?*  
 spoke-2SG.M hessah  
 ‘did you speak to Hessah?’

(76) *hā ya-bu maḥammad,*    *kallam-t*    *ḥeṣṣah ?*  
 Oh-father Muhammad, spoke-2SG.M Hessah?  
 ‘Abu Maḥammud, did you speak to Hessah?’

Interrogative statements of yes/no questions can be formed by raising the voice pitch of the conjunction ‘*aw lā*’.

(77) *ta-bi*                    *t-rūḥ*            *li-l-malāhi*    ‘*aw lā* ?’  
 2SG.M-want 2SG.M-go to-DEF-funfair or not  
 ‘do you want to go to the funfair or not?’

Yes/No questions can also be formed by sharply raising the intonation on the parts of the sentence to be questioned and repeat negatively the verb at the end of the sentence.

(78) *ti-dfāç ya-bu fahad 'aw ma ti-dfā'?*  
 2SG.M-pay Oh-father Fahad or NEG 2SG.M-pay  
 'Abu Fahad ,are you going to pay or not?'

Interrogative statements can also be formed by the use of interrogative elements. In such cases, the interrogative element is left in-situ and can occur in one of three positions within the sentence: initial (as in example (79)), medial (as in example (80)), or final (as in example (81)):

(79) *wiš 'axbār çali wa çyāla?*  
 what news Ali and children-POSS.2SG.M?  
 'how is Ali and his children?'

(80) *çali wiš 'axbār-a?*  
 Ali what news-POSS-2SG.M?  
 'how is Ali?'

(81) *çali rāh çnd mīn?*  
 Ali went.2SG.M to whom?  
 'to whom did Ali go?'

#### 4.6.1.3 Construct

An investigation of the construct structure in NA can be based on the CA theorizing of construct '*iḍāfah*' literally 'addition'. The construct structure signifies a genitival relationship where the first noun is considered to be 'possessed by' or 'attributed to' the second noun. Therefore, the relationship between the two nouns is a possessive one as in *bīat ar-rādžal* 'the house of the man'. In some cases, the

relationship between the two nouns in a construct structure may indicate some attribution as in *šāhar ramzān* ‘the month of Ramadan’.

In the construct structure, the modifier follows the head noun and they are both not conferential and stand in a genitive relationship with each other. This construct genitive structure can itself be complex as the string may have three nouns in which each one is the possessor of the preceding one. Constructs of more than three nouns are possible but not common.

- (82) *bāb sāyr-t oxu çali*  
 door-CONS car-POSS brother-POSS.Ali  
 ‘the door of Ali’s brother’s car.’

Possessed nouns are marked as definite or indefinite as in the following examples:

- (83) *hāḏa bīṭ rādžal*  
 this house-CONS man  
 ‘this is the house of a man.’

- (84) *hāḏa bīṭ ar-rādžal*  
 this house-CONS DEF-man  
 ‘this is the house of the man.’

#### 4.6.1.4 Modification

A number of nouns follow each other in a string with the first being the head noun and the following nouns being the modifiers. The recorded data and observation of NA shows that there is no limit on the number of the modifiers. The head and its modifiers agree in number, gender and definiteness. Therefore, if the head is definite (a proper noun or marked by the definite article -əl), the following modifiers will be definite. Example of masculine singular head and its modifiers is:

(85) *šif-na rādžal ḥabīb*  
saw-1PL man kind  
'we met a kind man.'

(86) *šif-na ar-rādžal il-ḥabīb*  
saw-IPL DEF-man DEF-kind  
'we met the kind man.'

Example of feminine singular head is:

(87) *šāry-it sāça ġālyia*  
bought-1SG watch expensive  
'I bought an expensive watch.'

(88) *šārryt as-sāça l-ġālyia*  
bought-1SG DEF-watch DEF-expensive  
'I bought the expensive watch.'

#### 4.6.1.5 Conditional clauses

Conditional structures are those having the particles *yom*, *ila*, *ila min*, *in*, *lo*, *kān*, *laman*, as clause initial markers. These particles are equivalent in meaning to the English 'if'. There is a strong preference in NA to begin with the conditional clause followed by the main clause. Some of these structures may have reference to the future by using *ila*, *in*, *lo* or to the present marked by *ila* with the main verb in the perfective form.

(89) *ila riḥ-nā džeddah, šāryi-nā wāḥid džidīd*  
when went-IPL Jeddah, buy-1PL one new  
'as soon as we arrive to Jeddah, we will buy a new one.'

The particles initiate the conditional clause may not be located at the beginning of a sentence.

(90) *il-‘akil    ıla    xalaş    bi-n-ak-la*  
 DEF-food if    finish.3SG.M eat-1PL-3SG  
 ‘if the food is ready, we will eat it.’

The use of the particle *in* is also frequent in NA. It only has the meaning of the English word ‘if’. It is commonly used with a future reference and followed by a verb in a perfective form.

(91) *In ligē-t    çali, gūl-a    yi-džīb    as-sāyara*  
 if meet-2SG.M Ali, tell.2SG.M-him 3SG.M-bring DEF-car  
 ‘if you meet Ali, tell him to bring the car.’

The particle *yom* literary ‘day’ initiates a time clause followed by a verbal clause in the perfective form. The particle *yom* is frequently preceded by the preposition – *wa* ‘and’.

(92) *wa yom daxal-at    l-bint, rāḥ-aw    kil-hum*  
 and when entered-2SG.F DEF-girl, went-3PL all-them  
 ‘and when the girl entered, they all went out.’

#### 4.6.2 Complex sentences

Complex sentences are those in which two or more simple sentences are combined to form one larger sentence. In complex sentences, one of the constituting simple sentences stands as an independent sentence while the other sentence functions as a subordinate part of it. For example,

(93) *xāll-a    ax-ū    y-stan-a*  
 left- 3SG.M brother-POSS.3SG.M 3SG.M-wait-him  
 ‘he left his brother waiting for him.’

The simple sentence ‘he left’ has for an object the entire sentence ‘his brother is waiting for him’.

Coordination and subordination are two kinds of complex sentences. The coordinating conjunctions used in NA are *w* ‘and’, *lammā* ‘when’, *ham* ‘too’, *mu bas..baṣād* ‘not only ..but also’, *ṣašān* ‘because’.

(94) *hawwa ṣand-a waladīn w ana ṣandī walad waḥid*  
 he POSS-3MSG two boys and I POSS-1SG boy one  
 ‘he has two boys and I have one.’

(95) *hāḏī il-bint mū bas t-dres fī ədž-džāamça*  
 this DEF-girl not only 3SG.F.study in DEF-university  
*baṣād ti-štiḡl bi əl-līl*  
 but also 3SG.F.work at DEF-night  
 ‘this girl is not only studying at university but also she works at night.’

A subordinate clause is marked by a conjunction at the beginning of the clause. Examples of a subordination category are relative clauses, complement clauses and adverbial clauses.

The relative clause is marked by the relative pronoun *əllī* initiating the clause.

(96) *il-bint əllī ta-dres fī ədž-džāamça ti-zwajāt*  
 DEF-girl REL 3SG.F-study in DEF-university 3SG.F-married  
 ‘the girl who studies at university has married.’

Complement clauses are of three kinds; factual, non factual and manipulate. A factual complement clause is introduced by the particle *ənna* ‘literally means that’ which is attached to enclitic forms of pronouns *-i, -nā, -k, -kum, -h, -ha, him, -ik, -kum*.

(97) *ṣaraf-t ənna-ha amṭārat aməs*  
 knew-1SG that-3FSG rained.3SG.F yesterday.  
 ‘I knew that it had been raining yesterday.’

Non-factual clause is also introduced by the particle *ənna*.

(98) *atwāqaṣ ʾanna-i ba-wṣal bukra*  
 ISG-expect that-ISG arrive.1SG tomorrow  
 ‘I expect that I will arrive tomorrow.’

Manipulate clauses are introduced by *illi*.

(99) *ana abġa-h huwwa illi y-ftaḥ aš-šubāk*  
 I 1SG-want-3SG.M he that 3SG.M-open DEF-window  
 ‘I want him to open the window.’

Adverbial clause is marked by adverbs initiate the adverbial clauses.

(100) *kallam-na-hum lamma səmaṣ-na l-xabar*  
 talked-IPL-3PL when heard-1LP DEF-new  
 ‘we talked to them when we heard the news.’

#### 4.7 Conclusion

The main concern of this thesis is an investigation of the representation of NA in synchronous CMC. It aims at answering the question of what it means to communicate in text-based CMC when Najdi internet users’ native language lacks an official standard orthography. Specifically, it aims at tracing different linguistic and literacy practices online and investigates their communicative functions in shaping IRC discourse. This chapter presents a descriptive analysis of spoken NA in terms of phonology, nominal and verbal morphology, and syntax.

There are different aspects that differentiate NA from MSA. Consider the following two examples.

NA *ar-redžaldžīl ma y-ḥbōn š-šīġel*  
 DEF-men NEG 3PL-like DEF-work  
 ‘The men do not like work.’

MSA    *'a-redžālu la      yu-ḥbōn      al-ḥamala*  
 DEF-men NEG 3PL-like      DEF-work  
 'The men do not like work.'

Lexically, we observe that the word for 'work' is *š-šigēl* in NA but *al-ḥamala* in MSA. The word for 'the men' has different pronunciations in NA and MSA. It is *ar-redžaldžīl* in NA but *al-redžālu* in MSA. There are typically also differences in function words, in the example of NA sentence, *ma* is used for 'not' while in the example of MSA *la* is used.

Phonologically, MSA has three short vowels, three long vowels and two diphthongs. NA displays a little variation from the MSA vowels. NA has five short vowels and five long vowels. There are 33 consonants in NA, five consonants more than MSA which are /p, v, tz, dz, g/. The first two consonants are borrowed from English and can be found only in loan words and foreign names. The three remaining consonants are productive.

The morphology of NA is similar to that of MSA in showing a morphological differentiation of different word categories indicated by early Arab grammarians (Ibn Aqeel 1980). These word categories include: (i) *'ism* 'nouns' which literally means 'name'; (ii) *fiʿl*, 'verb' which literally means 'action' and (iii) *ḥarf* 'particle' which literally means 'letter'. Though the nominal morphology in NA follows MSA in having *múfrad* (when there is one), dual *muḥanna* (when there are two), and plural *jamʿ*, the formation of dual in NA is different from MSA. It is only formed by adding the suffix- *-ayn* to the noun stem (e.g., *sāḥṭayn* 'two hour', *ktābayn* 'two books') in NA while in MSA, the dual is also formed by adding the suffix *-ān* to the noun stem (e.g., *sāḥṭātān* 'two hours', *ketabān* 'two books'). Another difference

between NA and MSA in this regard is that while the categories of singular and plural in NA are shown in the noun, adjective, verb and personal pronouns, the category of dual is shown only in the nouns and adjectives. On the other hand, all three categories of singular, plural and dual in MSA are shown in the noun, adjective, verb and personal pronouns.

Both MSA and NA have similar personal free and bound pronouns. However, the pronunciation of these pronouns is different (e.g. *anta* ‘you-2SG.M’ in NA is pronounced as *anta* in MSA). Another difference between NA and MSA is that there are no dual personal pronouns in NA. Duals are regularly referred to by plural forms. Bound pronouns in both NA and MSA are affixed to nouns indicating possession and affixed to verbs indicating direct object.

In regard of demonstratives, NA has different sets of demonstratives (e.g. *haðol* ‘those’ in NA stands for *ha‘wlā* ‘ in MSA). NA is also different from MSA in not showing dual demonstratives and in having short forms of demonstrative (e.g. *ða* ‘this’ stands for *haða* ‘this’ in MSA). As quantifiers are concerned, NA has some different lexical quantifiers items from those in MSA. Examples are *kill* ‘all, every’, *baçd* ‘some’, *wājid* ‘many, very’, *šwayy* ‘little’ *marrāh and bilmarrah* ‘very much’.

Like MSA, Adjectives in NA follow the nouns and are inflected for gender, number and definiteness. However, NA has adverb items that are not found in MSA (e.g., *baçdaīn, dāyām, badrī*) and other adverbs that are pronounced in a different way from MSA adverbs (e.g. *zāmān* in NA stands for *zamān* and *kiða* in NA stands for *hākaða*). NA prepositions indicating local relation that are not used in MSA include *gādām* ‘in front of’, *barra* ‘outside’ and *džwaa* ‘inside’.

The construct structure in NA follows that of MSA *'idāfah*' literally 'addition' that signifies a genitival relationship where the first noun is considered to be 'possessed by' or 'attributed to' the second noun. However, NA is different from MSA in showing possession only in the use of the possessive particle *ḥag* which is derived from the classical word *haq* 'right'. NA shows different lexical interrogative items (e.g., *wišu*, *wiš* 'what' in NA stands for *maḏa*, *warā*, *līš*, *wišula* 'why' all stands for the MSA item *limaḏa*). Other NA interrogative items are similar to those in NA with a different pronunciation (e.g., *tsīf* 'how' in NA is *kayfa* in MSA).

The formulation of the passive in NA, namely 'internal passive', reflects quite closely the *majhūl* form in MSA. However, the ways of forming 'internal passive' in NA is different from those in MSA where diacritical marks are used to change the verb stem into passive (e.g. *ḏarab* > *ḏureba*).

In MSA, particles like *laysa*, *lam* and *la* are used in the process of negation of verbal sentences. In NA, the two particles *ma* and *la* are basically used in the process of negation of verbal sentences. The item *la* is usually used with imperative verb forms while *ma* is used in all verbal statements. Some people of Riyadh use the particle *mū* and *mush* to negate verbal plain statements.

Syntactically, there are three differences between NA and MSA. First, in declarative clauses, the subject precedes the verb in NA (SVO order), but it follows the verb in MSA (VSO order). This is in fact not a strict requirement, but a strong preference: both varieties allow both orders. Second, both verbal and nominal sentences can be transformed into yes/no question by raising the voice pitch on the part of the sentence to be questioned in both NA and MSA. However, interrogative statements of yes/no questions in NA can be formed by raising the voice pitch of the

conjunction *'aw lā*. Third, unlike conditional clauses in MSA, the particles initiate the conditional clause may not be located at the beginning of a sentence in NA.

The following chapters are concerned with the representation of this spoken dialect online. Beginning with Chapter Five, the study investigates who use NA online as well as their general patterns of linguistic and literacy choices. A quantitative analysis of IRC participants' social profiles provides the initial empirical ground on which to build a more comprehensive examination of how NA is used online.

## CHAPTER FIVE

### IRC PARTICIPANTS: SOCIAL DIMENSIONS AND ATTITUDES TOWARDS USING NAJDI ARABIC ONLINE

The purpose of this study is to provide an empirical account of how NA is represented online. A preliminary investigation of the linguistic and orthographic practices that Najdi IRC participants employ to represent NA online reveals that CMC is best considered as a language form that is adapted to suit the particular settings in which it occurs. Herring (2001) claims that this new form of language used in CMC can be influenced not only by the CMC mode employed (i.e., multi-party vs. one-to-one, or synchronous vs. asynchronous), but also by the situational contexts and the social dimensions of internet users. In order to obtain a more complete picture of the representation of NA online, the researcher considers it necessary to investigate how NA is used in IRC in relation to the extra-linguistic factors of the IRC context.

Thus, the general aim of this chapter is to provide an account of the social profiles of IRC participants as well as their attitudes towards using NA online. In particular, the aim of this chapter is to present a quantitative, statistical study of IRC participants' choices of languages and scripts. In agreement with Herring and Paolillo, it is argued that "quantification provides the strongest basis for generalization across large data samples" (2006: 444). The chapter begins with an overview of CMC research that applied multivariate approaches to understanding online languages. Results of the "IRC Languages Questionnaire" will be presented and discussed.

## 5.1 Background CMC research

Although CMC research dates back to the early days of the technology in the 1970's, only a handful of researchers have begun to take internet users' social factors into account. The study of gender-related differences in CMC has a long and extensive history in the field, relative to CMC research as a whole. With a few exceptions (e.g., Herring 1993, Themistocleous 2008, Hinrichs (forthcoming)), investigation of CMC differences in relation to a set of internet users' social factors to see which factors correlated with these differences have been extremely under-researched.

Starting from the 1990s, numerous researchers have begun to research gender differences in CMC. What many researchers have been discovering and discussing is that the theories that have been developed to account for gender in face-to-face communication still apply in the new domain of CMC. Herring (1992, 1993, and 1994) studied gender differences in discourse style in posting messages to an academic forum. She found that females use more hedges and politeness markers. Compared with their male counterparts, they manifest more sympathetic attitudes towards their addressee. Male participants tend to make strong assertions and violate conventional norms of politeness. These findings resemble previously-detailed findings of gender-related differences in face-to-face discussions (Coates 1993, cited in Herring 2011). This designates the claim that internet users transfer their offline communication styles into CMC messages (Herring 2011).

Herring's claim has raised controversial arguments in regard to synchronous CMC. Some argue that previously-reported gender differences in face-to-face communication are breaking down in synchronous modes of CMC such as MUD

(Danet 1998) and IRC (Rodino 1997). On the other hand, Cherny (cited in Herring 2011) indicates that typical gender-related differences in face-to-face communication are replicated in playful language of a social synchronous MOO. In her study of ten IRC channels, Herring (2003) found that females smiled and laughed more than men, who were more aggressive in both style and attitudes towards addresses.

CMC studies that take participants' social dimensions into account within Arabic context are few in number. They are primarily concerned with the dominance of English in online communication and the use of colloquial vernaculars among young internet users. In a study of globalization and identity in Egypt, Warschauer *et. al.* (2002) investigated language choices among 43 "young professionals" in Cairo that were known to be working on the internet. Utilizing an email survey, he investigated what was the most common form of writing among young internet users in Egypt. English and Romanized colloquial forms of Arabic, which had very limited access and respect prior to the development of CMC communication, marked the most common choices of writing among participants.

Warschauer's findings were confirmed by Palfreyman and Alkhalil (2003), who investigated in a small-scale exploratory study how female university students use their non-standard vernacular online. Qualitative analysis of an IRC corpus and statistical analysis of seventy-nine email surveys revealed the emergence of a non-main stream orthographic system among female university students in United Arab Emirates (UAE). This emerging language form is marked by informal register, a fair amount of code-switching to English, and a Romanized form of UAE (non-standard vernaculars), and came to be termed "a funky language for teenzz to use", distinguishing the non-mainstream of teens' writing in UAE.

Themistocleous's (2008) study of Cypriot-Greek in synchronous CMC contributed greatly to the development of the "The IRC Language Questionnaire" used in this study. She investigated the use of Cypriot-Greek and Standard Greek in relation to age, gender, level of education, occupation, frequency of using the internet, and frequency of chatting through the use of an online questionnaire. She found that the choice between Cypriot-Greek and Standard Greek is only affected by age, occupation and frequency of chatting. The analysis reveals that the participants prefer to use the Cypriot-Greek variety online.

Hinrichs (forthcoming) treats the choice between standard and nonstandard spelling as a binary linguistic variable in exploring how the conditioning dynamics underlying spelling choices differ between Jamaican CMC writers at home and those in diasporas. He considered both external factors of gender and place of residency and internal factors of lexical meaning of the token and the linguistic environment (Creole versus English) in his study of Jamaican e-mails and blogs. The gender factor was found to have an insignificant effect on Jamaicans' choice between Jamaican English (JamE) and Jamaican Creole (JamC). Jamaican CMC writers living in diasporas were found to be much more likely than those living in their homeland to use nonstandard spellings for tokens that have a Creole-only function. Hinrichs explains that this means that diasporic writers tend to distinguish between the two codes JamE and JamC. Diasporic writers are less concerned about differentiating between them as they use nonstandard spellings to disambiguate between the two codes in instances where word forms are identical.

Though CMC research on gender-related differences is extensive, standard and non-standard spellings as well as languages choices are under-researched. That is what the present study aims to explore. The choice between MSA and NA both in

Arabic and Romanized script will be traced in this study. Moreover, the present study is significant in filling a gap in cross-cultural studies of CMC as it focuses on Arabic CMC and a vernacular that has not been researched yet. Studies of Warschauer (2002) and Palfreyman and Alkhalil (2003), though offering important implications about Arabic language in CMC, lack a methodological multivariate model that takes into account social dimensions in the new domain of CMC. Furthermore, in an Arabic context, participants' bilingualism and proficiency in English should be considered when dealing with language choices online. These two considerations will be taken into account in the this study.

The following sections report the findings obtained by the "IRC Language Questionnaire" and present the discussion of IRC participants' social profiles, as well as their attitudes towards using NA online.

## **5.2 Results of the "IRC Language Questionnaire"**

Two hundred and eleven questionnaires were collected and used for the analysis. Questionnaire results are presented in five main sections representing IRC participants' social profiles and their attitudes regarding the use of NA online, with reference to previous research and implications of the results. First, analysis of items 5, 6, 8 and 9 was used to establish IRC participants' profiles, including age, gender, nationality, level of education, and level of English proficiency. Item 7 was used only to exclude participants who threaten the homogeneity of sample representation; that is, participants who are not of a Saudi nationality. Second, items 10 and 11 were analysed to trace participants' internet use as well as their frequency of chatting. The third section - analysis of questionnaire items 12, 13, 14, 15, 16 and 17 - was used to

identify different patterns of language and script choice employed by IRC participants; that is, the kind of script they use to represent NA, their use of MSA and English. Questionnaire items 19 to 21 have been analysed to investigate the social context of using Romanised NA online and the different motivations underlying IRC participants' preference for representing NA in a Romanised script. Moreover, this section will explore how participants learned to write Arabic using Roman letters and numbers, and whether they use this kind of writing in offline settings like writing letters, notes, and SMS. The last section will be devoted to analysis of items 22 to 29, which are mainly concerned with the different attitudes IRC participants hold about writing NA online in relation to status of MSA.

### 5.2.1 Participants' social profiles

Table 5.1 IRC participants' social profile

<b>Age</b>	≤18	19-25	26-35	36-45	≥46
	24	83	82	22	0
<b>Gender</b>	Male	Female			
	87	124			
<b>Level of education</b>	Primary	Intermediate	Secondary	Undergraduate	Postgraduate
	1	10	35	140	25
<b>Level of English proficiency</b>	Beginner	Pre-intermediate	Intermediate	Upper-intermediate	Advanced
	13	39	32	82	45
<b>Total</b>	211				

As shown in Table 5.1, the majority of IRC participants belong to the 19-25 (83) age group, followed by 26-35 (82), then  $\leq 18$  (24), and finally the 36-45 (22) age group. None of the participants who responded to the questionnaire belong to the  $\geq 45$  age group.

In terms of gender, 124 of the total 211 subjects are female and 87 are male. The majority were in their secondary (35) and undergraduate (140) level of education. Twenty-five were studying for their postgraduate degree. Ten of the sample were in intermediate school and only one was in primary school.

The sample has different levels of English proficiency which mainly correspond with levels of education. The majority of participants had obtained an advanced level of English (45), and (82) were in the upper-intermediate level. This indicates that the majority of participants speak and write fluent English. A total of 39 participants were in pre-intermediate level and (32) were in intermediate one. Only (13) participants in the sample described themselves as beginners.

### **5.2.2 Participants' internet usage**

In terms of internet usage, 186 participants (88.2%) indicated they use the internet on a daily basis. Sixteen stated they use it 2-3 times a week, and only nine indicated they use the internet less than 2-3 times a week. Results are illustrated in Figure 5.1 below.

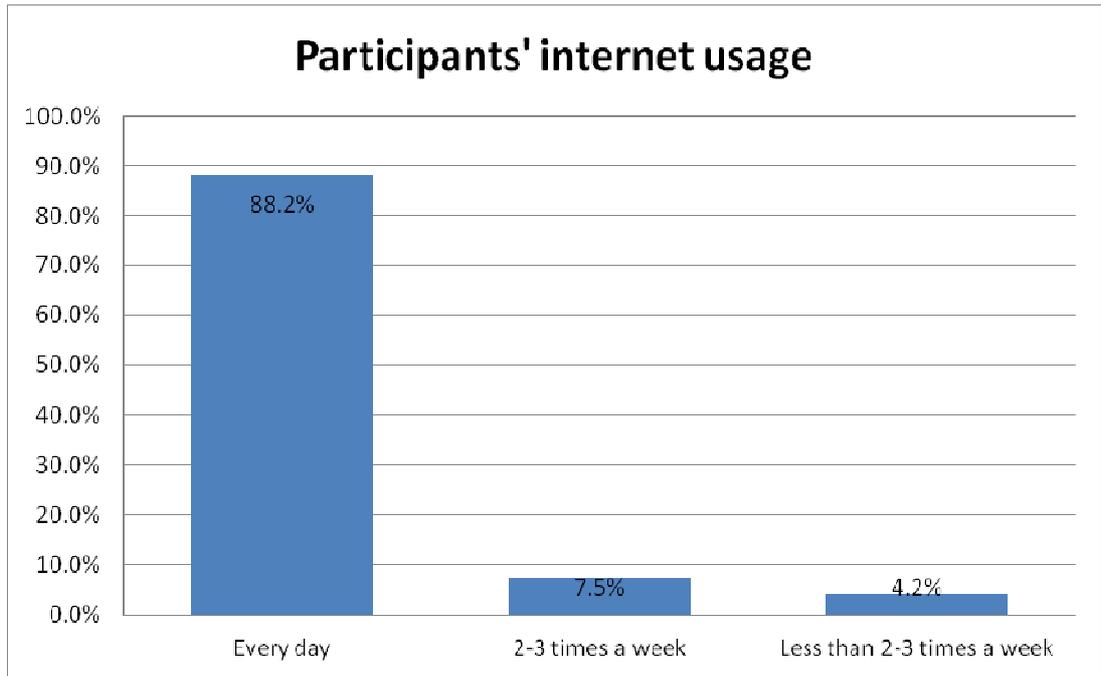


Figure 5.1 IRC participants' internet usage

The analysis shows similar results for the frequency of chatting, as the majority of participants (182) indicated they chat online every day. Fourteen participants indicated they chat 2-3 times a week, and 15 said they chat less than 2-3 times a week. Answers are illustrated in Figure 5.2 below.

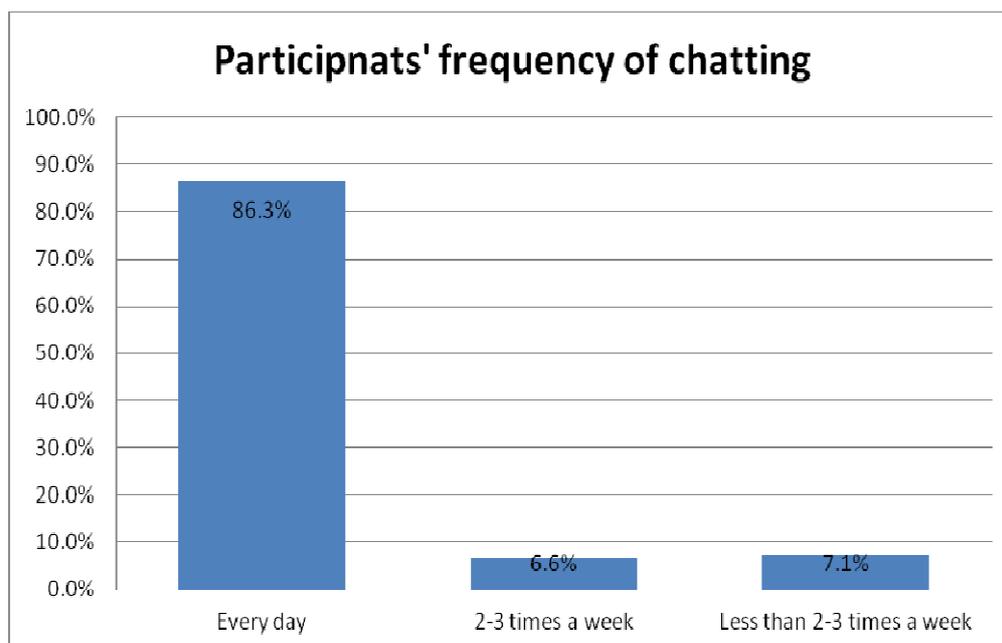


Figure 5.2 IRC participants' frequency of chatting

### 5.2.3 Participants' choice of script and language

The aim of this section is to explore the choice of language and script. The data reveals two levels of switching: switching between languages and switching between scripts.

An important aim of the present study was to explore IRC participants' preference of script. A preliminary investigation of the data illuminates unusual mixing of both Arabic and Roman script. Surprisingly, the sample reflects similar results. Seventy-six (36%) of the participants acknowledged they use Arabic script, 68 (32.2%) said they use English script (Roman letters and numerals), while the remaining 67 (31.8%) indicated they use both Arabic and English script online.

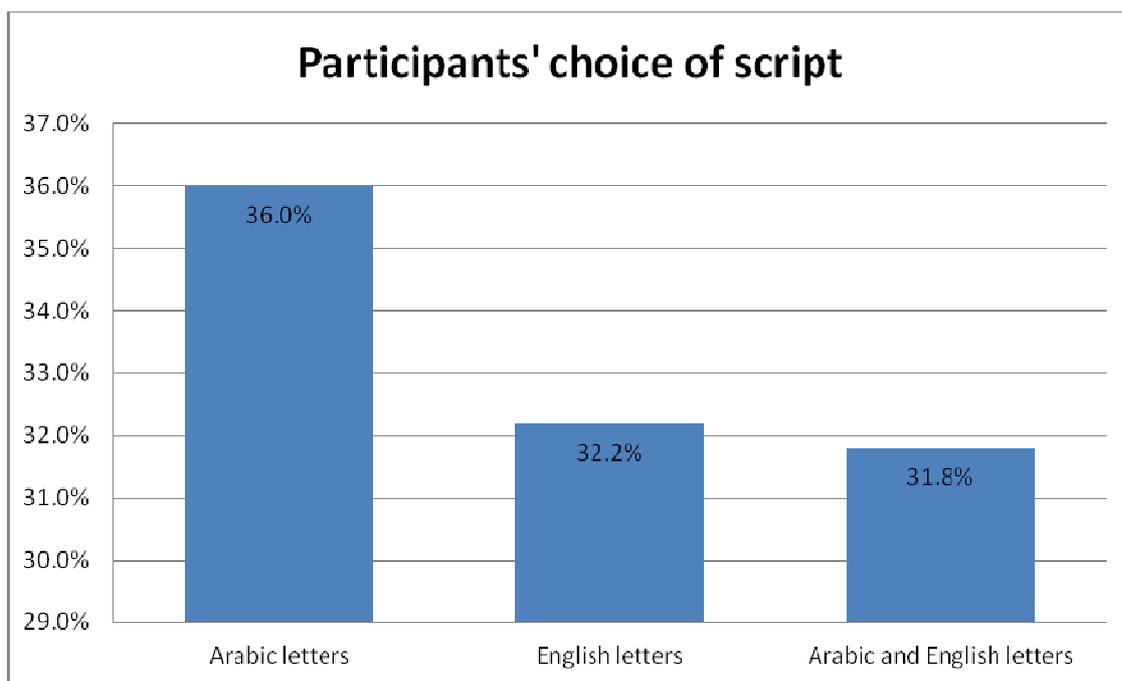


Figure 5.3 Participants' choice of script

Following the question about choice of script, IRC participants were given a series of different language choices and practices and asked to rate their authentic use of these choices using: “Never”, “Rarely”, “Sometimes”, “Often”, and “Always”. Since the IRC data reveals both alternational code-switching and insertional code-mixing<sup>12</sup>, it was necessary to assess both levels of code-switching. The statement ‘ I make use of a mixture of English and NA’ assess the intra-sentential level wherese the statement ‘I use code-switching to English in IRC’ assesses the inter-sentential code-switching. Arabic simple translation of these two statements are presented in the questionnaire. The language choices are:

1. I make use of MSA only in IRC;
2. I make use of NA only in IRC;

<sup>12</sup> Code switching occurs at the inter-sentential level, while code mixing occurs at the intra-sentential level (See Section 7.1.3.1).

3. I make use of a mixture of MSA and NA in IRC;
4. I make use of English only in IRC;
5. I make use of a mixture of English and NA in IRC; and
6. I make use of code-switching to English in IRC.

*The use of MSA in IRC:* MSA, the formal form of writing in Saudi Arabia, is seldom used by the 211 participants who completed the questionnaire, as no participants indicated that he/she used it all the time. The total of 115 participants (54.5%) responded that they never use MSA in IRC, while 61 participants (28.9%) indicated they rarely use MSA. Answers are illustrated in Figure 5.4 below.

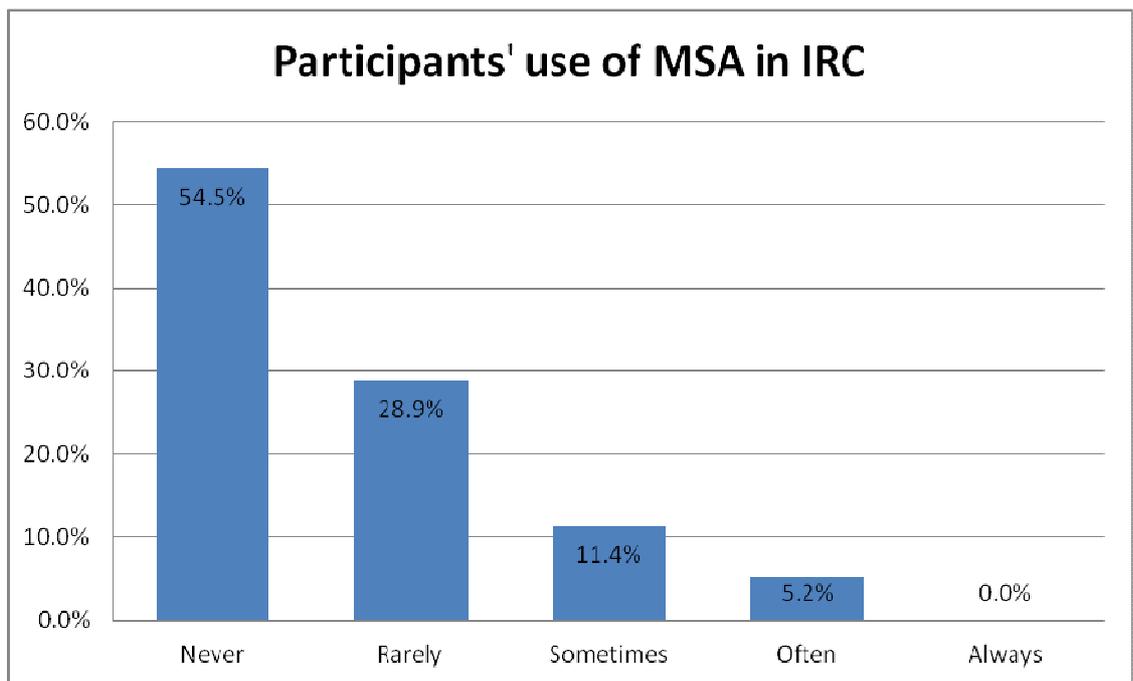


Figure 5.4 IRC participants' use of MSA in IRC

*The use of NA in IRC:* Contrary to the use of MSA in IRC, NA is shown to be used all the time by the majority of participants (58.3%). Sixty-eight (32.2%) participants indicated they use NA often, and only three participants (1.4%) said they

never use it in their IRC communication. Responses to this item are illustrated in Figure 5.5 below.

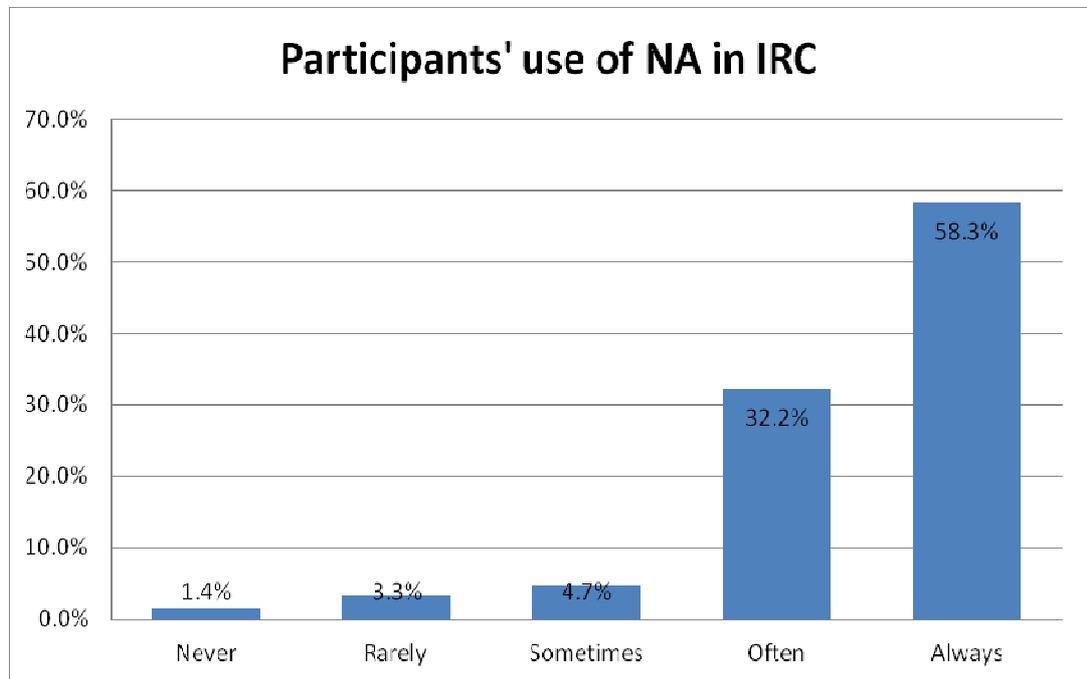


Figure 5.5 Participants' use of NA in IRC

*The use of both MSA and NA in IRC:* In regard to participants' use of a mixture of both MSA and NA, results of the questionnaire reveals a clear avoidance of using MSA online. Eighty-two participants (39.0%) acknowledged they rarely use both MSA and NA in their IRC communication, 47 (22.4%) said they never do, and only 11 (5.2%) indicated they always did. Answers to this item are illustrated in Figure 5.6 below.

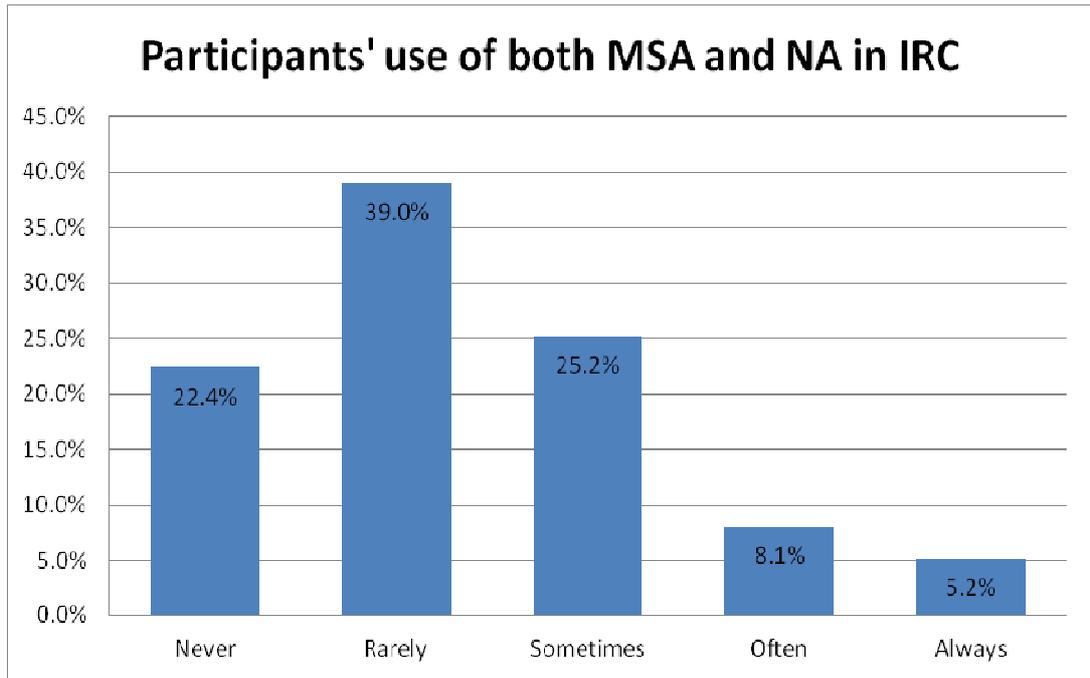


Figure 5.6 Participants' use of both MSA and NA

*The use of only English in IRC:* To ascertain the use of English in IRC, participants were asked to assess their use of English in IRC. Results indicate that the dominance of English is particularly strong in informal IRC communication. Fifty-eight participants (27.5%) indicated they sometimes use only English in IRC, 44 (20.9%) said they often use only English, and 16 (7.6%) confirmed they always use only English. On the other hand, 49 participants (23.2%) indicated they never use only English, and 44 (20.9%) agreed they rarely do. Results are illustrated in Figure 5.7 below.

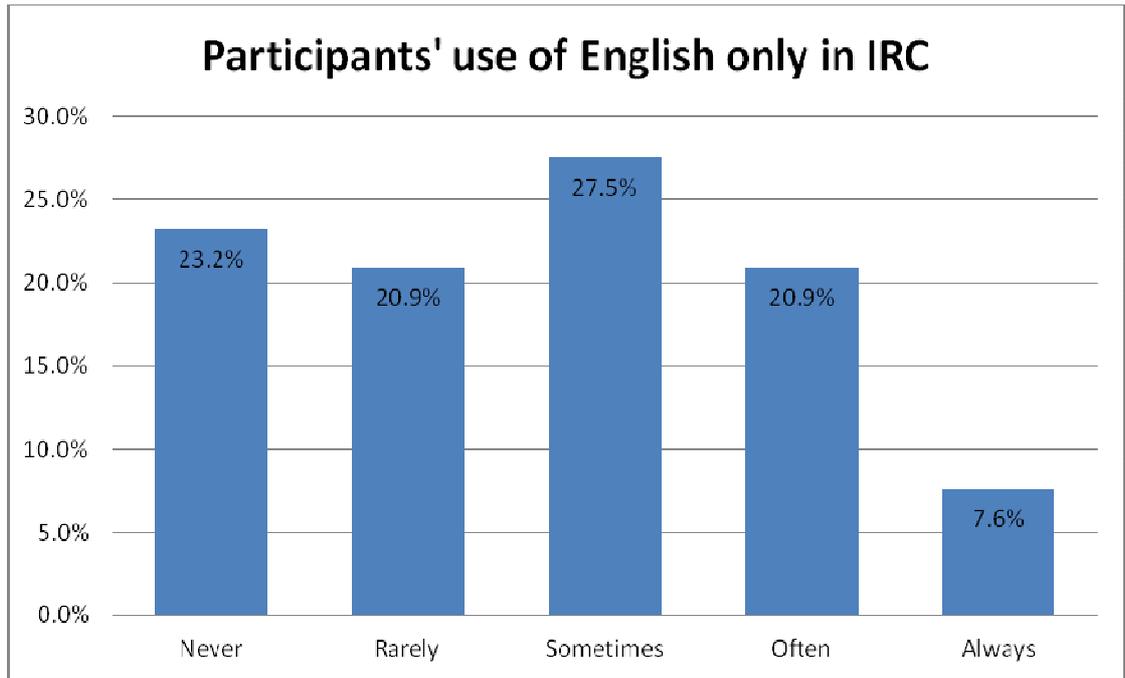


Figure 5.7 Participants use of English only in IRC

*The use of both English and NA:* Following the question of the frequency of using English only, it is important to assess participants' mixing of both English and NA. Only 21 (10.0 %) participants indicated they always do. Fifty-two (19.9%) said they often use English and NA, 58 (27.5%) said they sometimes do, 47 (22.3%) indicated they rarely do, and 43 (20.4%) never do. Results are given in Figure 5.8 below.

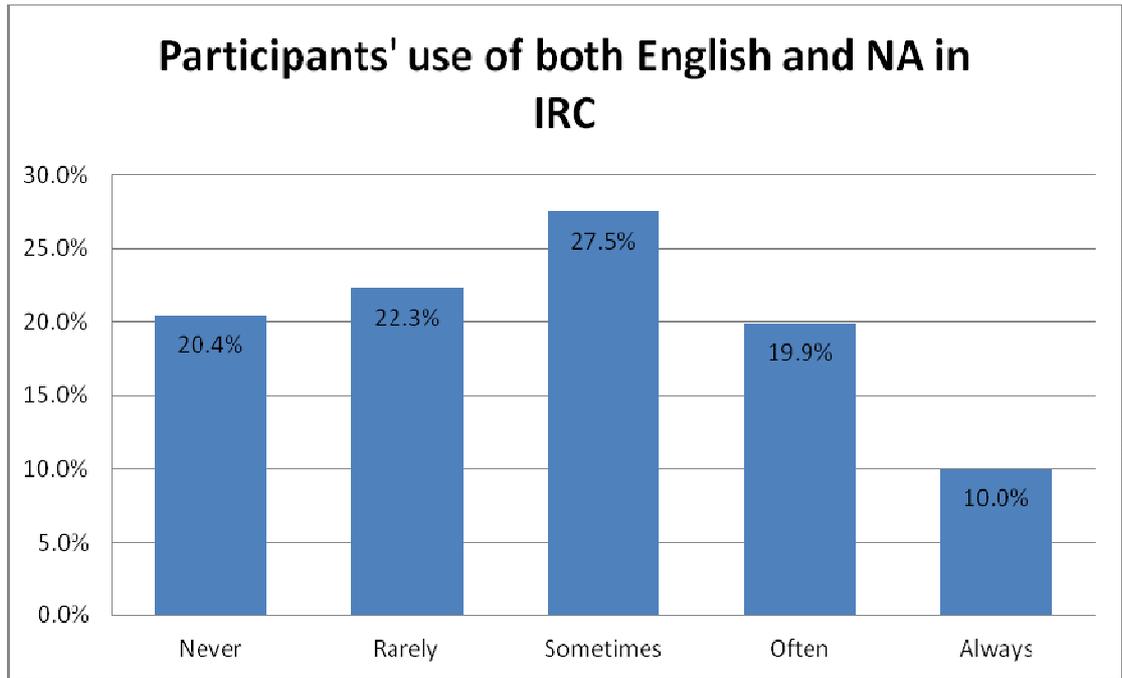


Figure 5.8 Participants' use of both English and NA

*Participants' code-switching to English:* To ascertain participants' code-switching practices in IRC, they were asked if they insert English items into their IRC communication. The clear majority (60) indicated they always do, and 49 (23.3%) said they often insert English items, while only 16 (7.6%) indicated they never do. Results are illustrated in Figure 5.9 below.

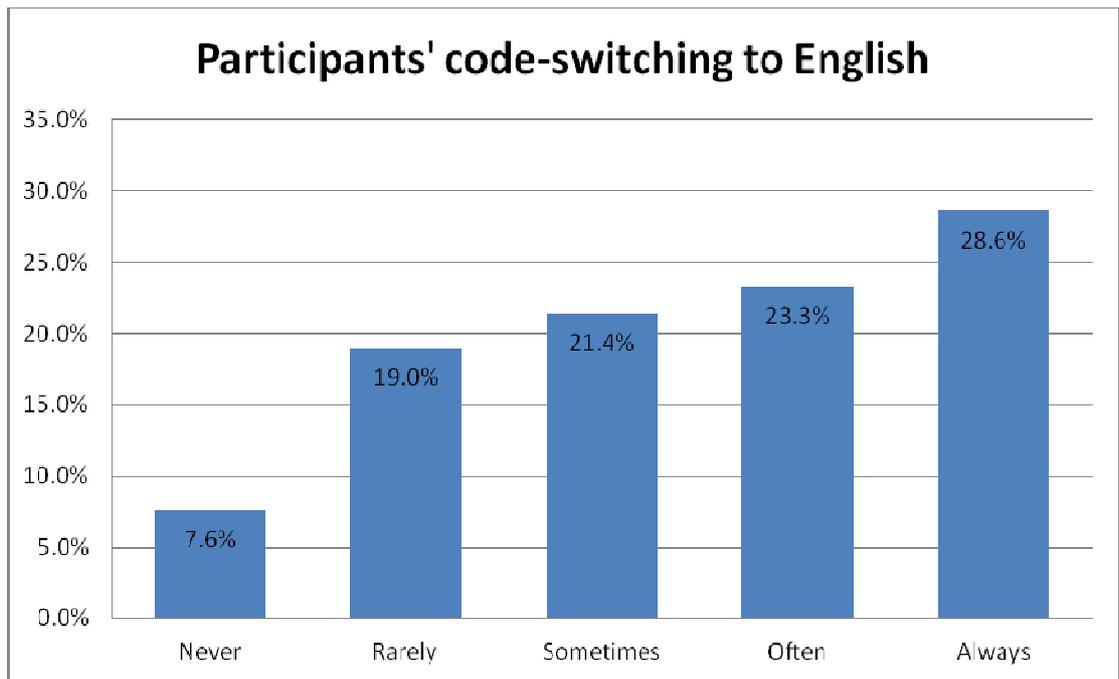


Figure 5.9 Participants' code-switching to English.

Results indicate that NA and English are the two main language forms used in IRC. On the whole, the presence of English in IRC is strongly manifested. The majority of IRC participants code-switch to English, with slightly less using either English only or NA only. The dominance of English is also reflected in participants' choice of Roman letters and numerals to represent NA. Nearly a third of participants indicated they only use Roman script to represent NA, and another third said they use both Arabic and Roman script. The establishment of NA in Romanised script as the language of IRC needs to be investigated in terms of reasons and motivations from the participants' perspective. The social context of Romanized NA is explored in the following section.

#### **5.2.4 Social context of Romanized NA in CMC**

IRC participants who indicated they use either Roman or a combination of Arabic and Roman scripts were required to answer three questions:

1. What are the reasons of using Romanized NA online;
2. Where did Romanized NA come from; and
3. Do they use Romanized Arabic in offline settings, such as writing down notes or letters?

*Reasons for using Romanized NA:* Participants were asked about the reasons underlying the use of this form of writing and were provided with three possible answers. They were also asked to provide other possible reasons and comments in order to expand data regarding this phenomenon. Surprisingly, ease of typing was chosen by the majority (91) as revealed by Figure 5-10 below. Participants reported that it is easier to write Roman letters and numbers on the keyboard than Arabic. This indicates participants' familiarity with the English keyboard layout. Although the Arabic keyboard is more accessible than before, internet users feel more comfortable typing English script. One participant indicated that he uses this form of writing as s/he types most of homework and other projects in English.

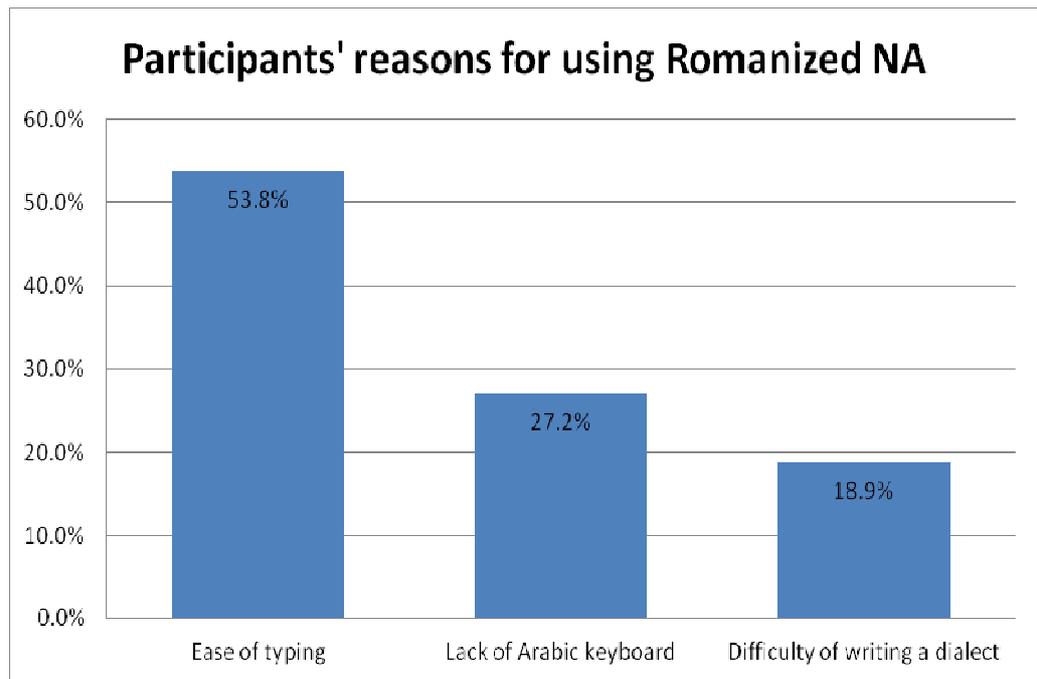


Figure 5.10 Participants' reasons for using Romanized NA.

Forty-six participants (27.2%) chose the technical factor, specifically, lack of Arabic keyboard. Several participants commented that this lack of technical support was truer in the past than now. Others noted they were abroad where the Arabic keyboard is not available, and they had to transcribe NA into symbols on their keyboards. Thirty-two participants (18.9%) indicated that the use of the Romanized form of NA is justified by the difficulty of writing NA down in Arabic script, which is related to MSA. Again, participants' familiarity with Arabic script, being the script of formal environment, together with technical factors shapes this new form of writing. A considerable number of participants commented explicitly on the positive connotations of Romanized NA. One indicated that this form of writing is used by educated people who use English in their daily life. Many commented that it is a prestigious way of writing which shows a great deal of elegance. Another participant commented that this form of

writing is a kind of a code s/he believes appropriate for his age, and that only young people could understand this way of writing. Others indicated it is the language commonly used within the IRC community. In other words, writing NA in Roman letters and numerals is an indication of belonging to the cyberspace community. All comments related Romanized NA to a positive connotation of IRC in-group belonging, values of education, proficiency in English and prestige.

Almost all comments provided by participants in this section were written in Romanized NA. Some inserted emoticons and multiple punctuations. Few comments were composed of English only. They believe that this form of writing is relevant for many kinds of online writing, even in responding to a formal questionnaire. Examples of IRC participants' comments are shown below:

1. “ma a2dree. bas e4tha nktem 3ala el,pc akeed el,ketaba t5telf” (“I do not know really. But if we write on the keyboard, the writing should be different”)
2. “3ashaan a7laa w marra funn w interesting :)” (“Because it is more elegant, funnier and more interesting”)
3. “it is more natural to communicate using thus way”
4. “i dont know really.. this s the way all people do when not face 2 face talking”
5. “may be because most of them do not have arabic keyborad. some get their devices and phones from abroad and so they should use english only”

6. “this is the way all my friends write and understand. it makes chatting very real and interesting”

7. “coz it s funn”

*Where did Romanized NA come from?* The majority of participants (126) indicated they learned Romanized NA through imitating others. Thirty-six (20.5%) said they do not remember how they learned this form of writing. Only (14) participants (8.0%) indicated that they learned to write Romanized NA by referring to rules published on the internet. Some indicated learning from other sources, such as TV and advertisements. Many commented that the main reason behind the emergence of Romanized NA is the lack of Arabic keyboards during the early stages of internet.

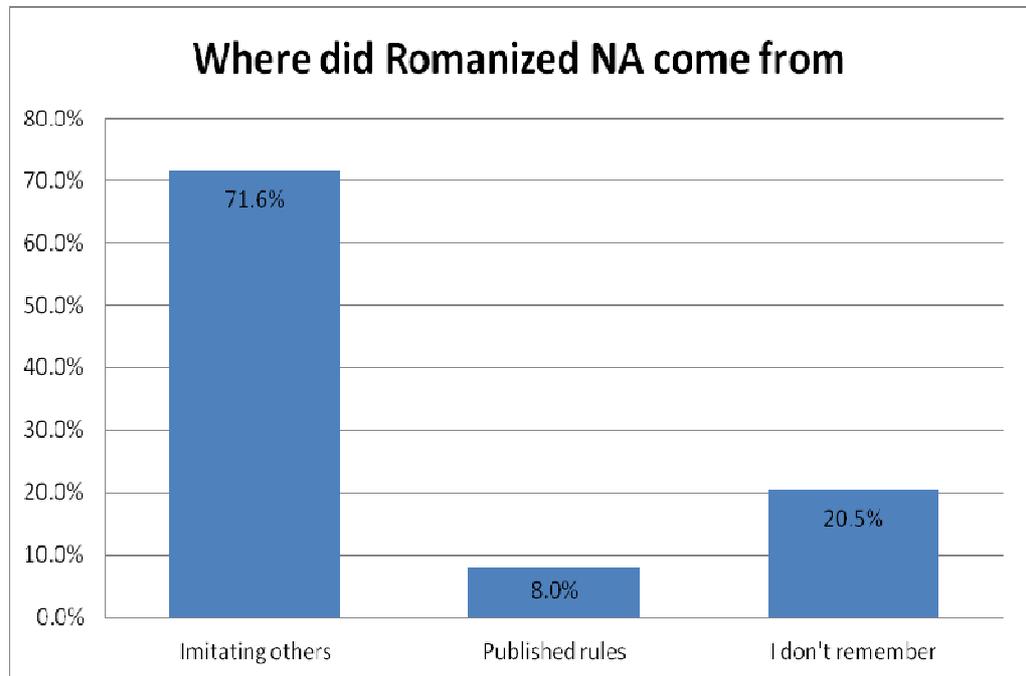


Figure 5.11 Where did Romanized NA come from?

*The use of Romanized NA in offline settings:* Participants were asked to assess their use of this kind of writing in mobile phone text messages, letters, and notes. The results that have been obtained show different attitudes in this regard. Fifty-eight participants (27.5%) indicated they often use Romanized NA in offline settings, and 28 (13.3%) indicated they always do. Fifty-seven (22.3%) said they sometimes do, and 53 (25.1%) indicated they never use it in offline texts, letters, or notes. Results are illustrated in Figure 5.12 below.

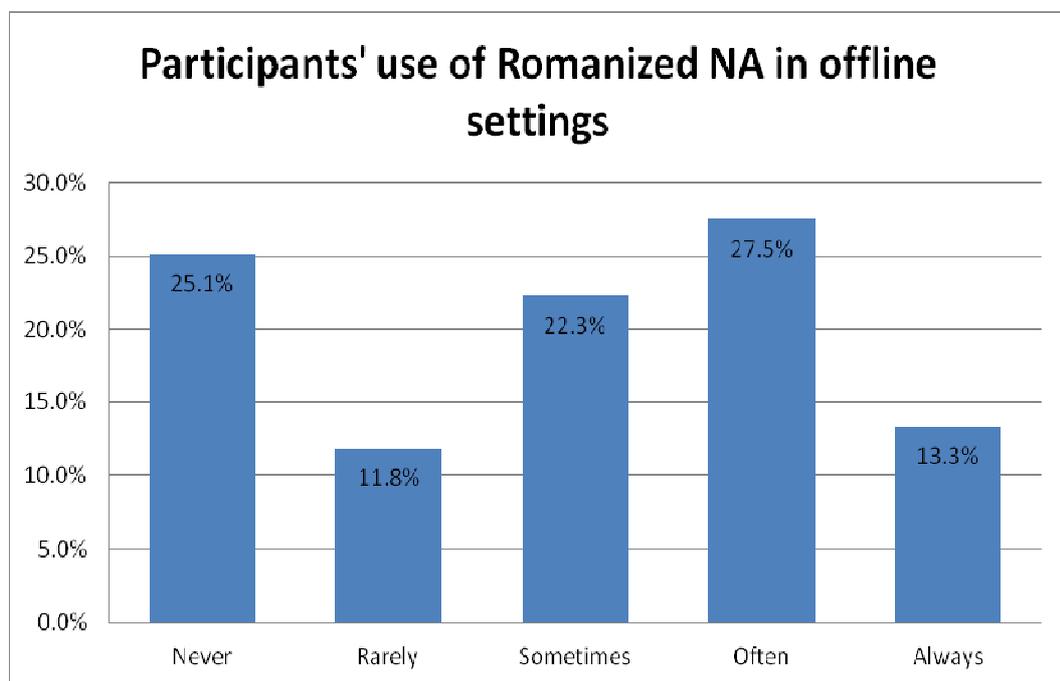


Figure 5.12 Participants' use of Romanised NA in offline settings

### 5.3.5 Participants' attitudes towards the use of NA online

An investigation of trends and patterns of the representation of NA online should take into account participants' perception of their literacy practices. This section aims to ascertain participants' general attitudes towards the use of NA online, as well as their perspectives about the status of NA in the social

community. Participants were given a series of statements implying different beliefs and attitudes, and were asked to evaluate these statements on the basis of agreement or disagreement. Below is a discussion and descriptive analysis of the results.

*Belief 1. Using NA online is damaging MSA status:* The aim of using this statement is to assess participants' perceptions about writing in NA. In the Arabic world, writing a vernacular has been always criticised as undermining the status of MSA. The question is: Does the same scepticism exist concerning the use of NA in IRC, where communication is always informal? Eighty-four (39.8%) participants strongly disagreed with the statement and 87 (41.2%) participants disagreed with that. The histogram below reveals very few participants agreed that using NA online is damaging MSA. This may account for participants' metalinguistic awareness that communicating in synchronous IRC is very different from formal contexts where MSA is supposed to be the language of communication.

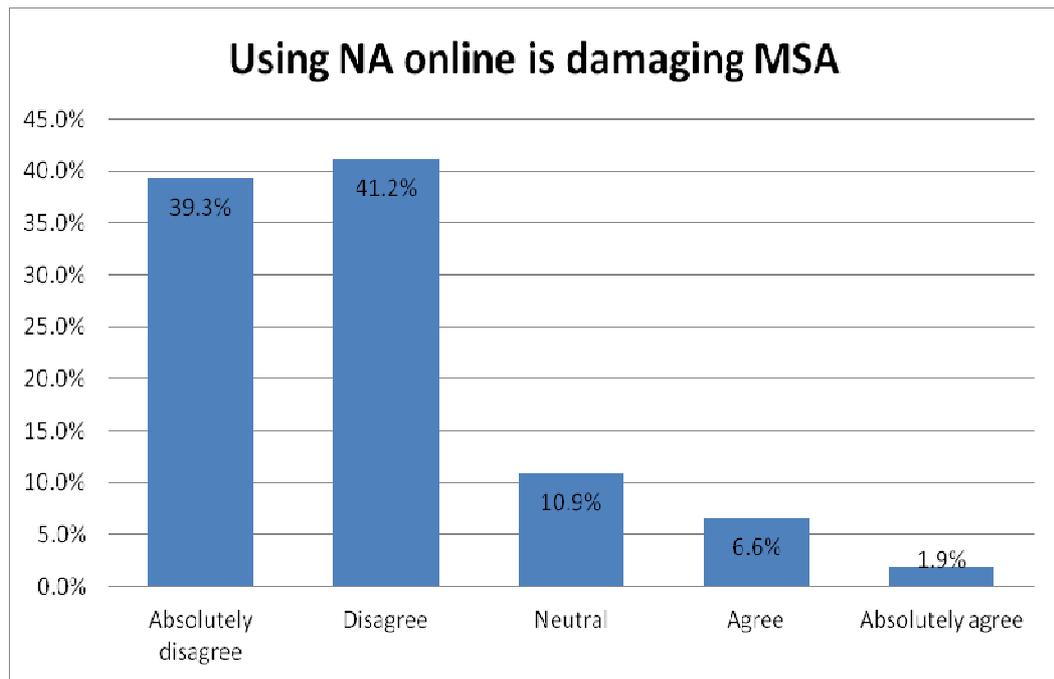


Figure 5.13 Histogram for Belief 1

*Belief 2. Reading and writing NA in Arabic script online is difficult for me:*

Participants' familiarity with Arabic script, being associated with MSA, has an effect on their choices when communicating in IRC. As participants are not generally familiar with reading texts in NA, it was interesting to assess their familiarity with reading and writing NA in Arabic script in IRC. The aim of this question was to uncover whether NA participants have difficulties reading and writing NA in Arabic script. Sixty participants (28.4%) strongly agreed they have difficulty reading and writing NA, and 34 (16.1%) agreed. On the other hand, 63 participants (29.9%) disagreed with the statement above, and 49 (23.2%) strongly disagreed. The remaining five participants held a neutral view. Answers to this question, as illustrated in the histogram below, indicate that participants' attitude in relation to writing NA in Arabic script is split in two contradictory attitudes. These two attitudes are reflected in their choice of script in IRC communication.

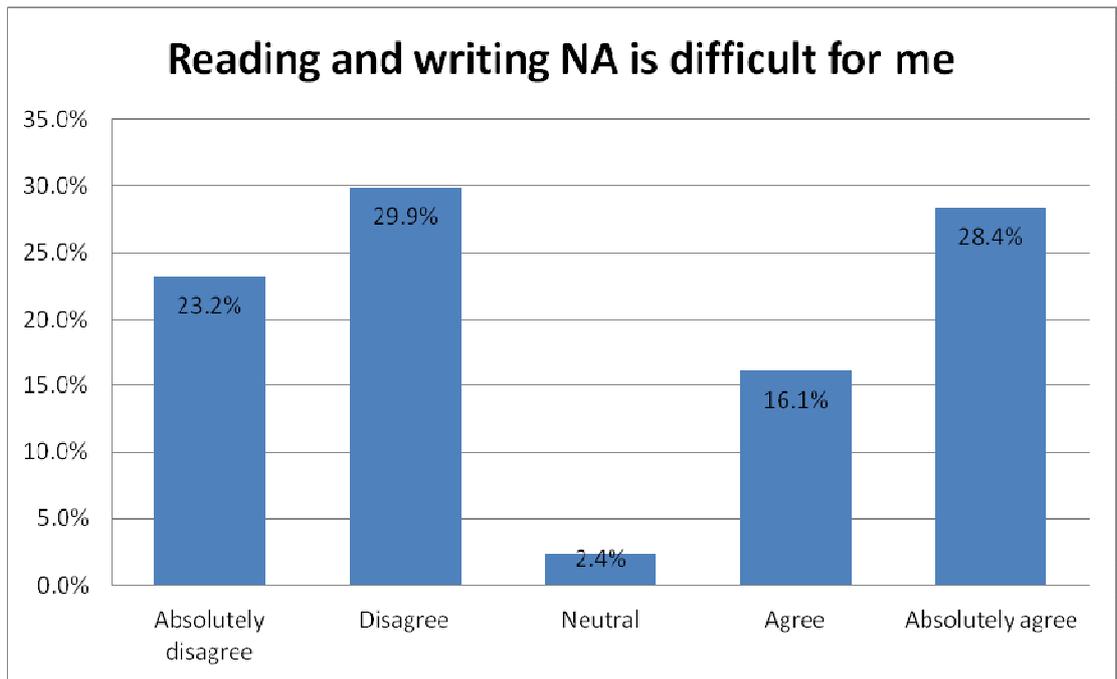


Figure 5.14 Histogram for Belief 2

*Belief 3. NA should not be used at all in written communication because it is a dialect not a language:* Following the aim of the previous two beliefs, Belief 3 directly asks participants to rate their dis/agreement with the above statement. In their literacy development, participants were expected to write in MSA, and were corrected if they misused grammatical rules. It is interesting to uncover their attitudes regarding NA in written communication where the audience is different from formal education. The clear majority of participants (121) disagreed with the statement and 47 strongly disagreed. 5 participants (2.4%) strongly agreed that NA should not be used in written communication and 14 participants (6.6%) agreed while 24 participants (11.4%) were neutral. It seems participants who do not feel comfortable reading and writing NA use English letters and numerals to write their non-standard variety. Answers are illustrated in Figure 5.13 below.

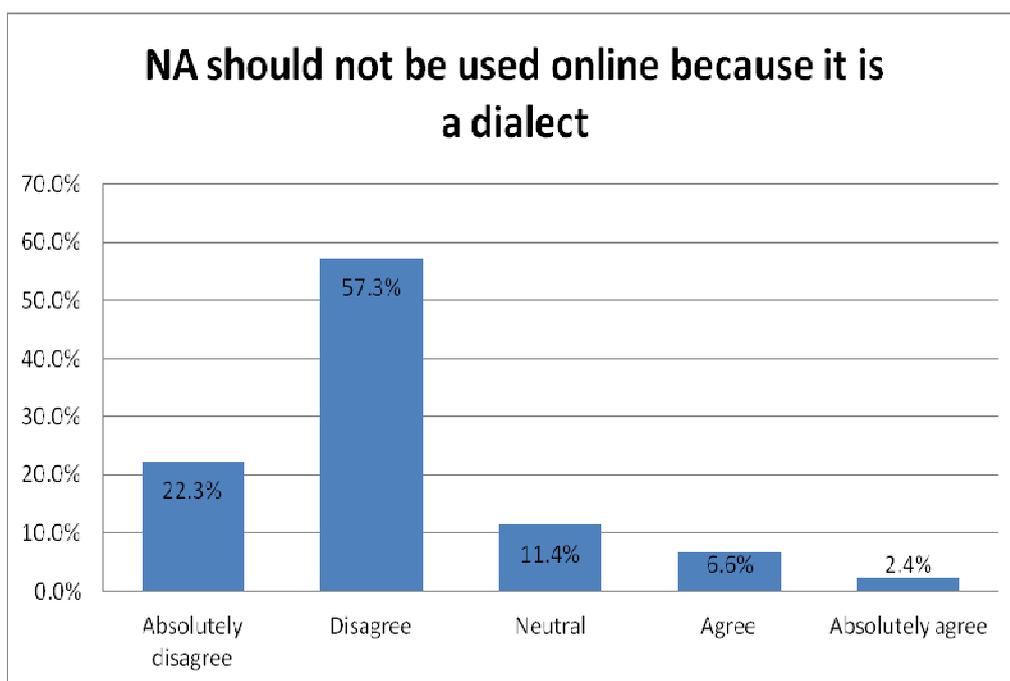


Figure 5.15 Histogram for Belief 3

*Belief 4. I use NA when chatting because I am actually speaking in IRC rooms while writing and it sounds more natural:* Being both interactive and written is what makes CMC resemble spoken language (Werry 1996, Schönfeldt 2001, Collect and Belmore 1993). The aim of this question was to find out whether participants use NA in IRC because it is the code of everyday oral communication, and therefore they find it suitable to use. The clear majority of participants strongly agreed with this assumption (107). Seventy-eight participants also agreed. Answers are illustrated below in the following histogram.

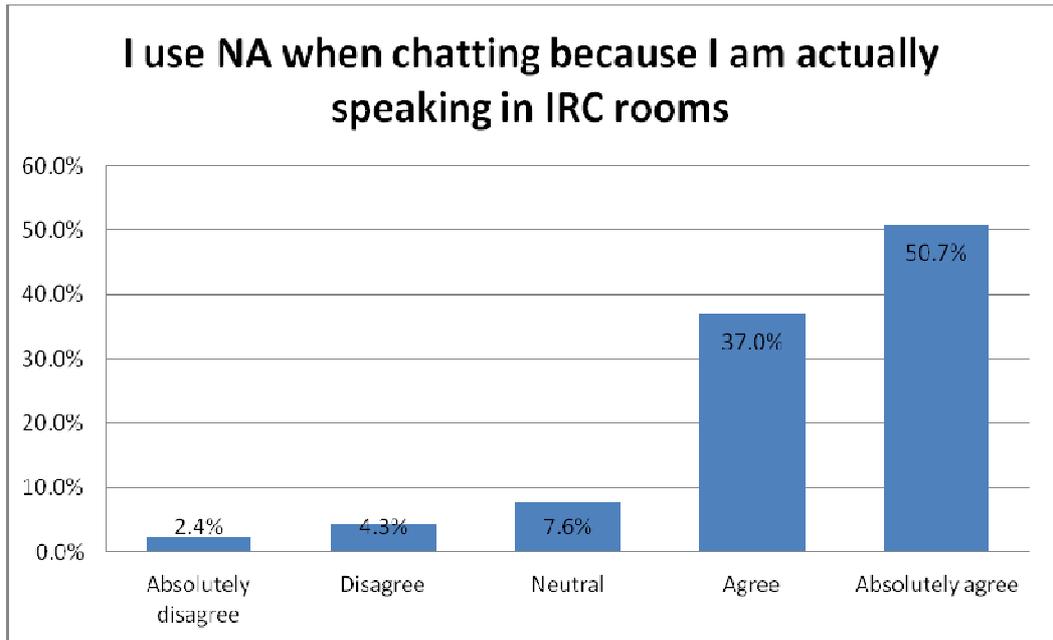


Figure 5.16 Histogram for Belief 4

*Belief 5. In the future, NA will be used in everyday “offline” written communication:* This question aims to investigate participants’ perspective about the future of NA with rapid technological development. Although the majority of participants prefer using NA online, very few agreed NA will be used in everyday written communication in future. However, most participants either disagreed with this belief (84) or were uncertain (68). This reflects participants’ perception about the association between MSA and offline written communication. Results are illustrated in the following histogram.

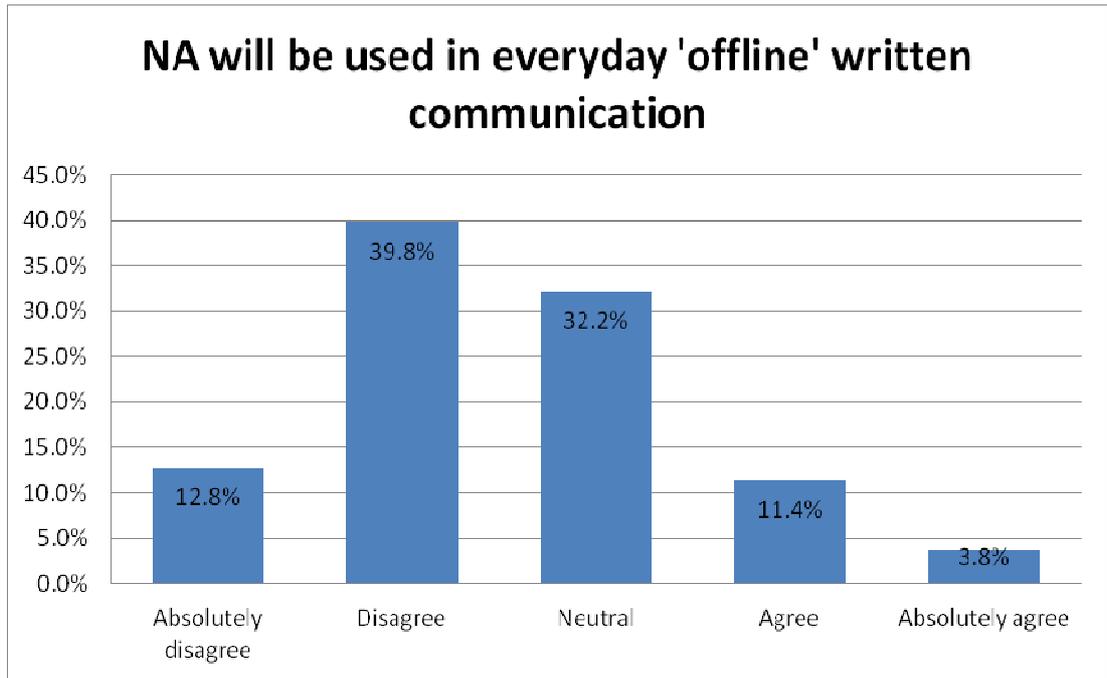


Figure 5.17 Histogram for Belief 5

*Belief 6. Writing NA in English letters and numerals is fun:* As the language of synchronous CMC is always described as being playful and creative (Danet 2001), it seems that representing NA online through English letters and numerals shapes connotations of fun and amusement, which is then associated with IRC informal communication. This new fun language has been invented, developed and identified by young internet users to be the code of online communication. This statement tests the participants' dis/agreement with this . Seventy-one participants (33.8) strongly agreed and 46 (21.9%) agreed with the belief. Details are illustrated in Figure 5.18 below.

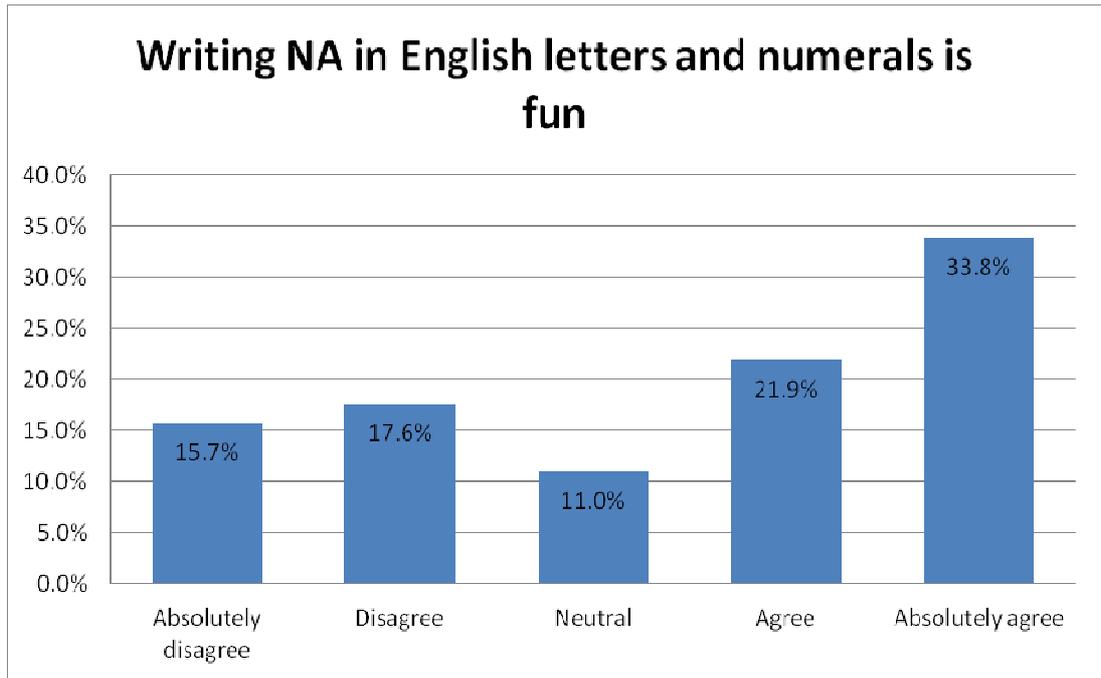


Figure 5.18 Histogram for Belief 6

*Belief 7. I find writing NA in Arabic script ugly:* This question is used in relation to Beliefs 2 and 3 which deal with the difficulty of reading and writing in NA, and whether it could be used in online communication. The current belief directly asks the participants to assess their attitude towards composing NA in Arabic script. Though a considerable number of participants do not feel comfortable reading and writing NA, the clear majority do not find it ugly. Ninety-two (43.6%) participants disagreed with the statement. Answers are illustrated in the following histogram.

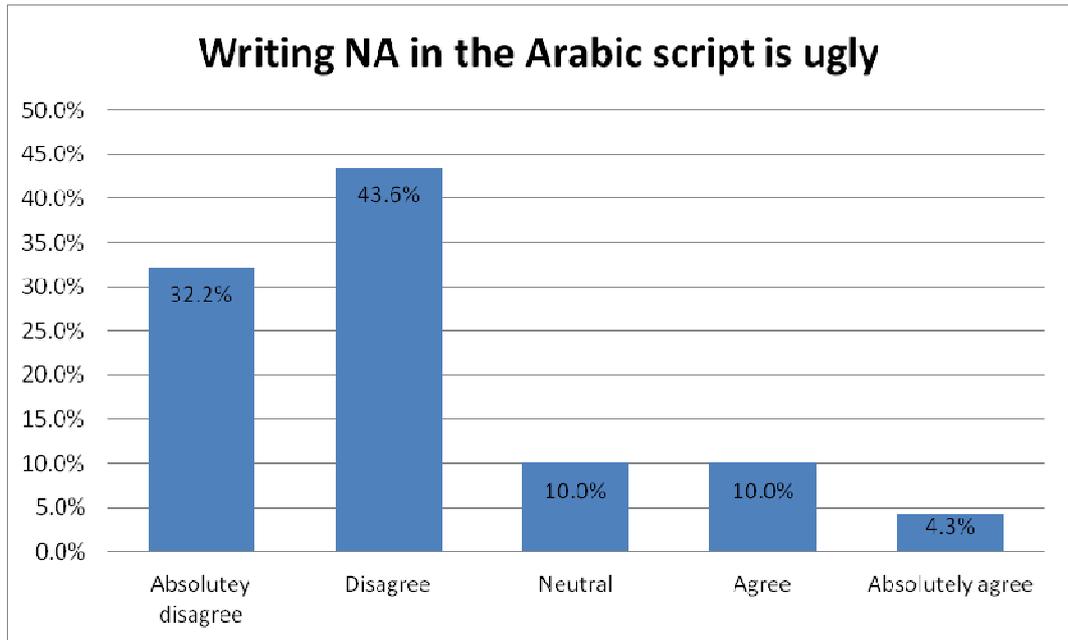


Figure 5.19 Histogram for Belief 7

*Belief 8. Writing NA in Roman script is a threat to MSA:* The aim of this question was to ascertain participants' perspectives about writing NA in Roman script, and whether or not it constitutes a threat to MSA. Although the majority of participants disagreed that using NA online is damaging MSA, 61 participants (28.9%) were uncertain and held a neutral attitude. The distribution of answers in the following histogram clearly shows that participants hold different positions on this belief.

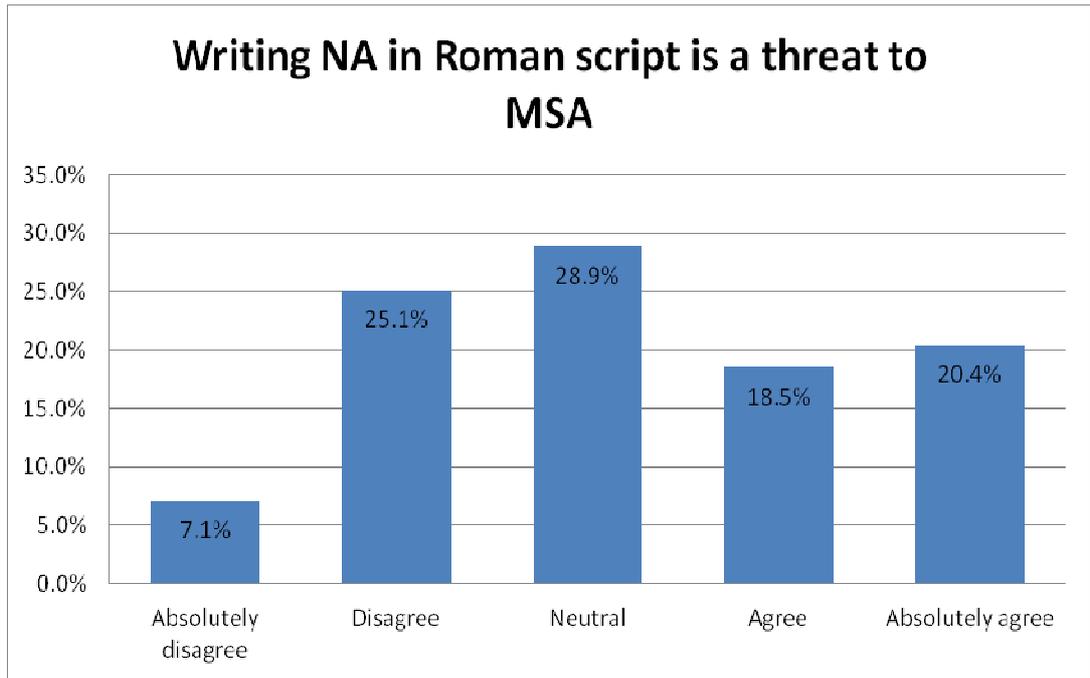


Figure 5.20 Histogram for Belief 8

#### 5.4 Conclusion

The predominant purpose of this study was to investigate the representation of NA in synchronous communication. This chapter aims to establish the social profile of IRC participants, their literacy choices, as well as their attitudes and values about how NA is used online. Results are quite suggestive of general trends to the use of NA online.

In regard to the social profiles of the IRC community, results reveal that the majority of IRC participants belong to the age group 19-25 and 26-35. They manifest high levels of education and proficiency in English. The majority of the participants use the internet almost every day and log in to IRC chat on a daily basis. Belonging to this age group agree with Androutsopoulos's (2000: 2) claim

about the “power of youth language” that creatively uses non-standard spellings in constructing anti-standard orthography that communicate social meanings and constitute a cyberspace code of communication. IRC participants in this study deliberately use innovative orthographic as well as graphemic resources mixed with English to represent a new form of NA online that is best considered to be the norm when communicating in text-based CMC.

The analysis of questionnaire responses reveals that there are two levels of code-switching: code-switching between languages and code switching between scripts. Basically, MSA is seldom used by IRC participants, while English and NA are commonly and frequently used. A Romanized form of Arabic, that had very limited usage prior to the advancement of technology, has been modified and adopted by IRC participants to transfer their non-standard variety into writing. The use of Romanized NA is justified by the ease of typing, lack of the Arabic keyboard and the positive prestigious connotation Romanized NA has.

As far as language attitudes towards using this Romanized form online mixed with English are concerned, IRC participants have a positive attitude towards using their non-standard variety online. The formation of this positive attitude could be connected largely with the participants’ perception of IRC communication as “talking in writing” (Collot & Belmore 1993: 14). It is interesting, however, to observe that the majority of IRC participants in this study not only have positive attitudes towards using NA online but also have a tendency to transcribe their non-standard dialect via Roman script. The strong presence of English online and technical constraints have both reinforced the emergence of Romanized NA online, which many participants find funny and

interesting. The tendency to use Romanized Arabic could also be related to the unfamiliarity participants encounter when writing NA in Arabic script. The majority of participants do not find Romanized NA corruptive to MSA. However, they do hold contradictory positions about the future of NA and whether it is a threat to MSA.

It is clear that the IRC participants in this study make a conscious literacy choice (using their non-standard dialect in online IRC interactions). This conscious choice of writing NA online can be best regarded as “emblematic” (Matras 1999: 495). By “emblematic”, the researcher means texts which are not intended to enhance the addressee’s knowledge in order to evoke actions on his part, but rather to elicit emotional identification on the part of the addressee with aesthetic symbolism of the texts in its particular language-external context. The choice of writing NA in IRC interactions is emblematic not only of the Najdi identity but also of the young Najdi internet users identity. Moreover, young Najdi internet users are aware that this literacy choice (writing NA online) allows them to transfer features of their non-standard dialect into online interactions and, therefore, allows them to differentiate themselves from non Najdi internet users within the global environment.

On the other hand, keeping in mind that for the majority of Najdi internet users, NA is a non-standard dialect that is seldom used in writing, deviating from the norms can be considered as identifying one-self with a particular social group. In this case, this social group is that of young internet users who are creating a novel code to be used in this new medium of communication.

## CHAPTER SIX

### ORTHOGRAPHIC REPRESENTATION OF NAJDI ARABIC IN IRC

Like all Arabic dialects, NA is not put into written form since to do so is viewed as undermining the image of CA. As the languages of the world came under pressure to meet the increased demands presented by technological advancement, the CMC platform has contributed to the maintenance of smaller languages and non-standard dialects by affording a space for their documentation and literacy support (Androutsopoulos 2004, Warschauer *et al.* 2002). These emergent practices of language use emphasize the argument that language in CMC is a language form isolated from both speech and writing, where writers have both the freedom and space to write their non-standard dialects. Thus, new orthographic systems are developed in order to represent these smaller languages and non-standard dialects.

The representation on non-Roman based languages has sparked interest in CMC research (e.g., Androutsopoulos 1998, 2000, 2004, Tseliga 2007, Palfreyman and Alkhalil 2007). It is the way internet users manipulate ASCII symbols to represent their non-standard dialects online that marks a major issue of investigation, and which shall be explored in the present chapter. IRC participants chose to deviate from the standard norms of MSA orthography in two ways: employing the Arabic orthographic system, which is most closely related to MSA; and relying on ASCII symbols, which is related to the lingua franca of the internet—English.

The aim of this chapter is to investigate the characteristics of Romanized NA online and the innovative means the IRC participants employ to represent their non-standard dialect in IRC communication. Specifically speaking, it aims at answering

the question of how IRC participants represent the consonants and vowels of NA via Roman characters.

This chapter is organized as follows: an overview of the theoretical background to the question of the orthographic characteristics of NA as represented in IRC will be provided first. Second, an account of empirical related CMC studies that synthesizes and summarizes significant findings will be given. Following this, methodological issues are presented. Then, discussion of the main findings will be given, and, finally, implications and conclusions will be drawn.

## **6.1 Orthography in CMC**

This overview begins with an outline of the relevant theoretical framework underlying the emergence of writing systems as well as sociolinguistic issues of writing. Terminological terms in the field of writing study research will be illustrated. A detailed review of CMC studies addressing the issue of orthography is conducted, with emphasis being given to relevant studies that contributed to the undertaking of the present study. To complete the background, an overview about the history of Romanized Arabic and its social significance will also be presented

### **6.1.1 Conceptual basis of orthography: A terminological framework**

The communication of meaning is the primary purpose of any writing system, and, in one way or another, conventional relationships between graphic and phonetic units are established to accomplish this (Coulmas 2003: 18). Meaning and sound are the two referential dimensions utilized by all writing systems. All writing systems

have phonetic and semantic interpretations; they differ in the importance attached to one or the other. Coulmas (2003) identifies four main dimensions of any writing system: visual perception, auto-indexicality, conventionality, and autonomy. Below is a brief explanation of these four dimensions.

- Visual perception: Across many theoretical differences, scholars of writing are agreed that writing is a form of communication created by the hand and appealing to the eye. It mainly consists of signs that have an assigned external referent. According to Coulmas (2003: 19), the ability of the central nervous system to “maintain a feedback circuit by using a visual input for controlling delicate manual movements [is] basic to writing”.
- Auto-indexicality: An important characteristic of writing is that written graphic signs, in addition to being codes learned by instruction, embody the principles of their learnability (DeFrancis 1989 cited in Coulmas 2003). Auto-indexicality refers to the systematic make-up of graphic signs that displays the procedures on the basis of which they can be learned and used (Coulmas 2003).
- Conventionality: Pictographs have communicative functions that might be associated with storytelling. Such an association is considered by Coulmas (2003) to be habitual but not conventional. Conventionality refers to the establishment of graphic signs into codes that assume an independent existence. The relationship between a code and its meaning is not ‘picturable’ (Coulmas 2003: 21), but rather can be deduced from the graphic composites and their arrangement.

- **Autonomy:** writing systems consist of units that, like sound systems, are structured and can be analyzed in terms of functional units and relationships. The distribution of units in writing are governed by their linear arrangement in forming larger units on the graphic level. In this sense, a writing system can be described as an autonomous linguistic system that can be studied independently of other levels of linguistic structure its units may refer to. The defining characteristic of autonomy of writing systems is that they are structured in a manner that interrelate with other levels of linguistic structures, such as phonemic and morphophonemic representation. These interrelations are often complex in nature and not always transparent.

To avoid confusion, it is necessary to clearly define references to symbols as well as the systems to which these symbols belong. To begin with a distinction between the terms ‘writing system’ and ‘orthography’, the definitions articulated by Baker (1997: 93, cited in Sebba 2009: 35) will be followed in this thesis. “Writing systems are graphic systems for representing a language, whereas orthography is a particular application of a writing system in order to represent a specific language, and which are either already used or proposed to be used for that purpose”. For the orthography of a language, a script must be chosen, devised, or developed (Sebba 2009). The term ‘script’ is used to refer to the graphic form of the units of a writing system such as Roman script and Cyrillic script (Coulmas 2003). The term ‘grapheme’ refers to the abstract type of a letter and its position in a given writing system, while ‘phoneme’ refers to the smallest segmental unit of sound used to form meaningful contrasts between utterances.

### **6.1.2 The sociolinguistics of writing systems**

Writing itself is best considered as a sociocultural phenomenon that reflects a language community's values. Social and cultural dynamics are involved at the development of a new writing system or the development of a new script. The three subsections that follow address, in turn, the sociolinguistic issue related to (1) the development of new writing systems, with particular reference to New Literacy Studies (NLS); (2) the choice between two scripts representing the same writing system; and (3) the sociolinguistic approaches to orthography.

#### **6.1.2.1 The development of new writing systems**

Barton (2000: 19) maintains that writing is “a culturally dependant variable” instead of statistic form of representation. This socially oriented view of writing is expressed by the work of Scribner and Cole (1981: 236), who maintain that:

“Instead of focusing exclusively on the technology of a writing system and its reputed consequences [...] we approach literacy as a set of socially organized practices which make use of a symbol system and a technology for producing and disseminating it. Literacy is not simply knowing how to read and write a particular script but applying this knowledge for specific purposes in specific contexts of use.”

Along the same line, Barton (2000) states that literacy practices are ways in which written languages reflect values, attitudes, and social relationships. He identifies six propositions about literacy:

- Literacy is best understood as a set of social practices; these can be inferred from events which are mediated by written texts.
- There are different literacy practices associated with different domains of life.
- Literacy practices are patterned by social institutions and power relationships, and some literacy practices become more dominant, visible, and influential than others.
- Literacy practices are purposeful and embedded in broader social goals and cultural practices.
- Literacy is historically situated
- Literacy practices change, and new ones are frequently acquired through processes of informal learning and sense making

(Barton 2000: 8)

This socially-oriented view of literacy is what NLS (Gee 1991, Street 1984, 1993) calls for. The assumption underlying the work of NLS begins with the notion of multiple literacies, which differentiates between an “autonomous” and “ideological” approach to literacy (Street 2003: 77). According to Street (1984), the “autonomous” approach to literacy focuses on the technical aspects of writing as well as its cognitive consequences (Olson 1988, Ong 1982). NLS represents the ideological approach to writing, which “does not attempt to deny technical skills or cognitive aspects of reading and writing, but rather understands them as

they are encapsulated within cultural wholes and within structures of power” (Street 1995: 161).

Sebba (2008), another exponent of the sociolinguistic approach to writing, views writing as involving conventional relationships at multiple levels. These conventional relationships must be initiated first among an entire language community and individual users. Later, a choice might be made by language users to follow or deviate from the established conventions of writing. According to Sebba (2008: 38), following the prescribed conventions of writing and deviating from it are both considered “literacy practices” with the former being unmarked. More pertinent to the present study is Sebba’s (2008) claim that opportunities for variation are multiplied by the existence of mixing languages and/or scripts.

The introduction of a writing system from another language, Sebba (2008: 41) argues, involves bilinguals who are familiar with the writing system of the other language. Those bilinguals act as intermediaries between the two languages and devise completely new writing systems. The choice of a script to represent a language is a powerful identity maker of association with particular religious, social, and cultural values.

#### **6.1.2.2 Digraphia**

It has been argued that the development of writing systems adds, in effect, a new variety to a language—a variety that is worth studying in itself and not only a reflection of spoken language (Dale 1980, Ferguson 1968). In this regard, Ferguson

noted that written varieties frequently develop their own characteristics that differentiate them from spoken varieties. The type and extent of these differences is often affected by the social structure of a language community. The case of written CA and the spoken non-standard Arabic dialects affords an example where differences between written and spoken languages are emphasized by the social structure exhibited by the Arabic community to the extent that they are almost regarded as separate languages. Ethnographic studies of writing have been addressing issues of understanding the social and cultural factors that influence the ways written codes are used . The current study is related to one topic within this area of investigation: digraphia. Below is a review of literature on the phenomenon of digraphia, including definitions, categories, and factors affecting the development of digraphia in a language community.

As sociolinguistic theory is mainly concerned with the interface of social and linguistic processes, an understanding of how social factors contribute to the choice of a script for a language is highly relevant to best understand a case of digraphia. It allows us to look at writing systems from a sociolinguistic point of view, where script choice is related to the social and cultural situations of a language community. The phenomenon of script choice is best understood in relation to religion, politics, and cultures. Digraphia is best defined as the use of two different scripts for the representation of the same language (Coulmas 1996: 129). The term is derived from the notion of diglossia, which is the use of two forms of a language in a community (Fishman 1967). Two categories of digraphia have been pointed out; “sequential digraphia” (when speakers of a language adopt different scripts over different times) and “genuine digraphia” (when speakers of a particular language use different scripts

contemporaneously) (DeFrancis 1984: 59). These two terms have been referred to as “diachronic digraphia” and “synchronic digraphia”, respectively, by Dale (1980: 5).

Dale (1980) provides perhaps the first general analysis of digraphia with a consideration of the cultural and political factors (e.g., religion, modernization) that underlie the development of digraphia in a language community. Dale claimed that the most important factor that affects the development of digraphia is the link between a given writing system and the religion of the speech community. Two decades after Dale’s claim, Grivelet (2001) made the following assumption in his relation of the choice of writing systems:

“The various questions related to writing, such as the choice of writing systems, the type or orthography, etc., are often understood as being obvious, based on two main assumptions: first, that the Latin script is the most suitable to form the base of a new writing system; and second, that a writing system should be phonemic. However, these answers are mainly based on linguistic observations, without much concern for the place and role of a writing system in society.”

(Grivelet 2001: 1)

Grivelet (2001), King (2001), Magner (2001) and other scholars synthesized different studies under the notion of digraphia providing an important contribution to the understanding of the role of writing systems from a sociolinguistic point of view. What all these studies have in common is that writing is carried out in a social context as the preference of a particular script is often related to considerations of ethnicity, religion, and politics. This clearly operates on both synchronic and diachronic cases of digraphia, as will be seen in the following paragraphs.

The Arabic script in Islamic countries is closely associated with the spread of Islam since the Arabic script is used for writing the Holy Qur'an. It acquired prestige and was felt to be the only appropriate vehicle for writing the language of a Muslim country. During the Medieval period, the Arabic script spread all over North Africa and the Middle East; Turkish is an example of a language that was written using the Arabic script. In 1920, due to the emergence of a new secular political system, Kamal Ataturk introduced a Romanized writing system as the official script to write Turkish (Grivelet 2001).

In "The poisonous potency of script: Hindi and Urdu", King (2001) emphasized that Hindi and Urdu are variants of the same language characterized by digraphia. Hindi is written using the Devanagari script (from left to right) whereas Urdu is written in a script derived from a Persian modification of Arabic script called Perso-Arabic script and is written from right to left. Such a case of digraphia, King indicated, is associated with the cultural and social structure of Indian society. Speaker of Hindi are generally Hindus, whereas speakers of Urdu are Muslims.

The graphemic difference between Hindi and Urdu is far more dramatic than the difference between the Cyrillic script of Serbian and the Roman script of Croatian (Magne 2001). Serbian and Croatian are two closely-related Southern Slavonic languages. The spread of the Roman script throughout Northern Europe in the middle ages is closely related to the spread of Roman Catholic Christianity during this period. The Russian Orthodox version of Christianity spread in Eastern Europe together with the Cyrillic script. In the SerboCroatian region, the digraphic situation was developed with the split between the Roman Catholic church and the Orthodox

Church. The two languages came to be known as Serbian and Croatian. Serbian, which has been written in the Cyrillic system since the tenth century, is spoken by those who converted to Orthodox whereas Croatian, which is written in Roman characters, is the language of those who converted to the Roman Catholic Church.

Another interesting situation of digraphia investigated by Grivelet is that of Mongolia where two writing systems existed diachronically for the Mongolian language; the Mongolian script which is the traditional script of Mongolia, which was replaced by the Cyrillic one, in the 1940s when Mongolia was a satellite country of the USSR. In the 1990s, a new democratic regime appeared, with emphasis on traditional values, including the Mongolian script. However, the Mongolian script did not succeed in replacing the Cyrillic script. Both scripts are used now in Mongolia and considered to be two different scripts and related to two different sets of history and values.

These different studies indicate that the choice of a particular script within a language community is affected by ethnic and religious ideologies. Thus, these studies draw attention to the ideological role of writing and writing systems in relation to the construction of social identities. With the growth of internet use and the spread of Roman characters of ASCII, digraphia is nowadays studied within virtual communities of the internet. Historically speaking, religion and politics are considered to be major factors in shaping digraphia within a language community. In the 21st century, communications of CMC should be viewed as another factor of the digraphic situations.

### **6.1.2.3 Sociolinguistic issues of orthography**

Taking the ideological perspective to literacy, not only the writing system but also orthography is a social practice. Since individuals deliberately choose particular scripts or deviate from standard spellings to mark their identities as well as cultural and religious values, orthography is also a social practice that is best analyzed in relation to the social and cultural considerations. The view of orthography as a social practice is proposed by Sebba (2007: 13), who argues that:

“...orthography is part of the ‘technology of a writing system’ but that writing system is itself a symbolic system embedded in a culture, shaping and yet also shaped by a set of cultural practices to which it gives, and by which it is given, meaning. While ‘orthography’ and ‘literacy’ are by no means synonymous, orthography is a fundamental element of written language; therefore, orthography too is situated in social practice. Orthography, too, needs and deserves a ‘practice account’.”

The term “orthographic practices” was devised by Sebba (2007: 26) to refer to practices that “represent reasonably coherent and consistent choices made by writers and printers in respect of how to write words”. He exemplifies this argument that orthographic practices reveal social meanings by the endings <-zing> and <-sing> in American and British orthography, respectively.

Non-standard spellings that deviate from the prescribed standard form also carry social meanings. In some cases, the use of non-standard spellings is used to manifest the phonological features of spoken language features (Berthele 2000). This is particularly clear in the case of minority languages where individuals use non-standard spellings to differentiate their spoken language from the majority

languages (Sebba 2000). Youth language is another use of non-standard spellings, marking youth identity by spelling rebellion (Androutsopoulos 2000). In investigating non-standard spellings in German fanzines, Androutsopoulos (2000) found that adolescents are free to use non-standard and creative spellings as the prescribed spelling norms are not imposed. The use of non-standard spellings is deliberately used by the young to manifest a non-mainstream identity. Another domain where the prescribed spelling norms are not imposed in writing is CMC. Below is a review of empirical CMC studies that address the orthographic representation of spoken languages.

### **6.1.3 Adaptation of writing systems in CMC**

In the early days of the internet, non-English speakers suffered from limitations of digital encoding possibilities that hinder both word processing and CMC communication. The text-transmission protocol was basically based on the (ASCII) character set. ASCII coding system is based on the Roman Alphabets and sounds of English. ASCII includes 128 seven-bit codes; 33 are non-printing control characters that affect how text and space is processed; 94 are printable characters, and the space is considered an invisible graphic.

With the massive increase of non-English internet users, Unicode protocol was developed with the explicit aim of transcending the limitations of traditional character encoding (ASCII). Accommodating 96,000 different characters, Unicode covers more than 90 scripts, including right-to-left ones. This expansion of digital encoding has led to the argument that the internet contributes to the maintenance of minority languages by providing practical opportunities for their literacy documentation (Warschauer *et al.* 2002). However, the impact of ASCII

protocol on the representation of non-Roman writing systems is still largely manifested in the Romanized transliteration of native scripts such as Greek, Arabic, and Persian online (e.g., Warschauer 2002 *et al.* , Palfreyman and al Khalil 2003, Tseliga 2004) despite the development of Unicode text-transmission protocol. These practices of transliterations are particularly clear when representing regional and non-standard varieties online, as the following case studies reveal.

In his investigation of code-switching to English online among Egyptian University Students, Warschauer and his colleagues (2002) found that a Romanized version of Egyptian Arabic is used extensively in informal e-mail and online IRC channels. Similar findings were reported by Palfreyman and al Khalil (2003) when investigating synchronous communication among female University students in Dubai. They investigated the ASCII-ized Arabic (AA), (i.e., a form of language in which ASCII characters set are used to represent Arabic in Instant Messaging (IM) and other electronic written communication). Their analysis of the corpus shows that “approximately” 25% of participants use mainly Arabic script in IM, 25% AA, and 50% English. A tendency was found among participants to employ Roman numerals to represent Arabic sounds that cannot be represented by the Roman characters (e.g., <7’> to represent /x/ and <2> to represent /‘/)

However, studies that investigated the representation of Greek online have occupied the centre position and yielded significant results in relation to the Romanized transliteration of native scripts. Analyzing the representation of Greek language through Roman characters, Androutsopoulos (2000, 2001) coined the term “Greekish”, referring to the transcription of Greek sounds in Roman characters. Among the most significant findings of his investigation is the fact that Greek internet users employ two kinds of representation: phonetic and orthographic. The

former refers to the use of Roman character to represent Greek sounds; the latter refers to the employment of signs and numbers to represent Greek characters.

Investigation into issues of how spoken languages are represented in the internet demonstrates that CMC may contribute to an increase in written language domains. Thus, minority languages and regional and non-standard dialects are afforded prestigious status by representing their compatibility with modern communication devices. Increasingly, limitations imposed by the ASCII text-transmission protocol resulting in “typographic imperialism” disappeared (Danet & Herring 2007). CMC profile, thus, accumulates orthographic documentation of a diversity of minority languages and non-standard varieties that were used mainly for oral communication.

Previous investigations on orthographic representation of non-standard dialects online are limited, with Greek studies representing the largest proportion (Androutsopoulos 2000, Androutsopoulos & Ziegler 2004). Only a handful of CMC studies have investigated the representation of Arabic non-standard dialects that lack official standardized writing systems (Warschauer *et al.* 2002, Palfreyman & al Khalil 2007). These few studies have been quantitatively oriented. That is, the corpus is analyzed mainly by counting instances of particular codes. Given the reliability of quantitative methodology that typically characterises a large proportion of empirical studies in this area, it takes more than code counts to understand how spoken language is shaped in CMC and how synchronicity fosters the emergence of new literacy practices online. Beyond the mere formal representation of NA online, the present study explores in an in-depth analysis the emergence of a new writing system online representing NA and how it constitutes a code of CMC. This new writing system carries social and pragmatic meanings and, thus, switching between

different writing systems will signify communicative functions similar to those achieved by language alternation in face to face communication. Investigating these communicative functions of script-switching constitutes an important aim of this study (Chapter Seven).

#### **6.1.4 Romanized Arabic Systems**

Romanized Arabic is the term alluding to the use of Roman characters as an alternative orthographic form of the Arabic language, which is normally written in the Arabic script. The phenomenon of Romanized Arabic dates back to the eighteenth century with early, albeit poorly documented, proposals calling for Romanizing Arabic for the sake making Arabic manuscripts readable for foreigners. They also facilitate teaching and learning Arabic as well as communicating with foreigners. (Yaghan 2008: 23). Different Romanized Arabic schemes (RASs) have been developed for the sake of rendering written and spoken Arabic into Roman alphabet. These schemes include United Nations Group of Experts on Geographical Names (UNGEGN), The International Standard Organization (ISO), International Phonetic Alphabet (IPA), Qalam (1985), Buckwalter transliteration and ALA-LC.<sup>13</sup> Buckwalter transliteration is the transliteration scheme developed and used by the *Encyclopaedia of Islam* which is regarded as the most authoritative conventions of Romanized Arabic.

---

<sup>13</sup> Details on transcription and transliteration of Arabic can be found in <http://open.xerox.com/Services/arabic-morphology/Pages/romanization>

In refereeing to these RASs, Beesley (1998) differentiates between transcription of sounds and transliteration of an orthography. The term transcription, according to Beesley, refers to the orthography devised to characterize the phonology or morpho-phonology of a spoken language whereas the term transliteration refers to the use of orthography of carefully substituted orthographic symbols in a one-to-one correspondence with the language's customary orthography. According to Beesley's account of orthographic transcription and transliteration of orthography, the conventions RASs are based on transcription of Arabic sounds rather than transliteration.

These different RASs differ in the way they address the inherent problems of rendering Arabic into Roman letters. For Arabic sounds that do not exist in English, digraphs or capitalization are used (For example, /x/ is represented by <kh> and /ʃ/ is represented by <S>. In other cases, Additional characters are used such as <ʃ> for the voiceless alveo-palatal /s'/ <ص>, and <ħ> for the voiceless pharyngeal fricative /h/ <ح>.

Because one-to-one correspondence is the case between Arabic and English and because Arabic is unvocalized, ambiguities and inconsistency often exist. Representing vowels reveals a considerable degree of variation. For example, <mina> and <meena> appear to represent the same word (Palfreyman and Alkhalil 2007). The Roman symbol <kh> has two different readings; it is either a representation of /x/, as in <Khaled> /xaled/ "a male name", or a cluster of two consonants /k/ and /h/.

With the emergence of the internet, which operates on the seven-bit ASCII character encoding sets that provide only Roman-script languages, Romanized Arabic was inevitably reinvented as the only option available to Arabic-speaking internet users. Since the 1990s, Unicode software development supporting non-

Roman scripts has been available. At that time Romanized Arabic was firmly established among early Arabic-speaking internet users. In this transitional period, both Arabic and Romanized scripts were available to internet users with the latter being the code of the internet. However, the form of Romanized Arabic in IRC is very different from its predecessors. Unlike RASs standards, Romanized Arabic in CMC is not aimed at teaching Arabic or making Arabic language readable for foreigners since it is used by speakers of Arabic to communicate with each other.

## **6.2 Methodology**

The aim of this chapter is to explore the case of Romanized Arabic in Najdi IRC. Therefore, it is necessary to critically review fundamental theoretical frameworks that relate to the explanation of these phenomena. Among these frameworks are the characteristics of Arabic orthography and transliteration and transcription approaches of RASs (see Section 6.1.4). Descriptive/structural analysis for the sake of identification of orthographic features of Romanized NA online was conducted. The IRC data was qualitatively analyzed with the aim of identifying the different representations of consonants and vowels of NA. The Romanized NA data in IRC was examined in terms of the above mentioned theoretical frameworks. The data was analyzed initially by counting instances of key symbols used to represent particular Arabic sounds. A qualitative analysis of the material was also conducted to investigate the linguistic resources that Najdi users draw on in formulating this form of writing. For the analysis of sources of variation, some numerical analysis was conducted.

### **6.3 Results: Romanized Arabic in Najdi IRC**

The discussion turns now to an investigation of the systematic patterns that IRC participants employ in representing this new developed form of Romanized Arabic. The investigation addresses the correspondence between Arabic sounds and Roman symbols in IRC interactions. Do they follow conventional schemes of sound transcription and orthographic transliteration adopted in the authoritative transliteration conventions of Romanized Arabic? What are the innovative ways IRC participants employ to represent their non-standard dialect?

#### **6.3.1 Romanized NA: consonants**

The investigation starts with an analysis of consonants that are represented by a single Roman character. Then, the researcher presents an in-depth analysis of consonants that are represented by more than one Roman representations with a discussion of frequencies of their occurrence in the IRC data. Special attention will be given to the representation of consonants that exist in NA but not in MSA such as /ts/ and /dz/ and the pharyngealized voiced interdental stop / d<sup>s</sup> /. Finally, an examination of the Roman representation of geminates and the definite article (-il) will be presented.

##### **6.3.1.1 One phoneme -to- one grapheme correspondence**

The fundamental principle in Romanized NA is based on the same principle of phonetic similarity in RASs. Arabic consonants are represented by Roman characters, based on the pronunciation of these Roman characters in English. Thus, the sound /b/ is represented by <b> and /k/ is represented by the letter <k>. Adopting this phonemic similarity seems acceptable and natural when Arabic consonants have

English counterparts. This is clearly a phonemic one-to-one mapping between phonemes and graphemes (Roman characters). However, the Arabic language has ten more consonants than English. In RASs, these sounds are capitalized, such as <H> for the voiceless pharyngeal fricative /ħ/ or are written as digraphs such as <gh> for the voiced velar fricative /ɣ/. In Najdi IRC, these sounds are represented by numerals that have visual similarities to the shape of Arabic alphabets that represent these sounds. This strategy is what Androutsopoulos (2006) refers to as orthographic transliteration (that is, the representation of Greek letters by visually similar Roman characters). The similarity between the shapes of these Arabic characters and numerals are strong to the extent that some may be viewed as mirror-image reversals of the Arabic characters or, in some cases, part of them. Below is a description of each Arabic consonant that is represented by a single Roman character.

Table 6.1 Representation of Arabic sounds that have English equivalents by a single Roman character

<b>Description</b>	<b>IPA</b>	<b>Arabic letter</b>	<b>IRC Phonemic representation</b>
voiced bilabial stop	b	<ب>	b
voiceless dental stop	t	<ت>	t
voiced dental stop	d	<د>	d
voiced palate-alveolar affricate	dʒ	<ج>	j
voiceless velar stop	k	<ك>	k
voiceless labiodentals fricative	f	<ف>	f
voiced labiodentals fricative	v	-	v
voiced uvular stop	q	<ق>	q

voiceless dental fricative	θ	<ث>	th
voiceless alveolar fricative	s	<س>	s
voiced alveolar fricative	z	<ز>	z
voiceless palatal fricative	ʃ	<ش>	sh
voiceless glottal fricative	h	<ه>	h
Voiced bilabial nasal	m	<م>	m
Voiced alveolar nasal	n	<ن>	n
Voiced dental lateral	l	<ل>	l
Voiced pharyngeal alveolar lateral	ℓ	<لّ>	l
Voiced dental-alveolar trill	r	<ر>	r
Voiced labio approximate	w	<و>	w
Voiced palatal approximate	y	<ي>	y
Voiceless dental affricate	-	-	tz
Voiced dental affricate	-	-	dz

Table 6.1 reveals that phonetic transcription is the norm with most of the Arabic consonants in Romanized NA. Most of the Arabic consonants (22 out of 29) are transcribed phonologically. The representation of these consonants is based on the phonetic similarity between the two sounds (Arabic and English consonants) rather than the visual resemblance between the shapes of the Arabic and English letters. The analysis indicates that IRC users employ the strategy of visual similarities only when an Arabic letter does not have a corresponding sound in English. For example, the dotted letter <ز> /z/ is represented by the corresponding English letter <z> in Romanized NA but never by <j>, which looks very similar to the Arabic letter <ز> . This reflects the background of IRC users as bilinguals who master the writing system of their second language.

Consonants /p/ and /v/ are borrowed from English and only used in loan words and foreign names (e.g. *pakistan* ‘Pakistan’, *kapris* ‘caprice’, *vayrus* ‘virus’, *nwvamber* ‘November’ ). Such words are pronounced as in the original language/p/ and /v/ and they are usually transcribed respectively as <p> and <v> in Najdi IRC. Arabic consonants that do not have English equivalents and are represented via orthographic representations are illustrated in Table 6.2.

Table 6.2 Representation of Arabic sounds that do not have English equivalents by a single Roman character

<b>Description</b>	<b>IPA</b>	<b>Arabic letter</b>	<b>IRC Orthographic representation</b>
voiceless pharyngealised dental stop	t <sup>ɕ</sup>	<ط>	6
voiced inter-dental stop	ð <sup>ɕ</sup>	<ظ>	6'
voiceless alveolar fricative	s <sup>ɕ</sup>	<ص>	9
Pharyngealised voiced inter-dental stop	d <sup>ɕ</sup>	<ض>	9'
voiced pharyngeal fricative	ʕ	<ع>	3
voiceless glottal stop	ʔ	<ء>	2
voiceless pharyngeal fricative	ħ	<ح>	7

Because in the Arabic alphabets several sounds are represented by the same symbol but with an additional dot on top, it was only natural that IRC chatters would represent them with the same number, only with an apostrophe or asterisk on top, as illustrated in Table 6.2. Examples found in IRC data that contain the numerals representing the above mentioned sounds are:

- (1) el,m6r 'DEF-rain'
- (2) 6'rf 'envelop'
- (3) 9rtee 'became.2SG.F'
- (4) wa9'7 'clear'
- (5) 3adi 'normal'
- (6) sw2al 'question'
- (7) 7jm 'size'

### 6.3.1.1.1 Digraphs

Digraphs regularly exist in the corpus. The digraphs that are represented phonologically by a single Roman correspondence are /θ/, represented as <th>, and /ʃ/, represented as <sh>. This clearly reflects the backgrounds of Najdi IRC participants as native speakers who have mastered the writing system of English. Examples of occurrences of these digraphs are given below:

- (8) ktheer 'a lot'
- (9) shhr 'month'

### 6.3.1.1.2 The pharyngealized voiced interdental stop /dʕ/

The merger of the pharyngealized voiced interdental stop /dʕ/ and the pharyngealized voiced dental stop /ðʕ/ to a pharyngealized voiced interdental stop /d/ is a feature shared by many Arabic non-standard dialects including NA. Looking at the Najdi IRC corpus as a whole, the sound /ðʕ/ <ظ> is regularly represented as <6'> whereas /dʕ/ <ض> is represented by the numeral <9'> although they are pronounced the same in NA (see table 6.2). In such cases, IRC participants are keen to maintain the difference between the two sounds reflected by Arabic orthography.

This is a transliteration of Arabic orthography rather than a transcription of spoken NA.

(10) el, 6'ann 'DEF-guessing'

(11) 9'reeba 'Tax'

#### 6.3.1.1.3 The dental-alveolar affricates /ts/ and /dz/

The variety of Arabic used in the present study is NA, and features of spoken NA affect orthographic symbols used in IRC conversations. One feature that distinguishes NA from other dialects in the Arabian Peninsula is the introduction of two new units [ts] and [dz], which result from the affrication of velar stops /k/ and /g/. These sounds occur only in final positions of words and are used by both men and women in addressing females. Such sounds do not occur in MSA and there is no standard Arabic letter to represent them nor IPA representations. When Najdi IRC participants want to represent these sounds they tend to transcribe them phonologically as <ts> and <dz> respectively.

(12) 3enits 'have-2SG.F'

(13) 9dz ? 'really'

#### 6.3.1.1.4 Geminates

These consonants are pronounced exactly like short consonants, but last longer. This difference in pronunciation is manifested in MSA writing by the use of a diacritic mark called shadda <◌◌◌>. It is written above the consonant which is to be doubled. This diacritic marker is commonly used in spelling to avoid ambiguity. (Example: <رر> /r/; مدرسة *madrasa* 'school' and مدرّسة *mudarrisa* 'teacher-F'). In IRC, due to the lack of diacritics in ASCII, geminates are represented by doubling

the letters that represent these sounds. In such cases, Najdi IRC participants are transcribing rather than transliterating their speech.

(14) **ummek** ‘mother- POSS.2SG.F’

(15) **elli** ‘which-REL’

(16) **enna** ‘that’

### 6.3.1.2 One phoneme -to- two or three graphemes correspondence

The analysis of IRC interactions reveals that four Arabic consonants are represented by more than one graphemes (including both phonemic and orthographic representations). A description of these consonants and their different Roman representation is given in Table 6.3 with reference to frequency of their occurrence in the IRC data in a sample of 8000 words of Romanized NA. Some of these representations are idiosyncratic (Androutsopoulos 1998) as they are used in few instances whereas other representations are frequently and regularly used by IRC participants. An investigation of sources of these variations will be also given.

Table 6.3 Representation of Arabic sounds by more than one Roman representation

Description	IPA	Arabic letter	IRC Orthographic representation			total
			Phonemic representation	Orthographic representation		
<b>Voiced uvular fricative</b>	ɣ	<غ>	<gh> N=8 %=11.6	<3’> N=61 %=88.4		69
<b>Voiced velar stop</b>	g	-	<g> N= 149 %=47	<8> N=170 %=53		319
<b>Voiced dental fricative</b>	ð	<ذ>	<th> N= 21 %=2	<4> N= 84 %=8		105
<b>Voiceless uvular fricative</b>	x	<خ>	<kh> N=15 %=.04	<5> N= 338 %= 94	<7’> N=6 %=.02	359

### 6.3.1.2 .1 The voiced uvular fricative /ɣ/

The voiced uvular fricative /ɣ/ is transcribed both phonetically <gh> (based on RASs and some relation with the sound /g/) and orthographically <3'> (based on the visual correspondence between the Arabic letter <ع > and a digraph composed of a numeral followed by an apostrophe <3'>). As illustrated in Table 6.3 which shows the number of uses per each representation, IRC participants have a great tendency to represent this sound orthographically rather than phonetically. The few occurrences of <gh> in IRC data are mainly names of persons or places showing the influence of RASs conventions. It is plausible that IRC participants have encountered these names in the Romanized form in academic context (such as books, newspapers.etc). Examples of both ways of representations of this sound are given below:

(17) **gh**ada 'Ghada. a female name'

(18) **gh**aleb 'Ghaleb. A male name'

(19) **3'**reeba 'strange'

### 6.3.1.2 .2 The voiced velar stop /g/

A major difference between MSA and NA is the realization of the sound /q/ <ق >. Like most dialects in the Arabian Peninsula, it is pronounced as /g/ in NA. This sound does not have a standard letter in MSA. Interestingly, the English orthography provides IRC participants with a ready-made letter <g>, which represents the same sound. Examples found in IRC data that contain this phonemic representation are:

(20) A: ma gelt shai? 'NEG said-1SG thing'

B: gooli li 'tell -2SG.F me'

On the basis of the visual similarity that underlies the use of numerals in IRC conversations, IRC participants have adopted the numeral <8> to represent the /g/ sound as it is visually similar to the letter <g>. As indicated in Table 6.3, the two different representations are used in the IRC in about the same amount.

The distinction between the two representations reflects a metalinguistic realization of a principle when writing sounds that do not have English counterparts. Namely, sounds that are untypical for English are represented by numerals (See Table 6.2). Since /g/ is, orthographically speaking, associated with the voiceless uvular /q/ <ق>, it is perceived as non-English sound that requires a numeral representative even though English does have /g/ <g>. The following examples provide illuminating examples in this regard.

(21) a8oolek 'ISG-tell- 2SGF'

(22) 8bl 'before'

(23) ma 8ed 'never'

There are few occurrences of the sound /q/ in the data . It appears only in classical or academic words such are *Qur'an*, *qan'a* 'channel' and *qalam* 'pen'. This reflects the IRC participants' awareness of the differences between the differences between NA and MSA and their tendency to maintain these differences in their transcription online.

(24) quran 'Quran'

(25) ttwaqa3een '2SG.F-think'

(26) q1m 'pen'

### 6.3.1.2 .3 The voiced dental fricative /ð/

The voiced dental fricative /ð/ has two representations; one being phonemic <th> and the other is neither phonemic nor orthographic <4>. Unlike the case of the voiced velar stop /g/, there is a general preference to represent this sound as <4> rather than <th>. There are only twenty one occurrences of <th> in the IRC data, but eighty one of <4> which manifests this trend. The reason of using <4> to represent this particular sound is not based on the principle of visual similarity but rather on the IRC participant's awareness about the difference between the English orthography and the Arabic orthography. The English digraph <th> is used to represent both /ð/ and /θ/. The distinction between the two pronunciations depends on familiarity with the context. In Romanized Arabic, the cluster <th> has two pronunciations; either as /ð/ as in example (27), or as a consequence of two consonants /t/ and /h/ as in examples (28) and (29). The ambiguity of reading such examples results in a general preference of using <4> to represent /ð/ as in example (30)

- (27) hatha 'this'
- (28) 5alatha 'aunt-POSS-3SG.F'
- (29) shan6atha 'bag-POSS.3SG.F'
- (30) ki4a 'this way'

### 6.3.1.2 .4 The voiceless uvular fricative /x/

The case of the voiceless uvular fricative /x/ is rather interesting, not only because it does not have an English equivalent that can be used for its phonemic transcription but also because it is represented in three different ways. Table 6.3 reveals the great tendency to represent this sound as <5>. This sound is represented as <5> 338 times

in the IRC sample while there are only 6 occurrences of <7'> and 15 occurrences of <kh>.

The use of the digraph <7'> to represent this sound is due to the fact that the Arabic letter representing this sound <ح> is very similar to the voiceless pharyngeal fricative /h/ <ح>. It was only natural that IRC participants would represent /x/ with the same number followed by an apostrophe. On the other hand, the use of <kh> to represent this sound shows an influence of RASs phonemic conventions which has some relation with the sound /k/. The occurrences of <kh> in the IRC data are mainly in names of persons or cities such as those in example (32) and example (33). Influence of official RASs spelling explains the use of <kh> in these cases as IRC participants have been exposed to these Romanized names in academic books. It is worth mentioning that the cluster <kh> has two possible pronunciations in Romanized Arabic; either as /x/ as in example (34), or as a consequence of two consonants /k/ and /h/ as example (35) which results in great ambiguity in terms of readability. This results in a general tendency to prefer the use of a distinct symbol <5> to represent /x/ as illustrated in examples (36) and (37). The use of <5> to represent /x/ is not arbitrary because it is derived from the Arabic word /xamsah/, 'five', which begins with this sound. Below are illuminating examples that contain the use of these different representations.

- (31) 7'ala 'aunti'
- (32) **Kh**aled 'a male name'
- (33) she**ikh**a 'a female name'
- (34) **kh**atm 'a ring'
- (35) mark**kh** 'a trade mark'
- (36) b**5**eer 'fine'
- (37) **5**la9 'finished'

The analysis of consonants in Romanized NA is, in many ways, divergent from conventional transliteration schemes which are based on a phonetic similarity between Arabic sounds and the English pronunciation of Roman characters. The fundamental logic in Romanized NA is based on the same principle of phonetic similarity. The divergent norms in Romanized NA become obvious with particular sounds that do not have corresponding sounds in English. In such cases, IRC participants do not use ready-made representations of RASs but rather tend to use distinct symbols based on the visual similarity between the shapes of Arabic letters representing these sounds and Roman characters. In other cases, they tend to rely on initials of a familiar spoken Arabic term such as <5>. The use of these symbols spread quickly among IRC participants to be part of the chatting process and an indication of belonging to cyberspace community. IRC participants share common knowledge about these representations which prevents ambiguity and provides a great amount of standardization.

### **6.3.1.3 Two phonemes -to- one grapheme correspondence**

The definite article (*'adātu t-ta 'rīf*) / əl-/ expresses the definite state of a noun of any gender and number (see section 4.5.1.4). It is assimilated before some consonants (/t/, /t̤/, /d/, /d̤/, /r/, /z/, /s/, /š/, /s̤/, /d̥/, /t̥/, /z̥/, /l/, /n/). Before the remaining 14 consonants that are termed 'lunar letters', the / əl-/ is pronounced as two phonemes /a/ and /l/. The Arabic orthography represents both different pronunciations consistently as <ﻻ >. This rendition of the definite article in orthography reflects the underlying grammatical characteristics of the word rather than its surface pronunciation.

The representation of the definite article / əl-/ in IRC data follows the Arabic orthography as it is represented consistently as <el,>, <el'>, <al-> and <el->. In this case, IRC participants are transliterating the Arabic orthography rather than transcribing their speech. The use of comma or hyphen after <el> and <al> indicates IRC participants' tendency to maintain the clarity of the underlying grammatical structure of the definite article and the words to which it is prefixed. Examples are given below:

(38) **el,jehaz** 'DEF-instrument'

(39) **el-m6ar** 'DEF-airport'

(40) **al,f6or** 'DEF-breakfast'

#### 6.3.1.4 No phoneme -to- one grapheme correspondence

Another example which shows effect of the Arabic orthography on Romanized NA is the maintenance of “*taa' marbūṭa*” in writing. The “*taa' marbūṭa*” is used to indicate feminine nouns and are attached to the feminine gender suffixes –a (always follows a fatha, short vowel /a/) ( for example *modaresa*, ‘teacher.F’ <مدرسة>, *madina*, ‘city’ <مدينة>). As these examples show, the “*taa' marbūṭa*” is not pronounced but rendered into writing. Following the same rule, this “*taa' marbūṭa*” is always rendered into Romanized NA; an evidence that suggests IRC participants' metalinguistic awareness in visualizing the Arabic orthographic writing system while transcribing feminine nouns in Romanized NA. Examples that show feminine nouns attached to the “*taa' marbūṭa*” are below.

(41) **madresah** 'school'

(42) **el mushkelah** 'DEF-problem'

### 6.3.2 Romanized NA: Vowels

With the exception of few consonants (/y/, /x/, /ð/, and /g/), representation of consonants in Romanized NA features consistent phonetic and orthographic transcription. However, vowels in Romanized NA present an entirely different case that exhibits a great deal of variation and makes use of a variety of Roman symbols. Whilst the shortage of Roman alphabet consonants resulted in a constraint that must be overcome by numerals, Roman vowel alphabets provide more options of letters and digraphs than the Arabic orthography does (Palfreyman and AlKhalil 2007: 57). The representation of vowels in the Arabic orthography deserves some comment as it affects the ways vowels are represented in Romanized NA. This section begins with a discussion of the representation of vowels in the Arabic orthography. Then a summary of the most observable patterns of vowels representation in Romanized NA will be given with a discussion of justifications of these patterns.

Vowel inventory in NA displays more vowels than MSA does (see section 4.4.1). In MSA, there are three short vowels, three long vowels and two diphthongs. In NA, there are five short vowels, five long vowels and two diphthongs. In the Arabic orthography, long vowels /ā/, /ī/ and /ū/ are transcribed as characters and written respectively as <ا>, <ي> and <و>. Short vowels are encoded by diacritical marks. Four main diacritics are used to indicate short vowels; <ـ> *fatha* for /a/, <ـَ> *kasra* for /i/, <ـُ> *damma* for /u/ and <ـْ > *sukūn* for no vowel. The representation of short and long vowels in Arabic orthography is illustrated below:

بُ	بُو	بِي	بَا	بُ	بِ	بَا
b	bū	bī	bā	bu	bi	ba

Figure 6.1 : Representation of short and long vowels in Arabic

Although the use of diacritical marks is taught at school as part of the educational curriculum, they only appear in formal and Qura'nic writing and are not used in the vast majority of writing to which Arabs are exposed to (for example, books, magazines, etc). The Arabic three-letter word <كتب> might be either 'books' pronounced as [kutub] or 'he wrote' pronounced as [katab]. The inference of these short vowels depends both on prior familiarity with the word and the context where they occur. The analysis now turns to the representation of these predictable vowels in Romanized NA followed by an analysis of long vowels and diphthongs.

### 6.3.2.1 Short vowels

The lack of orthographic representation of short vowels in the Arabic orthography makes them less likely to be represented in Romanized NA. This is clearly an effect of sound-to-letter transliteration of the Arabic orthography. As short vowels are typically omitted in writing, they are left out in Romanized NA following the convention of the Arabic orthography. This is especially the case with familiar words that are commonly used among IRC participants. Below are examples of short vowel omissions. The deleted vowels are illustrated in the transliterated gloss.

- (43) bs ‘bas’ ‘but’  
 (44) 93b ‘*šaḩab*’ ‘hard’  
 (45) tmr ‘*tamər*’ ‘dates’

In other cases, short vowels are rendered in Romanized NA. Three short vowels are represented by a single Roman symbol (/u/, /a/ and /e/). The other two short vowels, /i/ and /o/ are denoted by different Roman symbols. Table 6.4 illustrates the representation of short vowels in the corpus.

Table 6.4 Representation of short vowels in Najdi IRC

	vowels	IRC representation	Example
<b>Short vowels</b>	/i/	<e>	<e> (46) l7alek ‘alone’
		<i>	<i> (47) a8oolik ‘ISG-tell-3SG.F’
		<y>	<y> (48) enty ‘you.2SG.F’
			(49) e7ky ‘tell.2SG.F’
	/u/	<u>	<u> (50) shukran ‘thanks’
/a/	<a>	<a> (51) hala ‘hello’	
		(52) l7alek ‘alone’	
/o/	<o>	<o> (53) 3laikom elsalam ‘upon-2PL DEF-peace’	
		<u> (54) yu ‘oh’	
/e/	<e>	<e> (55) esem ‘name’	

### 6.3.2.2 Long vowels

The analysis of IRC data reveals that the most noticeable pattern in representing long vowels is the use of a single Roman symbol <e> ,<o> <a> and <u> . In some

cases, IRC participants tend to duplicate the Roman symbol to convey the length and stress of the vowel. Table 6-5 below illustrates with examples how long vowels are represented.

Table 6.5 Representation of long vowels in Najdi IRC

	vowels	IRC representation	Example
Long vowels	/ī/	<ee>	<ee> (56) kebeer ‘big’
		<e>	(57) re5ee9 ‘cheap’
		<e>	(58) sahranen ‘awake-2PL’
	/ū/	<oo>	<oo> (59) adoor ‘ISG-wander’
		<o>	<o> (60) 3’9on ‘Goson ‘a name of a girl’
		<ou>	<ou> (61) youm ‘day’
/ā/	<a>	<a> (62) kaif al,7al ‘how DEF-state’	
	<aa>	(63) anam ‘1SG-sleep’	
	<aa>	(64) salaam ‘hello’	
	<aa>	(65) sahraana ‘awake-2PL’	
/ō/	<o>	<o> (66) fo8 ‘above’	
	<u>	<u> (67) 7ul ‘almost’	
/ē/	<ee>	<e> (68) kbeer ‘big’	
	<ee>	(69) ee ‘yes’	

It is likely that the lack of orthographic representation of short vowels of NA in the Arabic orthography makes vowels less likely to be represented in Romanized Arabic. This results in leaving them out even when pronounced and adhere to the conventions of the Arabic orthography. Though vowel representation in Romanized

NA is unstable, certain trends are found to influence the rendering of vowels. These trends are discussed below.

### 6.3.2.3 Trends of vowel representation in Romanized NA

An examination of Romanized NA reveals certain trends that suggest the influence of RASs, English orthography as well as a vowel's position within a word and a vowel's morphological significance.

- **The influence of Romanized Arabic schemes:** The effect of RASs conventions appears clearly in names of persons and places. The only case where /ay/ appears in IRC data is the name of a city 'Riyadh'. The following examples are exact Romanized names that IRC participants have encountered in road signs and books. The conventional Romanized name will be given in the gloss.

(70) muhammad 'Muhammad'

(71) riyadh 'Riyadh'

(71) dammam 'Dammam'

- **The influence of the English orthography:** While the sound-to-letter transliteration of the Arabic orthography has an influence on vowel omission in Romanized NA, the English orthography appears to affect the maintaining of some vowels. Vowels are typically omitted in final position of a word except for the short vowel /i/ and its long form /ī/ as they consistently appear in the final positions.

(72) enty 'you-2SG.F'

(73) esm3ee 'listen--2SG.F'

The use of <oo> to represent the long vowel /ū/ also shows an influence of the

English orthography. Examples are:

(74) 7boob ‘pills’

(75) adoor ‘ISG-wander’

- **The influence of vowels’, position, length and morphological**

**significance:** The analysis reveals that a vowel’s morphological significance within a word influences the decision whether this vowel is omitted or retained. For example, short vowels tend to be retained in Romanized NA before the “*taa’ marbūta*” that is used to indicate feminine nouns (see Example (41) and (42) ). Despite the fact that this short vowel /a/ is short, unstressed and unrepresented in Arabic script. The reason why this short vowel is retained in Romanized NA is that it signifies the grammaticality of the feminine status of the noun.

Many Arabic suffixes end in long vowels and because of their orthographic representation, length, and the morphological significance they signify, these vowels are always retained in Romanized NA. The first personal possessive suffix [ī] is typically retained for these reasons.

(76) 5altee ‘aunt-POSS.ISG’

(77) examee ‘exam-POSS.ISG’

(78) 7abibte ‘darling-POSS.ISG’

#### **6.4 Conclusion**

Like all Arabic dialects, NA does not have a standard official orthography, and it is seldom put into writing. The analysis of IRC data reveals two crucial sociolinguistic occurrences: the representation of a spoken vernacular into a written form and the acceptability of colloquial usage in CMC. These findings confirm the claim made by Warschauer *et al.* (2002) about the development of small minority languages in CMC. According to Warschauer *et al.*, the unregulated CMC world

promotes the growth of Arabic dialects online, which will contribute to the linguistic balance of the Arab World. This chapter sets out to offer a comprehensive account of Romanized Arabic online using NA as a case study.

Regarding the characteristics of Romanized NA, the analysis yields significant interrelations among languages, scripts, and technology. The process of Romanizing NA has shown itself to be particularly consistent in the case of consonants. It involves a combination of both transliteration of Arabic orthography (e.g., the definite article (e.g., / əl-/ / d/) and transcription of spoken sounds (e.g., geminates). Romanized NA is mainly phonemic: either by adopting an equivalent English sound or following phonemic conventions of RASs. In cases where consonants do not exist in English, IRC participants tend not to use the ready-made forms of RASs. Instead, creative orthographic transcription is adopted, using ASCII symbols that either correspond visually to Arabic letters (e.g. <6>, <6'>, <3>) or derive from an initial of a familiar Arabic word (e.g., <5>). Another source of variation in Romanizing consonants is due to participants' awareness of the differences between the Arabic orthography and the English orthography. For example, the realization of the fricative /ð/ in English depends on prior knowledge. Moreover, the potential pronunciation of NA Romanized Arabic <th> as either /t+ /h/ or an interdental fricative /ð/ resulted in a distinction made between /ð/ and /θ/ in the process of Romanizing. Such distinction is significant to distinguish the sound /ð/ from a consequence of the two consonants /t/ and /h/, and, therefore, <4> is deliberately used to substitute /ð/.

Whilst the Arabic language has a number of consonants that do not exist in English and, consequently, do not have corresponding Roman characters, English provides a larger number of letters and digraphs representing vowels than the Arabic

orthography does. Because of the lacking of a standardized system of Romanized vowels (Aboelezz 2010), encoding vowels in Romanized NA shows a great deal of inconsistency. Short vowels, which are predictable in the Arabic orthography, are sometimes left out, following the convention of the Arabic orthography, and other times are spelled out in different ways. Long vowels are maintained in the Romanizing process by the use of single or multiple Roman symbols. Effects of RASs appear clearly in names, and certain trends suggest the background of IRC participants as bilingual who master the writing system of their second language. Vowels' positions within words as well as the morphological significance they contribute have influences on the rendering of these vowels in Romanized NA.

These systematic patterns of Romanized Arabic fit the four main dimensions proposed by Coulmas (2003) to identify standard writing systems: visual perception, auto-indexicality, conventionality, and autonomy (See Section 6.1.1). First, it consists of signs that have an assigned external referent. Second, writing graphic signs of Romanized NA embodies the principles of their learnability (DeFrancis 1989; cited in Coulmas 2003). They came to be acknowledged as part of the chatting process and spread quickly among internet users. Third, the relationship between graphic signs and their meaning in Romanized NA is conventional and can be deduced from the graphic composites and their arrangement. Last, Romanized NA is autonomous as it consists of units that, like sound systems, are structured and can be analyzed independently of other levels of linguistic structure its units may refer to.

To conclude, the present case of investigation is a contribution to the area of sociolinguistics of Arabic writing in CMC—an issue that has not been researched before. An understanding of the representation of NA online indicates that the orthographic practices in CMC are shifting along technological development. These

practices should be analysed in the light of evolution and social spread of technology within a language community as well as the linguistic situation within a language community. The above-mentioned results indicate that a clear incursion is taking place from a Romanized form of NA mixed with English into spaces that were previously filled by MSA. Romanized Arabic in CMC has changed from being a response to technical limitation into a code of the medium in which it is used. The analysis of the characteristics of Romanized NA illustrates some interrelations between characteristics of NA, conventions of Romanized Arabic, English language, and the Arabic orthography. The Romanized NA involves transcription of a non-standard dialect and mediation from properties of both Arabic and English writing systems.

CMC has provided NA with the forum it needed to become a viable choice for young bilingual Najdi speakers during text-based CMC. The way this form of communication developed and is used among IRC participants expresses the need for communicating in text-based CMC without the fear or thought of what Tagliamonte and Denis (2008: 12) call “language police”. The diglossic situation in the Arab world has made the need for this “language police” represented by educational or cultural institutions more serious. The role of these institutions is to verify that written Arabic does not reflect any transfer of the spoken dialect or the second language (English). The Romanized form of NA mixed with English allows young bilingual Najdi internet users to communicate in text-based CMC using their mother tongue without worrying about “language police”.

The systematic patterns of Romanized Arabic came to be acknowledged as part of the chatting process and a standardized norm of CMC. It forms the symbolic values of the Arabic code of internet communication and, thus, whoever masters this code is

shown to belong to the cyberspace community. The Romanized Arabic online with its letter-to-number correspondences seems to be used as an indication of a cyberspace culture even in offline settings in different domains. The use of this newly-invented form of Romanized Arabic persists among all Arabs and can be seen in domain names such as [www.otlob.com](http://www.otlob.com) and [www.bahth.com](http://www.bahth.com). More related to the present study is the fact that some domain names, including some numerals, exist: for example, [www.6arab.com](http://www.6arab.com), [www.ba7th.com](http://www.ba7th.com) and [www.a7la.com](http://www.a7la.com). The content of these websites is entirely Arabic. Clearly, romanizing these domain names has a purpose that extends beyond technological limitations. These website names are designed to attract the eyes of someone familiar with this form of Arabic. This implies that Romanized Arabic nowadays serves a function that extends beyond facilitating access for Arabic users to appeal to a wider audience of internet users.

The way Romanized NA is used, developed, and regarded in CMC supports Sebba's claim that:

“Orthography can be seen as the site of potentially intense struggles over identity and power, in which issues like the purpose of literacy and the status of languages are central, and orthographic characters may be imbued with a symbolic meaning that makes their phonemic symbolism and learnability of secondary importance”

(Sebba 1998:20)

## CHAPTER SEVEN

### CODE-SWITCHING AND SCRIPT-SWITCHING IN NAJDI IRC

Investigations into the code-switching phenomenon are based on the assumption that code-switching is not a random human activity, but rather a significant discursive and linguistic device that follows regular patterns and serves communicative functions. Over the last four decades, code-switching has changed from being a phenomenon that had been traditionally regarded as language corruption to an investigation field in its own right (Matras 2009). Many linguistic and sociolinguistic approaches have been proposed to understand both the grammar and function of code-switching (e.g., Gumperz and Poplack 1980, Myers-Scotton 1993, Auer 1995). Whether complementary or contradictory, these approaches have focused mainly on oral production of code-switching. The assumption underlying the focus on oral code-switching is that code-switching has been generally regarded as characteristic of informal speech and, therefore, written data is not a suitable site for studying code-switching. Another reason is that written language is constrained by the ideology of “correct” orthography, which strongly inhibits language alternation (Milroy & Milroy 1991: 66).

Unregulated and informal written language of CMC, where the standardization of language is less powerful, has defied the “tyranny of written monolingualism” (Sebba 2002: 66). The tendency for CMC written language to converge towards spoken form is well documented (Danet 2001, Herring 2006, Segerstad 2002, Alothman 2010). Another important feature of CMC is that it is multilingual. That is,

multilingual writers belonging to virtual networks communicate among themselves in (one or) more than one languages. As in face-to-face communication, they are tempted to switch between their overall repertoire of linguistic codes according to the needs of communication and the situational and contextual constraints. What is produced is not only multilingual but also multimodal. Written code-switching in CMC is not only uttered and heard, but, rather, written and read.

Synchronous IRC communication (the context of the present study) is conducted in ways which are best viewed as hybrid between written and spoken language (Werry 1996, Danet 2001). The reduced transparency of IRC messages and the lack of visual channels both heighten the IRC participants' meta-linguistic awareness and lead them to develop creative procedures to cope with these limitations. Consequently, they express their feelings and thoughts by written laughters, or verbal descriptions of actions and emotions. In this regard, Georgakopoulou (1997:158) claims that the lack of common contextualization cues, due to the absence of the visual channel, "results in an increased reliance on code-centred contextualization cuing, which would be otherwise delegated to different signals". Georgakopoulou's claim is especially relevant to the issue of code-switching online. IRC participants manipulate code-switching and script-switching, with other innovative written signs, in order to serve pragmatic functions, which, in ordinary spoken conversation, would be accomplished by other cues such as gazes, gestures, prosody, and phonological variations.

Before investigating the communicative functions of code-switching and script-switching in IRC data, a distinction has to be made between single-item code-

switching and lexical borrowing. Single-item code-switching “should show little or no integration into another language” (Poplack & Meechan 1998:129). Borrowing, by contrast, is the “adaptation of lexical material to the patterns of the recipient language” (Poplack & Meechan 1995: 200). This adaptation takes the form of morphological and phonological integration with different degrees into the speech of both bilingual and monolingual individuals, sometimes without awareness of their origin. These borrowed items, by definition, do not serve communicative functions. Since standardization of language in CMC is less powerful, CMC virtual communities may have different norms and perceptions of what is borrowed and should be nativised from the off-line community norms. Applying traditional criteria of code-switching research, the researcher will distinguish between borrowing and code-switching in both Arabic and Roman script in Najdi IRC.

In this chapter, the aim is to contribute new insights into grammatical and functional approaches to code-switching in synchronous CMC and to account for the phenomenon of script-switching in writing NA online. This follows a qualitative approach to analyze code-switching in the Najdi IRC corpus. As regards the analysis framework, the analysis begins by a distinction between borrowed items and code-switching (e.g., Poplack 1988, Muysken 2000, Myers-Scotton 1993) to exclude borrowed items from the contextual functionality of code-switching in IRC. Then, a heuristic approach is used in terms of Gumperz's (1982) notion of code-switching as a contextualization cue, with Auer's (1995, 1998, 1999, and 2000) theoretical framework of conversational analysis of code-switching and script-switching. The researcher argues that that both code-switching and script-switching are functional

and meaningful in IRC conversations. In line with Auer's theoretical framework, the researcher also argues that only by sequential analysis of IRC interactions can the issue of 'why' speakers switch codes in the same episode of conversation be tackled. It is argued here that Auer's Conversational Approach to code-switching can contribute valuably to the understanding of synchronous IRC since it is basically human interaction, which is handled very much like face-to-face interaction. Therefore, in Najdi IRC, interlocutors in IRC can perceive code-switching and script-switching as functional and meaningful.

Applying the theoretical apparatus of the Conversational Approach framework, the researcher establishes a productive methodological link between face-to-face interaction and synchronous CMC. Also, the application of the Conversational Analysis approach to code-switching in IRC can contribute to an understanding of the controversial issue of CMC as a third medium of communication. The interactive written discourse of IRC, whereby participants communicate through only writing, results in more reliance on code-switching for purposes of emphasizing, framing, and sarcasm than in face-to-face interaction. Being a key issue in the present study, it is necessary to establish a link between code-switching phenomenon in general and its communicative setting in CMC. Thus, the subsequent subsections give a brief review of the vast contributions in the literature on code-switching studies. Then, methodological issues will be addressed. Finally, results will be discussed and implications are drawn.

## **7.1 Code-switching and CMC**

This section begins with relevant theoretical background of the phenomenon of

code-switching in general, including its different terminological frameworks. Linguistic and sociolinguistic approaches to code-switching are detailed, with emphasis being given to Gumperz' (1982) notion of conceptualizing cue and Auer's (1984, 1995, and 1999) Conversation Analysis approach to code-switching (those being the main approaches adopted in the present study). Then a review of CMC studies addressing the phenomenon of code-switching with reference to recent and relevant Arabic CMC studies will be given.

### **7.1.1 Code-Switching: A terminological framework**

This section offers an account of the code-switching research agenda. Operational definitions of code-switching and related theoretical essential concepts will be reviewed. There is a noticeable lack of agreement on the exact definition of the essential concept "code". It is not in the scope of the present study to define this concept unambiguously; however, an account of how this term will be used in this study is offered. Then, code-switching research is reviewed from three main perspectives: structural approach, sociolinguistic approach (macro-level), and interactional approach (micro-level), from among which, the Conversational Analysis approach (1984, 1995, and 1999) to code-switching will be followed in the present study.

#### **7.1.1.1 Code**

According to Alvarez-Caccamo (1998:1), it is necessary for the analyst to identify the "code" that has been switched from before he/she can claim the occurrence of code-switching. Alvarez-Caccamo (1998:1) maintains that the concept of "code" should be understood as a "communicative code" and claims that "a code for

anthropology should be what it should also be for linguistics or for communication theory”. According to Alvarez-Caccamo, this communicative code is “a mechanism of transduction between intentions and utterances, and then between utterances and interpretations” (ibid: 38), such as prosody, gazes, and gestures. Alvarez-Caccamo’s interpretation of ‘code’ is extended to include any mechanism of switching between communicative intentions. Code-switching then will only include those instances that can be seen, as “an interactional achievement which is not prior to the conversation but subject to negotiation between participants” (Auer 1998: 15).

In written CMC conversation, this interactional achievement is accomplished by both language choice and script choice. Following this line of thinking, the code is used in this thesis in a broad sense to refer to both language (Arabic and English) and script (Arabic and Roman). The interactional achievement conceptualizing code-switching is interpreted in terms of switching between languages and switching between scripts. It is worth mentioning that any two juxtaposed languages or scripts are not necessarily two distinct codes. As a whole, they may be considered as being only one code if code-switching itself is the unmarked choice (Myers-Scotton 1992).

#### **7.1.1.2 Code-switching**

Code-switching is a common linguistic practice among bilingual and multilingual speakers that has received attention from several perspectives. In particular, code-switching is studied in psychology, general theory of language, first and second language acquisition. This multi-disciplinary nature of research that the code-switching phenomenon has received makes any attempt to provide a straightforward definition of what code-switching is inherently problematic. Nevertheless, early attempts to define code-switching date back to Weinreich’s (1953) description of

bilingualism as “the practice of alternately using two languages”. This is a very broad definition of code-switching, indicating the use of two or more languages.

In the course of the development of code-switching research, many specifications to this definition have been attempted for different purposes. One of the classic definitions that most CMC researchers adopt is Gumperz’s definition of code-switching as “juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or subsystems” (Gumperz 1982: 59). Myers-Scotton (1993: vii) gives a concrete definition of code-switching as “the use of two or more languages in the same conversation, usually within the same conversational turn, or even within the same sentence of that turn”.

### **7.1.2 Code-switching vs. borrowing**

The phenomenon of borrowing has been amply discussed in language contact literature (e.g., Haugen 1950, Weinreich 1953, Romaine 1989, Myers-Scotton 1992, Matras 2009). According to Muysken (1995: 189), borrowing is “the incorporation of lexical elements from one language in the lexicon of another language”. By language borrowings, Grosjean (1982) means lexical items which, though originally from language A, have come to be integrated into language B in a way that language B speakers use them without any awareness of their foreign origin. The language items that have undergone the process of borrowing are called “borrowings” or “loans”. For the sake of consistency, the term “borrowing” is used throughout this study.

Poplack, Sankoff, and Miller (1988) classified borrowed items, as opposed to single word code-switching, in two main categories: nonce borrowings and

established borrowings. Nonce borrowings differ from established borrowings in being singly-occurring in Embedded Language (EL) items that occur infrequently (or even only once) in a data corpus but show morphological integration into the Matrix Language (ML). (Sankoff, Poplack and Vanniarajan 1990). According to Poplack et al. (1988: 93), both established and nonce borrowings entail a lexical item from EL occurring in ML, and significantly adhere to the morphological and syntactic rules of ML. By contrast, single-word code-switching occurs when each monolingual item is lexically, morphologically, and syntactically grammatical in that language. The hypothesis that code-switching entails two grammars, whereas borrowing only entails one (Poplack *et al.* 1988: 93), remains an essential assumption of differentiating between borrowing and code-switching.

Distinction between code-switching, conscious single-word insertion, and borrowing (the diachronic process by which languages enhance their vocabulary through adopting lexical items from another language) is complex. Several criteria have been proposed to distinguish between code-switching and borrowing. Among these criteria is the phonological and morphosyntactical adaptation (Poplack et al. 1988, McClure 2001) which refers to the degree to which a foreign item has been integrated and adopted into the host language. Thus, Poplack and Meechan (1995: 200) indicate that:

“ ‘Code-switching’ may be defined as the juxtaposition of sentences or sentence fragments, each of which is internally consistent with the morphological and syntactic (and optionally phonological) rules of its lexifier language.... ‘Borrowing’ is the adaptation of lexical material to the

morphological and syntactic (and usually phonological) patterns of the recipient language.”

Borrowed material usually consists of cultural loans representing new concepts that have been integrated into a given language. Myers-Scotton (1993: 163) distinguishes two types of cultural borrowings: cultural forms and core forms. Cultural forms refer to concepts that are new to the ML culture and are easily adopted by monolingual speakers since equivalents in ML do not exist. Core forms, on the other hand, refer to items that have equivalents in the ML. The use of core forms presupposes bilingualism and code-switching at the initial process of borrowing.

The “Frequency Hypothesis” criterion, the degree of occurrence of a lexical item into the host language, has been also considered in distinguishing between code-switching and borrowing (Myers-Scotton 1987). Borrowed forms, Myers-Scotton (1987: 70) insists, “should be distinguishable set from the embedded language which speakers know in some abstract sense, borrowings are available to many (or all) speakers, while embedded forms in switching are not” in other words, borrowings tend to be part of the mental lexicon of the speakers, while code switched items are considered as part of the donor language.

The above mentioned criteria indicate that the distinction between borrowing and code-switching is not straightforward, and that they are best viewed as a continuum (Matras 2009). According to Matras (2009: 113), the prototypical kind of borrowings entails “the regular occurrence of a structurally integrated, single lexical item that is used as a default expression, often a designation for a

unique referent or a grammatical marker, in monolingual contexts”. On the other hand, the prototypical, least controversial code-switching is “the alternational switch at the utterance level, produced by a bilingual consciously and by choice, as a single occurrence, for special stylistic effects” (Matras 2009: 113).

In his analysis of language contact situations and mixed languages, Matras (2009: 288–307) emphasizes that these linguistic practices, language insertion – written and oral – is part of the overall bilingual activity. The emergence of such linguistic practices is due to conscious, indeed deliberate, acts by speakers who apply parts of the lexicon of an embedded language into the dominant matrix language. Matras also stresses the fact that mixed languages appear on a continuum, or functional cline. The bilingual speaker’ tendency to code switch and contrast is due to an effort to eliminate the burden of having to select the appropriate linguistic expression from among the two different languages. Bilingualism is responsible for a permanent tension whenever bilingual speakers must follow rules on the appropriateness of choices among sets within their overall linguistic repertoire. In order to decrease tension around contrast, the choice of a particular expression is eliminated through the process of fusion into a single repertoire.

Different analytical approaches have been proposed to understand both the grammar and functions of these consciously-produced language alternations. Below is an overview of the most influential approaches to code-switching.

### **7.1.3 Analytical approaches to code-switching**

The literature of code-switching falls into two main perspectives: grammatical (structural) and sociolinguistic (pragmatic). Theoretical models put forth within the linguistic approach attempt at categorizing code-switching patterns according to the levels where code-switching occurs in sentences (Poplack 1980), or identifying a base (matrix) language of code-switching according to syntactic and morphological criteria (Myers-Scotton 1993). Within the pragmatic sociolinguistic scope, code-switching is viewed as a communicative activity and has been assigned different functions in a discourse. The following two subsections are devoted to an overview of these predominant perspectives.

#### **7.1.3.1 Linguistic approaches to code-switching**

The first method of categorizing structures of code-switching is the difference between “alternational” code-switching and “insertional” code-mixing. Hoffman’s definition of alternational code-switching is “the alternate use of two languages or linguistic varieties within the same utterance or during the same conversation” (Hoffman 1991: 110). On the other hand, “insertional” code-mixing is defined as “the embedding of various linguistic unites such as affixes (bound morphemes), words (unbound morphemes), phrases and clauses from two distinct grammatical (sub-) systems within the same sentence and speech event” (Bokamba 1989: 278). In other words, code switching occurs at the inter-sentential level, while code mixing occurs at the intra-sentential level. Poplack (1980: 228) states that “code-switches will tend to occur at points in discourse where juxtaposition of L1 and L2 elements

does not violate the systematic rules of either language”.

Myers-Scotton’s (1993a, 2004) Matrix Language Framework (MLF) is based on the idea that the Matrix Language provides a grammatical framework into which the EL is inserted. The ML tends to provide system morphemes that are roughly equivalent to “grammar”, whereas EL inserts content morphemes which are roughly equivalent to “lexicon” (Matras 2009: 130). System morphemes, Callahan (2002) illuminates, are inflectional morphology like quantifiers and specifiers. Content morphemes have the feature of being “Thematic Role-Assigner or Thematic Role-Receiver” (Callahan 2002: 6), like adjectives, nouns and verbs. According to this framework, the inflectional bound morphemes are provided by the ML, while free morphemes are from the EL.

The MLF theoretical framework can determine ‘where’ code-switching may take place but is incapable of providing a concrete understanding of ‘why’ speakers switch codes while they are not obliged to. Auer (1984: 2-3) comments that “grammatical restrictions are but necessary conditions; code-switching is not merely a matter of well-formedness—it also has communicative content left unexplained by the analysis of syntactic surface constraints” In an attempt to explore the motivations and mechanisms that underlie the process of code-switching among bilingual speakers, a number of models have been formulated and developed from different sociolinguistic dimensions. Below is an overview of the three most influential sociolinguistic approaches to code-switching.

### **7.1.3.2 Sociolinguistic approaches to code-switching**

Early sociolinguistic investigation of code-switching reveals that code-switching in communication is best viewed as socially motivated linguistic practice, affected by the linguistic and social identity of the speakers, and constrained by the situational factors of the settings in which they take place. As Auer (1984:1) puts it, “ although the importance of the grammatical and sociolinguistic sides of the issue is beyond doubt, the analysis of the meaning of individual instances of language alternation seems to be the most basic and also the most unresolved question”. Investigations about why bilingual speakers code-switch and what the functions of code-switching are have been conducted by many scholars from different perspectives, and they have developed different theoretical models to explain this phenomenon. Studies by Bloom and Gumperz (1972), Gumperz (1982), Auer (1984, 1985, and 1999) and Myers-Scotton’s (1993b) are widely regarded as the most influential in code-switching from the interactional perspective. According to these studies, it is only with turn-by-turn scrutiny of the interactive conversation in which code-switching takes place that we can answer the questions of ‘why’ speakers use two codes within the same exchange passage (and so to achieve the local meanings of code-switching).

#### **7.1.3.2.1 Gumperz: Situation-Metaphor—We code and they code**

Perhaps no sociocultural linguist has been more influential in the study of code-switching than J. Gumperz. His explanation of code-switching as a contextualization cue has been influential in the fields of sociolinguistics, linguistic anthropology, and the sociology of language. Gumperz (1982) maintains that language is related to

social structures of its speakers and, thus, language alternation is usually associated with concepts of ethnicity and identity. In this regard, Gumperz (1982:66) argues that “the tendency is for the ethnically specific minority language to be regarded as the ‘we code’ [...] and for the majority language to be regarded as the ‘they code’”. In bilingual communities, Gafaranga (2007) maintains, languages form a structure and are functionally differentiated. One is the ‘we code’ while the other is the ‘they code’. According to Gafaranga (2005), bilingual speakers prefer to use one language over another in order to emphasize the corresponding identities in a society.

Relying on empirical data, Bloom and Gumperz (1972: 422-423) proposed two kinds of code-switching. The first kind is *situational non-conversational code-switching*, which is associated with situational factors such as a change in situation, interlocutors, content, and topic. It is also related to relationships between interlocutors, community norms, and values. In a bilingual community, only one language is considered by community norms to be appropriate in a particular situation. Thus, speakers tend to switch their language in order to maintain appropriateness. The second kind of code-switching proposed by Bloom and Gumperz (1972) is *metaphorical or conversational switching*, in which bilingual speakers switch their language though there is no change in the external situational factors. The rationale of such switching, Bloom and Gumperz argue, is to achieve particular communicative functions without a change in setting.

Gumperz’s approach to language study is sociolinguistic as he is interested in issues of linguistic variability. However, Gumperz’s approach is different from that of Labov (1972) and his followers, who were interested in accounting for the linguistic variability by reference to social factors such as gender, age, and ethnicity.

Rather, Gumperz is mainly interested in studying how linguistic variations are used as a linguistic resource on their own in actual face-to-face interaction. In this regard, Gumperz explicitly states that code-switching itself works as a contextualization cue in face-to-face communication to fulfil different communicative goals. He makes the following comment, which is particularly relevant to the purpose of the present chapter: “code-switching signals contextual information equivalent to what in monolingual settings is conveyed through prosody or other syntactic or lexical processes. It generates the presupposition in terms of which the content of what is said is decoded” (1982:98).

#### **7.1.3.2.2 Myers-Scotton’s Markedness Model**

Equally influential in code-switching research is the *Markedness Model* (1983, 1988, 1993b, 1998, and 1999) which was developed by Myers-Scotton in an effort to explore socio-psychological processes that underlie code-switching. The object of the *Markedness Model* is “the why” of language alternation (Li Wei 1998). Myers-Scotton and Bolonyai argue that language use follows principles of rational behaviour. They maintain that these principles are best explained as “cognitively based calculations that depend on the actor’s estimation of what action offers him/her the greatest utility” (2001:2). Myers-Scotton explains that the *Markedness Model* assumes that “all linguistic code choices are indexical of a set of rights and obligations holding between participants in the conversational exchange” (2000: 138). In other words, bilingual speakers make rational choices according to the rights and obligations they perceive in a communicative situation considering the costs and rewards associated with other

choices and, thus, acting to “optimise their returns is the premise of rational actor models” (Myers-Scotton 1999: 1259).

As for the choices themselves, Myers-Scotton categorizes them into two main categories; unmarked and marked choices. Matras (2009: 115) commented that “unmarked” choices are “an index of the default relations among participants”—their set of “Rights and Obligations” in that specific context. “Unmarked” is thus understood both as the “default” choice and as the most frequent one. “Marked” choices, based on this model, are those that are not expected. In other words, the unmarked choice is “the linguistic variety which is most expected, while the marked choice is most unusual” (Myers-Scotton 1993: 89). The *Markedness Model* assumes that speakers have a sense of markedness and are able to identify whether their linguistic choices are marked or unmarked, and that their choices are based on social-psychological factors (1993: 75).

#### **7.1.3.2.3 Auer’s Conversationalist approach**

In contrast to Gumperz’s approach to code-switching, and Myers-Scotton’s *Markedness Model*, is the conversation analytic approach to code-switching, pioneered by Peter Auer (1984), which aims to examine the functions of code-switching. This approach gives emphasis to the structure and sequential nature of the code-switching in conversation rather than relying on the external social factors. Auer’s Conversation Analytic approach to code-switching, which was developed further by Li Wei (1994) and Auer (1998) is based on the premise that the functions of code-switching emerge with the sequential development in

conversation. Thus, Auer (1995:116) argues for a sequential analysis of code-switching, saying:

“Any theory of conversational code alteration is bound to fail if it does not take into account that the meaning of code-alternation depends in many ways on its ‘sequential environment’. This is given, in the first place by the conversational turn immediately preceding it, to which code-alternation may respond in various ways.”

As detailed turn-by-turn analysis of language alternation is the starting point of identifying the meaning and functions of code-switching, Conversation Analysis is adopted as the basic methodology in this approach. Therefore, code-switching is considered to be part of “verbal action” and its significance can be identified by examining conversational structure which is dependent on both grammaticality and external societal factors, though both have relevance to code-switching (Gardner-Chloros 2009: 11). According to Auer (1984: 5-6), the Conversation Analytic approach to code-switching understanding has its own advantages over other approaches to code-switching study. First, it relies on the sequence of conversation, and emphasizes the influence of turns. Second, it decreases subjective influences as it focuses the analysis on turns of conversation.

Auer (1999) differentiates code-switching from other language alternation phenomena (i.e., language mixing (LM) and fused lects (FL)). According to Auer’s theoretical framework, code-switching cases are locally meaningful, and they derive their local functionality from interacting with its surrounding contexts

to achieve an interpretation as a whole. He argues that this local functionality distinguishes code-switching from other cases of language alternation. When juxtaposition[s] does not seem to have that local functionality, “from an interpretative point of view, they cannot be called code switching” (Auer 1999: 315). Following is Auer’s distinctions between CS, LM, and FL (1999: 328):

- Code switching (CS): referring to a phenomenon in which the juxtaposition of two languages is seen as locally meaningful activity by participants themselves.
- Language mixing (LM): used to depict the cases of language alternation in which the juxtaposition of two languages is seen as a recurrent pattern, and hence is meaningful in a more global sense.
- Fused lets (FL): related to a more stabilized situation in which the use of one language or the other, or certain constituents, is fossilized and thus is obligatory for speakers.

It is worth mentioning that Auer’s division between CS, LM, and FL is considered in the analysis of the present study. The use of English words like *ok*, *thanks*, *bye bye*, *hi*, and *hello* in face-to-face conversations and IRC is fossilized, and thus are not considered as instances of code-switching that are locally meaningful by IRC interlocutors.

The situated local functionality of code-switching as a contextualization cue, according to Auer, is strongly related to sequential patterns of language choice. Auer (1995) also distinguishes between discourse-related code-switching and participant-related code-switching. The former plays a role in structuring conversational acts,

including turn-taking and repair, while the latter is motivated by the language preferences, or the participants' competences (Auer 1984:12). Since these two patterns are adopted in the present study in order to understand the local functionality of code-switching in IRC, a brief summary of them follows:

- Discourse-related code-switching contributes to the organization of discourse in a particular episode. Discourse-related code-switching is the typical kind of code-switching that contextualizes features of the conversation, such as shift in topic, shift in participant constellation, shift in activity type and marking off of a side comment. Thus, discourse-related code-switching contributes to the organization of discourse (Auer 1995: 124).
- Participant-related code-switching is related to participants and their preference rather than redefining the discourse. According to Auer, participant-related code-switching allows assessments of/by participants. "Preference" here should be understood in the "conversation-analytic sense of an interactionally visible structure". However, deliberate decisions arising from political or other considerations might contribute to participant-related code-switching (Auer 1995: 124).

Myers-Scotton's *Markedness Model* has been criticized as having some limitations. It is not always possible, Li Wei (2005) argues, to identify and explain functions of code-switching by depending only on the rationality of speakers. Earlier, Stroud (1992) indicated that the difficulty in applying this approach lies in the difficulty of knowing to what extent the intentions and

meanings that the analyst assigns to switches are, in fact, intended by a speaker or perceived by interlocutors. On the other hand, the conversation analytic approach to code-switching has been criticized for excluding social context in interpreting functions of code-switching. Myer-Scotton and Bolonyai (2001: 4) argue that the structural analysis of any conversation constrains speakers “to view certain potential choices as preferred and others [as] not. However, the structural features studied by conversation analysis offer an exceedingly flat explanation of choices”.

A promising approach to code-switching, therefore, is identifying the signalling value and exploring the conversational meaning of code-switching. Such an approach requires a sequential analysis of code-switching, in which the code chosen for a particular speech activity must be considered in terms of the background of language choice in the preceding turns of conversation. Therefore, the object of this approach is not what verbal activities are associated with one choice over another, but instead, in which activities speakers switch from one language to another. From this perspective, researchers have been interested in developing typological frameworks of functions of code-switching, and their discussions are related to neither the Markedness approach nor the Conversational analytic one.

Their discussions of functions of code-switching can be called the checklist approach to code-switching. The methodology of such an approach is to develop elaborate typologies of certain conversational loci in which code-switching is notably recurrent. Myers-Scotton chose the label “better-taxonomy” approach

for these studies (1993: 63) as they use a checklist to identify functions of code-switching. In an effort to categorize functions of code-switching, Gumperz (1982) proposes his famous six functions of code-switching, which offered inspiration to different studies, including the present one. Gumperz's taxonomy of code-switching functions include *quotations*, *addressee specification*, *interjections*, *reiteration*, *message qualification*, and *personalization versus objectivation* (Gumperz 1982: 75-84). Grosjean (1982) presents a more comprehensive list of motivations of code-switching. Reasons for code-switching, according to Grosjean, are to:

“fill a linguistic need for lexical items, set phrase, discourse marker, or sentence filler; continue the last language used (triggering); quote someone; specify addressee; qualify message: amplify or emphasize (“topper” in argument); specify speaker involvement (personalize message); mark and emphasize group identity (solidarity); convey confidentiality, anger, annoyance; exclude some from conversation; change role of speaker: raise status, add authority, show expertise.”

(Grosjean 1982: 152).

Although checklists studies of code-switching are useful as they indicate that some conversational loci are particularly prone to code-switching, they have received much criticism from pioneering scholars in the field. Auer (1995) argues strongly against the mere listing of conversational loci of code-switching that lack the sequential account to indicate what exactly is meant by these categories. He also indicates that so-called typologies lack specific criteria of categorization as they often confuse conversational structures, functions of code-switching, and

linguistic forms (Auer 1995: 120-121). Myers-Scotton (1993: 93) maintains that the conversational categories of these typologies lack a coherent and comprehensive theoretical framework.

To sum up, interest in the sociolinguistic dimensions of code-switching has yielded different theoretical models that have been developed in attempts to explain the mechanism of code-switching from different perspectives. As code-switching is a frequent linguistic practice among bilingual speakers, it is employed by bi- and multi-lingual communities via digital media. The questions of how and why code-switching is used in CMC cannot be addressed independently of the sociolinguistic theoretical frameworks of code-switching offline discussed above. Having covered the literature of code-switching, the following section offers a critical synthesis of the available literature of code-switching in different CMC modes. The available literature on code-switching in CMC offers insights into how code-switching has been researched so far and into the promising perspectives for further study of code-switching.

#### **7.1.4 Code-switching in CMC**

Early linguistic investigation in CMC operated within a monolingual frame with an emphasis on issues of language/technology relation, such as the location of CMC discourse between spoken and written discourse. With the dominance of English functioning as the lingua franca in CMC, and the coexistence of different websites, channels, and forums that are composed of textual units of different languages, multilingualism became inevitable as a result of different purposes and audiences. “Whenever multilingualism exists, language choice becomes an

issue” (Danet & Herring 2007: 17). Code-switching in CMC attracted the attention of scholarly investigation from the 1990s onwards. Most code-switching research online has been conducted among diaspora communities from social and cultural perspectives utilizing a variety of research tools. Below is an overview of code-switching in CMC research with reference to the key methodological elements that CMC researchers appropriate and apply in different ways.

Paolillo’s (1996) early in-depth study of code-switching among Punjabi diaspora speakers sets the tradition of a large body of research into different diaspora communities. He examined code-switching practices in the Usenet newsgroup soc.culture.punjab using both quantitative and qualitative techniques. His aims were to find how frequently different languages are used in the Usenet, what patterns of code-switching Punjabi bilingual discourse online displays, and what factors motivate these patterns. His quantitative analysis reveals that the use of English is four times more frequent than the use of Punjabi. His type/token calculation reveals that there are no Punjabi function words among the most frequent forms and that the highest ranking English tokens are function words. Paolillo’s quantitative calculations reveal that English is used much more productively, and that the use of Punjabi is restricted to the insertion of loanwords and borrowing into English discourse. These loanwords and borrowing, according to Paolillo, are “necessitated by talking about the Punjab region, Punjab, and Punjab people” (Paolillo 1996: 23).

In order to arrive at an explanation as to why Punjabi is not used in a productive way in Usenet newsgroup soc.culture.punjab, Paolillo (1996) used

a qualitative approach to his data. He investigated communicative functions of both Punjabi and English in his data. Functions of Punjabi are of two kinds; fixed uses, which are phrases that are invariant, and creative uses, which are situationally relevant. Creative uses of Punjabi are found to occur in three types: as insults, as appeals, and as the climaxes of jokes. Political arguments are carried out wholly in English, an occurrence which, according to Paolillo, is due to the greatest prestige of English compared to the vernacular. Paolillo indicates that the functional marginalisation of Punjabi use in soc.culture.punjab is best considered as a consequence of different interrelated factors: the prestige status of English in South Asia, cultural ambivalence of expatriate Punjabis, and the specific soc.culture.punjab newsgroup norms, which favour English.

Shortly after Paolillo's investigation of Punjabi/English use in a Usenet newsgroup, Georgakopoulou (1997) investigated code-switching practices among Greek diaspora speakers living in the UK. More specifically, she examined how the self-presentation and interactional alliances in email communication are accomplished by style and code-switching. She explored social-cultural identities in CMC by going beyond what quantitative analysis of CMC can achieve and drawing on the methodologies of interactional sociolinguistics and ethnography of communication. In describing the style of email communication, Georgakopoulou suggests that the formation of interactional alliances is achieved by both style-shifting (from Katharevousa to Demotic<sup>14</sup>) and code-switching from Greek to English. In line with Gumperz (1982), she maintains that code-

---

<sup>14</sup> **Katharevousa** is a conservative form of the Modern Greek language conceived in the early 19th century as a compromise between Ancient Greek and Demotic of the time. **Demotic Greek** or **dimotiki** is the modern vernacular form of the Greek language.

switching in email data functions as major contextualization cue for deploying the “deprived” communicative effects of emails communication with respect to prosodic and paralinguistic features (Georgakopoulou 1997: 194).

Researching code-switching among diaspora speakers in an online setting has recently been conducted by Androutsopoulos (2007). He distinguishes between two main approaches to language alternation in CMC, both of which adapt sociolinguistic theories of code-switching (Gumperz 1982, Auer 1995) to CMC. The first approach is the study of code-switching (the use of more than one language in a single conversation), which typically draws on interactional sociolinguistic and the conversational analytic approach to code-switching (Gumperz 1982, Auer 1995). The second approach is the study of language choice (the distribution of languages use among bilingual and multilingual speakers), which operates on a macro level, drawing on the sociology of language (Li Wei 2005, Milroy & Pong 1992). The latter approach also focuses on the strong presence of English in multilingual CMC. English is observed to be preferred to Punjabi among expatriate South Asians in a synchronous mode of CMC (Paolillo 1997). It is also found to be frequently used for formal communication, whereas Egyptian Arabic is used for informal exchanges among Egyptian internet users (Warschauer *et al.* 2002).

Adopting both approaches to language alternation in CMC, Androutsopoulos (2007) examines language choice and code-switching among second- and third-generation immigrants (Persian, Indian, and Greek). Analyzing discussion forums of diasporic websites, he found that language use in diasporic communities in CMC is structured in complex ways. The choice between the community language and the dominant language is dependent on factors such as the historical depth of migration,

the size of an ethnic community, and the structure of each diasporic forum. A significant implication made by Androutsopoulos (2007) is that bilingual communication in diasporic forums is comparable to conversational code-switching in terms of its discourse functions.

The functional diversity of code-switching in asynchronous CMC mode is a significant observation in bilingual communication in diasporic forums examined by Androutsopoulos (2007). In contrast, Paolillo (2008) indicates that interactional code-switching should be more frequent in synchronous modes of CMC as they are more speech like than asynchronous ones. Paolillo's hypothesis of synchronicity is perhaps the only one in the literature of code-switching in CMC, and it originates from his recent study (2011) of English/Punjabi in IRC and Usenet. Paolillo (2011) finds that IRC interactions display creative conversational code-switching, whereas Usenet messages data is limited to formulaic code-switching. His findings support his early suggestion that synchronicity is a fundamental part of what it means for a communication mode to be "conversational" (Paolillo 2011: 1).

Based on the assumption of English preference among multilingual CMC, Warschauer *et al.* (2007) conducted a quantitative study among forty three young professionals in Egypt to examine language choice online. Warschauer's study of language choice online constitutes the first investigation of language contact in the Arabic context of CMC. Utilizing a written survey and examining emails sent by University students in UAE, Warschauer and his colleagues aimed to find out in what circumstances, and why, young Egyptian internet users use English and Arabic. The results of this study reveal two main findings: first, English is the dominant choice online among internet users in Egypt because it is the lingua franca of technological communication; second, a written form of Romanized Egyptian Arabic is

extensively used by young Egyptian internet users. MSA was found to be the least preferred choice.

Durham (2007) examined how external situations affects language choice in a mailing list of members of a Swiss medical organization. Examining language choice in that context provides a rich field of investigation due to the composite sociolinguistic situation that has four national languages (German, French, Swiss, and Romansch). English has been observed to be the lingua franca of Intra-Swiss communication among Pan-Swiss medical members. Thus, Durham investigated the members' use of English over four years to find out when and why language alternation occurred. The main findings emerged from this study is that English, though a non-native language for all, is the dominant language among participants of the study. English was found to be the most understood and accepted language in mixed language groups, mainly because it is a non-native language for all. Another implication of this study is the effect of the mode of communication (the mailing list) on language choice. As long as messages of the mailing list are directed to specific audiences, there is no need to worry about whether an entire group understands. Communicating to the multilingual audience via the internet, Durham claims, necessitates the use of English, as neither French nor German can serve as the main language in the Swiss context. On the other hand, English, the lingua franca of the internet serves the most understood language of communication.

Recent studies have been conducted to examine Arabic-French code-switching in CMC. Laroussi's (2011) book *Code-Switching, Language in Contact and Electronic Writing* gives answers to the question of testing the theoretical tools of oral code-switching in analysing written code-switching in CMC. Laroussi (2011) maintains that grammatical and sociolinguistic approaches to code-switching have focused

mainly on the analysis of spoken data. According to Laroussi, written code-switching does not show the same features of spoken code-switching, and therefore does not call for the same treatment. His investigation is based on the question, “are the theoretical tools that have been used for analyzing oral productions relevant for the analysis of code-switching in electronic writing?” (Laroussi 2011: 10). In his study of electronic Arabic-French code-switching in SMS, Laroussi (2011) applies the theoretical tools proposed by Poplack (1980, 1981, and 1988) to investigate syntactic-constraints of code-switching. As regards how constituents of the two languages fit together, the criterion of taking any sentence where the predicate is in Arabic as having an Arabic base does not always work. Laroussi (2011) found cases where it is not possible to classify sentences where the predicates are mixed.

The literature on code-switching in CMC reveals that a lot remains to be done in investigating code-switching in different CMC modes and different languages. Most code-switching research in CMC has focused mainly on examining language choice in diaspora communities (Paolillo 1997, Georgakopoulou 1997, Androutsopoulos 2007) or on the kind of relationship between the two languages and the effect of the language situation on language choice online (Durham 2007, Warschauer *et al.* 2007) . The grammar of code-switching in CMC has always been under researched. A valid general criterion of distinction between borrowed items and single-item code-switching is needed to understand CMC participants’ norms and perceptions about borrowed and foreign lexical items. As the IRC data displays code-switching to both English and Arabic, an explanation as to what motivates code-switching to English and Arabic should be obtained.

Most lacking in the literature is an investigation of code-switching in CMC within an Arabic context. There is also a noticeable shortage of research conducted to examine code-switching within synchronous CMC modes. Keeping the properties of Najdi IRC communication in mind, and utilizing both grammatical and sociolinguistic approaches to code-switching, this study will contribute to the understanding of code-switching practices in CMC within an Arabic context. CMC is not only multilingual, but also multimodal, and Laroussi's questioning about the relevance of the traditional approaches that have been utilized to understand oral code-switching for the analysis of code-switching in CMC should be considered. Li Wei et al. (2005) demonstrates that combined models synthesizing both aspects of micro and macro approaches would be fruitful in the study of language use among bilingual speakers. The researcher in the present study suggests, in line with Androutsopoulos (2007), that such a combination can be used in exploring code-switching in CMC and can be extended to explore other multilingual practices online such as script-switching.

## **7.2 Methodology**

A qualitative analysis was conducted to explore in great depth the occurrence and types of code-switching in IRC data. In this part of the analysis, the main priority was not the quantification of how frequent these cases of code-switching occur in the IRC data. Instead, the analysis aims at distinguishing cases of borrowing from code-switching. Descriptive/structural analysis for the sake of identifying general patterns

of grammatical integration of English lexical items in both Arabic and Roman script was conducted. Instances of English borrowings were structurally analyzed on the bases of morphological integration criteria (Poplack 1988, Muysken1990) for the sake of identifying patterns of morphological adaptations of English borrowed items in both Arabic and Roman scripts. In the process of analysis, borrowings are distinguished from code-switching following Poplack and Sankoff's (1984) and Muysken's (1990) accounts of the characteristics of borrowing. These characteristics include: (1) Borrowed lexical items are inserted in the native lexicon; (2) Borrowings show phonological, morphological, and syntactic adaptation; (3) Borrowings are used frequently in everyday speech; (4) Borrowings substitutes a language's own words; and (5) Borrowings are recognized and perceived by speakers as words in their own language. An element should not fulfil all these criteria to be considered as borrowing.

After distinguishing between instances of borrowings and instances of single-word code-switching, the communicative functions of code-switching to English/Arabic were analyzed. At this stage, the analysis was aimed at achieving a locally-situated interpretation of code-switching by interpreting both discourse-related patterns and participant-related patterns. The IRC examples were selected according to their representativeness in clarifying the local functions of code-switching and its correlation with immediate context. The qualitative analysis was oriented to categorize the types of code-switching identified in the Najdi IRC data and to investigate the most frequent functions of code-switching. The analysis concludes by creating a typological framework to illustrate in which conversational loci code-switching is particularly frequent in an IRC setting.

To answer the question of motivations influencing participants' switching between scripts, The IRC data was given then interactional discourse analytical procedure with emphasis on how interactions take place and how switching between different scripts contributes to the functionality of IRC texts. The analysis was aimed at achieving a locally-situated interpretation of script-switching by adopting Auer's Conversational Analysis approach to code-switching assuming that script alternation is not entirely arbitrary. Rather, it is driven by various discourse-related functions and participant-related functions. Instances where script switching exists provide important guiding concepts in analyzing data. However, these motivations were not immediately observable and identified. Rather, they had to be identified by an ongoing process of close interpreting and re-interpreting of IRC data. An effort was made to cover most instances of script-switching and to provide an account of their communicative functions

### **7.3 Morphological integration of borrowings**

Research into borrowing vs. code-switching has centred on the identification of foreign items that originated in one language and occur in the structure of another language. What characterizes this body of research is investigating the process of integrating borrowed items within the structure of the ML. If a foreign item is introduced into the ML, it will undergo integration processes before it becomes an acceptable item in the lexicon of ML. MacSwan (2000: 51) indicates that "the language faculty has a lexicon, which includes internal morphological rules of word formation". In cases of borrowing, these internal morphological rules of word formation influence the acceptability of foreign items into the ML.

In Najdi IRC, the influence of the Arabic morphology on English borrowed items is evident though different morphological rules<sup>15</sup>. As Arabic IRC data is composed in both Roman and Arabic scripts, it is significant to investigate and categorise the process of morphological adaptations in both scripts. To best understand the phenomena under investigations, findings will be compared to studies of borrowing in different spoken Arabic dialects.

### **7.3.1 Morphological integration of borrowings: Roman script**

In Romanized IRC, morphological adaptations of borrowed items are exhibited through the use of a variety of morphological bound and free morphemes. These include the use of the Arabic definite article (*al-*), prepositions, pluralisation, the *idafa* construction (possessive constructions), and verb conjugations. In some instances, modifications are undergone by more than one bound morpheme, indicating different degrees of modifications. Below is a discussion of these morphological adaptation processes. Arabic morphemes are given in bold. Original English words are provided in the gloss.

#### **7.3.1.1 The definite article**

The Arabic definite article (*al-*) is prefixed to nouns to express definiteness. It is also prefixed to each of that noun's modifying adjectives. When an English

---

<sup>15</sup> Since this study investigates written data, no reference will be made to phonological integration.

borrowed noun is used in Romanized IRC interactions, the definite article is prefixed. Examples below show this integration.

- (1) **el**,krismas        ‘DEF-Christmas’
- (2) **el** group        ‘DEF-group’
- (3) **el**, email        ‘DEF-email’
- (4) **el**-class        ‘DEF-class’
- (5) **el** pmessenger    ‘DEF-P Messenger’(a name of a software program)

The use of the Arabic definite article, (*al-*) before words such as (Christmas, group and email) indicates that these borrowed items have been integrated syntactically into the Arabic language. The insertion of a definite article is the most widely and easily adopted process of morphological integration in the IRC data. This contrasts Myers-Scotton’s (2002: 115) observation about borrowings of Dutch origin occurring in Moroccan Arabic as they were rarely attached to the definite article. On the other hand, she notes that French words occurring in North African Arabic appear with the French definite article “*le*” (2002: 116).

These borrowed items (Christmas, group, class, email) occur frequently in the recipient language and, therefore, can be established as loan words (Romaine 1994). Equivalents of these words do exist in the Arabic language, but the use of Arabic equivalents would sound very formal and would not fit the informal context of IRC interaction. In example (5), the Arabic definitive article is inserted before the English name of a programme, of which there is no corresponding Arabic name.

### 7.3.1.2 Preposition

Instances where the definite article is preceded by the prepositions such as (-*be*), meaning (*in*) and -*fee* meaning ‘in’ or ‘on’ indicate that these items have been integrated fully into Arabic syntax. Examples of borrowed items below show two levels of morphological adaptation: by the use of the definite article and the use of a preposition.

(6) **fee el** christmas ‘in DEF-Christmas’

(7) **bel** documents ‘in DEF-documents’

In example (6), the preposition and the definite article are separated from each other following norms of the Arabic writing. However, they are integrated in writing in example (7) to approximate speech.

### 7.3.1.3 Pluralisation

Pluralisation of English borrowed nouns into Roman IRC shows another case of morphological integration. The pluralisation in NA follows the rules of pluralisation of MSA. Sound plurals are formed by the addition of the suffixes (-*īn*) (masculine) and (-*āt*) (feminine). The other way of plural formation is what is known as the ‘broken plural’ formation (see section 4.5.1.2).

In Romanized IRC, usually the feminine sound plural marker - *āt* is used to pluralize English borrowed nouns. Broken plurals never exist in the IRC data. Examples are given below.

(8) **classaaat** ‘classes’

(9) **emilat** ‘emails’

(10) *el teacharat* 'DEF-teachers'

(11) *al essayat* 'DEF-essays'

The pluralisation process of English borrowed nouns in this study confirms findings of Hafez's (1996) study of English borrowings in Egyptian Arabic, and of Semeaton's (1973) study of English borrowed items in the Alhassa dialect. Both Hafez and Semeaton found that English borrowings are usually pluralised through the feminine sound plural marker (*-āt*), and broken plurals process never exists. In some cases, the process of pluralisation is found to be accompanied by the use of the Arabic definite article, as in examples (10) and (11). The use of both suffixes (the definite article and the plural marker (*-āt*) as lexical insertion features is also found by Owens (2005: 26) with English borrowings in Nigerian Arabic.

#### 7.3.1.4 Possessive constructions

Another morphological adaptation process found in Romanized IRC is the use of possessive construction. Possessive construction in Arabic is comprised of a minimum of two morphological units and is accomplished in two different ways:

- A: an indefinite noun+ a pronominal suffix (e.g., *saiyārt-i* 'car-POSS.ISG' (my car), *walad-hum* 'son-POSS.3PL' (their son)).
- B: an indefinite noun+ a definite noun -a construct of a possessor and a possessed nouns (e.g., *bab l-beīt* 'door DEF-house' (the door of the house)).

Both types of possessive constructions occur rarely in Romanized IRC to integrate the English borrowings in the syntax of Arabic. Below are the few examples that illustrate the occurrence of both types.

(12) classna                    ‘class-POSS.1PL’

(13) essayatna                ‘essays-POSS.1PL’

(14) link el jawwal        ‘link DEF-Jawwal ’ (a name of a telecommunication  
company)

Examples (12) and (13) exemplify the first type of possessive construction. Example (13) shows two levels of integrations: pluralisation and possessive construction. The occurrence of borrowed items within possessive constructions is in line with Owen’s (2005: 27) findings in Nigerian Arabic. He also noted that both constituents of the second type of possessive constructions can be borrowings, a case that never happened in Romanized IRC in the present study.

### **7.3.1.5. Verb conjugations**

English verbs are also borrowed and integrated into the structure of Arabic in Romanized IRC. Two kinds of integrations in relation to verbs are identified in the IRC data: the use of an auxiliary Arabic verb and the adherence to morphological verbal conjugation.

Monosyllabic English borrowed verbs are usually preceded by an Arabic auxiliary verb, typically *sawwa* ‘do’ that implies the action of the verb. Examples are given below.

- (15) *swwyat search* 'made-ISG search'  
 (16) *swwi copy and paste* 'make- 2SG.F copy and paste'

Example (17) displays an instance where the Arabic bound morpheme (*-ha*) (3FSG) is inserted after an English verb (*save*). In this case, the English verb (*save*), which occurs frequently in technical conversations, begins to gain the status of an established loan word and, therefore, is integrated morphologically into the Arabic language. In example (18), the bound morpheme (*-t*) (PAST.1SG) is inserted after the English verb (*cancel*).

- (17) *saivihā* 'save- 2SG.3SG.F' (save them)  
 (18) *kancalt* 'cancelled-ISG'

### 7.3.2 Morphological integration of borrowings: Arabic script

While integrating borrowings into Romanized IRC shows different types of adherence to Arabic morphology, borrowings occurring in Arabic script are morphologically integrated only by the use of the Arabic definite article in the case of nouns, and by the use of the auxiliary verb “*sawwe*” ‘do’ or ‘make’ in the case of verbs. The Arabic definite article is added to some English borrowings that are definite by nature (e.g., Facebook, YouTube) in an attempt to integrate these English items into the structure of Arabic. Some of these nouns are preceded by prepositions.

- (19) أكثر واحد في الكلاس  
*'akθar waḥid fī il-klās*  
 more someone in DEF-class  
 'the best one in the class'

- (20) اضغط اللنك  
*iḍgaṭ il-link*  
 click-2SG.M DEF-link  
 ‘click the link’
- (21) شفت الفيديو؟؟؟  
*shə-ft il-fedyō*  
 saw-2SG.M DEF-video  
 ‘did you see the video?’
- (22) هنا في الفايبيوك  
*hina fī l-faisbok*  
 here in DEF-facebook  
 ‘here in the Facebook.’

These examples illustrate the role of script in making these borrowed items nativised and made to fit the structure and form of Arabic. Though borrowings are adapted morphologically in Romanized IRC, this adaptation is achieved by transliterating these borrowed items into the writing system of Arabic. This graphic adaptation signifies the attribution of these borrowings to Arabic language. This adaptation works both on the dialectical level (NA) and the standard level (MSA). For example, the word “الفايبيوك” has occurred once in a formal IRC interaction that cites an advice written in MSA.

- (23) نحن هنا في الفيسبوك لتعلم كيف نتخاطب حين نختلف..  
*naḥnu huna fī l-faisbok li-nataṣallam kaifa na-taxāṭb ḥīn na-xtalif*  
 we here in DEF-facebook for-learn-1PL how discuss-1PL when disagree- 1PL  
 ‘we are here in Facebook learning how to negotiate when we disagree.’

The role of script in identifying these items as borrowed is reflected by the graphic representation of these items into the script of Arabic. This is clearly distinguished from instances of single-word code-switching where IRC participants shift to the Roman script, attributing these items to English. Examples of such cases of code-switching are given below. In example (24), participant A wants to quote his

instructor's announcement. He code switches to both English and Roman script. The same happens with participant A in example (25). They both switch their language to English and their script to Roman.

- (24) عندي **announcement**  
*çendi announcement*  
 have-ISG.POSS announcement  
 'I have an announcement.'

- (25) 1- A: تعب  
*taçab*  
 'difficult'  
 2- B: والتذاكر  
*w ə-taðāker*  
 and DEF-tickets  
 'what about the tickets?'  
 3- A: لا ..... يسوون لنا  
*lā y-sawūn l-na*  
 NEG .....3PL-make to-IPL  
 'they will make.....for us'  
 4- A: **refund**

English borrowed verbs are integrated into Arabic only by transliterating them into the Arabic writing system. Some of these are preceded by the auxiliary verb (*Swwa*). In the following example, the whole phrase (copy and past) is integrated into the Arabic sentence.

- (26) وبعدين كوبي اند بست  
*w baçdain copy and past*  
 and then copy-2SG.F and paste-2SG.F  
 'and then copy and past'

- (27) سويي داوون لود للصور  
*sawwī download li-şwar*  
 make-2SG.F download for-photos  
 'download the photos'

(28) سوي سيرش على النماذج الجديدة  
swwī            **search** *çala ən-namāðdž l-dždīda*  
make-2SG.F    search for    DEF-forms    DEF-new  
'search for the new forms'

Angermeyer (2005: 495) found that the script choice relates to the distinction between borrowings and single-word code-switching when he investigated bilingual American Russian advertisements. He found that writers of these advertisements choose the Cyrillic script to mark an item as borrowed, or the Roman script to mark it as foreign. His findings led him to conclude that in “biscriptual, bilingual writing, authors use script choice to attribute a lexical item to a particular language”. In the present study, the script is also found to be manipulated in order to serve adaptation of borrowed items, giving them new forms that fit the Arabic orthographic norms.

In the case of biscriptual, bilingual Najdi IRC, participants are found to use two different orthographies: each one is tied to a different language and different visual forms in which words are presented. When IRC participants integrate English items within Arabic discourse, they may either alternate between the two writing systems (in the case of code-switching), or they may transliterate English items into Arabic script associated with the Arabic language (in the case of borrowings). In the former case, integrated English items are perceived as foreign. In the latter case, they are perceived as native as they fit the form of the recipient language (Arabic).

#### **7.4 The communicative functions of code-switching in Najdi IRC.**

Following the Conversational Analysis approach , the researcher provides an account of the occurrence of code-switching attested by Najdi IRC participant. Adopting the theory of sequential analysis in Conversational approach, she analyzes code-switching within the interactive written discourse of bilingual speech patterns in IRC. A close scrutiny of Najdi IRC corpus elucidates a number of cases where code-switching signal certain functions induced by conversational analysis of contexts, and whose explanation does not require the invoking of an extra-linguistic factor (Auer 1998, Li Wei 1998). The occurrence of these cases of code-switching is associated with certain communicative activities, which might be either discourse-related or participants-related. IRC participants perceived these communicative functions of code-switching as meaningful since they mark certain points of IRC conversation in a functional way. Below is a summary of the main communicative functions of code-switching in Najdi IRC. The results include both communicative functions of code-switching to English and communicative functions of code-switching to NA as well.

##### **7.4.1 Discourse-related code-switching**

Discourse-related code-switchings are those associated with organizing conversational tasks such as turn-taking, preference marking, repair, and bracketing of side-sequences. Discourse-related code-switching is fundamental to the checklist approach to code-switching such as Gumperz's (1982) six conversational functions of code-switching. The methodology of such an approach is to develop elaborate typologies of certain conversational loci in

which code-switching is notably recurrent. Following the same vein, Auer (1988) proposed ten conversational functions of code-switching on the basis of a discourse-related pattern as opposed to participant-related one. In an attempt to make a typology of code-switching loci in the present study, instances of discourse-related code-switching are identified and analyzed in relation to their surrounding contexts.

#### **7.4.1.1 Code-switching for qualification**

In some cases, speakers feel the need to clarify and qualify their messages in a course of interaction. A frequent practice of code-switching is introducing a message in language A and then switching to language B for the sake of qualification and clarification (Gumperz 1982). Such code-switching is often found in IRC discourse where the dominant language is English and the elaborations are given in Arabic although the IRC participants seem to be highly proficient in English. The following is an extract where female participants are talking about dieting in English and code switch to NA when they want to qualify their messages. Participant A uses the Arabic word “bil3aks” meaning (on the contrary) (line 6) to clarify her point that relying on exercise is good building muscles even if it does not appear on the scale. The same participant gives details in Arabic “elli 3nd riadh el,hijen” (who works in Riyadh ElHejen clinic) (line 9) describing the address of her doctor. Participant C switches to NA to clarify what is meant by Atkins “7ag el,proteen” meaning (that of proteins) (line 10). This indicates that NA is considered the easier code for communicating and understanding in these contexts.

- (29)1- A: how is your running plan  
 2- B: I lost 6 pounds and then stuck :  
 3- C: in kilos????  
 4- B: may be 3  
 5- B: i hardly lose weight. But my size went from 14 to 12)  
 6- A: **bil3aks**.. 3 is good since your running you are building muscle so it won't show on the scale but your clothes will fit better or get bigger. 'on the contrary'  
 7- A: my nutritionist told me that  
 8- B: who is your nutritionist?  
 9- A: DR RANDA **elli 3nd riadh el,hijen**  
 Dr.Randa that-REL at Riyadh AlHijin 'a name of a clinic'  
 'Dr. Randa who works at Riyadh AlHijin'  
 10- C: she doesn't believe in atkins **7ag<sup>16</sup> el,proteen**  
 that-REL of DEF-protein-  
 'that of the protein'  
 11-A: i want to try running **wish raeek ya B????**  
 what view-POSS.2SG.F you-B?  
 'what do you think B?''  
 12- B: my plan is the 5k >>>>**ashall** one  
 'the easiest one.'

The use of the Arabic discourse "bil3aks" in the above extract illustrates Maschler's (1994) claim that bilingual speakers use code-switching to metalanguage the frame of their discourse. In this case, Maschler maintains that discourse markers are 'metalingual expressions' that occur at verbal activity boundaries. This verbal activity, according to Maschler, is highlighted not only by using a discourse marker but also by switching to another language in uttering this marker. On this point Maschler writes:

"Discourse markers are often highlighted by a language switch: the discourse they frame takes place mostly in one language, while the framing itself take place mostly in another. Furthermore, the frame often consists of clusters of switched discourse markers at these boundaries".

---

<sup>16</sup> The use of the particle 7ag which stands for ḥag is an indication that IRC participants tend to transfer features of their spoken dialect. NA is different from MSA in showing possession only in the use of the possessive particle ḥag (See Section 4.5.1.14 and Section 4.7)

#### 7.4.1.2 Code-switching for specifying an addressee

Gumperz (1982) included the function of specifying addressee in his typology. This kind of code-switching is particularly important in IRC rooms where there are a group of IRC participants interacting in a conversation and it is only through writing that a participant can request the attention of a specific addressee. In such cases, code-switching to Arabic is used to designate an addressee for IRC message. While both English and Arabic serve purposes of clarification, code-switching to Arabic, being the most informal code, has the connotation of friendly specification of a particular addresses. In example (29), participant A code switches to Arabic in line (11) to specify B as the recipient of her message. Other examples are the following.

(30) A: I want to try running **wish raek ya B????**  
what view-POSS.2SG.F  
'what do you think B?'

Other examples of code-switching for specifying an addressee are the following:

(31) A: everybody is celebrating your party. **wish hl7arakat B???**  
what this DEF-actions  
'that is wonderful'

(32) A: I am very bad with names. **6ab3an B** you ar an exception |:D  
'of course.'

### 7.4.1.3 Code-switching of tags

A very common code-switching in Najdi IRC is tag-switching or what Gumperz identified as ‘sentence fillers’ (1982: 77). According to Gumperz, sentence fillers refer to lexical items or phrases, which are bound by minimal syntactic constraints and, therefore, can be inserted easily at different positions in a sentence of another language. A lot of tags or sentence fillers, which are frequently used in spoken NA, are found, also, in Najdi IRC. They include (*ya3ny* ‘maybe’, *akeed* ‘of course’, *musta7eel* ‘impossible’, *walla* ‘really’, *t9adeg* ‘do you believe’, *el7emdllah* ‘thank to God’, *allah ya3een* ‘may God help’, *mashaallah* ‘wonderful’, *allah yesahhel* ‘may God help’). Matras (1998) examines the motivations of bilingual speakers to switch these sentence fillers in language contact situations. He argues that bilingual speakers’ motivation for code-switching is cognitive rather than strategic or intentional. In other words, the ‘cognitive pressure’ of having two linguistic systems at the disposal of bilinguals, leads to the non-separation of the systems at different points of the discourse. Examples of IRC messages where these tags are inserted in different positions of a sentence are as below:

(33) A: **el7emdllah** my plan is working properly:P

*əl-ḥamd li-lla*  
DEF-thank for-God  
‘thanks to God.’

(34) A: **ya3ny** they will not talk about it again???

‘that means,’

(35) A: u will give up **ya3ny?**

‘is that what you mean?’

(36) A: y r typing so quickly **mashallah**  
'wonderful'

(37) A: **eee wallah** I read a lot about food and fashion :")  
'yes indeed.'

(38): see you tomorrow **inshalla**  
'hopefully'

(39) hope ur work s **tammam**  
'ok'

#### 7.4.1.4 Code-switching for reiteration

Gumperz (1982) identified reiteration as one possible function of code-switching. According to him, bilingual speakers employ code-switching to repeat a message in order to elaborate, clarify or emphasize the content of their messages. In the current data, code-switching is found to serve the function of reiteration. IRC participants were found to write something in one language and then repeat it in another language for the sake of providing clarification, emphasis or humour. In example (40), participant A switches to Arabic for the sake of reiteration. The same case of reiteration justifies A's switching to English in example (41). In example (42), the participant switches to English (new) repeating the same Arabic one "jaded" (new).

(40) A: let's be realistic **5aleena wag3ee**  
let-2SG.F realistic  
'let's be realistic'

(41) A: shreet eliphone b 3200 >>>marā ra5ee9 woo **cheap.**  
bought-1SG DEF-i phone 3200. very cheap and **cheap.**  
'my iphone costs me 3200. It is very cheap.'

(42) A: el computer jeded w **new**  
DEF-computer new and new.  
'the computer is new.'

#### 7.4.1.5 Code-switching for conversational coherence

Achieving conversational coherence via language alternation is another discourse-related code-switching found in Najdi IRC. According to Auer (1998: 7), repeated use of the same words or phrases in the same language is employed by bilingual speakers in order to maintain coherence in their conversation. Such a discourse-related code-switching improves the organization of a piece of discourse (Auer 1995: 125). Different speakers repeating certain words or phrases in the same language is provoked, also, by convergence with the other speaker. In their explanation of Accommodation Theory, Gile & Coupland (1991:35) maintained that ‘convergence’ is:

“A strategy whereby individuals adapt each other’s communicative behaviours in terms of a wide range of linguistic/prosodic/non vocal features including speech rate, paused phenomena and utterance length, phonological variants, smiling, gazes, and so on.”

In the following example, three speakers are discussing the price of some tickets and they were arguing whether or not the price was non-refundable. Both B and C repeat the English phrase (non-refundable), which A has used previously. In this case, code-switching serves as a provider of conversational coherence.

- (43) 1- A: ARSELHUM EMAIL??  
          send-ISG-3PL email  
          ‘I sent them as an email.’  
      2- B: la la  
          ‘no no’

- 3- B: 3la 6ul henea  
immediately here'
- 4- B: a7ssn  
'better.'
- 5- C: kam 9araw?  
how much become-3PL  
'how much are they'
- 6- A: 2000 woo **non-refundable**  
'2000 and **non-refundable.**'
- 7- B: ana leggetum b 2000 wa mesh **non-refundable** ;p;p  
I found-1SG-3PL in 2000 and NEG-**non-refundable**  
'I also found them by 2000 but refundable.'
- 8- C: ween :O  
'where'
- 9- B: al6ayyar  
Al-tayyar (a name of an office).
- 10-B: w galeek enn elta4aker mosh **non-refundable??**  
and told-3SG.M that DEF-tickets NEG **non-refundable**  
'and did he tell you that it is refundable?'
- 11- A: 3'areeba wallah  
strange indeed  
'it is really strange'
- 12-C: t2akeedi. ygooloon **refundable** bs e4a df3tee  
make sure-2SG.F say-3PL **refundable** but if pay-2SG.F  
3'ayaro kalamhum  
change-3PL view-POSS-3PL  
'try to make sure, they always say it is refundable but when you  
pay in full, they change their mind.'
- 13-A: ana df3t 5alas  
I paid-1SG already  
'I have paid and it is over now.'

#### 7.4.1.6 Code-switching for personalization vs. objectiveness

Another discourse-related code-switching, identified by Gumperz (1982), is the juxtaposition of two codes in order to indicate personalization vs. objectiveness. According to Gumperz, code-switching can serve as an indication of the degree of personal involvement, to take authority of speech or to simply differentiate between personal opinions and others. IRC participants are found to apply such a type of code-switching. They switch to NA when they interacted in English to be less formal and more personal. They code switch to English when

conversing in NA to distinguish from their speech items of fact or to provide more authority to what they said. The distinction between personalization and objectiveness is what Wardhaugh termed the “affective dimension” of code-switching (Wardhaugh 1998: 103). In his elaboration on metaphorical code-switching, he maintains that the affective dimension of code-switching is where “you change the code as you redefine the situation-formal to informal, official to personal, serious to humorous, and politeness to solidarity”.

In the following example, an IRC participant switches to English to provide more authority to his message and to establish formality to his announcement. A is writing in Arabic about the time of the final exam. He gives details about the assessment being based on an assignment which is due in Rajab (a name of an Arabic month). S/he switches to English (no excuse) both to establish formality in his/her announcement and to provide more authority.

(44) A: el,final ra7 ye9eer fee awal rajab. Allh yaster  
 DEF-final will be- 3SG.M in first Rajab. God help-3SG.  
 el, tasleem ba7th. **no excuses**  
 DEF-assessment research  
 ‘the final will be on the first of Rajab (a name of an Arabic month). May  
 Gold help us. The assessment will be by by research’

In the following example, IRC participants are conversing formally in English about applying for a job and A switches to Arabic twice (in line 7 and line 8) to be less formal and more personal.

(45) 1- A: last time we spoke, he told me he’s submitted his cv and now  
 waiting for the reply.  
 2- A: I’m still waiting.  
 3- B: thank u sweetie  
 4- A: a friend of mine told me that the committee s from the whole

- school and they take like a month or 2 to respond to an application
- 5- C: why don't u apply 4 (name) clinic?
- 6- A: i am wasting my time.
- 7- A: **ell wallha**. I read a lot about food and nutrition.  
'yes indeed.'
- 8-A: but for a stable career they do not work  
**5leena 3ala el, tdrees y she5a**  
let- 2SG.F on DEF-teaching oh dear.  
'it is better to work in teaching, my dear.'

#### 7.4.1.7 Code-switching for requesting attention

Language change is found to mark a request for attention in face-to-face conversation (Li Wei & Milroy 1995: 286-287). In IRC rooms, where a group of interlocutors are communicating at once, code-switching is necessary in some situations as a request for attention. Often, juxtaposition of two codes is used to attract the attention of IRC interlocutors and make them respond quickly to an utterance. In the following example, many IRC participants are conversing and C repeats the phrase (Anybody here to chat?) many times in the course of Arabic conversations to attract the others' attention.

- (46) 1- A: sallam ya shabab  
hello oh guys  
'hi guys.'
- 2- B: maraa7eb  
'hello'
- 3- A: waaaaw
- 4- A: wish ha4th B  
'what is this'
- 5- C: **Anybody here to chat?**
- 6- B: mmmm. shukran  
mmmmm. 'thanks'
- 7-D: wallah ma fee a7ad  
indeed NEG in anybody  
'in fact, there is nobody.'
- 8- C: **Anybody here to chat?**

In the following interesting example, A and B are conversing in NA about being awake at a late hour of the day. C greeted them in NA first (*mar7aba*) *marḥaba* “hello” but when they did reply to him, C switched to English (anybody here???? where are the people???????) (line 16) after ten utterances of conversing between A and B. They keep ignoring C and s/he keeps using English (brb) “be right back” in an attempt to attract some attention.

- (47) 1- A: hi nouf  
 2- A: kaif el,7aaal?  
 how DEF-state  
 ‘how are you?’  
 3- B: ؟؟؟؟سهرانة  
 awake-2SG.F  
 ‘are you awake?’  
 4- A: يب  
 yeb  
 ‘yes.’  
 5- C: مرحبا  
 marḥaba  
 ‘hi’  
 6- B: 6ala3tee faj2ah kent afaker feek @@@  
 appeared-2SG.F suddenly was-1SG think-ISG in-you  
 ‘you appeared here suddenly. I was thinking of you.’  
 7- B: le7alek????  
 alone-2SG.F  
 ‘alone?’  
 8- A: la ma3 3’900n  
 NEG with Gosoon.  
 ‘no. with Goson.’  
 9- B: sahraneeen :o  
 awake-2PL  
 ‘are you both awake?’  
 10- A: laaa nmmt b4 al,f6wr w 97et  
 NEG slept-ISG before DEF-breakfast and woke up-ISG  
 al, fajer  
 DEF.noon.  
 ‘no, I slept after breakfast and I woke up in the morning.’  
 11- B: yuuuu  
 ‘oh.’  
 12- B: mooo 7looo ke4a  
 NEG good that  
 ‘this is not good.’  
 13- B: tjw3een

- hungry-2SG.F  
 'you will be hungary.'
- 14- A: ee mraaa :(  
 'yes indeed'
- 15- A: bs b7awl anam b3d shoi 2day  
 but will-ISG-try ISG-sleep after Shoge today  
 'but I am trying to sleep after Shoogi.'
- 16- C: **any body here???? where r people???????**
- 17- B: >>>A شفتي زواراة خميس  
*šefī zwarat xamīs*  
 saw-2SG.F Zwarat Xamees  
 'have you seen zwarat khamees?'
- 18- C:**brb**  
 'be right back'
- 19- A:eeee tjnnn @@@@ @@@@  
 'yes amazing'

#### 7.4.1.8 Code-switching for quotation

Among the six functions of code-switching identified by Gumperz (1982) is that of quotation. This function of code-switching, according to Gumperz (1982:76) is achieved when a speaker conveys a message in, either as a direct quotation or as a reported speech, another language. This occurs because "quotations are more effective if cited in the language originally used" (David 2003: 15).

This function is only achieved by code-switching to English in the IRC data. The following example describes this function. In the following IRC extract, two IRC participants are discussing fixing A's PC, which took a lot of time. A switches to English when s/he quotes the shopkeeper's response who seems to be a foreigner who can not speak Arabic (line 8) and, then, both A and B converse in Arabic. It is clear that A changes to English to make his description of the incident authentic and accurate and to mark the degree of his dissatisfaction with the poor service. A direct quote of the shopkeeper's response to A makes the conversation funny and colourful.

- (48) 1- A: salaam  
       'hi'
- 2- B: hala wallha  
       'hello'
- 3- B: wa 3aliakom el,salaaam  
       and upon-2PL DEF-peace  
       'hi'
- 4- C: ween el,nas :P:P:P  
       where DEF-people  
       'where are the people?'
- 5- B: A 3ndk el,mac?  
       have-2SG.M DEF-mac  
       'do you have Mac?'
- 6- A: shway  
       'little'
- 7- B: eish shway ;s  
       what- little  
       'what do you mean by (little)?'
- 8- A: galli el,ma7al **we will contact you soon.**  
       said-3SG.M DEF-store **we will contact you soon.**  
       'the shop keeper told me that he is going to contact me soon.'
- 9- B: kam 9ar lah?  
       when become-3SG.M  
       'how long has it been there?'
- 10- A: mmm ... esbo3en  
       mmm... 'two weeks'
- 11- A: eeee yt33333333bb  
       yes 3SG.M.be tired  
       'yes. It really takes a lot of time.'

#### 7.4.1.9 Code-switching institutional terms

Another function of code-switching is that of fulfilment of linguistic needs for a variety of purposes. In IRC contexts, where participants are exchanging technological knowledge by exchanging files, images and links, these linguistic needs are always of academic and technical nature. Code-switching in these cases is always to English. This can be attributed to the fact that technical learning in Saudi Arabia is conducted in English. Therefore, the English-based atmosphere of technological knowledge in Saudi Arabia urges IRC participants to use

English terms whenever referring to technical issues. This category serves, as Appel and Muysken (1987) put it, a referential function.

An illuminating example is the one below; B and C are explaining to A how to download a software programme to his blackberry phone. B switches to English in Roman script signalling foreign items.

- (49) 1- A: ما عرفت انزله  
*ma ʧaraft 'a-nazl-a*  
 NEG ISG-knew ISG-download-3SG.M  
 'I did not know how to download.'
- 2- B: **link** شفت  
*šif-t link*  
 saw- 2SG.M link  
 'did you see the link?'
- 3- A: اي  
 ʔ  
 'yes.'
- 4- C: **run** اضغط **save**  
*eḏghaṭ run mū save*  
 click-2SG.M 'run' NEG 'save'  
 'click (run) not (save)'
- 5- C: او كي  
*okai*  
 'ok.'
- 6- B: **next next** وبعدين بس  
*wa baṣdayn bas next next*  
 and then but (next) (next)  
 'and then, just click (next) (next).'

There are Arabic equivalents for words like “link” “run” “save” and “next” . Normally, these are considered to be very formal and IRC participants try to keep themselves more personal and informal. This example is, also, interesting with regard to script switching. To be exact, the switching is also in script signalling the attribution of these items to English. Whilst the IRC participants typed in Arabic, changing to English script requires the keyboard’s word processing

systems to be shifted. C and B both took the trouble in shifting script while typing to describe accurately how to download the software and to keep the mode of informality in IRC episode of discussion.

This particular case of code-switching, according to Angermeyer (2005), is an attempt to mark these items (link, next, save and run) as English. In his elaboration of the differences between lexical borrowing and code-switching, he suggests that lexical borrowing implies that a foreign word is written in the host script whereas code-switching occurs when the lexical item retains its native script as in the previous examples (See Section 7.3). Another example is the following IRC extract where two IRC participants discuss how to download some photos. The discussion is written in NA in Arabic script. B switches to English in line (2) when he wants to refer to a technological term ‘photo filter’. Translating this term into Arabic requires effort that would not allow the participant to continue negotiating the issue of downloading.

- (50) 1- A: نزلي الصور  
*nazz-lī* *əṣ-ṣwar*  
 download-2SG.F DEF-photos  
 ‘download the photos!’  
 2- B: ما عندي **photo filter**  
*mā ʕənd-ī* *photo filter*  
 NEG have-POSS-ISG photo filter  
 ‘I don’t have (photo filter )’

#### 7.4.1.10 Code-switching for signalling topic shift

As in face-to-face conversations, structure and topic are interrelated throughout IRC conversations. Brown & Yule (1983) indicate that “... there should be a point at which the shift from one topic to the next is marked”. In a

face-to-face conversation, these points of shifting are identified generally by an utterance-initial particle (e.g. *ah*), lexical indications (e.g. in fact, but, and) or intonation changes and body movement. In IRC, where visual and auditory channels of communicating are unavailable, code-switching, effectively, serves the function of changing topic.

Many studies of code-switching have proven the association between code-switching and changing topic in conversation (Gumperz 1982, Auer 1984, Myers-Scotton 1993, Alfonzetti 1998). The IRC data in the present study also demonstrates the coincidence between code-switching and introducing a new topic. In the following example, three IRC participants are negotiating in NA the price of a camera in both Arabic and Romanized scripts (line 1 to line 10). C seems to be disinterested in the discussion and he changes to English (Assignments were due on Saturday.) (line 10) both to change abruptly the topic of discussion and to give authority to his utterance.

- (51) 1- A: الكاميرا على حسب مواصفاتها  
*il-kamera çala hasab mwaşafat-ha*  
 DEF-camera on depend qualities-POSS-3SG.F  
 ‘It depends on the qualities of the camera.’
- 2- A: الايبود ما يصور؟؟  
*il-aibod ma-yi-şwer?*  
 DEF-ipod NEG-depict-3SG.M  
 ‘does the Ipod include a camera?’
- 3- B: la ana aby alcamera  
 no I want-1SG DEF-camera  
 ‘no. I want the camera.’
- 4- B: la shreeta bdoon elcamera  
 no bought-1SG without DEF-camera  
 ‘no, I bought it without the camera’
- 5- B: henna b3000.  
 here in-3000  
 ‘It is 3000SR, here.’
- 6- B: 8lt ymkn 3ndk ar59  
 guess-1SG maybe have-POSS-2SG cheaper

- ‘I guess it might be cheaper in your town.’
- 7- C: akeed ar59  
‘sure, cheaper.’
- 8- A: عطيني الموديل  
*ṣaṭīn-ī il-modail*  
give-2SG.F DEF-model  
‘tell me the model that you want.’
- 9- B: canon alkberah  
Canon DEF-big  
‘Cannon, the big one.’
- 10-C: **assignments are due on Saturday.**

#### 7.4.1.11 Code-switching for self-repair

The literature documented code-switching as serving a self-repair mechanism (Auer 1998, Li Wei & Milroy 1995, Milroy & Li Wei 1995, Alfonzetti 1998, Matras 2009). In face-to-face conversation, it is often used with other repair techniques such as hesitation pauses, vowel lengthening and other paralinguistic features such as body movement, nods and gestures. In his study of Italian-dialect code-switching, Alfonzetti (1998) indicates that code-switching could be used as a self-repair technique to solve problems of momentary lack of memory or those arising from making a mistake or facing a discourse problem. In such cases of self-repair situations, speakers frequently repeat their words in another code.

In IRC, code-switching is found, also, as a self-repair technique. In examples (52) and (53), the participants switch to Arabic after a pause to overcome their hesitation and try to reformulate the idea they want to communicate by making a transition to Arabic. They both insert repeated dots as an indication of a long pause. The participant in example (52) tries to repair his situation by providing justification for his contrary action (screaming). It seems that he is embarrassed by the act of screaming and tries to indicate (in Arabic) that he is just kidding. The



“(1) that, when alternative actions are open possibilities, one may be “preferred”, that is expected and chosen, if possible; and (2) that the difference between “preferred” and “dispreferred” alternatives is demonstrated in the turn shape chosen for doing one or the other.”

(Have 2007: 137)

Taking into account the Accommodation Theory, the IRC participants are generally found to be accommodating the language of their interlocutors. If the normal and most acceptable communicative behaviour is that of accommodation, then language alternation should be pragmatically significant. The pragmatic functions of juxtaposition of two codes might be one of the communicative functions listed above or signalling objection, disapproval and contrariness. In the following example, the three female participants are talking in English about food and dieting. C suggests trying a fad diet by continuing to eat proteins. Participant A disagrees with her suggestion (line 13). Her disagreement is marked both structurally (by using repeated dots indicating a long pause and preceding her utterance) and verbally (by switching to NA). Both of them go to English after A’s disagreement.

- (54) 1- A: my plan s working but im ruining it with my own hands.  
2- A: i lost 8 pounds up to now  
3- B: BTW winter clothes make u look fuller  
4- B: that’s why I hate wearing jackets  
5- B: and coats  
6- A: have u tried Paul Makenna????  
7- C: what’ s that??  
8- A: it’ s simply not a diet but a system to follow.  
9- C: i gain weight very fast.  
10- A: me 2222 :D:D.  
11-A: read about it....really amazing  
12- C: i m on atkins now. And I’m losing weight.

- 13- A:.....ya **she5a elli yroo7 bsr3a**  
 oh dear that-REL goes-3SG.M quickly  
**yrja3 bsr3a**  
 come-3SG.M quickly  
 ‘come on. the quicker you lose weight, the quicker you put it back  
 again.’
- 14- A: do u know what happened to atkins? the guy who invented atkins  
 diet?
- 15- B: died?
- 16- A: yup..... of high cholestrol

## 7.4.2 Participant-related code-switching

This section displays the participant-related code-switching in Najdi IRC. As has been indicated earlier, bilingual speakers produce participant-related code-switching in order to indicate their preference or competence (Auer 1984). In other words, speakers code switch to accommodate the language preferences of the participants in the conversation. In other cases, speakers want, simply, to avoid the language in which they feel uncomfortable and speak the one in which they have greater competence (Auer 1984: 125). Therefore, participant-related code-switching can be categorized into two main categories, which will be adopted in the following analysis. Preference-related code-switching which assists speakers to adapt to their own and interlocutors’ individual preferences and competence-related code-switching that helps speakers to adapt to language competence.

### 7.4.2.1 Preference-related code-switching

Technological constraints of word processing constitute a fundamental factor in setting a specific language as a preference in IRC. In some IRC rooms, high English proficient participants made considerable efforts to establish English as a preferred language of communication. In the following cases, a moderator gives

explicit instructions to communicate only in English. All IRC participants in the following extract do not have an Arabic keyboard and, therefore, prefer to type in English. They all seem to be highly proficient in English. When one of them (who seems to have Arabic keyboard support software) tried to write in Arabic, his writing is transformed badly on the screen. Consequently, an explicit instruction is given to write only in English (please write in English. Not this strange language. Or you can send me a key to help me read your messages) (line 4). The following utterance by participant D is, also, a request in Romanized Arabic not to type in Arabic as s/he is abroad where Arabic windows is not easily accessible.

(55) 1- A: it was easier for me to write in English.

2- A: all of us are fluent in English.

3- B: لا تتخيل معاناتي في ادراج الحروف العربية على الكمبيوتر

*lā t-txayal muṣanat-ī fī edrādž l-ḥarūf*

NEG 2SGM-imagin problem-POSS-ISG in inserting DEF-letters

*əl-ṣarabia ṣala əl-kimbuter*

DEF-Arabic on DEF-computer

‘you cannot imagine how difficult it is to insert the Arabic alphabets into my keyboard.’

4- C: **plz write in English. Not this strange language. Or you can send me a key to help me read your messages**

5- D: momken 6alab>> ya rait ma a7ad y akteb

possible demand>>Oh please NEG-anybody 3SG.M-write in

bl3arbi 3ashan ana barra wa ma fee pc 3ala windows 3arabii

DEF-Arabic because I abroad and NEG-in computer on windows Arabic.

‘can I ask you something? Would you please not write in Arabic as I

am abroad and do not have a computer with Arabic windows.’

#### 7.4.2.2 Competence-related code-switching

As the IRC participants differ in terms of their proficiency in English, they code switch in order to adjust their languages depending on their proficiencies (Auer, 1995). Participants, who are poorly proficient in English, often code switch to Arabic

when they struggle to understand and follow the conversation. Participants, who are highly proficient in English, also, code switch to Arabic in order to accommodate their communication depending on their interlocutors' proficiency. In the following extract, A and B are negotiating in good English how to activate a programme, which enables its users to make free calls. C, whose English is not good enough as seen in his first two utterances (line 3 & 6), tries to understand the installation process and ensure that it really offers free calls. Participant A code switches to Arabic (line 8 ,9 & 13) in order to clarify the point to C and make him understand.

- (56) 1-A: it's really amazing. Free international calls  
 2- A: i am going to send you all invitations as soon as downloaded.  
 3- C: explaining every detail pleeeez.  
 4- B: i' ve downloaded but, how can I get started?  
 5- A: check your contacts, you can see which of your iPhone contacts is already on Viber  
 6- C: can I put my phone on air plane mode??  
 7- B: yes, sure >>>C..Calling is via 3G or Wi-Fi.  
 8- A: **ya3nii 3ala elshabaka**  
 mean- 3SG on DEF-internet  
 'this means that calls are via internet or wi-fi.'  
 9- A: **msh 3ala el.telephon**  
 NEG on DEF-phone  
 'not the phone itself.'  
 10- B: does it work with blackberry?  
 11- A: not yet  
 12- A: A version for blackberry devices will be released soon.  
 13- A: **ra7 y6l3oon li blackberry 8reeb**  
 will 3PL-release for blackberry soon  
 'they will release a version for blackberry soon.'

## 7.5 The communicative functions of script-switching in IRC

The fact that Romanized Arabic in CMC has changed from being a response to technical limitation into a code of the medium in which it is used prompts investigation into the communicative functions associated with this code. An

investigation of IRC data reveals a considerable switching of scripts, both between and within sessions of conversations. That is, some participants switch script, alternating between the right-to-left Arabic script and the Romanized left-to-right script. This mixture of scripts is also evident in the results of the “IRC Language Questionnaire”; 76 (36%) of the participants acknowledged that they use Arabic scripts, 68 (32.2%) said that they use English scripts (Roman letters and numerals), whilst the remaining 67 (31.8%) indicated that they use both English and Arabic script online. Assuming that the script functions as a code of communication that carries social and cultural meanings, the researcher of the present study argues that switching between scripts is not entirely arbitrary. Rather, it is motivated by creative structuring of the discourse. In this section, the phenomenon of script-switching in this particular medium of communication is investigated with a focus on the case of NA online.

The emergent literacy practices in the CMC domain present new and innovative phenomena for investigation. The novelty of the script-switching phenomenon is reflected in the lack of any systematic definition of what script-switching is. Consequently, a systematic scheme of interpretation within which the occurrence of script-switching in a biscriptual CMC conversation can be identified and interpreted is still lacking. Androutsopoulos’s term (2004), computer-mediated digraphia, referring to the simultaneous use of two scripts representing one language in computer-mediated interactions, offers a practical framework to investigate script-switching online. Along this line of thinking, the notion of script-switching in the present thesis will be defined as the juxtaposition within the same written exchange of interaction belonging to different writing systems.

In classifying kinds of script-switching, the present study will follow Bloom and Gumperz's (1972: 422-423) classification of code-switching into situational non-conversational code-switching and metaphorical conversational code-switching. Situational script-switching is associated with situational factors such as a change in the situation, interlocutors, content, and topic. Most obvious of all in the CMC setting is the lack of software support. The second kind of script-switching is metaphorical or conversational switching, where writers switch their script though there is no change in the external situational factors. The rationale of such switching is to achieve particular communicative functions of discourse.

Conversational script-switching plays a similar role achieved by conversational code-switching in helping the writers to re-define and enhance the sequentially of discourse. Writers are motivated to initiate switches in their written discourse by stylistic effects and creative structuring of the discourse. In this respect, script-switching comprises functional and meaningful tools that help modifying the discourse. Switching in biscriptual written conversation, then, fit Gumperz's early (1972) description of code-switching as contextualizing cues. The functionality and meaning of conversational script-switching is achieved through the contrasting process between two visually and structurally different writing systems.

Following Auer's (1984) sociolinguistic approach to code-switching, the researcher interprets the meaningfulness of script-switching at the local level of IRC exchange. The local functionality of script-switching will be interpreted at the level of immediate interaction between writers rather than relating certain choices of script to its overall societal functions. These societal functions are, in fact, associations between the language of a particular script and social functions among the speakers of this language. Among these societal functions is the association between English

(Roman script) and technological terms, and the association between Arabic (Arabic script) and religious expressions.

The present study investigates the functionality of script switching between Roman script and Arabic script as a strategic mechanism “to enrich communicative potential” (Woolard 2004: 75). The juxtaposition of the two scripts is examined and evaluated in relation to its surrounding contexts for the sake of uncover local functionality of script-switching from two main dimensions: discourse-related and participant-related . Following Auer’s (1984) approach to code-switching, the occurrences of script-switching are oriented towards particular communicative effects developed through its connection with both discourse and participants in the IRC context. Those that are related to IRC discourse include the highlighting effects of reiteration, topic shifting, seeking emphasis, or change of mode (e.g., from informal to formal). Participants-related functions of script-switching include specifying of an addressee, participants’ incompetence, or inclusion and exclusion of other participants. Below is a description of these functions with examples from IRC data.

### **7.5.1 Discourse-related functions of script switching**

#### **7.5.1.1 Script-switching for signalling change of mode**

IRC participants’ familiarity with the use of both Arabic and Roman scripts in particular contexts directly affects their choice of script. The results showed that 32 participants (18.9%) indicated that the use of the Romanized form of NA is justified by the difficulty of writing NA using Arabic script. In assessing their familiarity with

the Arabic script<sup>17</sup>, 34 participants (16.1%) agreed and 60 participants (28.4%) strongly agreed that they have difficulty in reading and writing NA. This is related to the earlier discussion about MSA being reinforced within literacy education, in which students are often corrected when they use features of their dialect in their writing. IRC participants are not accustomed to reading or writing in their regional dialect, as they are exposed to MSA in almost all reading and writing contexts. They listen to and speak their native dialect, but read and write MSA. This issue of familiarity is clearly reflected when participants quote some formal sayings, poems, or religious quotations.

The following two extracts demonstrate the participants' perception of the Arabic script as the appropriate tool to write MSA but not their NA. In Ex (57), participant A is quoting a religious praying where s/he prays for dead Muslims to be peaceful and contented in their graves, and then s/he thanks God for the Eid celebration. The quote is written in the Arabic script. The participant immediately turns to Romanized NA in line 5 to resume conversation with friends online “*3eedkum mubarak w kil 3am w entm b5eer*”, to which B also replies using Romanized NA: “*bs shaklik 5las bad2ti...*” (line 6).

- (57) 1- A: جانت ليلة العيد....  
*džā't lailat l-çīd*  
 came -3SG. night DEF-Eid  
 'it is Eid's night.'  
 2... وتحت التراب أناس كانوا معنا العيد الماضي...  
*wa taħt l-turāb 'unās kānu maçana al-īd al-maḍī*  
 and under DEF-sand people were-3PL with-1PL DEF-Eid DEF-past  
 'there are people who were with us last Eid are now dead.,  
 3 - فأسألك رباه كما بشرتنا بقدم العيد ولك الحمد -  
*fa 'a-s'alu-ka rabā-h kamā bashart-anā bi-qdūm*  
 so, ISG-ask- 2SG.M God as enjoied-2SG.M-1PL with-coming

<sup>17</sup> The Arabic script is course familiar to the participants, but they are not used to writing their non-standards dialect with it.

*al-ḥād w lak al-ḥamd*  
 DEF-Eid and have-POSS-2SG.M DEF-thanks  
 ‘so I ask you God as you enjoyed us with the Eid..’

4- أن تسعد أموات المسلمين في قبورهم..

*‘an tu-sḥid ‘amwāt al-muslimīn fī qubūr-him*  
 that 2SG-enjoy deads DEF-Muslims in graves-POSS-3PL  
 ‘that you relief Muslim dead people in their graves.’

5- A: **3eedkum mubarak w kil 3am w entm b5eer**  
 Eid-POSS-2PL happy and every year and you-2PL in-good.  
 ‘happy Eid and may everyone be good every year.’

6-B:bs shaklik 5las bad2ti...  
 but seem-2SG.F already begun-2SG.F  
 ‘but it seems that you have already begun.’

Another participant in Ex (58) quotes some lines from a very well known writer.

The participant writes these lines in the Arabic script (line 1) and then turns immediately to the Roman script (line 2), where s/he writes in English “*Hope you all had a wonderful Eid celebration today ... I did :-D*”

(58) A 1- اقطف الفرحة ولو كانت صغيرة، وابتسم ولو من دون سبب، وتحدث إلى من حولك واسألهم عن أحوالهم

*eqṭef il-farḥa w law kān-at ṣaḡīra, w*  
 pick-2SG.M DEF-happiness even-if it was-3SM.F small and  
*əbtasem w-law min-dūn sabab, w taḥadaθ ela man*  
 smile- 2SG.M even without reason, and speak-2SG.M to those  
*ḥawl-ak w əs‘alhum ṣan ‘aḥale-him.*  
 around you and ask-2SG.M-3PL about states-POSS-3PL.

‘pick up any chance of happiness, smile even without reason, speak to people around you and ask them about their states.’

Dr.SalmanAlouda.

2-A **hope you all had a wonderful eid celebration today .. I did :-D**

Both extracts reflect participants’ identification of the Arabic script to be related to MSA, which is clearly separated from the IRC communication that is conducted in either NA or English. Those participants are not hindered by any technical constrains, as they clearly have an Arabic keyboard. They choose Arabic script to write their poetic quotations and prayers, which would be less attractive if written in Roman letters and numerals. This is an indication of participants’ metalinguistic awareness of the type of the language constitutes the code of IRC communication.

Such awareness prescribes that formal quotations and poems are an entirely written genre, which should be written in the Arabic script, whereas their IRC communication constitutes a hybrid genre between speech and writing, and should be written in NA.

Apart from the association between the Arabic script and MSA, occurrences of script-switching where the main intention is to change the mode of conversation from informal to formal are found in the IRC data. In the following extract, both A and B are conversing in Romanized NA at the beginning of the conversation. B suddenly switches to the Arabic script (line 9) asking A how she is doing at her work. A responds to B's question in the Arabic script, indicating that her work is really a great responsibility (line 10). They both continue using the Arabic script, discussing seriously how to cope with this work, which, apparently, is a new job for A. B suggests how to find new forms and how to use them. However, A indicates that she manages to make her own new forms. This work conversation is conceived as being serious by both writers and readers of this IRC exchange. This mode of seriousness is conveyed through the abrupt switch to Arabic script, which contrasts with the humorous and funny conversation at the beginning of the exchange.

- (59) 1- A: mar7aba :}  
hello  
2- A: kaif el7aal????  
how DEF-state  
'how are you?'  
3- B: ahlain :D  
'hello'  
4- A: BUZY :S  
5- C: ):):) سلامات والله  
*salamāt waḥḥa*  
wishes God  
'I hope you are fine.'  
6- C: :!!!! من زمان ما شفناك  
*min zamān mā shift-ik*

- from time NEG saw-ISG-2SG.F  
 'I have not seen you for a long time.'
- 7- A: 7ata ana, missing u ;-)
   
too I
   
'me too'
- 8- C: :D
- 9- B: كيف الدوام؟
   
*kaif əd-dawām*
  
how DEF-work
   
'how is your work?'
- 10 A: والله مسؤوليه
   
*walla mas'ōlaya*
  
and really responsibility
   
'it is a great responsibility.'
- 11- B: ما سويتى سيرتش على النماذج الجديدة؟؟؟؟
   
*mā sawīt-ī sairš çala ən-namaðedž dž-džedīda*
  
NEG did-2SG.F search on DEF-forms DE-new
   
'did you search for the new forms?'
- 12- B: وبعدين خلاص كوبي اند باست:
   
*w baçdain xalaş kubī and baist*
  
and then over copy- 2SG.F and past-2SG.F
   
'and then it is just copy and past.'
- 13- A: أنا غيرت النماذج كلها
   
*'anā ġair-t ən-namaðedž killaha*
  
I changed-ISG DEF-forms all-3PL.
   
'I changed all the forms.'

### 7.5.1.2 Script-switching for reiteration

Another communicative function of script-switching identified in the IRC data is script-switching for reiteration. That is, an IRC participant switches his script to repeat a point that has just been made in the other script. The occurrence of script-switching is accomplished by reiterating the same literal meaning in a different script. In this case, the reiteration facilitates the occurrence of script-switching, which fosters a smooth transition to another code. The transition from one code to another is sometimes multi-layered on the level of language as well as script.

Consider the following IRC extract, which illustrates the complexity and multi-layered nature of these kinds of switches. In the following IRC extract, three IRC participants are conversing in the Roman script both in English (lines 1, 3, 4, 5, 7, 8,

9, 10, 11, and 16) and Arabic (lines 2, 6, 12, 14 and 15) with insertional code-switching to Arabic in some English turns. B and C are welcoming A back and discussing B's proposal for a job. The conversation turns to kids issues (how is your kid? how are the girls (line 6, 7& 8). In line (10), A asks B if s/he is ready or not to meet the director using the Roman script. B replies with multiple questions (line 11), which either indicates her uncertainty about being ready for the meeting or her inability to understand A's question. A switches to the Arabic script in line (12) to reiterate and repeat the same question stated earlier in line (10) in the Arabic script. A's question is articulated twice for the sake of elaborating and clarifying her point. Line (12) acts as a literal translation of line (10), which evokes B's identification with the question who, in line (13), indicates her uncertainty and hesitation. In assuring B, A switches to English again in line (14).

(60)

- 1- C: r u going to meet el director?
- 2- A: i honestly do not know who s.... I will go 2morrow check.
- 3- B: allah y3een. Dr khaled replied my email saying they r still  
God help-3SG.M  
'may God help...'  
working on deciding a new director.
- 4- C: ahlaan :o :o :o :o  
ohhh
- 5- B: I really don't know :
- 6- A: >>>> C how s ur son?
- 7- C: 7emd ellah...he is doing great; he turned 1 two days ago.  
thank God  
'thanks to God'
- 8- C: how are the girls?
- 9- A: walla b5eer they went back to school  
really fine  
'they are really fine.'
- 10- A: esta3adeeeti ??  
prepared-2SG.F?  
'are you ready?'
- 11- B: ?????????
- 12- A: استعديتي؟  
*estaçadait-ī*

prepared-2SG.F  
'are you ready?'  
13- B: marrrra :OO  
'not really'  
14- A: do not worry.. el interview just to know you.

### 7.5.1.3 Script-switching for topic shifting

The occurrence of script-switching is also found to be associated with a change of topic in which transition from topic A to topic B takes place in IRC exchanges. The functionality of script-switching in this case is achieved by contrasting two different scripts. The following IRC extract is an example of an incident of script-switching which serves double functions. The central theme of the following IRC extract is the discussion of losing weight. Both A and B are discussing how Goson (a name of a girl) has lost weight. All the conversation is composed in Romanized NA. A expresses her intention of losing weight and adopting the strategy of walking 5Ks a day. In line (22), B asks A if s/he finds that useful or not and if s/he really lost weight. A replies affirmatively. However, B is eager to know exactly how much weight A has lost (line 24). A expresses reluctance in giving her exact weight by the use of a vocalizing strategy (mmmmm) in line (25). In line (26), B uses the same vocalization strategy to insist on knowing A's weight (laaa 9dddddggg kmmm). At this point in the conversation, A switches to the Arabic script, asking about the exact start of the second term (line 27). Apparently, A's attempt at shifting the theme of conversation is successful as B immediately stops asking about A's weight and turns to answer A's question about the second term.

In such incidents, script-switching achieves double functions. It is deliberately used for both shifting the topic and expressing A's disapproval/disinclination. This

disapproval is expressed both structurally by the use of the vocalization (mhhh) in the preceding turn and visually by shifting to the Arabic script.

- (61) 1- A: sha5barek?????  
 what news-POSS.2SG.F?  
 'how are you'
- 2- B: ahlaaaan :D  
 'hi'
- 3- B: 7mdllah b5eer  
 thanks to-God good  
 'thanks to God. I am fine.'
- 4- B: w ento sh5barkum ?  
 and you what news-POSS.2PL?  
 'and how are you?'
- 5 -A: b5eer  
 'good'
- 6- A: shloon el,derasah  
 what-colour DEF-study?  
 'how is your study?'
- 7- B: mashy 7alha  
 fine state-POSS.3SG.F.  
 'it is ok.'
- 8- A: wesh a5baar 3'9oon?????????  
 what news-POSS.3SG.F Goson?  
 'how is Goson?'
- 9- A: ma legatee 9oor ba3ad?  
 NEG found-2SG.F photos else?  
 'did you find more photos?'
- 10- B: 7mdllah b5eer  
 thanks-God good  
 'thanks God. She is fine.'
- 11-B: hahahahahahaha laaa 5laa9 b77 519u  
 hahahahaha no finished-3PL finished-3PL  
 'no.they finished.'
- 12-A: km 9arat 3'9oon?  
 how much become-3SG.F Goson?  
 'what is Goson's weight?'
- 13 - B: wsho km 9art ????  
 what how much become-3SG.F?  
 'what do you mean?'
- 14- A: waznha?  
 weight-POSS-3SG.F?  
 'her weight'
- 15- B: 55 ;p
- 16- A: wallah ??????????  
 'really?'

- 17- B: eee  
'yes'
- 18- A: waaaaaw :P:P
- 19- A: wish ha4a?  
what this?  
'this is amzing'
- 20- A: ana 9ert amshee kil yom 5 kilo  
I become ISG ISG-walk every day 5 kilo  
'I walk 5 kilos everyday.'
- 21- A: mashalllaaaa  
'good'
- 22- B: >>>A: w ma n7ftyy !  
>>> and NEG loaset-2SG.F weight  
'and did you lose weight?'
- 23- A: eylla  
'yes.'
- 24- B: km 9ar wznkk ;p  
how much become-2SG.F weight-POSS.2SG.F?  
'and what is you weight now?'
- 25- A: mmmmmm
- 26- B: laaa 9dddddggg kmmm ;p  
NEG really how much?  
'really, what is your weight?'
- 27-A: متى يبدأ الترم الثاني؟؟  
*metā yi-bdā at-term aθ-θānī*  
when start-3SG.M DEF-term DEF-second?  
'when will the second term start?'
- 28-B: yemkin fee shawal.  
'maybe in Shawal'

#### 7.5.1.4 Script-switching for requesting attention

In IRC rooms, where a group of interlocutors are communicating at once, the need to request attention is more frequent than face-to-face communication. The need to request attention from interlocutors is manifested by the lack of visual and auditory channels of communication. However, IRC participants manipulate both code-switching and script-switching for seeking attention. The juxtaposition of two scripts is used to attract the attention of IRC interlocutors and make them respond quickly to an utterance. In the following example, many IRC participants are conversing in Arabic about how to fill a certain form in university. The beginning of the IRC exchange is composed in the Arabic script, where A explains to both B and

C the importance of filling this form which is available on the university's website. B and C are asking about details (e.g., where to find it, what is the deadline, how to post it to friends...etc.). B asks A if it is possible to send it by post (line 3) but A doesn't reply to B's question and carries on giving instruction. Being very concerned about informing other friends to fill the form, B switches to the Roman script in an attempt to attract A' attention to his/her concern (the problem is nobody checks his emails these days) in line (7). A responds immediately to B's point, suggesting that it is possible to send it as a link. A's response to B's request for attention is in the Roman script, indicating that s/he is specifying B in this particular instruction.

- (62) 1- A: الاستماره الكترونية.. في موقع الجامعة  
*al-istemāra al-ktronia. fī mawqeṣ al-džāmṣa*  
 DEF-form electronic. In site-CONS DEF-university  
 'the form is electronic. It is in the university's site.'
- 2- B: يصير ارسله ايميل  
*yi-ṣīr 'a-rsela 'email*  
 3SG.M-can ISG-send-3SG.M email  
 'is it possible that I send it as an email?'
- 3- B: يعني ما نرسله بريد؟؟  
*yaṣnī ma n-arsela barīd*  
 mean-3MSG NEG send-IPL-3SG.M mail  
 'so can we send it as an email'
- 4- C: متى آخر موعد؟  
*metā axir mawṣid?*  
 when last appointment?  
 'what is the deadline?'
- 5- A: على الموقع وخلص يعبونها  
*ṣala l-mawqeṣ w xalāṣ yi-ṣabūn-ha*  
 on DEF-site and just 3PL-fill-3SG.F  
 'and they can fill it on the site.'
- 6- A: وارسلوه لكل اللي عندكم في اللسته  
*w arsl-ū l-kil illī ṣind-kum fī l-lista*  
 and send-2PL-3SG.M to-all those in-2PL in DEF-list  
 'and send it to all people in your list.'
- 7- B: **elmoshklah m7d yfta7 emilah**  
 DEF-problem nobody 3SG.M.open email-POSS-3MSG  
**bs ba7awel. killa 3ala el bībī**  
 but 1SG-try. All on DEF- blackberry.  
 'the problem is that nobody checks his emails..but I will try..all are

- using blackberry.'
- 8- A: **illa y9er arselah link.. w yfta7**  
 NEG 3SG.M-can ISG-send-3SG.M link. and 3SG.M-open  
**fi el bibi**  
 in DEF-blackberry  
 'on no. you can send it as a link and it will be opened in the  
 blackberry.'
- 9- B: ahh ok ok fahmt fahamt  
 ok ok understand-ISG understand-ISG  
 'ok.I understand.'

## 7.5.2 Participant-related functions of script switching

### 7.5.2.1 Script-switching for inclusion and exclusion of IRC participants

IRC communication is not limited to the exchange of information, but also aims at the shaping of interpersonal relationships among participants. A certain extent of negotiation is devoted to establish, maintain, and expand personal relationships among participants. In face-to face communication, speakers may choose to shorten the distance between them and a particular interlocutor, which might be accomplished by means of gesture or eye contact. Inclusion and disclusion of IRC participants might also be achieved verbally by alternating language (Auer 1984) in an attempt to accommodate to the addressee's preferred language. Lacking visual and auditory features of face-to face communication, switching between scripts can also contribute to the function of including or excluding other addresses in the conversation.

The following IRC exchange exemplifies this function. B's attempts to exclude C, who starts writing in the Arabic script, from the conversation begins in line (6) when s/he switches to the Roman script to direct his/her conversation to A. Both A and B are negotiating in the Roman script, being awake till a very late time in Ramadan. In an attempt to attract their attention, C switches to English with multiple

question marks in line (16) (any body here???? where r people????????). In another attempt to exclude him/her from the conversation and to specify A in the conversation, B switches to the Arabic script in line (17) directing her/his question to A and asking her/him if s/he saw a very popular series in Ramadan.

- (63)1- A: hi nouf  
 2- A: kaif el,7aaal?  
 how DEF-state?  
 'how are you?'  
 3- B: سهرانة؟؟؟؟  
*sahrāna?*  
 awake-2SG.F?  
 'are you awake?'  
 4- A: يب  
*yeb*  
 'yes'  
 5- C: مرحبا  
*marḥaba*  
 'hi'  
 6- B: **6ala3tee faj2ah kent afaker feek @@**  
 appeared-2SG.F suddenly was-ISG ISG-think about you-2SG.F  
 'you appeared suddenly..I was thinking about you.'  
 7- B: le7alek????  
 'alone?'  
 8- A: la ma3 3'900n  
 'no. with Gosoon.'  
 9- B: sahraneen :o  
 awake-2PL?  
 'are you both awake?'  
 10- A: laaa nmmt b4 al,f6wr w 97et  
 NEG slept-ISG before DEF-breakfast and woke up-ISG  
 al, fajer  
 DEF- noon.  
 'no..I slept before the breakfast and woke up in the very morning.'  
 11- B: yuuuu  
 12- B: mooo 7looo ke4a  
 NEG good that.  
 'it is not good to do that.'  
 13- B: tjw3een  
 hungry-2SG.F  
 'you will be hungary.'  
 14- A: ee mraaa :'  
 'yes indeed'  
 15- A: bs b7awl anam b3d shoi 2day  
 but ISG-try ISG-sleep after Shoge today

- ‘but I will try to sleep after Shoge today.’  
 16- C: any body here???? where r people??????  
 17- B: >>>A شفتي زوارَة خميس  
*šeft-ī      zwarat khamīs*  
 saw-2SG.F Zwarat Xamees  
 ‘did you see Zawart Xamees?’  
 18- C: brb  
 ‘be right back.’  
 19- A: eeee tjnnn @@@@ @  
 ‘yes amazing.’

### 7.5.2.2 Competence-related script-switching

Apart from specifying particular addressee in IRC conversation, another type of participant-related script-switching occurs when IRC participants try to adapt to script competence. IRC participants differ in their proficiency of the use of the Roman script. New IRC participants are usually poorly proficient in using numerals that stand for Arabic consonants and they, often, switch to the Arabic script when they struggle to understand or accommodate the speed of interaction.

In the following IRC exchange, A’s incompetence in using the Roman script and her struggle to follow the IRC conversation appears from the start of the conversation. In line (3) and (14), A uses <kh> to represent the Arabic consonant sound /x/; an indication of his poor proficiency of Roman numerals. In line (4) (nice meeting you), A does not use the well-known IRC abbreviation (u) for the word (you), and that is why she does not understand B’s question (gr8 n u?) (great and you?) and does not reply. Also, she struggles to understand C’s question in lines (7) and (8). In lines (17) and (19), A decides to switch to the Arabic script in order to make his IRC exchanges clear.

- (64) 1- A: hi  
 2- B: how r u  
 3- A: wish akhbarek  
 what news- POSS.2FSG?  
 ‘how are you’

- 4- A: nice meeting you
- 5- B: gr8 n u?  
'great, and you?'
- 6- B: >>>>A 3ndk black berry ??  
>>>>>have- 2SG.F blackberry?  
'do you have blackberry?'
- 7- C: u r maha' relative?
- 8- C: maha qashaam??
- 9- A: esmk maha qashaam  
name-POSS-2SG.F maha qashaam  
'is your name Maha Qashaam?'
- 10- C: laaa  
'no'
- 11- C: ana asaal?  
I ask- ISG.  
'I am asking you.'
- 12- A: you know maha qashaam??
- 13- A: ee  
'yes'
- 14- A: ana khalatha  
I aunt-POSS-3SG.F.  
'I am her aunt.'
- 15- C: ha ya halla walla  
ah.....hi indeed  
'hi...really'
- 16- B: 8altlii  
told-3SGF-ISG  
'she told me that.'
- 17- A: عندي بلاكيري بس مو عربي  
*çind-ī blāk bairī bas mū çarabi*  
have-POSS-ISG blackberry but NEG Arabic.  
'I have blackberry but it is not Arabic.'
- 18- C: ana fe riydh now  
I in Riyadh now  
'I am in Riyadh now.'
- 19- A: لا  
*lā*  
'really?'

## 7.6 Conclusion

This chapter aims to provide an empirical account of patterns of multilingual practices in Najdi IRC. Specifically speaking, the present chapter sets out to seek answers to the question of the role of script in integrating borrowings and the functionality of code-switching and script-switching in IRC. Applying

Conversational Analysis approach, the aim was to achieve local interpretations of IRC participants' code-switching and script-switching in order to understand why IRC participants code switch and under what circumstances they often code switch.

In Romanized NA, morphological integration of borrowed items shows complete adherence to the Arabic morphological rules. These borrowed items are integrated into the Arabic data mainly by transliteration into the Arabic script. The role of script in attributing a given English item to either English or Arabic parallels the distinction between single-word code-switching and borrowings. It can be hypothesized that script choice can provide a new perspective on the borrowing vs. code-switching debate in multilingual CMC. Because each language is tied to a particular script and a particular writing system, associations between languages and their writing systems are utilized to attribute a given language item to a particular language. Romanized NA, which shows interrelations between Arabic and English writing systems, minimizes boundaries between languages in contact and, therefore, facilitates the process of integrating English borrowed items into the Arabic lexicon.

The analysis confirms early claims about English's dominance in CMC settings. According to Androutsopoulos (2007), English is favoured over indigenous or minority languages in situations of language contact. It is found, too, that English is preferred for formal interaction amongst Egyptian professional internet users (Warschauer et al. 2002). In Switzerland, English serves, also, as the lingua franca of a multilingual professional network (Durham 2003). The evidence from quantitative analyses indicates that, though the Arabic's eminence

in Najdi IRC is stronger than English, switching to both languages (Arabic and English) serves parallel communicative functions.

In order to find explanations for these general patterns among IRC participants, a qualitative analysis of IRC data that describes functions of code-switching has been conducted. Conley and O'Barr (1990) indicate that "...linguistic variation in any setting is not random, but socially patterned" (cited in David 2003:8). IRC data reveals that code-switching, whether from NA to English or from English to NA, has patterns and serves communicative functions in relation to both the discourse and the participants of IRC conversation. The interactional analysis of code-switching yields a comprehensive typology of IRC discourse functions based on Gumperz's (1982) basic functional framework of code-switching. In Najdi IRC, the main communicative functions achieved by code-switching are:

- Code-switching for qualification;
- Code-switching for specifying an addressee;
- Code-switching of tags;
- Code-switching for reiteration;
- Code-switching for personalization vs. objectiveness;
- Code-switching for self-repair;
- Code-switching to correspond to the language preference;
- Code-switching for conversational coherence;
- Code-switching for requesting attention;
- Code-switching for quotation;

- Code-switching for fulfilment of the linguistic needs;
- Code-switching for signalling topic shifting; and
- Code-switching to correspond to participants' competence.

The variety of these communicative functions confirm the assumption of the functionality of code-switching in spoken discourse. Many studies (Gumperz 1982, Milroy & Muysken 1995, Myers-Scotton 1993, Poplack 1980; Romaine 1995, Matras 2009) argue that bilingual speakers employ code-switching as a valuable linguistic resource to achieve certain communicative functions. These communicative functions of code-switching are required in IRC to “transgress the level of single turns or posts and examine the sequential organization of code-switching within threads of dialogically related posts or messages” (Androutsopoulos forthcoming). Analyzing code-switching in Najdi IRC within the sequential frameworks in which they occur, reveals creative use of code-switching for a variety of purposes. However, the lack of visual channels in synchronous CMC with rapid transitions and short exchanges “results in an increased reliance on code-centred contextualization cuing, which would be otherwise delegated to different signals” (Georgakopoulou 1997: 158).

Georgakopoulou's notion of code-centred contextualization cuing is relevant to script-switching online too. IRC participants manipulate script-switching in order to serve pragmatic functions, which, in ordinary spoken conversation, would be accomplished by other cues such as gazes, gestures, prosody, and phonological variations. Applying the theoretical tools of Auer's Conversational Approach to script-switching, it is found that alternating between scripts online is employed by IRC participants to achieve different communicative functions. Some of these

functions are related to structuring the discourse (such as topicalization, change of mode, requesting attention, and reiteration) and, thus, classified as discourse-related functions. Other functions are related to IRC participants' relationships (inclusion or exclusion) as well as IRC participants' competence, and fall under the category of participant-related functions. At the same time, it is found that script-switching is multi-layered in that it can serve various functions in the same incident of conversation. These findings support the view that, in biscriptual written CMC discourse, script is better understood as part of an overall repertoire of linguistic structures, from which writers choose to suit their contextual and situational needs.

Related to the previous assumptions is the position of code-switching in CMC in relation to code-switching research. The literature on CMC has assessed the language of CMC as a hybrid between spoken and written discourse. In CMC, written language has its own characteristics: often, it is informal and vernacular. It is usually a multi-party conversation with a variety of semiotic symbols. An understanding of CMC will "contribute to a deconstruction of spoken/written dichotomies and a move beyond the assumption that only spoken conversational code-switching constitutes 'authentic' code-switching" (Androutopoulos forthcoming). This criterion of theorizing code-switching emphasizes CMC as a new domain of communication, where discourse is shaped by technological features of CMC as well as innovative characteristics of this hybrid medium of communication. The following chapter is concerned with investigating the specific characteristics of Najdi IRC that make it a hybrid between spoken and written languages and proposing a taxonomy of its innovative linguistic and paralinguistic features.

## CHAPTER EIGHT

### THE SPECIFIC CHARACTERISTICS OF NAJDI IRC LANGUAGE

What has been intuitively known for a long time but is only now being openly stated by CMC researchers is that as CMC has come to dominate human communication. It has encouraged a dramatic expansion in the variety and creativity of language. What makes CMC so appealing as a medium of communication is the way it draws on characteristics belonging to both speech and writing. In his concise definition of CMC, December (1996: 2) states that CMC is ‘the asynchronous and synchronous creation and transmission of messages using digital techniques’. The linguistic features vary in both forms of CMC according to the type of messaging system used (Herring 2002). At one extreme is the asynchronous communication, which in many of its functions (e.g. archiving, publishing, data basing) is no different from traditional writing. On the other hand, communication in chat-groups and instant messaging, though expressed through the medium of writing, shows many of the core characteristics of speech.

At the onset of research on CMC, the aim has been to investigate the idiosyncratic features of this new form of language. The early CMC research has been mainly concerned with providing descriptive accounts of the linguistic features of text-based CMC (Herring 1996, Danet 2001, Werry 1996, Ferrara, et.al 1991). All these studies conclude that CMC is best viewed as a hybrid between speech and writing that is developing into a new medium which shows language users at their most innovative, adapting a variety of linguistic and technical resources for a variety of purposes.

Internet users draw on existing language varieties, often their native ones, and on technical resources, to formulate this new form of communication.

Drawing on both existing language varieties and technical resources, CMC is altering not only the way people communicate, but also the ways they use written language, compose in it, and interact through it. Being simultaneously “interactive” and “written” is what makes CMC like both written and spoken language. Herring (2004: 66) went to emphasise that the language of CMC is different from spoken and written language in its great variability. According to Herring (2004:67), this variability is caused by a variety of technological and situational variables influencing the context of text-based CMC. Among these technological variables are synchronicity, the high degree of anonymity and the persistence of CMC transcript. The situational variables includes participation structure, purpose of communication, topics of discussions, norms as well as participants’ background and amount of experience. (2004: 67). Along the same lines, Danet (2001: 21) argues that as a ‘cross’ between letters writing and conversation, e-mail has changed traditional-letter-writing conventions, and chat groups have made visual aspects of spoken language of greater significance than ever before. New communication codes, according to Danet, continue to develop with unprecedented speed.

The adaptation of keyboard characteristics to suit these new codes of communication is appealing in its creativity and spontaneity. The most obvious features of this adaptation are the development in word-formation. A huge number of new acronyms and abbreviations had emerged recently, taking place of both common conversational phrases and phrases referring particularly to the surroundings of CMC. Furthermore, internet users develop innovative ways to articulate various

paralinguistic features in a written form in order to compensate for visual and auditory aspects of spoken communication.

Earlier scholarly research published in English on CMC has addressed issues of internet communication that employ Roman-based scripts, neglecting CMC in languages not utilizing Roman-based characters (Danet & Herring 2007). Most of these researches published in English described computer-mediated English in different CMC modes and, thus, their results have been generalized about the language of CMC. Recently, researchers have turned their attention to other languages using CMC modes, often their native language, which has contextualized text-based CMC within the history of these languages (for example, Tseliga 2002, Nishimura 2007 & 2008 and Lee 2002).

The main concern of this study is to investigate NA in online communication, an area that has never been explored before. In this chapter, the researcher attempts to explore the characteristics of Najdi IRC language that make it a hybrid between spoken and written language, providing evidence from IRC discussion. The researcher applies qualitative analysis to investigate the linguistic resources that Najdi IRC participants draw on in formulating these features. A general taxonomy of Arabic CMC characteristics is proposed. To be able to characterize its typical features, the following subsection is dedicated to the definition of spoken and written communication. Following this section, a review of accumulated CMC research under this topic will be given. Then, the methodology will be explained and results will be discussed.

## **8.1 The relationship between speech and writing: A theoretical framework**

This section aims to provide a theoretically based account of the several approaches to the study of speech and writing that have been taken throughout the history. It also presents a contextualizing background of the different stages of previous approaches to speech and writing. Most of these traditional approaches have influenced how the issue of identifying the language of CMC is approached. In so doing, this section begins by reviewing Saussure (1919: cited in Hughes 1996:130) and Bloomfield (1933), who form rational hypotheses about speech and writing on the basis of what constitutes the proper object of language study. Then, linguists such as Halliday (1985) and Hass & Hornby (1970) are addressed, whose assumptions are based on the actual manifestations of speech and writing, and who assume that speech and writing are complementary systems. The review also includes linguists who worked on identifying social functions of speech as opposed to writing with a focus on cognitive consequence of writing such as Olson (1988) and Ong (1982). Finally, the section moves to linguists who view the relationship between speech and writing as taking place across a continuum rather than a divide (Biber 1999, Crystal 1995).

In attempting to answer the question of how language should be studied, Saussure (1983) made the assumption that the spoken form is primary, and the written form is just a tool of representing the spoken form. He dismisses the idea of studying language through written forms, claiming that “it is rather as if people believed that in order to find out what a person looks like it is better to study his photograph than his face” (Saussure 1983: 24; quoted in Hughes 1996: 128). For Saussure, the actual examples of speech as used by people are the real object of interest rather than the

abstract system of writing. The written forms are not considered, by Saussure, as part of the language, but as a means of manifestation of the sound system that makes up the language. The argument on which Saussure built his hypothesis of excluding written forms from the study of language is that the spoken form is affected by accidental factors such as the context in which the speech occurs, while written forms are permanent and highly-standardized.

Adopting a behavioral approach to language study, Bloomfield (1933) shares many of the attitudes expressed by Saussure about the study of speech and writing. Not only does he regard writing as a means of recording speech and not an independent of language, but he also argues that the speech-writing relationship is rather a straightforward mechanical process. According to Bloomfield (1933), “writing is not part of language, but rather a comparatively modern invention for recording and broadcasting what is spoken” (Bloomfield 1942: 126; quoted in Hughes 1996: 133). Further, he asserts that the only function of writing that enables linguistic study is that it maintains a fossilization of earlier sounds of a spoken language, and therefore facilitates diachronic studies. Bloomfield’s argument about the primacy of speech over writing is built on a number of assumptions: speech precedes writing chronologically in both the history of a language community and an individual’s language acquisition; writing is an arbitrary representation of a language; decoding written forms is unattainable without knowledge of spoken language; and that written forms hinder our understanding of actual speech (Bloomfield 1953: 21; quoted in Hughes 1996: 133).

A growing body of research has dealt with the differences between spoken and written forms of a language from a functional perspective (Halliday 1985, 1989;

Hass 1970, Vachek 1976). What these studies have in common is the premise that the written form is a product of the functions it serves in actual language use. Unlike the purer theorists reviewed above, the assumption that speech is the true object of linguistic study is not central in this approach. For example, Hass (1970) states explicitly that written forms are not secondary in either nature or language study. He moves to explain how it is possible to pass from phonemes to graphemes and vice versa, and further criticizes the assumption that graphemes- phonemes correspondence is a chaotic one. Along the same line, Vachek (1976) argues that speech and writing are best viewed as twin domains with complementary language functions. The speech forms carries out dynamic functions, writing accomplishes static ones.

Halliday's approach to speech vs. writing keeps a balance between the two forms in regard to each forms' communicative functions. He considers speech and writing as two separate mediums that manifest language and that they evolve differently with typical features of their own. He goes on to describe the differences in forms and functions which affect the natures of the two manifestations (Halliday 1985: 169). The intrinsic nature, according to Halliday, leads to further distinction between the two forms in that each is serving purposes that the other could not. For example, though writing has greater lexical density and possibility of planning and editing than speech does, it does not capture prosodic and intonational features. He agrees with Vachek (1976) that the nature of writing as opposed to speech influences the forms in which human ideas are formed; writing presents ideas in terms of static 'products' and speech presents them as 'process' (Halliday 1989: 81).

David Olson (1988) and Walter Ong (1982) assume the writing supreme in linguistic theory and emphasize the cognitive consequences of writing within society. They state that the shift from speech ‘orality’ to writing ‘literacy’ is interrelated with social and cultural developments within a language community (Ong 1982: 172). From this has evolved the assumption that the shift from speech to writing enhances critical thinking and that substantial logical thoughts can only be carried out in literate societies. Olson (1988: 28) indicates that there are different values placed on kinds of human thoughts in literate cultures compared to primarily oral cultures. Justifying this assumption, Olson (1988) maintains that when writing began to serve the functions of recording human memory, the human mind began to carry out more analytic activities and deliver logical implications. According to him, ‘it is the availability of an explicit written record and its use for representing thought that impart to literacy its distinctive properties’(1988: 28). The assumptions made by linguists of this camp are based on a binary divide between spoken and written language.

In the course of time, a growing body of literature has proposed that this clear-cut dichotomy between spoken and written forms of language is based on idealized analysis of the two forms. Linguists of this camp (such as Biber 1988, Baron 2007, Hyland 2002, Crystal 1995) suggest that the actual use of language “contains a mix of oral and written features and that the two modes overlap and coexist in very complex patterns” (Hyland 2002: 50). Most compelling among such proponents is Biber (1988), who, in his attempts to identify the underlying functional dimensions of writing, clearly argues that “text structure is multifaceted and that no single dimension of comparison can separate speech and writing” (Hyland 2002: 51).

Biber's modal, by which "spoken and written texts are compared along dimensions of linguistic variation" (Biber 1988: 9) draws on the fact that these dimensions are in a continuum relationship rather than as separate phenomena. For example, the "involved vs informational" dimension is found to vary among instances of written language. A particular writing may indicate a great deal of involvement if it includes verbs like "feel", "believe", and "think", and a considerable information if it has many adjectives and nouns. In analyzing genres of both speech and writing, Biber (1988) indicates that some written forms, such as personal letters, are different from academic forms and similar to spoken language. He made the following comment, which attributes the evolving of speech and writing to perform language:

"There is no linguistic or situational characterization of speech and writing that is true of all spoken and written genres. On the one hand, some spoken and written genres are very similar to one another. On the other hand, some spoken genres are quite different from one another, as are some written genre. The relations among these genres are systematic, but must be specified in a multi-dimensional space."

(Biber 1988: 36)

The differences between speech and writing are generally specified in the literature either according to the functions that these two forms of language perform (Halliday 1989; Biber 1988), or the degree of attachment and reflection that each permits (Crystal 1995). Among the latter approaches, Crystal's (1995) *The Cambridge Encyclopedia of the English Language* systematically differentiates between speech and writing. Examples of differences between speech and writing are:

- speech is “time-bound”, “dynamic” and “transient”. It is part of an interaction in which both participants are usually present, and the speaker has a particular addressee in mind. Writing is space-bound, static, and permanent. It is the result of a situation in which the writer is usually distant from the reader, and often does not know who the reader is going to be.
- In speech, there is no time-lag between production and reception, unless one is deliberately introduced by the recipient. The spontaneity and speed of most speech exchanges make it difficult to engage in complex advanced planning. Intonation and pause divide long utterances into manageable chunks, but sentences’ boundaries are often unclear. By contrast, there is always a time-lag between production and reception in writing. Writing allows repeated reading and close analysis, and promotes the development of careful organisation and compact expression, with often intricate sentence structure. Units of discourse are usually easy to identify through punctuation and layout.
- In speech, participants typically rely on extralinguistic cues such as facial expressions and gesture to aid meaning. With writing, lack of visual contact means that participants cannot rely on context to make their meaning clear; nor is there any immediate feedback. Most writing, therefore, avoids the use of deictic expressions, which are likely to be ambiguous.
- Many words and constructions are characteristic of speech, such as contracted form (*isn’t, he’s*). Lengthy co-ordinate sentences are normal, and are often of considerable complexity. There is nonsense vocabulary, obscenity, and slang, some of which does not appear in writing, or occurs only as graphic euphemism. Some words and constructions are characteristic

of writing, such as multiple instances of subordination in the same sentence, elaborately balanced syntactic patterns, and the long sentences found in some legal documents. Certain items of vocabulary are never spoken, such as the longer names of chemical compounds.

(Crystal 2006: 29-29)

Along with Biber (1988), Crystal (2011) maintains that it is more realistic to think of speech and writing as being the products of a multidimensional continuum. Within this continuum, a particular spoken or written language can be located as being more or less like speech and more or less like writing. This multidimensional continuum approach to the differences between speech and writing underlies the statement that the language of CMC is best viewed as a mixed medium of communication.

## **8.2 The language of CMC: A mixed medium of communication**

Although writing is the primary medium of CMC, “the question of how speech is related to writing is at the heart of the matter” (Crystal 2001: 18). In the early 1990s, attempts were made to classify CMC interaction compared with other language modalities. It has been argued that CMC exhibits features of both written and spoken language as it is typed by keyboard characters and read as a text on screen, and as it displays features of spoken language such as informality and quick message exchange (Ferrara et al. 1991). Thus, CMC has been classified as a third hybrid modality between spoken and written modalities and characterized by unique features. In this regard, various terms have been coined in order to describe this hybrid modality online, some of which include “interactive written discourse”

(Ferrara et al. 1991), “electronic language” (Collot & Belmore 1996) and “Netspeak” (Crystal 2001). In describing this type of language, Herring indicates that it is “less correct, complex, and coherent than standard written language” (2001: 616). This may be due to the fact that internet users employed a variety of orthographic strategies to impose intonations and paralinguistic cues into interactive written discourse online.

The substantial body of linguistic research into language of CMC falls into two main approaches. The first approach is concerned with locating the language of CMC within the speech vs. writing dichotomy (Collot & Belmore 1996, Yates 1996). An important aim of these studies is to find out how text-based CMC is different from other language modalities (i.e., spoken and written language). These studies are based on the assumption that the relationship between speech and writing is a continuum rather than a divide (Halliday 1985, Biber 1988). The second approach of research is concerned with identifying the medium-specific features of CMC by analyzing the idiosyncratic linguistic features that are assumed to be specific to online communication. A common concern of this research is investigating abbreviations, capitalization, and emoticons that characterize synchronous and asynchronous CMC. The following sections provide an overview of the linguistic research of CMC language, with emphasis being given to studies that contributed to the formulation of the research questions investigated in the present study.

### **8.2. 1 The language of CMC between speech and writing**

The issue of CMC language having both spoken and written aspects is discussed quantitatively by Collot and Belmore (1996), Yates (1996), and Nishimura (2008), each applying different approach. The study conducted by Collot and Belmore (1996) is summarized first.

Collot and Belmore (1996) quantitatively studied aspects of spoken and written hybrid nature of what they call “Electronic Language” compared to face-to-face settings. Applying Biber’s (1988) multidimensional-multi-feature (MD-MF) approach in establishing a descriptive framework, they investigated lexical and grammatical features of BBS discourse compared with computerized corpora of spoken and written English. Their corpus consists of more than a million words gathered from BBS websites and analyzed by using (MD-MF) explanatory factor analysis that identifies six dimensions. Each of these dimensions is related to linguistic functions. Correlation between communicative functions and linguistic characteristics are analyzed in Biber’s continuum. These dimensions are:

- Informational vs. involved production;
- Non-narrative vs. narrative;
- Situation-dependant vs. explicit;
- Over expression vs. persuasion;
- Non-abstract vs. abstract; and
- On-line informational elaboration

Utilizing the MD-MF functional approach, Collot and Belmore found that BBS messages are closer to interviews and personal and professional letters in Biber’s continuum.

Yates's (1996) quantitative study is also concerned with comparing online corpora with computerized corpora of spoken and written English. His online corpus is based on message exchanges through a computer conferencing system at the Open University in the United Kingdom. There were 152 conference topics with 50 messages or more. From each of these, the first 50 messages are taken for analysis. The factors around which comparisons between online corpus and computerized corpora of spoken and written English are made are drawn from Halliday's (1978) model of language use. This model focuses on textual, interpersonal and ideational aspects of language modalities. Yates found that text-based CMC exhibits aspects of written language in regard to vocabulary use (textual aspect). However, he found that the CMC corpus is more like spoken language with respect to the use of personal pronoun (interpersonal aspect). Another important finding in Yates's study is that CMC employs greater use of modal auxiliaries (ideational aspect) than both spoken and written computerized corpora.

Drawing on Yates's study (1996), Nishimura (2008) investigated the linguistic aspects of Japanese BBS communication. She created small corpora of written and spoken Japanese and a synchronous Japanese CMC, and regarded the morpheme as a basic meaningful unit of quantitative analysis. She argues that the decisive factor that differentiates between speech, writing and CMC in her data is interjections. Applying Halliday's model of language use, Nishimura found that, though spoken qualities have been incorporated in Japanese Bulletins, CMC users fail to reproduce interjections that are produced unconsciously in speech. Significant differences in the use of auxiliaries are also found among speech, writing, and CMC corpora. Nishimura concludes that the language of synchronous CMC as a whole is basically written especially in terms of the overall particle usage.

### **8.2.2 The specific characteristics of CMC language**

Studies of Collot and Belmore (1996), Yates (1996), and Nishimura (2008) reveal that CMC is best viewed as a hybrid form of language displaying qualities of both spoken and written qualities. A growing body of empirical research has been conducted to investigate the idiosyncratic characteristics of this new form of language with attempts to propose taxonomy of characteristics of CMC language.

Ferrara's (1991) pioneering study is one of the first that appeared in an academic journal, and, thus, set the tradition of CMC research. In this study, Ferrara and her colleagues investigated experimentally stylistic features of CMC among online interaction of 23 computer professionals and their spouses who were asked to do travel arrangements with the travel agent. Ferrara and her colleagues investigated the messages sent to the travel agent and the responses to these messages, which reveal peculiar linguistic features to what was later termed "Interactive Written Discourse" (IWD). Among these characteristics are unconventional grammar, omission of subject pronouns, and reduction of syntactic structures. Ferrara and her colleagues claim that the characteristics of both written and spoken language formulate an emergent register that accounts for stylistic and linguistic features of CMC.

Far from being experimental, as Ferrara and her colleagues were in their approach to online interaction, Werry (1996) investigates existing CMC by analyzing linguistic features of synchronous CMC. Werry's study has significantly contributed to the undertaking of investigation in the present project. He conducted his investigation in a synchronous IRC, providing a classification of features that

characterize this little-investigated genre (Herring 1996). Moreover, Werry indicates that orality is an essential concept in synchronous modes of CMC as participants were found to imitate aspects of spoken language through an extensive use of linguistic, graphic, and orthographic representations to create the effects of voice and to entertain other users. Examples of such innovative linguistic and orthographic devices that have been found among internet users in Werry's study include:

- The use of abbreviations (e.g., “u” for “you”, “r” for “are”)
- The use of paralinguistic cues in order to create the effects of voice, tone and gestures (e.g., duplicated letters “loool”, “coool”, “mmmmmm”, and capitalization (e.g., “LOSER”)).
- The use of emotions created by keyboard characters to express facial expressions (e.g., :-) “smile”, :S “confused”, :-( “sad”)
- The use of acronyms (e.g., FAQ “Frequent Asked Questions”)
- The use of punctuation to create the effects of spoken delivery. Periods and hyphens are employed to create pauses and to indicate tempo (Frouser, 2000).
- The use of non-standard spelling (e.g., “l8er” for “later”)

(Werry 1996: 18)

Studies that investigate typographical and visual aspects began to emerge in the 2000s. Danet's (2001) book *Cyberpl@y* is the pioneering in this field of investigation, and it is based on ethnographic observations and researching visual and playful elements in different CMC settings, including emails, IRC, and the World Wide Web (WWW). Two facets of typography in CMC are drawn in Danet's book and have contributed to the undertaking of many CMC studies,

including the present one. These facets are the visual and playful aspects of typography in CMC. The visual aspect, which is of particular interest to the present study, is mainly concerned with visual practices employed online. Najdi users of the internet are found to employ creative visual devices and non-linguistic symbols to represent their non-standard dialect online. The aspect of playfulness is concerned with word play that is used for enjoyment. Unconventional literacy, emoticons and acronyms are enjoyed by internet users as they enhance belongings to the group and, hence, contribute to creating community identity. Danet's nine-year research in the phenomena of typography in CMC resulted in the development of Danet's inventory of digital writing. In Table 8.1 below, Danet's inventory of digital writing is provided, with supplemented examples from different sources.

Table 8.1 Common features of digital writing (Danet 2001:17).

	<b>Examples from several sources</b>
<b>Multiple Punctuation</b>	type back soon!!!!!! (Danet 2001) annny problems???? (Werry 1996)
<b>Eccentric spelling</b>	type back sooooooon! (Danet 2001)
<b>Capital letters</b>	I M REALLY ANGRY AT YOU (Danet 2001)
<b>Asterisks for emphasis</b>	I'm really *angry* at you. (Danet 2001)
<b>Written-out laughter</b> <b>Other vocalizations</b> <b>Music/noise</b>	Hehehe hahahaha (Danet 2001) Mmmmmmm poc poc poc (Werry 1996)
<b>Description of actions</b>	*grins* <grin>

	<g> (Danet 2001)
<b>Emoticons</b>	:-) (smile) ;-) (wink) :-( (frown) (Danet 2001) (@_@) (I am very surprised)
<b>Abbreviations</b>	Pls. [ please]; tks [thanks] ( Werry 1996)
<b>Rebus writing</b>	LOL; BRB [laughing out loud], [be right back] (Danet 2001) CU [see you] (Danet 2001)
<b>All lower case</b>	hi, how are you? (Danet 2001)

Applying Danet's (2001) inventory of features of digital writing, Nishimura (2007) investigated linguistic innovations and interactional features of casual online communication in an online Japanese context. Using Japanese BBS messages as the primary source of data, Nishimura investigated linguistic innovations and interactional features among young Japanese BBS users. She found that young Japanese BBS users employ colloquial language online and they use spoken features such as final particles in order to interact with other users online as if they were in face-to-face conversation. She also found that the varying degrees of style mixture in Japanese text-based CMC are related to the extent to which the users feel involvement with and closeness to each other.

To sum up, while research on linguistic features of English CMC has accumulated over the past 20 years (Werry 1996, Herring 2004, Crystal 2001, 2004, 2008, 2011, Danet 2001), investigation of languages other than English has been limited (Danet & Herring 2007). To the present researcher's knowledge, there is no single study in CMC literature that has been conducted to identify systematically the specific characteristics of Arabic CMC language compared to spoken and written

communication. The present study is important because it aims at investigating NA use online, an area that has not yet been explored in this area of inquiry. Thus, this study attempts to investigate language use online within a different cultural context in order to avoid ethnocentric bias regarding the linguistic features that are shown to be characteristics of CMC as investigated in earlier studies. Arabic writing conventions are different from those of English (for example, in Arabic writing, there is no lower case and upper case), it is interesting to compare the use of these characteristics to their English counterparts.

The present study follows Werry's (1996) qualitative approach in analyzing the language of Najdi synchronous CMC with Danet's (2001) "common features of digital writing" as a framework of analysis. The researcher of the present study argues that, in line with Crystal (2001) and Herring (2004), the immediacy of CMC makes it best seen as a new hybrid language that is compiled by linguistic and technical resources. More specifically, this chapter is mainly concerned with what exactly the linguistic and paralinguistic features of the so-called "a hybrid medium" (Danet 2001: 17) developed and employed by Najdi users, are. Another aim of this study is to find out what are the linguistic resources that Najdi internet users draw on in developing these features. It is also intended that a comprehensive taxonomy of characteristics identified in Arabic CMC is proposed.

### **8.3 Methodology**

To seek answers for the generated questions in this chapter, the IRC data was given both qualitative and quantitative analysis. The aim of the qualitative analysis

was the identification of linguistic and paralinguistic features for the sake of developing a general taxonomy of features of Arabic IRC. In the manner of the traditional qualitative coding method, the themes of analysis were not decided beforehand but rather emerged as the analysis proceeds (Strauss & Corbin 1999).

Analysis began with identification of the general recurrent features emerging from the raw data. During this stage of analysis, the researcher identified and tentatively named the conceptual categories into which the phenomena observed will be analyzed (such as ‘capitalization’, ‘letter acronyms’, abbreviations, etc). The goal was to create descriptive categories which form a preliminary framework for analysis.

The next stage of analysis involved re-examination of the categories identified to determine how they were formed and how they were linked to each other. This stage involved identifying one or two core categories to which all other sub-categories relate and building a conceptual taxonomy. For example, categories such as ‘letter acronyms, phonetic acronyms, abbreviations, letter homophones and numeral homophone are grouped under the core-category ‘E-abbreviation’. The final stage in qualitative data analysis was to display the data resulting from the three stages of analysis in an organized and compressed manner that allowed conclusion drawings (taxonomy of linguistic and paralinguistic features of Arabic CMC<sup>18</sup>).

---

<sup>18</sup> Before presenting the inventory, it must be noted that the researcher had to delete facial icons representing different emoticons such as ☺ and ☹ for technical reasons. Although the use of these facial emoticons was clearly abundant in the IRC data, the copy and paste function used to save the IRC sessions to word documents automatically deleted these different facial icons. Thus, this CMC technique will not be included in CMC inventory in the present study as it cannot be measured.

The present analysis is concerned with investigating features that are found to be frequently employed by Najdi participants and are commonly identified in the previous research. In order to obtain an account of the frequency of certain characteristics in the IRC data, the data was given quantitative analysis. Occurrences and frequency of each characteristics is counted and summarized.

The researcher proposed taxonomy of features of CMC appearing in the data. The CMC features investigated in this study are based on an investigation of IRC data and previous literature (Tseliga 2002, Nishimura 2002 & 2004, Lee 2002, Herring 1996, Danet, 2001, Werry 1996). Danet's "common features of digital writing" (2001: 17) contributes mainly to the developing the CMC inventory in the present study. Danet's taxonomy of digital writing is based on the assumption that synchronous CMC has been argued by CMC scholars to be a hybrid of speech and writing. Thus, the first seven features in Danet's taxonomy are devices to compensate for the nature of the medium as attenuated "speech", whereas the last two are devices that help the user to type as fast as possible. Consequently, they are writing-specific devices.

## **8.4 Results**

### **8.4.1. E- abbreviation**

The issue of how language is abbreviated and shortened is perhaps the most interesting feature in terms of analyzing synchronuous CMC. Traditionally, acronyms and abbreviations are formed to stand for compound words that describe technical jargon (e.g. *HTML* for Hyper Text Markup Language'). Short forms for

sentences or frequently-used expressions were not very common until acronym emerged in different modes of CMC. The immediacy of IRC conversation, and the creative use of keyboard characters, resulted in new unconventional abbreviations used in a rather innovative way that soon spread among all IRC users, and came to be acknowledged as part of the chatting process. This subsection investigates the approaches that Najdi users adopt in abbreviating their language in IRC. A comparison with the most commonly used abbreviations in English CMC was also conducted.

Lexical reductions in CMC are no longer restricted to traditional classifications of abbreviation in language such as acronyms (e.g. *NATO*) and initialisms (e.g. *DVD*). New strategies with regard to forming shortened lexical expressions in CMC have been identified (Crystal 2001). Appendix D displays the most common abbreviation in English CMC, some of which were already in use in traditional writing before the advent of the internet e.g. FYI for “For Your Information”. Others seem to have been developed specifically to suit the needs of online communication. IRL is an example of one that emerged in digital writing, meaning “In Real Life”. As in both spoken and written communication, abbreviations promote efficiency. In synchronous CMC, they save valuable time.

Looking at the Najdi IRC corpus as a whole, Najdi users were found to be using creative and innovative adaptations of the common abbreviation in English CMC to suit the needs of synchronous IRC. Najdi internet users have probably observed these common abbreviations in chat rooms, where the English language dominates, and then have adopted them to Arabic IRC. The efficiency of these common abbreviation and demands for abbreviating the user’s own local dialect have both made Najdi users create new abbreviations based on Arabic

orthography. Other abbreviations and acronyms are made up by analogy with what those based on English look like. The methods of lexical reductions that have been employed by Najdi IRC participants in Najdi IRC are as follow:

#### 8.4.1.1 Acronyms of sentence

Acronyms of sentences are word forms that are encoded by stringing the first two or even three initial letters together in a sequence of words (Crystal 2001:41). What happens in IRC is that everyday words are reduced. The meaning of an acronym in IRC may not be understood by other participants, unless they are already familiar with the language of IRC. In Romanized Najdi IRC, examples of Najdi acronyms are the exact counterparts of the most common acronyms in English CMC. These examples include (LOL ‘laughing out loud’, BRB ‘be right back’, B ‘back’, WB ‘welcome back’, TUT ‘take your time’). These English CMC acronyms are inserted frequently into Romanized Najdi IRC. They appear almost in every single conversation between IRC participants. Table 8.2 below indicates the number of occurrences of each acronym. The use of these borrowed elements demands that specific background knowledge has to be shared by both sender and receiver, in order for them to interact simultaneously in a quick and easy way, as well as saving time typing long utterances. The following extract illuminates how these acronyms are used in Najdi IRC. (Acronyms are illustrated in bold)

- (1) A: eee/ re7na e lmarsa marrra bard  
yes went- IPL DEF-Marsa very cold  
‘yes... we went to Al-Marsa and it was very cold.’  
A: **brb**

'be right back'  
**B: tut**  
 'take your time'  
**C: B>>>erselee le 9or elmarsa plz**  
 >>> send-2SG.F to-ISG photos DEF-Marsa please  
 'send me Al-marsa's potos please!'  
**C: saivthum killhum**  
 saved-ISG all-3PL  
 'I saved all of them.'  
**B: :(**  
**A: B :) :)**  
 'back'  
**C: WB**  
 'welcome back'

Acronyms identified in Najdi IRC might be categorized into conventional letter acronyms based on English CMC, and unconventional phonetic acronyms based on rendering these acronyms into Arabic script. The following subcategory illustrates how Najdi users adapt Arabic orthography to form innovative acronyms that suit their online activity in terms of communicating in NA.

#### 8.4.1.2 Phonetic acronyms

These acronyms are made up from Arabic letters that resemble the sounds of conventional English acronyms. Najdi users employ these acronyms frequently, and the rules of using them are based on the shared knowledge that exists between participants. Thus, these acronyms function as the code of the IRC community and only used in context of IRC messages. Examples include *برب* *brb* (Be Right Back), *تيت* *tīt* (Take your time), *باك* *bāk* (Back), and *وب* *web* (Welcome Back).

### 8.4.1.3 Abbreviations

The Arabic language has developed conventional abbreviations for different purposes. The most authoritative conventions for abbreviation are given by The Library of Congress and by the Encyclopedia of Islam. An abbreviation of a word is made by omitting certain portions of a phrase, and stringing some of its initials as separate letters. Examples are الخ *'elāx'* which stands for إلى آخر الكلام *elā axer al-kalām* 'etc.', Q.ق.هـ. which stands for قبل الهجرة *qabel al-hejrā* 'before Hejra' and Q.M.ق.م. standing for قبل الميلاد *qabel al-melād* 'before birth of Essa'. In IRC, Najdi users are not found to use these conventional abbreviations. Instead, innovative abbreviations have emerged and have spread among all IRC users. These abbreviations are not found neither in dictionaries nor in Arabic books and it seem that these abbreviations are specific to IRC. The abbreviation بي بي *bī bī* standing for 'Blackberry' appears twenty three times in the IRC data, The Arabic abbreviation الخ is never used in the corpus, but the English abbreviation etc occurred seven times in the IRC data. The English CMC abbreviations 'thanz' and 'plez' are also frequently used in Romanized NA.

(2) A: شفتها فالبي بي  
*šifta-ha fī-l-bī bī*  
saw-1SG-3SG.F in-DEF-blackberry  
'I saw her in the blackberry.'

(3) A: my bb **pin** is 345612

(4) A: laaa ma a8wem ba arselha I bader  
no NEG ISG-resist will ISG-send-3SG.F to Bader  
'no...I cannot resist. I will send it to Bader.'

B: ARSALT....lbader w mishary w maha...etc.  
ISG-sent- to-Bader and Mishary and Marha.. etc  
'I sent it to Bader, Mishary and Maha..etc.'

#### 8.4.1.4 Letter homophone

Letter homophone is an innovative method of abbreviation that has emerged in CMC (Crystal 2001, Danet 2002). Letter homophone, termed as ‘Rebuse Writing’ by Naishimura (2002), is the use of a single letter to represent a word with a similar sound. The most common letter homophones in English CMC are ‘u’ for ‘you’, ‘r’ for ‘are’, ‘b’ for ‘be’ and ‘c’ for ‘see’. These letter homophones are copied by Najdi users into their Romanized NA IRC and are rendered by other users into Arabic script such as (سي يو *sī yū*). Exaple (5) below illustrates the use of these English common letter homophones. Example (6) illustrates rendering these English letter homophones into Arabic script.

- (5) B: ok im feeling tired now  
A: its 7'arban  
‘it is not working.’  
B: ok momo  
B: nice meeting u  
A: U 2

- (6) A: سي يو  
*sī yū*  
‘see you’  
B: باي  
*buay*  
‘bye.’

#### 8.4.1.5 Numerals

In an analogy with the method of letter homophone, international IRC participants employ the use of numerals for creating new words that save both time and the effort of typing, and make fun within the IRC community. For example, 88 (‘bye bye’ in English resembles the pronunciation of ‘8’ in Cantonese), and 99 (‘Nite Nite’ ‘good night’]) (Lee 2002: 13). The influence of the users’ native tongue i.e. Cantonese in this case, claims Lee, resulted in the emergence of these innovative

words. Numeral homophone strategy is also employed by Najdi users when interacting using Romanized NA, but in a rather different way from Lee's study. Numerals have a significant value when writing in Romanized NA in IRC. Najdi users have deliberately made use of a number of numerals to represent sounds of NA that do not have English counterparts (See Chapter Six). IRC participants use these numeral conventions to represent certain Arabic words. The most common two numeral homophones in Romanized NA IRC extracts are '97' and '78'. According to the conventions created in IRC, the numeral 9 stands for the voiceless alveo-palatal fricative sound /s/ , whereas the numeral 7 is used to represent the voiceless pharyngeal fricative /ħ/. Thus, a combination of the two numerals results in the formation of the word (صح) *ṣaḥ* 'right'. Along the same lines, the numeral 8 represent the voiced velar stop /g/ which does not have a representative in the Arabic letters. The numeral homophone 78 represents the word *ḥag* (right). Below are examples of these numeral homophones in the Najdi IRC corpus. The numerals combination <97> occurs thirty four times in the IRC data and the numerals combination <78> occurs eightee times.

(7) A: kent ashooof el, bluetooth **78** el,jehaz  
 was-ISG ISG-seeing DEF-bluetooth of DEF-computer  
 'I was fixing the bluetooth in my computer.'

(8) A: wish esm el, madeena????  
 what name-CONS DEF-city?  
 'what is the name of the city?'

B: corvallis

A: fee Oregon???  
 'in Oregon?'

B: **97**  
*ṣaḥ*  
 'that is right.'

### 8.4.1.6 Combination of initial letters and numeral homophones

Combining letter and numeral homophones is an innovative strategy for abbreviating English CMC (Crystal 2001, Danet 2002). Examples include ‘b4’ standing for the word ‘before’ and ‘2day’ standing for the word ‘today’. These English homophones that have been created in English IRC and then adapted to other CMC contexts where the dominant language does not use the Roman alphabet system (Lee 2002). The Najdi IRC corpus reveals an extensive use of these English homophones in Romanized NA. IRC extracts below include some of these homophones.

(9) A: 5ala9 barselhum **2morrow** :)  
 ok ISG-send-3PL tomorrow  
 ‘ok..I will send them tomorrow.’  
 B: :O

(10) A: sawee save **b4** run  
 make-2SG.F save before run  
 ‘click on ‘save’ not ‘run’.’

Table 8.2 The most common abbreviations and acronyms in Najdi IRC

Category	Token	No. of occurrence
Letter acronyms	lol	43
	brb	112
	tet	62
	b	19
	wb	12
Phonetic acronyms	لول	154
	برب	287
	تيت	216
	باك	19
	وب	11
Abbreviations	etc	7
	bb	23
	بي بي	47
Letter homophones	u	91
	r	72

	c	84
Numerals	97	34
	78	18
Initial letters and numeral homophones	b4	18
	2day	71

### 8.4.2 Addressitivity

Addressitivity refers to the different ways by which IRC participants indicate the intended IRC addressee in the course of conversation. Indicating the address's nickname is an attempt to avoid ambiguity as this discourse features is typically negotiated in face-to-face communication by means of paralinguistic cues such as gaze, voice and intonation (Werry 1996: 52). The lack of simultaneous non-verbal cues in IRC communication (such as gaze, gesture and nodding) which signal active attention and guides structures of exchange in face-to face communication resulted in a high degree of addressitivity in IRC. It has become conventional in IRC communication to indicate the intended addressee's by writing his/her name at the beginning of a new IRC utterance, preceded by multiple angel brackets. Below are examples.

(11) A: >>>A shiftee      zwarat 5amees  
>>> A saw-2SG.F Zwarat Khamees  
'did you seeZwarat khamees?'

(12) A: everybody is celebrating your party. **wish hl7arakat >>>B???**  
what this DEF-actions B  
'that is wonderful.'

### 8.4.3 Paralinguistic cues

As IRC lacks visual and auditory channels, innovative set of linguistic and orthographic devices have to be developed to compensate for the lack of paralinguistic cues of face-to-face communications. These innovative features are made up by English letters, punctuation marks and ASCII characters and produce the effect of voice, intonation and gestures that are lacking in the interactive written discourse of IRC. Analysis of IRC data reveals five fundamental paralinguistic features that both compensate for the lack of auditory and visual channels and enhance playfulness of IRC communication. These are: the use of emoticons, creative use of capitalization, multiple punctuations, reduplication and vocalization. Below is a description and examples of each.

#### 8.4.3.1 Emoticons

Emoticons are devices which supply information about non-verbal aspects of communication pictorially, and are composed of clusters of ordinary typographic symbols such as commas, periods and parentheses<sup>19</sup>. The word is a conflation of “emotions” and “icons” (Raymond 1996: 173) and came to be associated with casual and humorous writing. When viewed with the head tilted towards the left shoulder, they form ‘faces’ that indicate facial expressions. They are often used to alert a reader to the writer’s mood, and can modify the interpretation of a plain

---

<sup>19</sup> The emergence of emoticons in writing dates back to the Eighteenth Century when Abraham Lincoln inserted “;)” in his 1862 speech. In “Is That an Emoticon in 1862?” in the *New York Times*, there is a controversy as to whether it is an authentic punctuation cluster or an emoticon (Danet, 2001). In 1881, typographical emoticons were published by the U.S. satirical magazine, *Puck*. The creation of :- ) and :-( symbols with a specific suggestion were initiated in 1982 by Scott Fahlman (Raymond, 1996).

text. Nowadays, emoticons are used frequently in web forums, instant messengers and informal emails. The two basic emoticons :) and :( express positive attitudes and negative attitudes respectively. Appendix E illustrates the most commonly used emoticons in English CMC.

The variety of attitudes and meanings that emoticons convey has led to the emergence of a widespread view in the CMC field that emoticons are the paralanguage of the internet (Dery 1993). However, they have to be concisely added to a text. They can indicate a misperception of a user's mood, but a single emoticon still allows a huge number of interpretations (happiness, joke, sympathy, good mood, delight, amusement, etc.). Their absence does not mean that the user lacks the emotion conveyed. Internet users who routinely use emoticons find their unmarked messages misinterpreted precisely because they have no emoticons attached to them (Allothman 2010)

In analyzing the Najdi IRC corpus, Najdi users are found to use emoticons rendered in Roman script which look, in many cases, similar to those found in English CMC. Emoticons that are frequently used in the corpus of the present study are the simplest ones that are likely to be understood by most IRC users and which can also be typed quickly. The two most common emoticons in the Najdi IRC corpus are simple eyes and a smiling or sad mouth :) and :( . Sometimes, the nose element is added which seems to be a solely a personal taste :-) and :-(. Table 8.3 identifies the most commonly used emoticons among Najdi users.

Table 8.3 The most frequent emoticons in IRC (Alothman 2010)

Token	Meaning	No. of occurrences
:), :-), :}, :], :-}, :-]	happiness, pleasure, smiling, etc	312
:D, :P, (^-^), ;D, :)), :}}, :]]	big smile	114
@@@@	Smiling	62
:(, :-(, :{, :-{, :[, :-{ ,	Sadness	402
;(; ;-(; ;-[; ;p; ;-{; ;'('	Crying	84
:O, :-O, : , :-0, ;- ,	Astonishment, surprise	53
;), ;-), ;-}, ;D, :P, ;-P, :*D, ;')	Wink	113
:S, ;S,	Hesitation	32
:-{, :-O	frown, shouting	42
;*), :*}, :* , :* ,	embarrassment, shyness	34

As revealed by the table above, Najdi users have adopted unaltered feature emoticons and international emoticons used in English CMC in their Romanized NA. Emoticons in Najdi IRC are composed of a range of different punctuation marks, keyboard characters and the Roman alphabet, but no single Arabic letter. Najdi users tend to reduplicate emoticons for the sake of emphasizing their feelings, opinions and attitudes. The following extract is taken from IRC and illustrates how emoticons are used more extensively by Najdi users when chatting in Romanized NA.

(13) A: mar7aba **:P:P:P**  
 'hello.'  
 A: kaif el7aal??????  
 what DEF-state?  
 'how are you?'

B: ahlain :D  
 'hello'  
 A: BUZY :S  
 C: سلامات والله  
*salamāt waḥḥa*  
 peaces really  
 'it is good that you are fine.'  
 C: :!!!!من زمان ما شففتك  
*min zamān mā šift-ik*  
 from long time NEG saw.ISG-2SG.M  
 A: 7ata ana, missing u ;-)  
 'me too  
 C: :D:D:D

Emoticons are one of the most distinctive features of CMC that convey the kinesic features typical of face-to-face conversational speech. However, they are not the only mechanism devised to compensate for the lack of features of face-to-face speech. Creative use of capitalization is also used by Najdi users to display the fluidity of face-to-face spoken interactions.

#### 8.4.3.2 Capitalization

Capitalization is not used for proper nouns or at the beginning of a sentence. Rather, it is used as a convention for expressing prosodic effects. As Arabic orthography does not have lower and upper cases, capitalization is only found in Romanized NA. In such cases, it conveys prosodic features such as giving more emphasis and stress to words. in Ex (14), the IRC participant uses captialization to add impact of emphasis (he has a serious headache) while the participant in Ex (15) captializes his/her whole utterance to make it clear that he/she does not have enough money.

(14) A: sorry kent      adawwer      elmalf      78      bokrah  
 sorry was-1SG ISG-search      DEF-file      for tomorrow  
 w rasi                      **DAYER**                      fee eljehaz.  
 and head-POSS 1SG wandering-3SG.M in      DEF-computer  
 ‘sorry. I was searching the file for tomorrow and I have a serious  
 Headache.’  
 B: lol  
 ‘laugh out loud.’

(15) A: sheft              iphone 4????  
 saw-2SG.M i.phone 4?  
 ‘did you see i.phone 4?’  
 B: **MA 3ENDI**      \$\$\$\$\$\$\$  
 NEG have-POSS-1SG “money”  
 ‘I do not have Money.’

### 8.4.3.3 Multiple punctuation

Punctuation marks are conventional norms that help to indicate the structure and organization of written language. They are used to express questions, exclamations and indicate statements. Arabic employs much the same punctuation as English. The similarity is so strong that very few variations may occur. However, written from right to left, Arabic uses a reversed question mark (؟), and a reversed comma (،). Multiple punctuation that comes at the end of a sentence, such as a question mark, period or exclamation mark, are extensively used by users of IRC in many languages (Nishimura 200, Danet 2001, Werry 1996).

Communicating in a synchronous mode of CMC, punctuation marks are used in unconventional norms that help to both organize written language and compensate for non-verbal clues that are lacking in text-based IRC. They are deliberately multiplied to express attitude (surprise, anger, happiness, etc.), ask questions, give emphasis or ‘make oneself heard’ (Werry 1996: 19). The analysis of the IRC corpus reveals an extensive employment of punctuation marks. Table 8.4 below reveals the

most common punctuation marks that are multiplied for purposes of emphasizing , showing interest, and expressing anger.

Table 8.4 The most frequent multiple punctuation marks in IRC

Token	No. of occurrences
multiple question mark	211
multiple exclamation marks	243
multiple angle brackets	62
multiple dash	84

Example (16) below is taken from an IRC chat log, and illustrates this use. A similar use of question mark to express astonishment and surprise is also found in example (17).

(16) A:salaaaam  
       'hi.'  
       B: weeeen elnas!!!!!!!  
           where DEF-people?  
           'where are the people?'

(17) A: sheft       iphone 4?????  
       saw-2SG.M i.phone 4?  
       'did you see i.phone 4?'

#### 8.4.3.4 Reduplication

Reproducing spoken pronunciation in typed IRC messages through unconventional spelling is a feature shared by a number of international IRC rooms (Danet 2001, Lee 1996, Nishimura, 2007 and 2008). Attempts to compensate for speech in typed messages are the major reason for producing unconventional spelling. For extended and tuned pronunciation, letters are



exclamation and various other sounds. The analysis of the IRC data in the present study reveals extensive visualization of mimetic and onomatopoeic expressions, such as ‘mmm’ and ‘ohhhh’, in both Arabic and Romanized scripts. Common vocalizations in English CMC have been adopted in Najdi IRC. Some of them are adjusted for Arabic script. The insertion of such devices in IRC arouses a high degree of closeness and involvement.

Examples below show how Najdi users devise unconventional features of typing in order to convey auditory features of informal face-to-face conversations. Most of these devices are duplicated in an attempt to convey prolonged expressions that match those produced in real face-to-face conversations.

(21) B: يووووووووو

yuuuuu

‘oh’

C: :O:O:O

(22) A: shareet      iphone 4 ;p

bought-ISG iphone 4

‘I bought i phone 4.’

B: **waaaaaaaaaaw**

## 8.5 Conclusion

An important aim of this chapter is to analyze the specific features of Najdi IRC compared to spoken and written communication. According to Werry (1996: 53), “factors such as screen size, average typing speed, minimal response times, competition for attention all contribute to the emergence of certain characteristic properties”. This section summarizes these specific characteristics of Najdi IRC.

Interactions in synchronous IRC take place in real time through the medium of written language. IRC participants need to say in writing what they have been used to saying in speech. The IRC interactions demands to be read with ‘simultaneous involvement of the ear and eye’ (Werry 1996: 59). The reduced transparency of messages heightens the participants’ meta-linguistic awareness, and leads them to treat words, letters and keyboard characters as objects in a rather unconventional way. They evolve a number of innovative linguistic features that simulate of face-to-face communication. This includes conveying phonetic effects (e.g. vocalizations), animating prosodic cues (e.g. emoticons and reduplications) and organizing discourse (e.g. addressivity). On the other hand, spatial constrains of IRC screen size, minimal response time and competition for attention resulted in innovative sets of abbreviations (e.g. letter acronyms and phonetic acronyms). These linguistic innovations shows interrelations of English (e.g., abbreviations ), English CMC (e.g., letter homophones and emoticons), NA (e.g numerals ), and Arabic orthography (e.g., rendering English homophones into the Arabic script).

These creative unconventional abbreviations, and other strategies for manipulating communication in IRC, seem to function as a marker of the participants’ ability to master IRC communication. All IRC participants are literate people; they know how to spell, write and punctuate their writing. The fact is that they chose not to use conventional writing techniques in order to present themselves as having mastered the IRC code, and thus belonging to the IRC community. Their messages often bear more resemblance to IRC code than to

standard language. This code is specific to the IRC environment and is not employed in traditional spoken or written forms outside this environment.

These specific features of IRC language shows interrelation between languages (Arabic-English) and scripts (Arabic-Roman). Some of these linguistic features are copied from English chat rooms such as emoticons and letter acronyms suggesting a transfer of English CMC features. Other devices have been altered to the Arabic orthography such as phonetic acronyms. In other cases, Najdi IRC participants, like Japanese CMC participants (Lee 2002) are inspired to create new CMC features based on the Arabic orthography in an analogy with English CMC features (such as numerals). Such practices suggest cross linguistic trends among internet users using different languages to accommodate their text-based CMC to the English CMC norms. This confirms Paolillo's (1999) claim that since IRC messages are typed on a keyboard, there is an inclination to use the conventions of English spelling or create innovative ones in an analogy with them. Table 8.5 below presents a taxonomy of IRC specific features found in the Najdi IRC.

Table 8.5 Taxonomy of Najdi IRC features

	CMC feature	Explanation of CMC Technique	Example
<b>1</b>	<b>E-abbreviation</b>	Any reduced form of language. There are six different types of reductions	
	(a)Letter acronyms	Reducing a sentence by stringing the first two or even three initial letters together in a sequence of words.	lol brb tet b wb
	(b)Phonetic acronyms	Using Arabic letters that resemble the sounds of conventional English acronyms	لؤل <i>lūl</i> (laughing out loud) برب <i>brb</i> (be right back) باك <i>bāk</i> (Back)
	(c)Abbreviations	Abbreviating a phrase by omitting certain portions of a phrase, and stringing some of its initials as separate letters	etc bb(blackberry) بي بي <i>bī bī</i> (Black berry)
	(d)Letter homophones	The use of a single letter to represent a word with a similar sound.	u (you) c(see) r(are)
	(e)Numeral	The use of numerals to represent words	4 2 97 78
	(f) Initial letters and numeral homophones	Combining letter and numeral homophones.	b4 u2
<b>2</b>	<b>Addressitivity</b>	Naming the person or thing addressed to avoid confusions	>>>how are u hanouf? lool almarsa??
<b>3</b>	<b>Paralinguistic Cues</b>	A set of features used to create the effect of voice, emotion, gestures and tone in CMC	
	(a) Emoticons	Clustering ordinary symbols such as commas, periods and parentheses to creat devices that convey non-verbal aspects of communication pictorially	:) ;P :O :( :D
	(b) Capitalizations	Capitalizing whole words or sentence.	WOW HAY WHAT
	(c) Multiple Punctuation	Exaggerated use of punctuation marks	weeeen elnas!!!!!!!!!! <i>wīn e-nnās?</i> ( where are the people)  NO!!!!!!!!!!
	(d)Reduplication	Reduplicating letters in words	loooooo waaaaaaw salaaaaaaaaaam
	(e) Vocalizations	Producing discursive forms that approximate speech	“mmm” “ahhhh:

## CHAPTER NINE

### CONCLUDING REMARKS AND IMPLICATIONS

This research is a response to the lack of English-language publications on Arabic CMC, as much as to a wider lack of research on multilingualism of written communication in the Arabic world. As such, the purpose of the present study is to offer an investigation into the use of NA in synchronous CMC. The investigation starts with the hypothesis that NA online is a specific form of language that is composed of three resources: languages (Arabic /English), scripts (Arabic/Roman), and specific graphemic, semiotic, and linguistic features of CMC (e.g., emoticons, abbreviations). Accordingly, the study has attempted to answer a group of research questions in relation to a) the orthographic representation of NA online, (b) code-switching online, and c) the specific CMC features of NA online. The investigation yielded significant results demonstrating the universality of cross linguistic trends among internet users and innovative strategies for dealing with technological marginalization.

The purpose of the present chapter is to emphasize the importance and contribution of this study in the field of CMC research in general and sociolinguistic study of Arabic CMC in particular. This chapter begins with an overall summary of the whole study. It sums up the general description of each key issue of the study and offers a general discussion of the study's findings. Based on these findings, this chapter presents the study's implications, and recommendations for further research identified through the findings.

## 9.1 Summary

In this section, each of the previously proposed research questions will be stated and discussed in terms of the relevant results that contribute to the understanding of how NA is represented online. Such a discussion will be integrated by the relevant data analysis results and an interpretation of these results. References will be made to related research cited in the literature review section when needed.

Research Question 1 of the present study was: what are the characteristics of spoken NA in terms of phonology, morphology and syntax? In order to answer this question, a sample of five young Najdi informants was asked to respond to the RMS survey. Recordings were analyzed and a complete description of NA was obtained. Such a description was deemed necessary to understand NA online. The phonology of NA displays more vowels and five more consonants than MSA does . The fact that short vowels are not represented in the Arabic orthography and these five additional consonants do not have alphabetical counterparts raises the question of how these vowels and consonants are represented online—most importantly, how they are represented by the orthography of another language (English). The morphology of NA is similar to that of MSA in showing a morphological differentiation of different three word categories: nouns, verbs, and particles. The analysis of syntax of NA also displays syntactic categorisations parallel to those of MSA with its basic verbal and nominal sentences.

Research Question 2 was targeted at establishing social profiles of IRC participants in terms of age, gender, and education. Their general linguistic and literacy choices were traced, as well as their general attitudes about using innovative linguistic practices to represent NA online. A total of one hundred and eleven IRC

participants responded to the “IRC Language Questionnaire”, which was designed via Internet Survey Packages, and posted in IRC chat rooms. The IRC participants were asked to report on their social background (age, gender, level of education, English proficiency, and frequency of using the internet and frequency of chatting online). They were also asked to report on their linguistic and literacy choices and attitudes towards using and representing NA online. Descriptive analysis was obtained in order to see the overall patterns of choices and attitudes held by those participants.

In general, results of the “IRC Language Questionnaire” confirm Androutsopoulos’s (2000) claim about the power of youth language that creatively uses non-standard spellings in constructing anti-standard orthography that constitutes a cyberspace code of communication. The majority of IRC participants in the study belongs to the age group 19-25 and 26-35 and have different levels of education and proficiency in English. This suggests that teens and young adults are leading the way towards both the representation of NA online and the social acceptability of this representation. Representing NA online by a mixture of Arabic and Roman scripts mixed with English with extensive graphemic resources (numerals and ASCII symbols) and non-standard spellings seem to function as the code of IRC communication.

Moreover, this specific use of NA is the language IRC participants prefer to use while communicating online compared to other standard language choices (e.g., English only, MSA, or NA in Arabic script). IRC participants have a positive attitude towards using their non-standard variety online, which is largely

associated with IRC participants' perception about the instantaneous synchronous nature of CMC that has changed the "monologue of writing" into a "dialogue" (Baron 2002: 410). The majority of participants do not find the use of NA corruptive or damaging to MSA. This contradicts the (offline) widely held view that writing down Arabic dialect corrupts MSA (Platform and Al Khalil 2007). The formation of this positive attitude towards writing NA in online settings is related partially to the wide increase of NA use in the media, domain names, and websites. It is interesting that a large number of IRC participants not only hold positive attitudes about using NA online but also prefer writing NA via Roman script. This tendency to use Romanized Arabic could also be related to the unfamiliarity participants encounter when writing NA in Arabic script. It is also related to the linguistic freedom those participants feel when writing in Roman script compared to situations under the "language police" (Tagliamonte & Danis 2008:12) associated with the Arabic script and represented by educational institutions.

Research Question 3 is mainly concerned with the question of how to communicate in text-based CMC when one's native language lacks an official standard orthography. Chapter Six has shed some light on the ways that Najdi IRC participants use Roman characters to represent their spoken non-standard dialect in online chat. The major characteristics of this tendency are the use of creative orthographies and the employment of unconventional spellings in an effort to represent sounds that do not exist in MSA. With respect to these sounds, namely dental-alveolar affricate /ts/ and /dz/, velar stop /g/, labio-dental fricative /v/, and labial stop /p/, IRC participants use orthographic conventions that originate both

from MSA literary writings and transcribing these sounds into Roman script. By using the Roman-based script, IRC participants in this study are able to capture the distinctive phonological features of their spoken language, in cases where the Arabic script does not offer such opportunities.

This practice could also be connected with shedding the demonstration of loyalty to the norms imposed by societal institutions, thus a kind of rebellion against the standard orthography. As mentioned in section 6.4, students in the Arabic schools are corrected when they use features of the regional variety in their writings, as this is considered to be a 'mistake' and therefore the use of correct Arabic (i. e. the standard orthography) is reinforced. But here we have a case where adolescents use creative, expressive and oppositional language in an effort to diverge from the adult 'correct' form, promoting therefore their youth identity. The quantitative study, presented in Chapter Five, has revealed that the majority of the subjects who chat on the internet are teenagers and young adults. Since spelling rules are not imposed in the domain of the internet, adolescents are free to use whichever conventions they wish by choosing non-standard spellings. The quantitative analysis also reveals that NA is shown to be used all the time by the majority of participants (58.3%). On the other hand, MSA is seldom used by the 211 participants who completed the questionnaire, as no participants indicated that he/she used it all the time and The total of 115 participants (54.5%) indicated that they never use MSA in their IRC interactions.

The phenomenon of Romanized NA in CMC has apparently emerged as a response to technical limitations to enable Najdi internet users to use their mother tongue online. Romanized NA has established itself as the language of the of

internet, and its use indicates membership in the cyber community. The persisting use of this form of Arabic in CMC when software support for Arabic became available is an interesting phenomenon. The continuing use of Romanized NA in IRC indicates IRC participants' tendency to avoid using the Arabic script and draw on letters of the English language, which has an outstanding status in CMC communication.

In order to answer Research Question 3, a large corpus of IRC data was collected and subjected to textual analysis. The analysis was targeted towards the question of : how do Najdi IRC participants represent the consonants and vowels of NA in Romanized symbols? The analysis reveals interrelations among languages (Arabic/English) and ASCII-characters. The process of Romanizing NA involves a combination of both transliteration of Arabic orthography (e.g., the definite article (e.g. / əl-/ , and / d/) and transcription of spoken sounds (e.g., geminates). The transcription process underlying Romanized NA is mainly phonemic: either by adopting an equivalent English sound or following phonemic conventions of RASs. In cases where consonants do not exist in English, IRC participants tend to adopt orthographic transcription using ASCII symbols that either correspond visually to Arabic letters (e.g. <6>, <6'>, <3>) or derive from an initial of a familiar Arabic word (e.g. <5>).

As far as vowels are concerned, encoding vowels in Romanized NA shows a great deal of inconsistency in IRC. Short vowels, which are predictable in the Arabic orthography, are sometimes left out following the convention of the Arabic orthography, and other times are spelled out following the conventions of the English orthography. Influence from RASs appears clearly in names. The vowels'

positions within words as well as the morphological significance they contribute have influence on the rendering of these vowels in Romanized NA.

The present study reveals that rapid growth of CMC has changed the ways that people use language as new orthographic practices has been developed. The researcher suggests that CMC has a unique capability to change sociolinguistic perspectives in regard to script choice. Thus, technological development should be added to factors of religion and nationalism in shaping and developing cases of digraphia within a community. In the present study, CMC has developed a new synchronic digraphic situation within a CMC community where internet users use two different scripts. These scripts are visually different from each other. Romanized NA is read from left to right, whereas Arabic script is read from right to left. They diverged from different languages, yet represent the same spoken dialect online.

Research Question 4 is concerned with investigating communicative functions of code-switching and script-switching in IRC discourse, as well as grammatical manners of integrating borrowed items into both Arabic and Roman scripts. Quantitative results in Chapter Five indicate that the IRC participants have a tendency to mix NA with English. They also acknowledge their tendency to use code-switching to English. The clear majority (60) indicated they always do, and 49 (23.3%) said they often insert English items in their IRC interactions. Following Poplack and Sankoff's (1984) and Muysken's (1990) accounts of characteristics of borrowing, the present study investigated and categorized the process of morphological adaptations of borrowings in both scripts. In Romanized IRC, morphological integrations of borrowed English items are exhibited through the use of a variety of morphological bound and free morphemes. The integration includes

the use of the Arabic definite article (*al-*), prepositions, pluralisation, the *idafa* construction (possessive constructions), and verb conjugations. By contrast, borrowings occurring in Arabic script are morphologically integrated only by the use of two free morphemes: by the Arabic definite article in the case of nouns, and by the use of the auxiliary verb “*sawwe*” in the case of verbs. However, transliterating borrowings into the Arabic script plays a visual role in integrating these lexical items into the Arabic lexicon.

In other cases, code-switching to NA/English is so excessive to the extent that it would be difficult to decide what is the base language and what is the matrix language. In this particular case, it is the specific factor of bilingualism<sup>20</sup> in English that directs the speaker to the strategy of borrowing. Following the argument presented by Matras (2009), the researcher in the present study assume that one way for the bilingual speaker to prioritise contrast is to eliminate the burden of having to select the appropriate linguistic expression from among the two sub-components (two different languages) within their linguistic repertoire. This tension, according to Matras, is related to the surrounding the planning of an utterance in those situations in which the tension arises due to the clash between anticipated hearer-speaker expectations, and the nature of the link between propositions which the speaker is about to present. In order to ease tension around contrast, the choice among sets is eliminated through the process of fusion into a single repertoire, in which a single set of expressions is generalized for use in both languages.

---

<sup>20</sup> The quantitative study in Chapter Five reveals that the majority of IRC participants are proficient in English.

Applying Auer's conversational analytical approach to code-switching in the IRC data, code-switching has been found to serve communicative functions similar to those of face-to-face conversation. These functions are related to both the discourse and the participants of IRC conversation. However, the lack of visual channels in synchronous CMC with rapid transitions and short exchanges results in an increased reliance on code-switching (Georgakopoulou 1997: 158). This supports Paolillo's (2008) early hypothesis that interactional code-switching should be more frequent in synchronous modes of CMC as they are more speech-like than an asynchronous one.

Assuming that Romanized NA functions as a code of communication that carries social and cultural meanings, it is argued here that switching between scripts is not entirely arbitrary. Rather, it is motivated by creative structuring of the discourse and lack of visual and auditory channels of IRC discourse. The reduced transparency of messages and IRC participants' tendency to write as they talk heighten the IRC participants' meta-linguistic awareness, and lead them to switch between Arabic and Roman scripts. This script-switching enables IRC participants to accomplish "code-centred contextualization cuing, which would be otherwise delegated to different signals" (Georgakopoulou 1997: 158). Applying the methodological tools of Auer's approach to script-switching online, it is found that alternating between scripts online achieves different communicative functions that structure the IRC discourse. Some of these functions are discourse-related functions (such as topicalisation, change of mode, requesting attention and reiteration). Other functions are participant-related, such as those related to IRC participant's relationships (inclusion or exclusion) as well as IRC participants' competence.

Research Question 5 was concerned with investigating the specific features of Najdi IRC that make it a hybrid between speech and writing. The investigation reveals another level of language alternation between common English CMC features identified in the literature and creative innovative CMC based on the Arabic orthography. Several distinctive features were found to contribute to the innovations of IRC communication. Spelling and punctuation marks rules were not followed. New types of abbreviations, multiple punctuation expressions, and ASCII-characters have been used abundantly. Creative and new devices of various kinds based on Arabic have been used (e.g., numeral homophones), in addition to those copied from international chat rooms (e.g., emoticons). Some of these devices have been altered to the Arabic orthography and, thus, better suit the participants' needs (e.g., numeral and phonetic homophones).

## **9.2 Implications**

The findings of the study suggest that CMC has the power to shape innovative linguistic and literacy practices online in regard of Arabic non-standard dialects. Most obvious of all is the use of non-standard dialects that were not previously used for written purposes. The teenagers and young adults are mainly leading the way in both using these dialects online and accepting this use in writing.

Another important finding of the present study is that synchronicity is an important factor in shaping IRC discourse and a fundamental part of what it means for a communication mode to be “conversational” (Paolillo 2011:1). IRC participants need to say in writing what they have been used to saying in speech. This interactive written discourse, to use Ferrara’s (1996) term, results in a hybrid

medium of communication that falls between speech and writing. The findings in this study have profound consequence regarding theorizing on differences between spoken and written communication—an area of research that has provoked strong debate (see section 8.2.1). Researching the specific features of CMC provoked other dimension of the debates: CMC allows informal, rapid, and regional languages to be found in speech. IRC communication is clearly a form of writing formed into letters and words that we read and write. At the same time, IRC communication is dynamic, interactive, and ephemeral.

The reduced transparency of messages heightens the participants' meta-linguistic awareness, and leads them to treat words, letters, and keyboard characters as objects in a rather unconventional way. Internet users have adapted written languages to the needs of the new modes of CMC by producing written, text-based CMC that fall somewhere between the written (at distance) and spoken (face-to-face) modes. In other words, the instantaneous synchronous nature of IRC that has changed the “monologue of writing” into a “dialogue” (Baron 2002: 410) and has led to the appearance “new stylistic conventions” (Baron 2002:404). These stylistic conventions include the following: (a) conventions that are of a writing-specific nature (such as abbreviations) and (b) conventions that are speech-related and compensate for auditory features (such as vocalizations) and paralinguistic features of spoken discourse (such as emoticons and multiple punctuations).

This hybridization of features of written and spoken language minimizes the boundaries that separate speech from writing. IRC participants' tendency to communicate in what Tagliamonte & Denis (2008:8) call “written speech” allows the co-existence of more than one language and more than one script, revealing a

much more complex process of hybridization of text-based CMC compared to those monolingual CMC texts. With the resulting fall of the boundaries between speech and writing, the language alternation (English/colloquial NA) that exists in the speech of bilingual Najdi internet users is reflected in their synchronous CMC writing with alternation between Arabic and Roman scripts. By convention, languages are strongly tied to particular writing systems, and standard orthographies produce recognizable visual forms of these languages (Sebba 2007: 30). Consequently, orthography contributes to the identification of a given text to a particular language. In situations of computer-mediated digraphia, simultaneous use of two different scripts minimizes the distinctions and boundaries between languages that are in contact.

In the bilingual diglossic case of NA, the emergence of Romanized NA is an expression of the need for communicating without the fear of the “correct” orthography, which strongly inhibits both informal features of colloquial spoken and code switching to English (Milroy & Milroy 1991: 66). The ideology of “correct” orthography is represented in educational and cultural institutions whose main role is to confirm that written Arabic language does not reflect any of the informal features of spoken NA and does not include language alternation. Romanized NA allows Najdi internet users to continue expressing themselves in their spoken dialect without worrying about the “correct” norms of writing. In this case, using Roman script in written CMC texts acts as “a venue in which teenagers are free to use all these features together” (i.e., features of formal and informal languages (Tagliamonte & Denis 2008: 38)). Romanized NA is best considered as what Bruthiaux (cited in Elessawi 2011: 274) described as a language variety with “weak

political and administrative control over form and usage;” as well as “the ability to enjoy freedom to accommodate unplanned user-driven change leading to both structural simplification and a degree of creolization as the language adapts to local conditions in a multiplicity of sociolinguistic settings”. Due to the above diglossic consideration, CMC has provided Arabic (including NA) with the form it needed to become a viable choice for young Najdi internet users for communicating in spoken NA in text-based CMC.

This form of text-based CMC differs from face-to-face communication in regard to theorizing on criteria of distinction between borrowed items and code-switching. In contrast to earlier debate on this issue (see Poplack 1988; Muysken 2000; Myers-Scotton 1993), these criteria are mainly based on written instances of borrowing. The role of script in attributing a given English item to either English or Arabic is significant in distinguishing between single-word code-switching and borrowings. It can be hypothesized that script choice can provide a new perspective on borrowing vs. code-switching debate in multilingual CMC. Romanized NA, which shows interrelations between Arabic and English writing systems, minimizes boundaries between languages in contact and, therefore, facilitates the process of integrating English borrowed items into the Arabic lexicon. This integration is shown by a variety of morphological adaptations.

Investigating code-switching in the present study tests the theoretical tools of oral code-switching in analysing written code-switching in CMC. Though the theoretical tools that have been used for analyzing oral productions of code-switching proved to be relevant for the analysis of written code-switching, written code-switching does not show the same features of spoken code-switching. The

occurrence of written code-switching in IRC discourse is shaped by social and pragmatic factors, as well as technological features of CMC synchronicity (Androutsopoulos forthcoming). Lack of visual and auditory channels of spoken discourse resulted in more reliance on code-switching to achieve functions that could be typically accomplished by facial expressions, tone of voice, and gestures.

An important finding that emerged from the study is that CMC demonstrates the universality of cross linguistic trends among internet users using different languages. These trends include adaptation of writing systems to suit the mode of CMC, language alteration, innovative strategies for dealing with technological marginalization, and innovative strategies to compensate for the medium lacking visual and auditory channels. These trends demonstrate the fact that CMC has the potential to shape language use in new ways.

Regarding the bilingual diglossic situation of the Arab world, though CMC fosters the use of non-standard dialects that were not previously used for writing, it adds to the complexity of the linguistic situation there. It combines these emerging hybridized written texts with the already existing multiplicity levels that mark the diglossic situation in the Arab world. Reasons of the emergence of this process of hybridization are communicative convenience (as revealed by textual analysis) and social prestige (as reported by the “IRC Language Questionnaire”).

### **9.3 Suggestions for further study**

The research methodology followed in this study helped in exploring the representation of NA online, mapping out interrelation among languages, scripts, and

technology. However, there is yet little evidence bearing on the question of how CMC has “radically changed the face of literacy” (Merchant 2001: 294) within the Arabic context, which might remain an issue for future research. The following recommendations are made for further research:

- It would be advantageous to conduct similar studies to the present one with some key modifications. Succeeding research would benefit from, for example, investigating turn-taking, coherence, and CMC discourse structuring. Then, results would have to be compared to face-to-face interaction and written interactions to see whether there are significant differences in these respects.
- An analysis and investigation of the use of NA online could be further confirmed by the replication of the procedures of the present study with another use of non-standard Arabic dialect online. Indeed, further research along this line of enquiry would benefit from duplicating CMDA adopted in the present study. General characteristics of Arabic CMC would be further revealed.
- As all participants in this study were anonymous, an area worth exploring would be language variation online as far as social backgrounds of participants is concerned. Experimental CMC studies comparing male and female language use could reveal interesting findings that might enrich the sociolinguistic CMC area of investigation.

- This study has highlighted the representation of NA online through using online survey questionnaire and textual IRC data analysis. While replication studies are certainly needed, further studies in this area could be conducted by utilizing a variety of data sources (e.g., social networks websites (e.g., Facebook, Twitter, BBS, or emails). A comparison of synchronous and asynchronous communication might be carried out.
- After obtaining sufficient knowledge on how non-standard Arabic dialects are used online from these suggested recommendations, it would be worth recommending a framework for the study of multilingual practice in Arabic CMC and exemplifying it with findings from different case studies. What is meant here by framework is a systematic CMC treatment that takes into account properties of digital media, including literacy competences, constraints of keyboard production, and visual language.

## REFERENCES

- Abboud, P. 1978. The vowel of the imperative prefix in Najdi Arabic. In: Jazayeri, E. and Polome, W. eds. *Linguistics and Literary Studies in Honour of Archibad A. Hill, ii Descriptive Linguistics*. 1st ed. Paris: The Hague.129-138.
- Abboud, P. 1964. *The syntax of Najdi Arabic*. PhD Dissertation. University of Austin, Texas.
- Abod, P. 1979. The verb in northern Najdi Arabic. *Bulletin of the School of Oriental and African studies* 4, 647-499.
- Aboeazz, M. 2010. A Latinised Arabic for all? Issues of representation, purpose and audience. *Proceedings of the International Symposium on Arabic Transliteration Standard: Changes and Solutions*. Abu Dhabi, UAE.
- Alfonzetti, G. 1998. The conversational dimension in code-switching between Italian and dialect in Sicily. In: Peter, A. eds. *Code-switching in conversation*. 1st ed. London: Routledge. 18-214.
- Alothman, E. 2008. *Vocabulary learning strategies in correlation with proficiency level*. Germany: VDM.
- Alothman, E. 2010. Computer mediated communication: New varieties of language. *Proceedings of 4th Saudi International Conference - SIC 2010*. University of Manchester , UK.
- Alothman, E. forthcoming. Code-switching in social media: The case of English/Arabic Facebook. *Proceedings of Language and Social Media Seminar- BAAL*. University of Leicester, UK.
- Alothman, E. forthcoming-a. The impact of global communication on dialect use in Saudi Arabia: A linguistic investigation of Najdi Arabic in Internet Relay Chat. Paper presented at the Third Gulf Research Meeting (GRM), Cambridge, 11 - 14 July 2012, organized by the Gulf Research Center Cambridge (GRCC) at the University of Cambridge, UK.
- Alothman, E. forthcoming-b. Romanized Arabic: Past and present. *Proceedings of the 43th Seminar for Arabian Studies*,UK.
- AlSweel, A.1992. Case and mood in Arabic. *King Saud University* 4, 38-56.
- AlSweel, A. 1981. *The verbal system of Najdi Arabic. A morphological and*

- phonological study*. MA Dissertation. University of Washington.
- Alvarez-Caccamo, C. 1998. From “switching code” to “code-switching”: Towards a reconceptualization of communicative codes. In: Auer, P. eds. *Code-switching in conversation: Language, interaction and identity*. 1st ed. London: Rutledge. 29-50.
- Androutsopoulos, A. and BeiBwenger, M. 2008. *Introduction: Data and methods in computer-mediated communication*. Available at: [www.languageatinternet.de/urn:nbn:de](http://www.languageatinternet.de/urn:nbn:de) [Accessed: 4 November 2011].
- Androutsopoulos, J. forthcoming. Code-switching in computer-mediated communication. In: Herring, S., Stein D & Virtanen T.eds. *Handbook of the pragmatics of CMC*. 1st ed. Mouton de Gruyter.
- Androutsopoulos, J. 2011. Language change and digital media: A review of conceptions and evidence. In: Coupland , N. and Kristiansen, T. eds. *Standard languages and language standards in a changing Europe*. 1st ed. Oslo: Novus.145-161.
- Androutsopoulos, J. 2010. The study of language and space in media discourse. In: Auer, P. and Schmidt, J. eds. *Language and space: An international handbook of linguistic variation. Volume I: Theory and methods*. 1st ed. Berlin: de Gruyter.740-758.
- Androutsopoulos, J. 2009. ‘Greeklish’: Transliteration practice and discourse in a setting of computer-mediated digraphia. In: Georgakopoulou, A. and Michael, S. eds. *Standard languages and language standards: Greek, past and present*. 1st ed. London: Farnam. 221-249.
- Androutsopoulos, J. 2007. Language choice and code-switching in German-based diasporic web forums. In: Brenda, D. and Herring, S. eds. *The multilingual internet*. 1st ed. Oxford: Oxford University Press. 340-361.
- Androutsopoulos, J. 2006. Introduction: Sociolinguistics and computer-mediated communication. *Journal of Sociolinguistics* 10(4), 419-438.
- Androutsopoulos, J. 2004. Exploring language variation on the internet: Regional speech in a chat community. In: Gunnarsson, B. eds. *Language variation in Europe. Papers from ICLaVE* . 1st ed. Uppsala: Uppsala University Press. 99-111.
- Androutsopoulos, J. 2001. Orthographic variation in Roman-alphabetized Greek. In:

- Agouraki, Y. eds. *4th International Conference on Greek Linguistics, Nicosia*. University Studio Press . 1st ed. 383-390.
- Androutsopoulos, J. 2000. Non-standard spellings in media texts: the case of German fanzines. *Journal of Sociolinguistics* 4(4), 514-533.
- Androutsopoulos, J. 1998. Orthographic variation in Greek e-mails: a first approach. *Glossa* 46, 49-67.
- Androutsopoulos, J. and Georgakopoulou, A. 2008. Youth, discourse, and interpersonal management. In: Gerd, A. and Ventola, E. eds. *Handbook of interpersonal communication*. 1st ed. Berlin: de Gruyter. 2-20.
- Angermeyer, P. 2005. Spelling: bilingualism: Script choice in Russian American classified ads and signage. *Language in Society* 34 (4), 493-531.
- Anis, I.1952. اللهجات العربية *fī Ilhajaat I çabiyyah* The Arabic dialects. Cairo: Maktabat lanj lo almaSriyah.
- Apple, R. and Muysken, P. 1987. *Language contact and bilingualism*. London: Edward Arnold.
- Auer, P. 2000. A conversation analytic approach to code-switching and transference. In: Wei, L. eds. *The bilingual reader*. 1st ed. London and New York: Routledge. 115- 145.
- Auer, P. 1999. From code-switching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* 3, 309-332.
- Auer, P. 1995. The pragmatics of code-switching: A sequential approach. In: Milroy, L. and Muysken, P. eds. *One speaker, two languages: Cross-disciplinary perspectives on code-switching*. 1st ed. Cambridge: Cambridge University Press .115-135.
- Auer, P. 1988. *Code-switching in conversation: Language, interaction and identity*. London: Routledge.
- Auer, P. 1984. *Bilingual conversation*. Amsterdam: John Benjamins .
- Badawi, B. 2012. *An intonational study of colloquial Riyadhhi Arabic*. Ph D Dissertation. UCL, London.
- Baron, N. 2005. Instant messaging and the future of language. *Communication and the ACM* 48 (7), 29-31.
- Baron, N. 2001. Who sets the e-mail style? Prescriptivism, coping strategies and

- democratizing communication access. *The Information Society* 18 (5), 403-413.
- Barton, D. 2000. *Situated literacies: Reading and writing in context*. London and New York: Routledge.
- Barton, D. 2007. *Literacy: An introduction to the ecology of written language*. 2nd ed. Oxford: Blackwell.
- Beesley, K. 1998. *Romanization, transcription and transliteration*. Available at: <http://www.xrce.xerox.com/competencies/content-analysis/arabic/Romanization/> [Accessed: 10 May 2010].
- Biber, D. 1999. *Variation across speech and Writing*. Cambridge: Cambridge University Press.
- Bloom, J. and Gumperz, J. 1972. Social meaning in linguistic structures: Code switching in Northern Norway. In: Gumperz, J. and Hymes, D. eds. *Directions in sociolinguistics: The Ethnography of Communication*. 1st ed. New York: Holt, Rinehart, and Winston. 407-434.
- Bloomfield, L. 1933. *Language*. New York: Holt.
- Bogdan, R. and Biklen, S. 1982. *Qualitative research for education: An introduction to theory and methods*. Available at: <http://www.edu.edts.org.sg/> [Accessed: 20 May 2009].
- Bokamba, E. 1989. Are there syntactic constraints on code-mixing? *World Englishes* 8 (3), 277-292.
- Callahan, L. 2002. The matrix language frame model and Spanish/English code-switching in fiction. *Language and Communication* 22, 1-16.
- Chun, C. 2007. *Code-switching in conversation: A case study of Taiwan*. PhD Dissertation. University of Manchester.
- Collot, M. and Belmore, N. 1996. Electronic language: A new variety of English. In: Herring, S. eds. *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*. 1st ed. Amsterdam: John Benjamins Publication Company. 13-27.
- Coomber, R. 1997. Using the internet for survey research. *Sociological research online* 2 (2), 10-19.
- Cornelius, C. and Boos, M. 2003. Enhancing mutual understanding in synchronous computer-mediated communication by training: Trade-offs in judgmental tasks. *Communication Research* 30 (2), 147-177.

- Coulmas, F. 2003. *Writing systems: An introduction to their linguistic analysis*. Cambridge: Cambridge University Press .
- Coulmas, F. 1996. *The Blackwell encyclopedia of writing system*. Oxford: Blackwell.
- Couper, M. & Traugott, M. 2001. Web survey design and administration. *Public opinion quarterly* 65 (2), 230-254.
- Creswell, J. 2003. *Research design: Qualitative, quantitative and mixed methods approaches*. Upper Saddle River, NJ: Merrill Press.
- Crystal, D. 2011. *internet linguistics*. London: Routledge.
- Crystal, D. 2008. *Txting: the gr8 db8* . Oxford: OUP.
- Crystal, D. 2001. *Language and the internet*. Cambridge: Cambridge University Press.
- Crystal, D. 1995. *The Cambridge encyclopedia of the English language*. Cambridge: Cambridge University Press.
- Davies, A. 2004. General introduction. Applied Linguistics: Subject to discipline? In Davies, A & Elder, C. eds. *The Handbook of Applied Linguistics*. 1st ed. Oxford: Blackwell. 1-16.
- Dale, I. 1980. Digraphia. *International Journal of the Sociology of Language* 25 (1), 5-13.
- Danet, B. 2001. *Cyberpl@y: Communicating online*. New York: BERG.
- Danet, B. 1998. Text as mask: Gender and identity on the internet. In: Jones, S. eds. *Cybersociety 2.0*. 1st ed. Sage: Thousand Oaks, CA. 139-149.
- Danet, B. and Herring, S. 2007. *The Multilingual internet: Language, culture and communication online*. New York: Oxford University Press.
- December, J. (1996) Units of analysis for internet communication. *Journal of Computer-Mediated Communication* 1 (4), 1-20.
- December, J. (1992) *Characteristics of oral cultural in discourse on the net*. Available at: <http://www.december.com/john/papers/psrc93>. [Accessed: 1 September 2008].
- DeFrancis, J. 1984. Digraphia. *Words* 35 (1), 59-66.
- Durham, M. 2007. Language choice on a Swiss mailing list. In: Danet, B. and Herring, S. eds. *The multilingual internet: Language, culture and communication online*. 1st ed. New York: Oxford University Press. 319-337.
- Dutton, H. 1998. The first amendment online. Santa Monica's public electronic

- network. In: Tsagarousianou, R. and Tambini, D. eds. *Cyberdemocracy: Technology, Cities and civic networks*. 1st ed. New York: Routledge.125-151.
- El Essawi, G. 2011. Arabic in Latin script in Egypt: Who uses in and why? In: Allssa, A. and Laila, S. eds. *Global English and Arabic: Issues of language, culture and identity*. 1st ed. Oxford: Peter Lang.253-280.
- Ferguson, C.1968. Language development. In: Fishman, J. and Ferguson, C. eds. *Language problems of developing nations*. 1st ed. New York: Wiley.27-35.
- Ferguson, C. 1959. Diglossia. *Word* 15, 325–340.
- Ferrara, K. and Brunner, H. 1991. Interactive written discourse as an emergent register. *Written communication* 8 (1), 8-34.
- Fielding, N.G., Lee, R.M., & Blank, G. 2008. *The SAGE handbook of web-based research methods*. Thousand Oaks, CA: Sage Publications.
- Fishman, J. 1967. Bilingualism with and without diglossia; Diglossia with and without bilingualism. In: W, L. eds. *The bilingual reader*. 1st ed. London and New Rutledge.81-88.
- Fishman, J. 1965. Who speaks what language to whom and when? *Linguistics* 2 (1), 67-88.
- Fouser, Robert J.,Inoue, N., and L., Chungmin. 2002. The pragmatics of orality in English, Japanese and Korean computer-mediated communication. *Words on the Web* 5, 52-55.
- Gao, L. 2006. Language contact and convergence in computer-mediated communication. *World Englishes* 25(2), 299–308.
- Gardner-Chloros, P. 2009. *Code-switching*. Cambridge: Cambridge University Press.
- Gee, P. 1991. *Social linguistics and literacies: Ideology in discourses*. London: Taylor & Francis.
- Georgakopoulou, A. 1997. Self-presentation and interactional alliances in email discourse: the style- and code-switches of Greek messages. *International Journal of Applied Linguistics* 7 (2), 141-164.
- Grenfell, M. and Harris, V.1996. *Modern languages and learning strategies: In theory and practice*. London: Rutledge.
- Grivelet, S. 2001. Digraphia: Writing systems and society. *International Journal of the Sociology of Language* 150, 150 (1), p.1-50.
- Grosjeans, G. 1982. *Life with two languages*. MA Dissertation. Cambridge

University.

Gumperz, J. 1982. *Discourse strategies*. Cambridge: Cambridge University Press.

Habash, N. 2010. *Introduction to Arabic natural language processing*. Morgan & Claypool: Morgan & Claypool.

Halliday, M. 1985. *Spoken and written language*. Oxford: Oxford University Press.

Haspelmath, M. 1997. *Indefinite pronouns*. Oxford Studies in Typology and Linguistic Theory. Oxford: Oxford University Press.

Hass, W. and Hornby, P. 1970. Use of contrastive stress by preschool children. *Journal of Speech and Hearing Research* 1, 395-399.

Hatch, E. and Farhady, H. 1983. *Research design and statistics for applied linguistics*. New York: Newbury House.

Haugen, E. 1950. The analysis of linguistic borrowing. *Language* 26, 210-231.

Haven, P. 2007. *Doing conversation analysis: A practical guide*. London: Sage Publication Ltd.

Herring, S. 2008. Virtual community. In: Given, L. *Encyclopedia of quantitative research methods*. 1st ed. SAGE.

Herring, S. 2004. Computer-mediated discourse analysis: An approach to researching online behavior. In: Barab, S. and Kling, R. eds. *Designing for virtual communities in the service of learning*. 1st ed. New York: Cambridge University Press. 338-376.

Herring, S. 2003. Gender and turn allocation in a Thai chat room. *Journal of Computer-Mediated Communication* 9(1), 56-98.

Herring, S. 2001. *Computer-mediated discourse*. Available at: <http://ella.slis.indiana.edu/~herring/cmd.pdf> [Accessed: 25 March 2009].

Herring, S. 1996. *Computer-mediated communication: Linguistic, social and cross-cultural perspectives*. Amsterdam: John Benjamin Publication Company.

Herring, S. 1994. Politeness in computer culture: Why women thank and men flame. *Cultural performances: Proceedings of the Third Berkeley Women and Language Conference*. 278-294.

Herring, S. 1993. Gender and democracy in computer-mediated communication. *Electronic Journal of Communication* 3(2), 14-34.

Herring, S. 1992. *Gender and participation in computer-mediated linguistic*

*discourse*. Available at:

<http://ella.slis.indiana.edu/~herring/participation.1992.pdf>

[Accessed: 20 May 2010].

- Herring, S. and Paolillo, J. 2006. Gender and genre variation in weblogs. *Journal of Sociolinguistics*, 10 (4), 339-459.
- Herring, S. and Kapidzic, S. 2011. Gender, communication, and self-presentation in teen chatrooms revisited: Have patterns changed? *Journal of Computer-Mediated Communication* 17(1), 39-59.
- Hewson, C. 2003 Conducting Psychological Research on the Internet. *The Psychologist*, 16(6), 290-293.
- Hewson, C., Yule, P., Laurent, D. & Vogel, C., 2003. Internet research methods: A practical guide for the social and behavioural research. London, Thousand Oaks, New Delhi: Sage.
- Hinrichs, L. forthcoming. How to spell the vernacular: a multivariate study of Jamaican e-mails and blogs. In: Sebba, M. and Androutsopoulos, J. eds. *Orthography as Social Action: Scripts, Spelling*. 1st ed. Berlin-New York: Mouton de Gruyter. 13-27.
- Hoffman, C. 1991. *Introduction to bilingualism*. New York: Longman.
- Hughes, R. 1996. *English in speech and writing*. London: Routledge.
- Hyland, K. 2002. *Teaching and researching writing*. London: Routledge.
- Ibn Aqeel, I. (1980). شرح ابن عقيل *šarḥ ibn ʿaqaḗl*. Cairo: Maktabat lanj lo almaSriyah.
- Ingham, B. 1996. *Najdi Arabic: Central Arabian*. Amsterdam: John Benjamin Publication Company.
- Johnstone, T. 1967. *Eastern Arabian dialect studies*. London: Oxford University Press.
- Johnstone, T. 1963. The affrication of Kaf and Gaf 'in the Arabic dialects of the Arabian Peninsula. *Journal of Semitic Studies* 8(2), 625-678.
- Joope, M. 2000. *The Research Process*. Available at:  
<http://www.ryerson.ca/mjoppe/rp.html> [Accessed: 22 May 2009].
- Khayat, M. 2004. في سبيل العربية *fī sabīl al-ʿarabiyya*. *Towards Arabic*. 4th ed. Cairo: Wahba Library.
- King, R. 2001. The poisonous potency of script: Hindi and Urdu. *International Journal of the Sociology of Language* 150, 43-59.

- Labov, W. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.
- Laroussi , F. 2011. *Code-switching: Languages in contact and electronic writing*. Peter Lang: Frankfurt am Main.
- Lee, C. 2007. Linguistic features of email and ICQ instant messaging in Hong Kong. In: Danet, B. and Herring, S. eds. *The multilingual internet: Language, culture and communication Online*. 1st ed. New York: Oxford University Press.184-208.
- Lee, C. 2002. Literacy practices in computer-mediated communication in Hong Kong . *The Reading Matrix* 2 (2), 1-25.
- Lee, K. 2011. Affordances and text-making practices in online instant messaging. *Written communication* 24(3), 223-249.
- Lehn, W.1967. Vowel contrasts in Najdi Arabic. In: Stuart, D. eds. *Linguistic studies in memory of Richard Slade Harrell*. 1st ed. Washington D.C: Georgetown University Press. 123-131.
- Li, Wei. 2005. 'How can you tell?' Towards a common sense explanation of conversational code-switching. *Journal of Pragmatics* 37(2), 375-389.
- Li, Wei. 2002. 'What do you want me to say?' On the conversation analysis approach to bilingual interaction . *Language in Society* 31, 59-180.
- Li, Wei. 1994. *Three generations, two languages, One family: Language choice and language shift in a Chinese community in Britain*. Clevedon: Multilingual Matters.
- Li, Wei. and Milory, L. 1995. Conversational code-switching in a Chinese community in Britain: A sequential analysis. *Journal of Pragmatics* 23, 281-299.
- MacSwan, J. 2000. The architect of the bilingual language faculty: Evidence from intrasentential code-switching. *Bilingualism: Language and cognition*, 3(1), 37-54.
- Magner, T. 2001. Digraphia in the territories of the Croats and Serbs. *International Journal of the Sociology of Language* 150 (1), 11-26.
- Mann, C. & Stewart, F. 2000. *Internet communication and qualitative research: A handbook for researching online*. London, Thousand Oaks, New Delhi: SAGE.

- Maschler, Y. 1994. Metalanguaging and discourse markers in bilingual conversation. *Language in Society* 23, 325-366.
- Matras, Y. 2009. *Language contact*. Cambridge: Cambridge University Press.
- Matras, Y. 2002. *Romani: A linguistic introduction*. Cambridge: Cambridge University Press.
- Matras, Y. 1999. Writing Romani: The pragmatics of codification in a stateless language. *Applied Linguistics* 20(4): 481- 502.
- Matras, Y. 1998. Utterance modifiers and universals of grammatical borrowing. *Linguistics* 36. 281-331.
- Meecham, Marjorie and Janie Rees-Miller. 2001. Language in social contexts. In W. O'Grady, J. Archibald, M. Aronoff and J. Rees-Miller (eds) *Contemporary Linguistics*. Boston: Bedford/St. Martin's. 537-590.
- Merriam, S.B. 2001. *Qualitative research and case study application in education*. San Francisco, CA: Jossey-Bass Publishers.
- Milroy, J. and Milroy, L. 1991. *Authority in language: Investigating language prescription and standardization*. London: Routledge & Kegan Paul.
- Milroy, L. and Gosdon, M. 2003. *Sociolinguistics: method and interpretation*. Oxford: Blackwell.
- Milroy, L. 1987. *Observing and analyzing natural language: a critical account of sociolinguistic method*. Oxford: Blackwell.
- Myers-Scotton, C. 2002. *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, C. 1999. Explaining the role of norms and rationality in code-switching. *Journal of Pragmatics* 32, 1259-1271 .
- Myers-Scotton, C. 1993a . *Duelling languages: grammatical structure in codeswitching*. Oxford: Oxford University Press.
- Myers-Scotton, C. 1993b . *Social motivations for code switching: evidence from Africa*. New York: Oxford University Press.
- Myers-Scotton, C. 1992. Code-switching as a mechanism of deep borrowing, language shift, and language death. In: Brenzinger, M. eds. *Language death. factual and theoretical explorations with special reference to East Africa*. 1st ed. Berlin: Mouton de Gruyter. 31-58.
- Myers-Scotton, C. 1988. Self-enhancing codeswitching as interactional power.

- Language and Communication* 8(3), 199-211.
- Myers-Scotton, C. 1989. Code-switching with English: Types of switching, types of communities. *World Englishes* 8(1), 333-346.
- Myers-Scotton, C. 1987. Code-switching and types of multilingual communities. In: Lowenberg, P. eds. *Georgetown University Roundtable on Languages and Linguistics*. 1st ed. Washington : Georgetown University Press.61-82.
- Myers-Scotton, C. 1983. The negotiation of identities in conversation: A theory of markedness and code choice. *International Journal of the Sociology of Language* 44, 115-136.
- Myers-Scotton, C. and Bolonyai, A. 2001. Calculating speakers: Codeswitching in a rational choice model. *Language in Society* 30, 1-28.
- Nishimura, N. 2008. *Aspects of Japanese computer-mediated communication: Linguistic and socio-cultural perspectives*. PhD Dissertation. Sheffield Hallam University.
- Nishimura, Y. 2007. Linguistic innovations and interactional features in Japanese BBS Communication. In: Danet, B. and Herring, S. eds. *The multilingual internet: Language, culture and communication online*. 1st ed. New York: Oxford University Press.163-180.
- Olson, D.1988. Mind and media: The epistemic functions of literacy. *Journal of Communications* 38 (3), 45-53.
- Ong, J.1982. *Orality and literacy: The technologizing of the word*. London and New York: Methuen.
- Owens, J. 2005. Bare forms and lexical insertions in code-switching: A processing based account. *Bilingualism: Language and cognition* 8(1), 26-49.
- Palfreyman, D. and Alkhalil, M. (2007) A Funky Language for Teenzz to Use: Representing Gulf Arabic in Instant Messaging. In: Danet, B. and Herring, S. eds. *The multilingual internet: Language, culture and communication Online*. 1st ed. New York: Oxford University Press. 43-60.
- Paolillo, J. 2011. Conversational code-switching on usenet and Internet Relay Chat. *Language@Internet* 1, 1-21.
- Paolillo, J. 1999. Language variation on Internet Relay Chat: A social network approach. *Journal of Sociolinguistics* 5(2), 180-211.
- Paolillo, J. 1996. *Language choice on soc.culture.punjab*. Electronic Journal of

- Communication 6(3), Available at:  
<http://www.cios.org/EJCPUBLIC/006/3/006312>. [Accessed: 23 May 2009].
- Parker, L. 1992. Collecting data the e-mail way. *Training and development* 23(52), 55-64.
- Poplack, S. and Meechan, M. 1995. Patterns of language mixture: Nominal structures in Wolof-French and Fongbe-French bilingual discourse. In: Milroy, L. and Muysken, P. eds. *One speaker, two languages*. 1st ed. Cambridge: Cambridge University Press.
- Poplack, S. and Sankoff, D. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26, 47-104.
- Prochazk, T. 1988. *Saudi Arabian Dialects*. London: Kegan Paul.
- Rabin, C. 1951. *Ancient West Arabian*. London: Taylor's Foreign Press.
- Rajah-Carrim, A. 2007. Mauritian Creole and Language Attitudes in the Education System of Multiethnic and Multilingual Mauritius. *Journal of Multilingual and Multicultural Development* 28 (1), 51-60.
- Rampton, B. 1998. *Speech community*. Handbook of Pragmatics. Amestrdam/Philapdelphia: John Benjamins.
- Raymond, E. 1996. The new Hacker's dictionary. In: Raymond, E. eds. *Ethical standards in social work: A critical review of the NASW codes of ethics*. 1st ed. Washington D.C: NASW Press.
- Reid, R. 1991. *Electropolis: Communications and community on Internet Relay Chat*. PhD Dissertation. University of Melbourne, Australia.
- Richard, B. 2003. *Designing Virtual Worlds*. New York: New Riders publishing.
- Rheingold, H. 1993. *The virtual community: Homesteading on the electronic frontier*. New York: Routledge.
- Rodino, M. 1997. Breaking out of binaries: Reconceptualizing gender and its relationship to language in computer-mediated communication. *Journal of Computer-Mediated Communication* 3 (3), 56-81.
- Romaine, S. 1994. *Bilingualism*. 2nd ed. Oxford: Blackwell.
- Schonfeldt, J. 2001. Chat-communication. Interaction, society and identity. In: BeiBwenger, M. *Synchronous computer-mediated communication*. 1st ed. Sweden: Stuttgart.
- Schuldt, P. and Totten, J. 1994. Electronic Mail vs. Mail survey response rates.

- Marketing Research* 6, 36-44.
- Sebba, M. 2009. Sociolinguistic approaches to writing systems research. *Writing Systems Research* 1 (15), 35-49.
- Sebba, M. 2007. *Spelling and society: The culture and politics of orthography around the world*. Cambridge: Cambridge University Press.
- Sebba, M. 2002. Orthography and ideology: Issues in Sranan spelling. *Linguistics: An interdisciplinary journal of the language science* 38 (5), 925-949.
- Sebba, M. 1998. Phonology meets ideology: The meaning of orthographic practices in British Creole. *Language problems and language planning* 22(1), 19-47.
- Segertad, S. 2002. *Use and adaptation of written language to the conditions of computer-mediated communication*. PhD Dissertation. Goteborg University, Sweden.
- Sinclair, J. and Coulthard, M. forthcoming. *Towards an analysis of discourse*. London: OUP.
- Siebenhaar, Beat (2006). Code choice and code-switching in Swiss-German Internet Relay Chat rooms. *Journal of sociolinguistics* 10(4): 481-506.
- Smeaton, B. 1973. *Lexical expansion due to technical change*. Bloomington: Indiana University Press.
- Smith, A. and Leigh, L. 2012. Virtual subjects: Using the internet as an alternative source of subjects and research environment. *Behavior Research Methods, Instruments, & Computers* 29, 496-505.
- Snakoff, D., Poplack, S. & Vanniarajan S. 1990. The case of the nonce loan in Tamil. *Language variation and Change* 2, 71-101.
- Schonlau, M., Fricker, R. & Elliott, M. (2002). *Conducting research surveys via E-mail and the Web*. Santa Monica: Rand.
- Storrer, A. 2001. Sprachliche Besonderheiten getippter Gespräche. In: Beibwenger, M. eds. *Synchronous computer-mediated communication*. 1st ed. Sweden: Stuttgart.3-24.
- Strauss, A. and Corbin, J. 1999. *Basics of qualitative research Grounded theory procedures and techniques*. Newbury Park: Sage Publications.
- Street, B. 1995. *Social literacies :Critical approaches to literacy in development, ethnography and education*. Harlow: Longman.
- Street, V. 1993. *Cross-cultural approaches to Literacy in development, ethnography*

- and education*. Harlow: Longman.
- Street, V. 1984. *Literacy in theory and practice*. Cambridge: Cambridge University Press.
- Stroud, C. 1992. The problem of intention and meaning in code-switching. *Text* 12, 127-155.
- Tagliamonte, S. and Denis, D. (2008) Linguistic ruin? LOL! Instant messaging and teen language. *American Speech* 83 (1), 3-34.
- Tashakkori A. & Teddlie, C. 1998. *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks: Sage Publication
- Themistocleous, T. 2008. *The use of Cypriot Greek in Synchronous computer-mediated communication*. PhD dissertation. University of Manchester.
- Troest, M. 1998. *Computer-mediated communication: lingua ex machina*. Available at: <http://www.sprog.auc.dk/~motr96/marvin/www/library/> [Accessed: 2 September 2009].
- Tse, A .2012. Comparing two methods of sending out questionnaires: E-mail versus mail. *Journal of the Market Research Society* 37, 441-446.
- Tseliga, T. 2007. It's All Greeklish to me! Linguistic and sociocultural perspectives on Roman-Alphabetized Greek in asynchronous computer-mediated communication. In: Danet, B. and Herring, S. eds. *The multilingual internet: Language, culture and communication online*. 1st ed. New York: Oxford University Press.116-140.
- Tseliga, T. 2004. Exploring language variation on the internet: Regional speech in a chat community. In: Androutsopoulos, J. and Ziegler, E. eds. *Language variation in Europe: Papers from ICLaVE 2*. 1st ed. Uppsala University Press: Uppsala.
- Vachek, J. 1976. *Selected writings in English and general linguistics* . Prague: Academia.
- Versteegh, K. 2001.*The Arabic Language*. Edinburgh: Edinburgh University Press.
- Wardhaugh, R. 1998. *An Introduction to sociolinguistics*. 3rd ed. Oxford: Blackwell.
- Wardhaugh, R. 2006. Linguistic insights into the reading process. *Language Learning* 18 (3), 235-252.
- Warschauer, M., ElSaid G. and Zohry A. 2007. Language choice online: Globalization and identity in Egypt. In: Danet, B. and Herring, S. eds. *The*

- multilingual internet: Language, culture and communication online*. 1st ed. New York: Oxford University Press.303-317.
- Warschauer, M., ElSaid G. and Zohry A. 2002. *Language choice online: Globalization and identity in Egypt: Journal of Computer-Mediated Communication*7(4), Available at: <http://jcmc.indian.edu/vol7/issue4/warschauer.html> [Accessed: 3 August 2009].
- Watson, J. 2002. *The phonology and morphology of Arabic*. Oxford: Oxford University Press.
- Weinreich, U. 1953. *Languages in contact*. The Hague: Mouton.
- Werry, C. 1996. Linguistic and interactional features of Internet Relay Chat. In: Herring, S. eds. *Computer-mediated communication: Linguistic, social and cross- cultural perspectives*. 1st ed. Amsterdam: John Benjamin Publication Company.47-64.
- Wright, K. 2005.Researching internet-based populations: Advantages and disadvantages of online survey research, online questionnaire authoring software packages, and web survey services. *Journal of Computer-Mediated Communication* 10 (3), 11-26.
- Yaghan, M. 2008.'Arabizi': A Contemporary style of Arabic slang. *Design Issues* 24 (2), 39-52.
- Yates, S. 1996. Oral and written linguistic aspects of computer conferencing. In: Herring, S. eds. *Computer-mediated communication: Linguistic, social and cross- cultural perspectives*. 1st ed. Amsterdam: John Benjamin Publication Company.29-46.
- Yu, G. 2001. *A Dictionary of China's Internet language*. China: Beijing.

## Appendices

### Appendix A. The Arabic Version of the Romani-Morph-Syntax Dialect Survey

(RMS)

ID Number	English Phrase	Arabic Phrase
1	person	شخص
2	girl	بنت
3	boy	ولد
4	baby	وليد
5	man	رجل
6	woman	إمرأة
7	father	أب
8	mother	أم
9	sister	أخت
10	brother	أخ
11	child	طفل
12	grandfather	جد
13	grandmother	جدة
14	wolf	ذئب
15	donkey	حمار
16	fish	سمكة
17	cow	بقرة
18	fox	ثعلب
19	bull	ثور
20	camel	جمل
21	animal	حيوان
22	worm	دودة
23	bird	طير
24	goat	مغزاة

25	cat	قِطْ
26	dog	كَلْب
27	gazelle	غَزَال
28	cattle	عَنَم
29	pigeon	حَمَامَة
30	rabbit	أَرْنَب
31	sheep	خَرُوف
32	heart	قَلْب
33	blood	دَم
34	leg	سَاق
35	face	وَجْه
36	tail	ذَيْل
37	ear	أُذُن
38	head	رَأْس
39	tongue	لِسَان
40	chest	صَدْر
41	tooth	سِنَّ
42	belly	بَطْن
43	foot	قَدَم
44	mouth	فَم
45	arm	ذِرَاع
46	eye	عَيْن
47	finger nail	ظَفْر
48	bones	عَظْم
49	guts	أَحْشَاء
50	body	جِسْم
51	skin	جِلْد
52	back	ظَهْر
53	hand	يَد
54	day	يَوْم

55	night	ليلة
56	water	ماء
57	fire	نار
58	stars	نَجْمَة
59	palms	نخلة
60	river	نَهْر
61	mosquito	بعوضة
62	wood	خَشْبَة
63	inlet	خَوْر
64	sea	بَحْر
65	tree	شَجَرَة
66	mountain	جَبَل
67	earth	تُرَاب
68	dust	عُبار
69	seed	بَذْرَة
70	bark	قَشْرَة
71	moon	قَمَر
72	well	بئر
73	sand	رَمْل
74	rain	مَطَر
75	wind	رِيح
76	wheat	قَمْح
77	root	جَنْر
78	flower	زَهْرَة
79	Sun	شَمْس
80	sky	سَمَاء
81	winter	الشَّتَاء
82	rock	صَخْرَة
83	cockroach	صُرْصُور
84	north / east / south / west	شَمال / شَرْق / جَنُوب / عَرَب

85	fog	ضباب
86	beach	شاطئ
87	wood (place)	غابة
88	cloud	غيمة
89	summer	الصيف
90	spring	الربيع
91	autumn	الخريف
92	dawn	فجر
93	crescent moon	هلال
94	milk	لبن
95	egg	بيضة
96	dates	تمر
97	cheese	جبنه
98	meat	لحم
99	bread	خبز
100	oil	زيت
101	olive	زيتون
102	honey	عسل
103	rice	ارز
104	cream	قشدة
105	breakfast	إفطار
106	salt	ملح
107	pepper	فلفل
108	butter	زبدة
109	flour	طحين
110	silver	فضة
111	officer	ضابط
112	land	أرض
113	morning	صباح
114	friend	صديق

115	shape	شَكْل
116	chair	كُرْسِي
117	poet	شَاعِر
118	year	سَنَة
119	name	إِسْم
120	bracelet	سِيوَار
121	husband	زَوْج
122	wife	زَوْجَة
123	weight	وَزْن
124	time	وَقْت
125	balance	مِيزَان
126	vomit	قَيْء
127	livelihood	رِزْق
128	thing	شَيْء
129	fishing net	شَبْكَة
130	problem	مُشْكَلَة
131	king	مَلِك
132	boat	مَرَكَب
133	idea	فِكْرَة
134	drop	قَطْرَة
135	mind	عَقْل
136	division	فِـرْقَة
137	market	سُوق
138	money	فَلُوس
139	countryside	رِيف
140	cover	عِطَاء
141	room	عُرْفَة
142	necklace	عِقْد
143	car	سَيَّارَة
144	hour	سَاعَة

145	song	أغنية
146	outside	خارج
147	street	شارع
148	airplane	طائرة
149	restaurant	مَطْعَم
150	doctor	طبيب
151	system	نظام
152	midday	ظَهْر
153	shadow	ظِلّ
154	road	طريق
155	perfume	عطر
156	Dhofar	ظفار
157	president	رئيس
158	part	جُزء
159	family	أسرة
160	port	ميناء
161	pocket	حيب
162	village	قَرْية
163	member	عَضو
164	paper	وَرَقَة
165	fasting	صَوْم
166	month	شَهْر
167	gift	هَدِيَّة
168	goal	هَدَف
169	light	نور
170	prison	سِجْن
171	self	نَفْس
172	right	يَمِين
173	left	يَسَار
174	bath	حَمَام

175	dream	حُلْم
176	colour	لَوْن
177	condition	حَالَة
178	clothes	مَلَابِس
179	God	الله
180	tribe	قَبِيلَة
181	house	بَيْت
182	window	نَافِذَة
183	door	بَاب
184	gold	ذَهَب
185	grave	قَبْر
186	frankincense	أَبْنَان
187	trader	تَاجِر
188	blessing	بَرَكَة
189	book	كِتَاب
190	ruins	أَثَار
191	inheritance	إِرْث
192	mosque	جَامِع / مَسْجِد
193	hunger	جُوع
194	smoke	دُخَان
195	school	مَدْرَسَة
196	minute	دَقِيقَة
197	line	خَط
198	incense	بُخُور
199	city	مَدِينَة
200	town	بَلَدَة
201	ring	خَاتِم
202	plough	مِحْرَاث
203	war	حَرْب
204	spirit/soul	رُوح

205	harvest	حصاد
206	truth	حَقّ
207	luck	حَظّ
208	government	حكومة
209	council	مجلس
210	thirsty	عطشان
211	long/tall	طويل
212	middle	وَسَط
213	great	عظيم
214	at/by	عِنْدَ
215	about/from	عَنْ
216	after	بَعْدَ
217	wide	واسع
218	on	على
219	far	بَعِيدَ
220	tomorrow	غداً
221	expensive	غالي
222	closed	مغلق
223	angry	غاضب
224	in	في
225	poor	فَقِيرَ
226	single	وَحِيدَ
227	boiled	مَسْلُوقَ
228	only	فَقَطَ
229	old/ancient	قَدِيمَ
230	near	قَرِيبَ
231	fried	مَقْلِيّ
232	all	كُلّ
233	early	بِأَكْرَأَ
234	How...?	كَيْفَ...؟

235	How many days...?	كم يوم...؟
236	many	كثير
237	heavy	ثقيل
238	tired	مُتعب
239	this book / this city / these boys	هذا الكتاب / هذه المدينة / هؤلاء الأولاد
240	that book / that city / those girls	ذلك الكتاب / تلك المدينة / أولئك البنات
241	good	جيد
242	bad	سئء
243	new	جديد
244	true	صحيح
245	behind	خلف
246	light	خفيف
247	smooth	ناعم
248	straight	مستقيم
249	fat	سمين
250	thin	رقيق
251	thick	سميك
252	dry	جاف
253	wet	رطب
254	special	خاص
255	never	أبداً
256	cold	بارد
257	inside	داخل
258	yes	نعم
259	no	لا
260	easy	سهل
261	beautiful	جميل
262	high	عالٍ
263	forbidden	ممنوع
264	with	مع

265	in front of	أمام
266	yesterday	أمس
267	pregnant	حامل
268	what is your name / what is happening...?	ما اسمك / ماذا يحدث؟
269	next to	جانب
270	now	الآن
271	clean	نظيف
272	roasted	مشوي
273	you (m.s)	أنت
274	you (f.s)	أنتِ
275	he	هو
276	she	هي
277	we	نحن
278	you (dual)	أنتما
279	I	أنا
280	you (m.pl)	أنتم
281	you (f.pl)	أنتن
282	they (m.pl)	هم
283	they (f.pl)	هن
284	they (dual)	هما
285	necessary	ضروري
286	and	و
287	or	أو
288	strong	قوي
289	when is your birthday...?	متى عيد ميلادك...؟
290	better	أحسن
291	worse	أسوأ
292	here	هنا
293	there	هناك
294	where...?	أين...؟





338	Come	أجىء / تجيء / تجيئين / يجيء / تجيء / نجيء / تجيئون / تجئن / تجيئان / يجيئون / يجيئن / يجيئان / تجيئان /
339	Came	جئتُ / جئتَ / جئتِ / جاء / جاءت / جئنا / جئتم / جئتن / جئتما / جاؤوا / جئنا / جاء / جاءنا
340	Doubt	أشك / تشك / تشكين / يشك / تشك / تشك / تشكون / تشككن / تشكان / يشكون / يشككن / يشكان / تشكان
341	Doubted	شككتُ / شككتَ / شككتِ / شك / شك / شككت / شككتنا / شككتكم / شككتن / شككتما / شكوا / شككن / شكنا / شكنا
342	Begin	أبدأ / تبدأ / تبدئين / يبدأ / تبدأ / تبدأ / تبدؤون / تبدأ / تبدأ / تبدآن / تبدؤون / يبدأ / يبدأ / يبدأ / بدأت / بدأت / بدأت / بدأ / بدأت / بدأنا / بدأت / بدأت / بدأت / بدأت / بدأنا / بدأنا
343	Began	أروي / تروي / تروين / يروي / تروي / تروي / تروين / تروون / تروين / تروين / ترويان / يروون / يروين / يرويان / ترويان
344	Tells / relates	رويبتُ / رويبتَ / رويبتِ / روي / روي / روي / روي / رويتم / روين / روين / رويتما / رويوا / روين / روي / روي
345	Told / related	أدرّسُ / تدرّسُ / تدرّسين / يُدرّسُ / تدرّسُ / تدرّسُ / تدرّسون / تدرّسن / تدرّسان / يُدرّسون / يُدرّسن / يُدرّسان / يُدرّسان
346	Teaches	درّستُ / درّستَ / درّستِ / درّس / درّس / درّس / درّسنا / درّستم / درّستن / درّستنا / درّستوا / درّسن / درّسنا
347	Taught	أسافرُ / تُسافرُ / تُسافرين / يُسافرُ / تُسافرُ / تُسافرُ / تُسافرون / تُسافرن / تُسافران / يُسافرون / يُسافرن / يُسافران
348	Travels	سافرتُ / سافرتَ / سافرتِ / سافرت / سافرت / سافرت / سافرتنا / سافرتتم / سافرتن / سافرتما / سافرتوا / سافرتنا
349	Travelled	أخبرُ / تُخبرُ / تُخبرين / يُخبرُ / تُخبرُ / تُخبرُ / تُخبرون / تُخبرن / تُخبران / يُخبرون / يُخبرن / يُخبران
350	Informs / thank	أخبرتُ / أخبرتَ / أخبرتِ / أخبرت / أخبرت / أخبرت / أخبرتنا / أخبرتتم / أخبرتن / أخبرتما / أخبرتوا / أخبرتنا
351	Informed / thanked	أتكلّمُ / تتكلّمُ / تتكلّمين / يتكلّمُ / يتكلّمُ / يتكلّمُ / يتكلّمون / يتكلّمن / يتكلّمان / يتكلّمون / يتكلّمن / يتكلّمان
352	Speaks / talks	تكلّمْتُ / تكلّمْتَ / تكلّمْتِ / تكلّم / تكلّم / تكلّم / تكلّمنا / تكلّمتم / تكلّمتن / تكلّمتما / تكلّموا / تكلّمنا
353	Spoke / talked	أبداّلُ / تبدّلُ / تبدّلين / يتبدّلُ / يتبدّلُ / يتبدّلُ / يتبدّلون / يتبدّلن / يتبدّلان / يتبدّلون / يتبدّلن / يتبدّلان
354	Exchange	تبادلتُ / تبادلتَ / تبادلتِ / تبادلت / تبادلت / تبادلت / تبادلتنا / تبادلتتم / تبادلتن / تبادلتما / تبادلتوا / تبادلتنا
355	Exchanged	أنصرفُ / تنصرفُ / تنصرفين / ينصرفُ / ينصرفُ / ينصرفُ / ينصرفون / ينصرفن / ينصرفان / ينصرفون / ينصرفن / ينصرفان
356	Leave	



377	I go to the market every day	أذهب إلى السّوق كل يوم
378	He might travel to Muscat tomorrow	قد يسافر إلى مسقط غداً
379	We have just come back from town	رجعنا من المدينة تَوّاً
380	She must have arrived there by now	لا بدّ أنّها وصلت هناك الآن
381	Why did you forget to return the shirt yesterday?	لماذا نسيتَ أن ترجع القميص أمس؟
382	It is necessary to construct new houses	يجب إنشاء بيوت جديدة
383	We should have visited her family's house	كان يجب علينا أن نزر بيت عائلتها
384	They had to travel	وجب عليهم السفر
385	You must always be careful	عليك أن تكون حذراً دائماً
386	Don't go to the market today!	لا تذهبوا إلى السّوق اليوم!
387	Look for the old books!	ابحثوا عن الكتب القديمة!
388	Meet (f.s) her tomorrow!	قابلها غداً!
389	Where did you lose (2 m.s) your keys?	أين فقدت مفاتيحك؟
390	She started to tell a long story	بدأت تروي قصّة طويلة
391	Can you come with me?	هل من الممكن أن تجيء معي؟
392	She became scared when she saw him	أصبحت خائفة عندما رآته
393	If I had some money I would give it to you	لو كان لي بعض المال لأعطيته لك
394	My clothes are not clean	ملابسي ليست نظيفة
395	If you had arrived yesterday, you would have seen her	لو وصلت أمس لرأيته
396	After five minutes he started to talk	بعد خمس دقائق بدأ يتكلم
397	I walked with your friend's brother	مشيت مع أخي صديقك
398	There are only twenty-four hours in one day	هناك أربع و عشرون ساعة في اليوم فقط
399	If he had found the house, he wouldn't be here now	لو وجد البيت لما كان هنا الآن
400	We will travel to Dubai in our father's car	سوف نساfer إلى دبي في سيارة أبينا
401	Respect those who are older than you!	احترموا من هم أكبر منكم سناً!
402	These are the bags that you gave me yesterday	هذه هي الشنط التي أعطيتها لي أمس
403	The man's house was always very cold	كان بيت الرجل بارداً جداً دائماً
404	I do not want to talk about that subject	لا أريد أن أتكلّم في ذلك الموضوع
405	You don't have to go to the cafe	لا يجب أن تذهب إلى المقهى

406	He must not remain at home	يجب أن لا يبقى في البيت
407	I have two sons and a daughter	عندي ولدان و بنت
408	If you want to find her, you will find her	إن أردت أن تجدها وجدتها
409	I do not know whether I was going to send them or not	لا أعرف إن كنت سأرسلها أم لا
410	If I go I will meet him	إذا ذهبتُ قابلته
411	If he studies he will succeed	إذا ذاكرَ ينجح
412	You will not be allowed to go unless you finish your work	لن يُسمح لك بالذهاب إلا إذا أكملتَ عملك
413	My father's office is far away	مكتب والدي بعيد من هنا
414	Your sister's book is valuable	كتاب أختك ثمين
415	His cousin (paternal) is a merchant	أبن عمه تاجر
416	Your mother is a doctor	أمك طبيبة
417	These pens are mine	هذه الأقلام لي
418	These animals are theirs	هذه الحيوانات لهم
419	We always drink tea at the cafe after work	بعد العمل نشرب شاي في المقهى دائماً
420	This is the man from whom I received a letter	هذا هو الرجل الذي سلمني رسالة
421	You walk in the mountains in the summer, but in the winter you always walk on the beach	تمشون في الجبال خلال الصيف ولكن في الشتاء تمشون على الشاطئ دائماً
422	If I were you I would not do that	لو كنتُ مكانك لما فعلتُ ذلك
423	When will they return from their trip?	متى يعودون من رحلتهم؟
424	What is the name of those plants?	ما اسم تلك المزروعات؟
425	What are the names of those young girls?	ما اسماء أولائك الفتيات الصغيرات؟
426	Her son and her daughter are always busy at school	ولدها و بنتها مشغولان في المدرسة دائماً
427	Their daughters are skilled cooks	بناتهن طبّاحات ماهرات
428	How much are these delicious dates?	بكم هذه التمور اللذيذة؟
429	He asked who went to town on Tuesday.	سأل عمّن ذهب الى السوق يوم الثلاثاء
430	Which friend came to your party?	أي صديق جاء إلى حفلتك؟
431	Which girl did you see?	أيّة بنت رأيتها؟
432	I know the man who lives here	أعرف الرجل الذي يسكن هنا
433	She knows the woman who lives there	تعرف المرأة التي تسكن هناك

434	He knows the men who live there	يعرف الرجال الذين يسكنون هناك
435	We know the women who live here	نعرف النساء الآتي يسكنن هنا
436	He is coming from Muscat	يجيء من مسقط
437	There was a man asking for you yesterday	كان رجل يسأل عنك أمس
438	The camels usually stand beneath the tall palm trees	تقف الجمال تحت النخيل الطويلة عادةً
439	He is still alive	ما زال حيّاً
440	He was still alive	كان ما يزال حيّاً
441	I still don't understand what you mean	ما زلتُ لا أفهم ما تقصده
442	Her mother is still undecided	أمّها ما زالت مُتَرَدِّدَةً
443	This old building is still standing	ما تزال هذه العمارة القديمة قائمة
444	I did not see the houses that you were talking about.	لم أشاهد البيوت التي تتكلم عنها
445	We started to cry	أخذنا نبكي
446	I have no paper with me	ليس معي أيّ أوراقٍ
447	Are you not hungry?	ألسن جائعة؟
448	He is not in the desert	هو ليس في الصحراء
449	I was not hungry	لم أكن جائعاً
450	We were not in the desert	لم نكن في الصحراء
451	They are not busy	ليسوا مشغولين
452	Their baby was born at seven o'clock in the morning	وُلِدَ وليدهم في الساعة السابعة صباحاً
453	The glass was broken on the floor	كسِرَ الكأس على الأرض
454	This article was written five years ago	كُتِبَت هذه المقالة منذ خمس سنوات
455	A meeting is held in this room monthly	يُعقد اجتماع في هذا الغرفة شهرياً
456	My car will not be repaired until next week so I will be forced to hire a car	إصلاح سيارتي لن يتم قبل الأسبوع القادم و لذلك سأضطر إلى استئجار سيارة
457	These books were written in English originally, then translated into Arabic	كُتِبَت هذه الكتب بالإنكليزية أصلاً ثم تُرجمت إلى العربية
458	I am going to sleep now because I am tired	سأنام الآن لأنني تعبان
459	He/she has not arrived yet	لم يصل حتى الآن - لم تصل حتى الآن
460	I ask her to tell me her favourite story	أسألها أن تحكي قصتها المفضلة
461	The manager must be at home	لا بد أن المدير في البيت
462	The door opened suddenly	فُتِحَ الباب فجأةً

463	In the morning we drink coffee but in the evening we always drink tea	نشرب القهوة صباحاً ولكن دائماً نشرب الشاي في المساء
464	You must be more careful in the future	يجب عليك أن تكون أكثر حذراً في المستقبل
465	He might go home tomorrow	ربما سيرجع إلى البيت غداً
466	I will give it to them because I do not need it now	سأعطيهم لهم لأنني لا أحتاجه الآن
467	When I got there I saw that she wasn't home	عندما وصلتُ هناك رأيتُ أنّها لم تكن في البيت
468	You shall meet the President tonight!	لسوف تُقابل الرئيس هذه الليلة
469	She smiled at me when I gave her all the presents	ابتسّمت لي عندما أعطيتها كل الهدايا
470	He told me a story about his brother who was a soldier	حكى لي قصة عن شقيقه الذي كان جندياً
471	I think he arrived last night but I am not sure	أظنّ أنه وصل أمس ولكنني غير أكيد
472	I still remember the first time I travelled to that city	ما زلت أذكر أول مرّة سافرتُ فيها إلى تلك المدينة
473	A man approached me who I had not seen before	اقترب منّي رجل لم أراه من قبل
474	It is possible that it will rain in July	من الممكن أن تمطر في يوليو
475	You must understand	عليك أن تفهمين
476	That road is not used during the winter	لا يُستخدم ذلك الطريق خلال الشتاء
477	The mosque was built many years ago	بُني المسجد منذ سنوات طويلة
478	The newspaper is printed and published in Muscat	تُطبع و تُنشر الجريدة في مسقط
479	This girl is sicker than the man	هذه البنت مريضة أكثر من الرجل
480	The two girls work inside the house	البنتان يشتغلان داخل البيت
481	The two brothers fight outside the office	الأخوان يتضاربان خارج المكتب
482	This is a question which has no answer	هذا سؤال لا جواب له
483	His friend's father arrived from the field	وصل أبو صديقه من الحقل
484	It is said that you are unlucky	يُقال أنّك غير محظوظة
485	Mountainous regions are found in the north of this country	في شمال هذا المدينة توجد مناطق جبلية
486	The truth is known	عُرِفَت الحقيقة
487	We gave the trader some of the money	أعطينا التاجر بعض المال
488	He would write to them every day and tell them his news	كان يكتب اليهم كل يوم و يخبرهم بأخباره
489	I want to work in the market because I like it a lot	أريد أن أشتغل في السوق لأنني أحبّه كثيراً

490	I did not tell her that I would go	لم أقل لها أنني سأذهب
491	I do not like what was mentioned in the newspaper	لا يُعجبني ما ذكر في الصحيفة
492	We were forced to wait because of the rain	اضطررنا أن ننتظر بسبب المطر
493	He comes from another village	يجيء من قرية أخرى
494	The sound of the engine was unusual	كان صوت المحرك غير عادي
495	It is impossible for her to see the moon tonight	لا يمكن لها أن ترى القمر هذه الليلة
496	There was no one there to hear us	لم يكن هناك من يسمعنا
497	I am not the one who does that	لست أنا من يفعل هذا
498	I am granting you two wishes which I will carry out for you	أهديك أمرين أحققهما لك
499	I still remember the first time I travelled	ما زلت أذكر أول مرة سافرت فيها
500	Something is better than nothing	شيء أحسن من لا شيء
501	That picture was more beautiful than this one	كانت تلك الصورة أكثر جمالاً من هذه
502	I arrive in less than ten minutes	أصل في أقل من عشر دقائق
503	She wants to travel for two months or a little longer	تريد أن تسافر لشهرين أو أكثر بقليل
504	It is one of the oldest cities in the world	إنها من أقدم المدن في العالم
505	We study at a different university	ندرس في جامعة مختلفة
506	They teach at the same university	يدرسون في نفس الجامعة
507	She broke the window accidentally	كسرت النافذة عن غير قصد
508	He smashed the glass on purpose	كسر الكأس قصداً
509	I am expecting bad news	أتوقع أخبار سيئة
510	They warned him to stop before he hurt himself	حذروه أن يتوقف قبل أن يجرح نفسه
511	I remember the story now	أتذكر القصة الآن
512	The story reminded me of my childhood	ذكرتني القصة بطفولتي
513	Tomorrow I will know the truth	سأعلم الحقيقة غداً
514	He taught us the rules of football	علمنا قواعد كرة القدم
515	They are separating the sheep and the goats	يفرقون بين الخراف و المعز
516	She scattered the seeds on the ground	رمت البذور على التربة
517	You are in a hurry today because you are late	أنت متعجلة اليوم لأنك متأخرة

518	I hurried him to leave the room	استعجلته أن يخرج من الغرفة
519	I wash my face every morning	أغسل وجهي كل صباح
520	We must scrub this floor quickly	يجب علينا أن نغسل هذه الأرضية بسرعة
521	You returned to your house	رجعت إلى بيتك
522	You returned a book that you had bought last week	أعدت كتاباً اشتريته في الأسبوع الماضي
523	I usually listen to the radio	أستمع إلى الراديو عادةً
524	I let him hear the music	أسمعه الموسيقى
525	She reminded him of his duties	ذكرته بواجباته
526	We negotiated in order to reach a new agreement	فاوضنا لنصل لاتفاقٍ جديدٍ
527	She sits alone on the beach	تجلس وحيدةً على الشاطئ
528	He sat with his wife on the beach	جالس زوجته على الشاطئ
529	I wrote a long letter	كُتبت رسالةً طويلةً
530	She has been writing to her old friends for a long time	كانت تكتب صديقاتها القديمات لمدةً طويلة
531	During the summer we work in the city	خلال الصيف نعمل في المدينة
532	They do business with many foreign companies	يتعاملون مع شركات أجنبية كثيرة
533	That team defeated all its rivals last year	تغلب ذلك الفريق على كل منافسينهم في السنة الماضية
534	They will try to defeat a new team tomorrow	سيحاولون التغلب على فريقاً جديداً غداً
535	I tried to call you but you did not reply	حاولت أن أتصل بك ولكن لم ترد عليّ
536	I am furious about what you said	أنا غاضب مما قلته
537	The boys fought until their father arrived	تعارك الأولاد حتى وصول أبوهم
538	They informed me about their family problems	أخبرتني عن المشاكل في عائلتهما
539	He always makes me happy each time he visits my house	يُسعدني كل مرةٍ يجيء فيها إلى بيتي دائماً
540	I made the people enter the room	أدخلت الناس الغرفة
541	After that we locked the door	بعد ذلك أغلقنا الباب
542	The weather forced them to return to their village	أرجعهم الطقس إلى قريتهم
543	If you are ill I will feed you	إذا كنت مريضاً سأطعمك
544	You dressed your daughter this morning	ألبست بنتك هذا الصباح
545	The farmer informed the whole village	أعلم المزارع القرية كلها

546	The driver let them down at the side of the road	أنزلهم السائق علي جانب الطريق
547	His story made her laugh	أضحكتها قصته
548	The teachers presented their students at the conference	المعلمات قدّمنَ طلابهنَّ في المؤتمر
549	We always set up our camp near a well	كنّا نقيم مخيّمنا قرب بئر دائماً
550	This company published four newspapers weekly	كانت تصدر هذه الشركة أربع جرائد أسبوعياً
551	They learned these poems at school	تعلّموا هذه القصائد في المدرسة
552	The train departed from the station	تحركَ القطار من المحطة
553	Their reasons were explained in the report	فُسرّت أسبابهم في التقرير
554	After the storm our things were scattered over a wide area	بعد العاصفة تبعثرت أشياءنا عبر منطقة واسعة
555	We used to speak with each other every day	كنّا نتكلّم مع بعضنا كل يوم
556	You will not get married?	ألا تتزوّج؟
557	The bus timetable is changed in the summer because the nights are shorter	يتغيّر جدول الأوتوبيس في الصيف لأنّ الليالي أقصر
558	His poems were influenced by his experiences in another country	تأثرت قصائده بتجاربه في بلدٍ آخر
559	The caravans roamed around in the distant mountains	تجوّلت القوافل في الجبال البعيدة
560	The two friends corresponded for many years	الصديقان تراسلا لسنوات كثيرة
561	The two workers co-operated in the factory	العاملان تعاونوا في المصنع
562	They dealt with the matter	تناولوا الأمر
563	The brothers met together outside the mosque	تقابل الأخوان خارج الجامع
564	The traders argued together about the price	تجادل التجار على السعر
565	I pretended to know nothing, but she knew that I was lying	تظاهرتُ أنني لا أعرف شيء ولكنّها عرفتُ أنني كنتُ كاذباً
566	His son is pretending to be ill because he does not like to go to school	يتمارض ولده لأنه لا يحبّ الذهاب إلى المدرسة
567	I pretended to be busy, but I was very lazy today	تظاهرتُ أنني مشغول ولكن كنتُ كسلان جداً اليوم
568	The ministers agreed with one another about the report	الوزراء اتفقوا على التقرير
569	The windows were completely broken	تكسرت الشبّابيك تماماً
570	The ruins were revealed next to the old port	اكتشفت الأثار قرب الميناء القديم
571	The electricity to the village was cut off	إنقطعت الكهرباء عن القرية
572	These flowers are blossoming today	تنتفّح هذه الزهور اليوم

573	The meeting was held last year	انعقد الاجتماع في السنة الماضية
574	He joined them in the team	انضم اليهم في الفريق
575	We moved around the city looking for work	تجولنا حول المدينة و نحن نبحث عن العمل
576	She approached the angry dog with caution	اقتربت من الكلب الغاضب بحذر
577	The crowd assembled in the square	اجتمع الناس في الميدان
578	He waited for the bus whilst talking to his friend	انتظر الأوتوبيس و هو يتكلم مع صديقه
579	The dome of the mosque is supported by large pillars	دُعِمَت قبة الجامع علي الأعمدة الكبيرة
580	They will not be united until they agree	لن يتحدوا حتى يتفقوا
581	Her voice is well-known everywhere	صوتها مشهور في كل مكان
582	I bought those vegetables this morning	اشتريت تلك الخضراوات هذا الصباح
583	He blushed when he saw her	احمرّ وجهه عندما رآها
584	The hills turn green during the winter	تخضرّ الجبال خلال الشتاء
585	His face was tanned from the sun	اسمرّ وجهه من الشمس
586	His back was bent from a long and hard life	اعوجّ ظهره من حياة طويلة و شاقة
587	The officer asked them about the accident they had witnessed	استعلمهم الشرطي عن الحادث الذي قد شاهدوه
588	They used the plough to cultivate their land	استخدموا المحراث ليعلموا أرضهم
589	Did you enjoy the film that you watched last night?	هل استمتعت بالفيلم الذي شاهدته ليلة أمس؟
590	We will hire a car from this company because it is trustworthy	سنستأجر سيارة من هذه الشركة لأنها موثوقة بها
591	I borrowed his bicycle because my car was broken	استعرت دراجته لأن سيارتي كانت معطلة
592	If he leaves in the morning he will be here before the evening	إذا غادر في الصباح سيكون هنا قبل المساء
593	I have seen her only once or twice	لم أرها إلا مرة أو مرتين
594	They rarely go to the doctor	لا يذهبون إلى الطبيب إلا نادراً
595	I have read this book apart from one chapter	قرأت هذا الكتاب إلا فصلاً واحداً
596	She has only ridden in a car a few times	لم تركب سيارة إلا مرّات معدودة
597	How do they know that you are busy?	كيف يعرفون أنك مشغولة؟
598	We visited only our relatives	لم نزر إلا أقاربنا
599	You have nothing to do but talk to me	لا عمل لك غير التكلّم معي
600	The weather in the evening was hot but the garden was cool	كان الطقس حاراً في المساء إلا أنّ الحديقة كانت باردة

601	This is unreasonable	هذا غير معقولٍ
602	He left the house for an unknown reason	ترك البيت لسبب غير معلومٍ
603	Muhammad! Mind those fast cars!	يا محمد احذر تلك السيارات المسرعة!
604	Do you know where you are going?	أتعرفين أين أنتِ ذاهبة؟
605	While I was asleep in bed I heard a noise	و أنا نائم في الفراش سمعتُ ضجيجاً
606	We were sitting near his father's office	كنا جالسين قرب مكتب أبيه
607	Their mother entered the room but they were not sleeping	دخلت أمهنَّ الغرفة ولكن لم تكن نائمات
608	Her daughters are employees at the post office	بناتها موظفاتان بمكتب البريد
609	He came wearing a blue shirt	جاء لابساً قميصاً أزرقاً
610	She sat, putting her head in her hands	جلست هي واضعة رأسها بين يديها
611	I cannot drive along this road because it is always closed	لا أستطيع أن أقود على هذا الطريق لأنه مُغلق دائماً
612	I was afraid of them	كنتُ خائفاً منهم
613	All of the shops will be locked up tomorrow	ستكون جميع المحلات مغلقة غداً
614	I remained, fascinated by the story she was telling me	ظللتُ منبهراً بالقصة التي كانت تحكيها لي
615	I mentioned to him a teacher with whom he was familiar	ذكرتُ له معلماً كان مألوفاً لديه
616	He is certainly over seventy years old	من المؤكد أنه تجاوز السبعين
617	They are expected to graduate from university	كان من المتوقع أنهم سيتخرجون من الجامعة
618	I read the article that had been published in the newspaper two months ago	قرأتُ المقالة التي قد نُشرت في الصحيفة قبل شهرين
619	He gave his brother's telephone number to his friend	أعطى رقم هاتف أخيه لصديقه
620	Her mother gave the bracelet to her	أعطتها والدتها السوار
621	I understand what she was trying to explain	أفهم ما حاولت أن تبيّنه
622	Nothing like this has happened to me	لم يحدث لي شيء كهذا
623	You surely know that	أنت بالتأكيد تعرفين ذلك
624	When my two daughters grow up, will they do the same as me?	عندما تكبر بنتاي هل ستفعلان مثلي؟
625	The two of them study Arabic literature	الاثنان يدرسان الأدب العربي
626	The market traders do not know that you are my cousin	تجار السوق لا يعرفون أنك بنت خالي
627	He crosses the street when he wishes	يعبر الشارع وقتما يريد
628	I rarely drink coffee twice after supper	نادراً ما أشرب القهوة مرتين بعد العشاء

629	I am often unable to sleep	كثيراً ما أعجز عن النوم
630	Perhaps he knows better than me	ربما يعرف أحسن مني
631	I will listen to your story, and perhaps I will change my mind	سأستمع إلى قصتكم و ربما غيّرت رأبي
632	I help her as much as I can	أنا أساعدها بقدر ما أستطيع
633	Whenever he left the house, she saw him to the door	كان كلما خرج من البيت ودّعه حتّى الباب
634	The manager is not present at the meeting	المدير ليس موجوداً في الاجتماع
635	I remembered this later	تذكّرت ذلك فيما بعد
636	It was four o'clock in the morning when we began	كانت الرابعة صباحاً عندما بدأنا
637	If we knock on the door of your house, be ready!	إذا طرقنا باب بيتك فكوني مستعدّة!
638	I will stay with you wherever you may go	سأبقى معك أينما ذهبت
639	The storm had calmed down completely after a couple of hours	هدأت العاصفة هدواً كاملاً بعد ساعتين
640	You will stay here for three days	ستبقين هنا ثلاثة أيام
641	The bus stops every two hundred metres	يتوقّف الأوتوبيس كل منتي متر
642	I had lots of reasons	كانت لدي عشرات الأسباب
643	He walked a few steps in front of me	مشى بضع خطوات أمامي
644	It weighs no more than a few grammes	لا يزن سوى عدّة غرامات
645	He was a few years older than us	كان يكبرنا بعدّة سنوات
646	The clock struck twice then it stopped	دقّت الساعة دقتين ثم توقفت
647	After a few minutes we became alone	بعد دقائق أصبحنا وحدنا
648	We found several old chairs	وجدنا بضعة كراسي قديمة
649	If he did not come I would go to sleep early	إن لم يأتِ نمت مبكراً
650	If they are not at home they will be at their neighbour's house	إن لم يكونوا في بيتهم يكونوا في بيت جارهم
651	I don't know whether I am able to fulfil the demands	لا أعرف إن كنت قادراً على الوفاء بالمتطلبات
652	I could not remember whether I had given her the things	لست أذكر إن كنت أعطيتها الأشياء
653	I will meet his family if I visit Salalah	سأقابل عائلته إذا زرت صلالة
654	Come to our house tomorrow, if you wish	احضر إلى بيتنا غداً إذا أردت
655	I will discuss that subject with him if he wants to	سوف أبحث معه ذلك الموضوع إذا رغب في ذلك
656	If you are students then you are not rich	إن كنتم طلاباً فلستم أغنياء

657	If you delay you will lose your place	إذا تأخرتم خسرتم مكانكم
658	As for her companion, she was a doctor	أما صاحبته فهي طبيبة
659	As for the others, they remained	أما الآخرين فإنهم ظلوا
660	Winter has now arrived	جاء الشتاء الآن
661	As for me, I froze to the spot	أما أنا فجمدتُ في مكاني
662	The Middle East is a region full of history	الشرق الأوسط منطقة عامرة بالتاريخ
663	This is what I imagined in the beginning	هذا ما تصوّرته في البداية
664	This is what happened with me	هذا ما حدث معي
665	I saw both men and women there	رأيتُ كلا الرّجلين و كلتا المرأتين هناك
666	This is a man whom I like and respect	هذا رجل أحبّه و أحترمه
667	I didn't pay attention to what he said	لم أهتمّ بما قاله
668	I contacted everybody I knew	اتّصلتُ بكلّ مَنْ أعرفهم
669	The man sat in front of the television watching a football match	جلس الرجل أمام التّلفزيون يتابع مباراة كرة قدم
670	Why didn't you tell me this when I was in the market?	لماذا لم تقولي لي هذا عندما كنتُ في السّوق؟
671	I sat on the balcony watching it	جلستُ على الشّرفة أراقبها
672	He married her whilst he was young	تزوَّج بها و هو صغير
673	My friends are waiting for me there in the cafe	ينتظرنني أصدقائي هناك في المقهى
674	We ate our meal whilst sitting in the room	أكلنا وجبتنا و نحن نجلس في الغرفة
675	We sat in the cafe drinking coffee and talking	جلسنا في المقهى نشرب القهوة و نتكلّم
676	The people around me are laughing and chatting	الناس حولي يضحكون و يُثرثرون
677	Eat the bread and drink the water, Karima!	كّلي الخبز و أشربي الماء يا كريمة!
678	One day we went to the desert together	ذهبنا معاً ذات يوم إلى الصّحراء
679	Don't drink coffee after ten o'clock in the evening!	لا تشرب القهوة بعد السّاعة العاشرة مساءً!
680	She was too young to understand	كانت أصغر سنّاً من أن تفهم
681	I won't stay longer than a few months	لن أبقى أكثر من شهور قليلة
682	The distance is much further than that	المسافة أبعد من هذا بكثير
683	It was the most modern and biggest building in the town	كانت أحدث و أكبر عمارة في المدينة
684	Finish your work!	أكمل عملكم!
685	Most of the employees are still students in	أغلب الموظّفين لا يزالون طلاباً في الجامعة

	the university	
686	We travelled the distance in twelve hours	قطعنا المسافة في اثنتي عشرة ساعة
687	Out of all of his brothers, Waleed loves sport the most	وليد أكثر إخوته حباً للرياضة
688	It was the biggest tree I had seen in my life	كانت أكبر شجرة رأيتها في حياتي
689	The important thing is that you return quickly. Do you agree?	المهم أن ترجع بسرعة. هل توافق؟
690	The three employees arrived at the company	وصلت ثلاث موظفات الى الشركة
691	I want to teach reading and writing in the school	أريد أن أدرس القراءة و الكتابة في المدرسة
692	I thought that the house was burning	اعتقدت أن البيت يحترق
693	You ran across the square as fast as you could	ركضت عبر الميدان بأقصى ما تستطيع
694	He has been teaching children since he was a young man	هو يُعلم الأطفال منذ شبابه
695	He said that they came through the mountains	قال إنهم جاءوا من خلال الجبال
696	I will relax at home tonight	سأستريح هذا المساء في المنزل
697	After that we might go to a restaurant	بعد ذلك قد نذهب إلى مطعم
698	Don't be frightened sister!	لا تخافي يا أختي!
699	Why have I come to you instead of you coming to me?	لماذا جئت إليك بدل أن تأتي أنت إلي؟
700	Rather than travelling to Abu Dhabi, he set out with us to Muscat	بدلاً من أن يسافر أبو ظبي رحل معنا إلى مسقط
701	The view was really wonderful and I don't know how to describe it to you	كان المنظر رائعاً فعلاً و لا أعرف كيف أصفه لكم
702	I agreed with them to meet in front of the shop	اتفقت معهم على أن نتقابل أمام المحل
703	The boy began to cry when he didn't find his mother beside him	جعل الطفل يبكي حين لم يجد أمه بجانبه
704	In the morning we left the house and headed north.	في الصباح خرجنا من البيت و اتجهنا شمالاً
705	I knew that I would not find my friend	كنت أعرف أنني لن أجد صديقي
706	Don't say anything about your friends!	لا تقل شيئاً عن أصدقائك!
707	We will meet tomorrow as you wish	نلتقي غداً كما تريد
708	How much money do you have with you?	كم معك من المال؟
709	I tried to remember exactly what happened	حاولت أن أتذكر ما حدث بالضبط
710	Each of them lives with two or three friends	يعيش كل منهم مع صديقين أو ثلاثة

711	I read two or three letters	قرأتُ رسالتين أو ثلاثة
712	Come on let's get out of here	هيا بنا نخرج من هنا
713	The sun appeared, and so the temperature rose	طلعت الشمس فارتفعت درجة الحرارة
714	He received the message yesterday, then left the capital	استلم الرسالة أمس ثم غادر العاصمة
715	When we heard that news we rejoiced greatly	لما سمعنا ذلك الخبر فرحنا فرحاً عظيماً
716	We started to laugh when we heard his joke	بدأنا نضحك عندما سمعنا نكتته
717	My mother is the woman who is standing in the queue	أمي هي السيدة الواقفة في الطابور
718	There is a big celebration in our village every Eid Fitr	هناك حفلة كبيرة في قريتنا كل عيد فطر
719	He arrived in January	وصل في يناير
720	My father and my brother went to market on Saturday	ذهب أبي و أخي الى السوق يوم السبت
721	She kissed the baby on its cheek	قبلت الوليد على خده
722	There is not much room between the table and my bed	لا توجد مسافة كبيرة بين المنضدة و سريري
723	There is a mosque opposite the school	هناك مسجد مقابل للمدرسة
724	There is a picture above the bed	هناك صورة فوق السرير
725	The little boy hid behind the tree	اختفى الولد وراء الشجرة
726	His house is behind the mosque	بيته وراء المسجد
727	I injured myself with the knife	جرحت نفسي بالسكين
728	The man's head was covered with a kaffiyeh	كان رأس الرجل مغطى بكوفية
729	Do you work with a hammer?	هل تعمل بمطرقة؟
730	Do you know how to use this machine?	هل تعرف كيف تستخدم هذه الآلة؟
731	He slaughtered the chicken with a knife	ذبح الدجاجة بسكين
732	This chair is made of wood	هذا الكرسي مصنوع من الخشب
733	That chair is made of metal	ذلك الكرسي مصنوع من المعدن
734	She is as big as you	هي في نفس حجمك
735	Although he is quiet, he is not shy	مع أنه صامت ليس خجولاً
736	Although he was born in Oman, he does not speak Arabic	مع أنه وُلد في عُمان لا يتكلم اللغة العربية
737	The market is in the middle of the village	السوق في وسط القرية
738	She sat between them and spoke with them	جلست بينهم و تكلمت معهم

739	A man came towards me hiding something behind his back	اقترب رجل مني يُخفي شيء خلف ظهره
740	Waiting for him we drank another cup of coffee	كنا ننتظره و نحن نشرب فنجان قهوة ثاني
741	Singing she was washing the clothes	كانت تُغني و هي تغسل الملابس
742	My son is crying because his friends are laughing at him	يبكي ابني لأن أصدقاءه يضحكون عليه
743	They managed to arrive here early by following the desert road	استطاعوا أن يصلوا هنا باكراً باستخدام الطريق الصحراوي
744	Take any five apples and give them to me!	خذ أي خمسة من التفاحات و أعطها لي!
745	Is there any body at home at noon?	هل هناك أي شخص في البيت ظهراً؟
746	I don't believe that anyone knows the truth	لا اعتقد أن أي شخص يعرف الحقيقة
747	She remembered the story better than anyone else	تذكرت القصة أحسن من أي شخص آخر
748	She runs faster than anyone else	تجري أسرع من أي شخص آخر
749	He runs faster than anyone else	يجري أسرع من أي شخص آخر
750	He could be anywhere!	من الممكن أن يكون في أي مكان!
751	Did you buy yourself anything yesterday?	هل اشتريت لنفسك أي شيء أمس؟
752	Have you seen anyone there?	هل رأيت أي شخص هناك؟
753	I do not like the food. Give me something different!	لا أحب هذا الأكل. أعطيني شيء آخر!
754	Has anything happened?	هل حدث أي شيء؟
755	If you see something, tell me!	لو رأيت أي شيء أخبرني!
756	Have you ever lived in Muscat?	هل سكنت في مسقط من قبل؟
757	I will help you anytime	سأساعدك في أي وقت
758	There are no more seats next to you	ليست هناك أية مقاعد أخرى بجانبكم
759	I didn't get money from anybody	لم أستلم مال من أي شخص
760	No one uses this book now	لا أحد يستخدم هذا الكتاب الآن
761	I don't know anybody here	لا أعرف أي شخص هنا
762	Somebody is living in this house but I do not know who	يسكن شخص في هذا البيت ولكن لا أعرف من هو
763	However, nobody is living in that house	مع ذلك لا أحد يسكن في ذلك البيت
764	She said that she did not know anyone	قالت أنها لا تعرف أي شخص
765	No-one forced you to do that	لم يُجبرك أحد على فعل ذلك
766	Except for our grandmother, no one in the village knows the song	لا أحد يعرف الأغنية في قريتنا ما عدا جدتنا

767	He did not buy anything in the market	لم يشتري أي شيء من السوق
768	She did not want anything to drink	لم ترد أن تشرب أي شيء
769	What did he give you? Nothing!	ماذا أعطاك؟ لا شيء
770	I never go to the market to buy bread	لا أذهب أبداً للسوق لأشتري خبزاً
771	He visits me often but he never lets me know when he is coming	يزورني كثيراً ولكن لا يخبرني أبداً متى يأتي
772	They have never lived here	لم يسكنوا هنا أبداً
773	I knew a man who never had any money with him	عرفت رجلاً لم يمتلك مالا قط
774	Perhaps I can give this man some money	ربما يمكنني أن أعطي هذا الرجل بعض المال
775	I want to visit Paris some day	أريد أن أزور باريس يوماً ما
776	I heard that other people live here as well	سعدت أن أشخاص آخرين يعيشون هنا أيضاً
777	We went to the local restaurant with some of the guests yesterday	ذهبنا إلى المطعم المحلي مع بعض الضيوف أمس
778	I gave some water to your four camels	أعطيت بعض الماء لجمالكم الأربعة
779	I must go to the shop to buy some food tonight	يجب عليّ أن أذهب إلى المحلّ هذا المساء لشراء الأكل
780	In the mosque some of the people are kneeling, some are standing	في الجامع بعض الناس يركعون و بعضهم يقفون
781	The mother gives her child some food	الأم تعطي طفلها بعض من الأكل
782	Every evening he wanted to go somewhere	كل مساء كان يريد الذهاب إلى مكان ما
783	I want to go somewhere nice and peaceful for a week	أريد الذهاب إلى مكان لطيف و هادئ لمدة أسبوع
784	Somehow he left without paying the bill	بطريقة ما غادر دون أن يدفع الحساب
785	He wants to meet someone here	يريد أن يقابل شخصاً ما هنا
786	We sent someone to ask for the manager	أرسلنا شخصاً ما يطلب المدير
787	He got angry with me because of something	غضب مني لسبب ما
788	Every morning I wake up my child at seven o'clock	أوقظ طفلي في الساعة السابعة كل صباح
789	I found something. Guess what is is!	وجدت شيئاً . خمن ما هو!
790	You should visit me sometime	يجب أن تزورني وقتاً ما
791	Our grandmother sometimes forgets our names	تنسى جدتنا أسماننا أحياناً
792	Everybody except the grandfather left	غادر الجميع إلا الجدّ
793	Everyone seems to be thinking only of themselves	يبدو أن كل الناس يفكرون بأنفسهم فقط
794	She woke up everybody with her screaming	أيقظت كل الناس بصراخاتها

795	Except for an old man everyone was smoking	الكل كان يُدخّن ما عدا رجل كبير
796	I understand every word he says	أفهم كل الكلمات التي يقولها
797	My child makes me angry every day	يغضبني طفلي كل يوم
798	The teacher teaches his pupils a new letter every day	المدرّس يُعلّم طلابه حرف جديد كل يوم
799	My children make me happy every day	يسعدني أطفالي كل يوم
800	I wanted to travel to town with those three men	أردت أن أسافر إلى المدينة مع أولئك الرّجال الثلاثة
801	Do you want to talk to these four boys?	هل تريد أن تتكلّم مع أولئك الأولاد الأربعة؟
802	I do not want to hear about it any more	لا أريد أن أسمع عنه بعد الآن
803	The father is trying to teach his sons	يحاول الأب أن يُدرّس أبنائه
804	He wants to become famous	يريد أن يصبح مشهوراً
805	You can believe him, his words are true	يمكنك أن تثق به. كلامه صحيح
806	Can you mend these holes with a thread?	هل يمكنك أن تصلحي هذه الثّقوب بالخيط؟
807	The water is drinkable	الماء صالح للشرب
808	Who can build a house without nails?	مَن يستطيع أن يبني بيت بدون المسامير؟
809	I can open the door	أستطيع أن أفتح الباب
810	We can go to the market tomorrow in the afternoon	نستطيع أن نذهب إلى السّوق غداً بعد الظّهر
811	If you want I can help you to paint your house	أستطيع أن أساعدك في دهان بيتك إذا أردت
812	We cannot go back until she gets well	لا نستطيع أن نرجع حتى تتحسن
813	I couldn't open the door	لم أستطع أن أفتح الباب
814	I will close the window so that he can't hear us	سأغلق الشّباك لكي لا يستطيع أن يسمعنا
815	She walked around the house without finding an open door	مشّت حول البيت بدون أن تجد باب مفتوح
816	I couldn't convince her to come with me	لم أستطع أن أقنعها أن تأتي معي
817	We couldn't find my grandmother's silver ring	لم نستطع أن نجد خاتم جدّتي الفضي
818	He can't leave until he finds the key	لا يمكن له أن يغادر حتى يجد المفتاح
819	I cannot study because I have to help my mother	لا أستطيع أن أدرس لأنّه يجب عليّ أن أساعد أمي
820	I can't repair the car on my own so my brother repairs it for me	لا أستطيع أن أصلح السّيارة بنفسي لذلك أخي يصلحها لي
821	You have to go there even if you don't want to	يجب أن تذهب هناك حتى إن لم ترد ذلك
822	After I left school I started to work in the	بعد تخرّجي من المدرسة بدأت أن أعمل في السّوق

	market	
823	Finally he has started working	أخيراً بدأ يعمل
824	My skin became red because of the sunshine	أصبح جلدي أحمر بسبب الشمس
825	It started to rain while we were waiting	بدأت تمطر و نحن ننتظر
826	Old people like to tell stories about their younger days	الكبار يحبون أن يحكوا قصص عن شبابهم
827	My daughter is scared to go across any bridge	بنتي الصغيرة تخاف من عبور أي جسر
828	I like to drink a cup of coffee in the morning	أحب أن أشرب فنجان قهوة في الصباح
829	The old women like to sit in front of the house and talk for hours	النساء الكبيرات يحببن أن يجلسن أمام البيت و يتكلمن لساعات
830	I want him to go away	أريده أن يغادر
831	He asked me to give him the money	سألني أن أعطه المال
832	I told her to buy some tea and sugar	طلبتُ منها أن تشتري بعض الشاي و السكر
833	He told me to sell the car next year	طلب مني أبيع السيارة في السنة القادمة
834	I demanded from her to leave immediately	طلبتُ منها أن تغادر فوراً
835	She made him leave	جعلته يغادر
836	He ordered her to open the door	أمرها أن تفتح الباب
837	He came to my house to see me	جاء إلى بيتي ليزورني
838	I came home to talk to you	رجعتُ إلى البيت لأتكلم معك
839	She sits the child on the chair to feed it	أجلستُ الطفل على الكرسي لتطعمه
840	He pushes the people aside to get on the bus	يدفع الناس جانبا ليركب الحافلة
841	To go to the shop you have to walk towards the mosque	لكي تذهب إلى السّوق يجب عليك أن تمشي نحو المسجد
842	I have nothing left to give you	ليس لدي أي شيء أعطه لك
843	I gave the man some money so he could buy coffee	أعطيتُ الرّجل بعض المال لكي يمكنه شراء القهوة
844	It is known that he is very rich	من المعروف أنه غني جداً
845	I know that he said this to you	أعرف أنه قال هذا لكم
846	I heard that you got a good job in town	سمعتُ أنّك حصلت على وظيفة جيّدة في المدينة
847	I think that he has lived here since June	أعتقد أنه قد سكن هنا منذ يونيو
848	I said to him that the shop opens at 10 in the morning	قلتُ له أنّ المحلّ يفتح السّاعة العاشرة صباحاً
849	The woman told me that she had just arrived in the village	قالت لي المرأة أنّها قد وصلت في القرية توّأ

850	I hope it stops raining soon	أرجو أن يتوقف المطر قريباً
851	He had already gone before we got there	كان قد غادر بالفعل قبل أن نصل هناك
852	Before we reached the house she had already gone	كانت قد غادرت البيت بالفعل قبل أن نصل
853	When I was young I lived in a big house in a town	عندما كنتُ صغيراً سكنتُ في بيت كبير في مدينة
854	When I came home I found a pile of rubbish in the street	عندما رجعتُ إلى بيتي وجدتُ كومة زباله في الشارع
855	When I opened the door, the cat hid under the bushes	عندما فتحتُ الباب اختبأت القطه تحت الشجر
856	When I was young I used to go to the market very often	عندما كنتُ صغيراً كنتُ أذهب إلى السوق كثيراً
857	Whenever he laughs the whole neighbourhood can hear him	متى يضحك يستطيع أن يسمعه الجوار كله
858	By repairing his car for him we saved him a lot of money	باصلاح سيارته وفرنا عليه الكثير من المال
859	I did not go out of the house for four days because it was raining	لم أخرج من البيت لأربعة أيام بسبب المطر
860	I went home because I was tired	رجعتُ إلى بيتي لأنني كنتُ تعبان
861	I have covered myself with a blanket because it is cold	غطيتُ نفسي بغطاء لأن الطقس بارد
862	I gave it to them because I do not need it anymore	أعطيتُهم لهم لأنني لا أحتاجه بعد الآن
863	There were so many people at the funeral	كان يوجد كثير من الناس في الجنازة
864	Therefore I didn't see her there	و لذلك لم أرها هناك
865	He told him how to sell his land	أخبره كيف يبيع أرضه
866	That man knew how to repair cars	عرف ذلك الرجل كيف اصلاح السيارات
867	She taught her daughter how to boil an egg	علّمت بنتها كيف تسلق بيض
868	How can we get to town without asking for my brother's car?	كيف نساfer إلى المدينة بدون أن نطلب سيارة أخي؟
869	They walked all the way without speaking to each other	مشوا الطريق كله بدون أن يتكلموا مع بعضهم البعض
870	When are we going to meet?	متى سنقابل؟
871	What food does your brother like?	أي أكل يحبّه أخيك؟
872	Oh father! What have you done?	يا أبي! ما فعلته؟
873	What will I be when I become old?	ما سأكون عندما أصبح كبيراً؟
874	How many days?	كم يوم؟
875	How many remained?	كم بقي؟
876	How many years have you lived here?	كم سنة سكنت هنا؟

877	Who is that person coming towards us?	مَنْ الشَّخْصَ الَّذِي يَتَّجِهَ نَحُونَا؟
878	Who did this?	مَنْ فَعَلَ هَذَا؟
879	Who is coming with you?	مَنْ يَجِيءُ مَعَكَ؟
880	Where does he live until the summer?	أَيْنَ يَسْكُنُ حَتَّى الصَّيْفِ؟
881	Hey Karima! Where are you going?	يَا كَرِيمَةَ! أَيْنَ تَذْهَبِينَ؟
882	Where is he during the day?	أَيْنَ هُوَ خِلَالَ النَّهَارِ؟
883	I would like to ask him why he did this?	أُرِيدُ أَنْ أَسْأَلَهُ لِمَاذَا فَعَلَ هَذَا؟
884	Why is she happy?	لِمَاذَا هِيَ سَعِيدَةٌ؟
885	Why did she do this?	لِمَاذَا فَعَلَتْ هَذَا؟
886	He was not at home	لَمْ يَكُنْ فِي الْبَيْتِ
887	The bottle was not in the bag when I looked for it	لَمْ تَكُنِ الزَّجَاجَةُ فِي الشَّنْطَةِ عِنْدَمَا بَحِثْتُ عَنْهَا
888	There was nothing there	لَمْ يَكُنْ أَيُّ شَيْءٍ هُنَاكَ
889	He is not at home	لَيْسَ فِي الْبَيْتِ
890	There is nothing there	لَيْسَ هُنَاكَ أَيُّ شَيْءٍ
891	He is not sad	لَيْسَ حَزِينًا
892	Tomorrow I will not be at home	لَنْ أَكُونَ فِي الْبَيْتِ غَدًا
893	Tomorrow you will not be at home	لَنْ تَكُونَ فِي الْبَيْتِ غَدًا
894	I was not at home	لَمْ أَكُنْ فِي الْبَيْتِ
895	You were not at home	لَمْ تَكُنْ فِي الْبَيْتِ
896	You came to my house but I was not at home. I had gone out	جِئْتَ إِلَى بَيْتِي وَلَكِنْ لَمْ أَكُنْ هُنَاكَ. كُنْتُ قَدْ خَرَجْتُ
897	I am not at home	لَسْتُ فِي الْبَيْتِ
898	You are not at home	لَسْتَ فِي الْبَيْتِ
899	I knew that you were not at home yesterday	عَرَفْتُ أَنَّكَ لَمْ تَكُنْ فِي الْبَيْتِ أَمْسَ
900	You know that I will not be here tomorrow	تَعْرِفُ أَنَّي لَنْ أَكُونَ هُنَا غَدًا
901	Those animals will not be here in the evening	تِلْكَ الْحَيَوَانَاتُ لَنْ تَكُونَ هُنَا فِي الْمَسَاءِ
902	Neither he nor his brother have enough money	لَيْسَ لَهُ هُوَ أَوْ أَخُوهُ الْمَالُ الْكَافِي
903	I don't have much money with me right now	لَيْسَ لِي مَالًا كَثِيرًا فِي هَذِهِ اللَّحْظَةِ
904	She does not have a house	لَيْسَ لَهَا بَيْتٌ
905	She does not have a brother	لَيْسَ لَهَا أَخٌ

906	I do not have a car	ليست لي سيارة
907	I do not have a sister	ليست لي أخت
908	There was only a little hair on his head and no teeth in his mouth	كان يوجد شعر قليل في رأسه و لا توجد أية أسنان في فمه
909	She did not have the strength to walk back home	لم تكن لديها القوة لتمشي الى بيتها
910	I would not ask you if I knew where it is	لم أكن لأسألك لو كنت أعرف أين هو
911	One boy's shirt was torn	كان قميص ولد واحد ممزقاً
912	Young men are very courageous	الشباب شجعان جداً
913	The door was painted black	كان الباب مطلياً بالأسود
914	The bicycle was stolen	كانت الدراجة مسروقة
915	The dress is washed	الفستان مغسول
916	She is dressed	لبست ملابسها
917	The meat is roasted	اللحم مشوي
918	He is full	هو شبعان
919	The house is painted	البيت مطلي
920	He came himself to see me	جاء بنفسه ليزورني
921	He bought himself another coffee	اشترى فتجان قهوة آخر لنفسه
922	He suddenly saw himself in the mirror	رأى نفسه في المرآة فجأة
923	We ourselves will know the news soon	سنعرف بأنفسنا الأخبار قريباً
924	Is this egg boiled?	هل هذا البيض مسلووق؟
925	I don't know what to do at all	لا أعرف ما أفعله إطلاقاً
926	Tell me what you have done today!	أخبرني ما فعلته اليوم
927	He talked so much, he made me forget what I wanted	تكلم كثيراً لدرجة أنه جعلني أنسى ما أريده
928	I bought the golden bracelet, that I will give to my mother	أشتريتُ السوار الذهبى الذي سأعطه لأمي
929	He asked me what to do to earn some more money	سألني عما يفعله لكسب المزيد من المال
930	This book is mine	هذا الكتاب لي
931	That bicycle is mine	تلك الدراجة لي
932	Those pens are mine	تلك الأقلام لي
933	That book is yours	ذلك الكتاب لك
934	This bicycle is yours	هذه الدراجة لك

935	This car is ours	هذه السَّيَّارة لنا
936	These seats are ours	هذه المقاعد لنا
937	That pan is ours	تلك المقلاة لنا
938	This book is not mine	هذا الكتاب ليس لي
939	That bicycle is not yours	تلك الدَّرَاجة ليست لكِ
940	Those seats are not ours	تلك المقاعد ليست لنا
941	I send you a letter every week (m/f) (s/pl)	أرسل لك رسالة كل أسبوع / أرسل لك رسالة كل أسبوع / أرسل لكم رسالة كل أسبوع / أرسل لكنَّ رسالة كل أسبوع
942	I send him a letter every week	أرسل له رسالة كل أسبوع
943	I send them a letter every week (m/f)	أرسل لهم رسالة كل أسبوع / أرسل لهم رسالة كل أسبوع
944	I send her a letter every week	أرسل لها رسالة كل أسبوع
945	They send me a letter every week	يرسلون لي رسالة كل أسبوع
946	She sends you a letter every week	ترسل لك رسالة كل أسبوع
947	This is my brother	هذا هو أخي
948	This is my sister	هذه هي أختي
949	They are his parents	هما والداه
950	I was at home	كنتُ في البيت
951	I came home	رجعتُ إلى البيت
952	I did not come home	لم أرجع إلى البيت
953	You were at home (m/f) (s/pl)	كنتَ في البيت / كنتِ في البيت / كنتم في البيت / كنتنَّ في البيت
954	You were not at home (m/f) (s/pl)	لم تكن في البيت / لم تكوني في البيت / لم تكونوا في البيت / لم تكنَّ في البيت
955	He was at home	كان في البيت
956	He was not at home	لم يكن في البيت
957	We were at home	كنا في البيت
958	We were not at home	لم نكن في البيت
959	They were at home (m/f)	كانوا في البيت / كنَّ في البيت
960	They were not at home (m/f)	لم يكونوا في البيت / لم يكنَّ في البيت
961	Tomorrow I will be at home	سأكون في البيت غداً
962	Tomorrow you will be at home	ستكون في البيت غداً
963	I know	أعرف

964	I do not know	لا أعرف
965	I am not going to town	أنا لا أذهب إلى المدينة
966	I do not want to go to town	لا أريد أن أذهب إلى المدينة
967	You are not going to town (m/f) (s/pl)	أنت لا تذهب إلى المدينة / أنت لا تذهبن إلى المدينة / أنتم لا تذهبون إلى المدينة / أنتن لا تذهبن إلى المدينة
968	You do not want to go to town (m/f) (s/pl)	لا تريد أن تذهب إلى المدينة / لا تريدين أن تذهبي إلى المدينة / لا تريدون أن تذهبوا إلى المدينة / لا تردن أن تذهبن إلى المدينة
969	She doesn't want to go to town	لا تريد أن تذهب إلى المدينة
970	He doesn't want to go to town	لا يريد أن يذهب إلى المدينة
971	We do not want to go to town	لا نريد أن نذهب إلى المدينة
972	They do not want to go to town	لا يريدون أن يذهبوا إلى المدينة / لا يريدن أن يذهبن إلى المدينة
973	I might go to town	من الممكن أن أذهب إلى المدينة
974	He didn't see me and I didn't see him	لم يرني و لم أره
975	He didn't see you (m/f) (s/p)	لم يرك / لم يرك / لم يركم / لم يركن
976	He didn't see him	لم يره
977	He didn't see her	لم يرها
978	He didn't see us	لم يرنا
979	He didn't see them (m/f)	لم يرههم / لم يرهن
980	I am happy today	أنا سعيد اليوم
981	You are happy today (m/f) (s/pl)	أنت سعيد اليوم / أنت سعيدة اليوم / أنتم سعدون اليوم / أنتن سعدات اليوم
982	He is happy today	هو سعيد اليوم
983	She is happy today	هي سعيدة اليوم
984	We are happy today	نحن سعدون اليوم
985	They are happy today (m/f)	هم سعدون اليوم / هن سعدات اليوم
986	Yesterday I was sad	كنت سعيداً أمس
987	Yesterday you were sad (m/f) (s/pl)	كنت سعيداً أمس / كنت سعيدة أمس / كنتم سعدين أمس / كنتن سعدات أمس
988	Yesterday he was sad	كان سعيداً أمس
989	Yesterday she was sad	كانت سعيدة أمس
990	Yesterday we were sad	كنا سعدين أمس
991	Yesterday they were sad (m/f)	كانوا سعدين أمس / كن سعدات أمس

992	He is not as much a man as his father	هو ليس رجل مثل أبيه
993	I am not as clever as you	أنا لست ذكياً مثلك
994	She walks exactly like her brother	تمشي مثل أخيها بالضبط
995	Her age is the same as yours	عمرها نفس عمرك
996	He doesn't work as much as his colleague	لا يعمل بقدر ما يعمل زميله
997	Football I like it a lot!	كرة القدم أحبها كثيراً
998	Whether you agree or not I don't care	توافق أو لا ، أنا لا أعبأ
999	Why did the people come? They came to mourn	لماذا جاء الناس ؟ جاؤوا ليشاطروه العزاء
1000	I could not find my keys anywhere	لم أستطع أن أجد مفاتيحي في أي مكان
1001	He suddenly appeared from nowhere	ظهر فجأة من حيث لا أدري
1002	There was nowhere else to go	لم يكن هناك مكان آخر يذهب اليه
1003	My brother made me bring it back	جعلني أخي أرجعه
1004	I got these things from my sister	أستلمت هذه الأشياء من أختي
1005	I need new clothes for my brother's wedding	أحتاج إلى ملابس جديدة لعرس أخي
1006	I have a few things to do here before I can go back home	يجب عليّ أن أعمل قليل من الأشياء قبل أن أرجع إلى البيت
1007	I am sorry I lost your keys. Please don't be mad with me	أسف أنني أضعت مفاتيحك . لا تغضب مني
1008	I will wait for you in front of the mosque	سأنتظرك أمام المسجد
1009	He bought new clothes so that he could go into town	اشترى ملابس جديدة لكي سيستطيع أن يذهب إلى المدينة
1010	I have already said this to you before	قد قلت لك هذا من قبل
1011	We said to one another: Where are we?	قلنا لبعضنا البعض : أين نحن ؟
1012	He bought himself another coffee	اشترى لنفسه قهوة أخرى
1013	I introduced this man to all the men in the village	قدّمت هذا الرجل إلى كل الرجال في القرية
1014	Later I sold the animals to my neighbour	بعث الحيوانات إلى جاري في ما بعد
1015	After work he usually goes to the local shop to buy some sweets for his children	بعد العمل يذهب إلى البقالة عادةً لكي يشتري بعض الحلويات لأطفاله
1016	He is lying to us	هو يكذب علينا
1017	She is showing her daughter her new ring	تُري بنتها خاتمها الجديد
1018	Her friend is making her stay for another day	تجعلها صديقتها تبقى ليوم ثاني
1019	Today we are feeding the animals early	اليوم نُطعم الحيوانات باكراً

1020	The fire made us leave our houses	الحريق جعلنا نغادر بيوتنا
1021	I do not open the door to strangers	لا أفتح الباب للأغرب
1022	His father makes him send the goods back	أبوه يجعله يرجع البضائع
1023	Soon I will have enough money to buy a car	سأحصل على المال الكافي لأشتري سيارة قريباً
1024	We told our son to become a doctor	قلنا لولدنا أن يصبح طبيب
1025	I expect that I will be able to pay you tomorrow	أتوقّع أنني سوف أستطيع أن أدفع لك غداً
1026	I told my youngest son to talk to you	طلبتُ من إبني الأصغر أن يتكلم معك
1027	Maybe he was telling lies	ربّما كان كاذباً
1028	She said to her friend: I want to be young again	قالت لصديقتها : أريد أن أصبح صغيرةً مرّة ثانية
1029	She thanked me for the beautiful earrings	شكرتني على الأقراط الجميلة
1030	That teacher showed pictures to the students	ذلك المعلم عرض صوراً على الطلاب
1031	He left this food for you and your brothers	ترك هذا الأكل لك و اخوانك
1032	I heard the news from the boys	سمعتُ الخبر من الأولاد
1033	I got the coat from him	حصلتُ على المعطف منه
1034	He stole it from them	سرقه منهم
1035	I ate some bread	أكلتُ بعض الخبز
1036	I read some of the newspaper	قرأتُ بعض الصّحيفة
1037	I saw him at the party with two of his friends	رأيتُه في الحفلة مع إثنين من أصدقائه
1038	She stained the cloth with some juice	لطختُ المفرّش ببعض العصير
1039	Instead of bread my son bought sweets	اشتري إبني حلوى بدلاً من الخبز
1040	Why didn't you ask me, instead of that man?	لماذا لم تسألني بدلاً من ذلك الرّجل؟
1041	I have two sisters	عندي أختان
1042	I have two brothers	عندي أخوان
1043	I have only one son	عندي ولد فقط
1044	Where is the money? With him	أين المال؟ معه
1045	They have their books with them	كتبهم معهم
1046	The man who came to the wedding has a new car	الرّجل الذي جاء الى الزفاف لديه سيارة جديدة
1047	The men who came to the wedding have new cars	الرّجال الذين جاؤوا الى الزفاف لديهم سيارات جديدة
1048	The boy's father had a big beard	كان لأبي الولد لحيّة كبيرة

1049	The girl's mother had a beautiful gold necklace	كان لأمّ البنت عقْد ذهبي جميل
1050	I do not have much furniture in my room - only a chair, a bed and a table	ليست لي مفروشات كثيرة في غرفتي - هناك كرسي وسرير و طاولة فقط
1051	You have a gold ring too	عندك خاتم ذهبي أيضاً
1052	So you have two cars	و كذلك عندكم سيارتين
1053	My father always has some money with him	أبي معه بعض المال دائماً
1054	I usually go for milk in the morning	أذهب لشراء اللبن في الصباح عادةً
1055	An old man went to the well for water	ذهب رجل كبير الى البئر من أجل للماء
1056	They fight for the water	يتعاركون عن الماء
1057	My son asked me about his grandfather	سألني إبني عن جدّه
1058	I went across the bridge	ذهبتُ عبر الجسر
1059	She saw an accident with her own eyes	شاهدتُ حادث بعينها
1060	She still sews with a needle	ما زالت تخطيط بإبرة
1061	I am lonely without you	أنا وحيد بدونك
1062	It takes hours to reach the city without a car	تستغرق ساعات لتصل المدينة بدون سيارة
1063	I forced myself to eat the food	جعلتُ نفسي أن أكل الطعام
1064	He made himself get up, even though he was tired	أجبر نفسه على النهوض على رغم أنه كان مُتعب
1065	She made herself sick by eating all the sweets	جعلتُ نفسها على مريضة بسبب أكل كلّ الحلويات

## Appendix B. Informed Consent Form & Voluntary Consent Form

Dear \_\_\_\_\_,

My name is Ebtesam S AlOthman and I am a linguist at the University of Manchester, School of Languages, Linguistics and Translation. I am working on Computer-Mediated Communication research and my interest is to investigate how language online is different from face-to-face communication and how technological factors foster the development of a new varieties of languages online. My PhD research is entitled "**Digital Vernaculars: An Investigation of Najdi Arabic in Synchronous Computer-Mediated Communication** "

To help me fulfil the aims of the research, you are invited to participate in the ( Morpho-Syntax Dialect Survey) which is part of my PhD research - a scale survey that based itself on the identification of phonological, morphological and syntactic aspects of a dialect. The aim of this instrument is to help me in conducting a complete description of Najdi Arabic. The questionnaire includes 1065 items. They will be read by me in Modern standard Arabic and you are required to provide the Najdi equivalent of each item. You will also be asked to provide your own comments when necessary.

If you accept to participate in this interview, the researcher will use assumed names so that no one else would know about your participation and your responses will be held in the strictest confidence. The information you provided will be used only for the purpose of the research and will not be available to a third person. Moreover, your participation in this study is on voluntary basis. You do not have to participate in this study if it seems inconvenient to your opinions, and you may withdraw from participating at any time without consequence.

Attached is a voluntary consent form. If you choose to participate in the interview part of the study, please sign the attached voluntary consent form. Please do not hesitate to contact the researcher in case you have any questions or comments concerning the study. Your comments are very important to me.

The Researcher,  
Ebtesam AlOthman  
University of Manchester  
Manchester,UK  
School of Languages, Linguistics and Cultures  
[Ebtesam.Alothman@postgrad.manchester.ac.uk](mailto:Ebtesam.Alothman@postgrad.manchester.ac.uk)

## Voluntary Consent Form

Please read the statements below, think about your choice, and sign if and when you are ready to agree, or take this form home and discuss it with anyone you wish to and then return it to me later if you wish to participate in this research:

**Mrs. Ebtessam S. Al-Othman** has fully explained to me the nature and purpose of this research project in a way that I have understood. She has encouraged me to be actively involved during the dialect survey and has responded to all of my questions and concerns in a satisfactory and respectful way.

I hereby give my agreement to participate in this study research **Digital Vernaculars: An Investigation of Najdi Arabic in Multilingual Synchronous Computer-Mediated Communication** conducted by **Mrs. Ebtessam S. Al-Othman** at the **University of Manchester**

Printed Name \_\_\_\_\_

Signature \_\_\_\_\_ Date \_\_\_\_\_

Contact numbers \_\_\_\_\_ E-mail \_\_\_\_\_

Address \_\_\_\_\_

Signature of the researcher \_\_\_\_\_

## Appendix B. Informed Consent Form (The Arabic version)

بسم الله الرحمن الرحيم

أخي-أختي

يسرني دعوتكم للمشاركة في إجراء دراسة تحليلية كجزء من متطلبات اطروحة الدكتوراة الخاصة بي في علم اللغويات في كلية اللغات واللغويات والثقافة- جامعة مانشستر- المملكة المتحدة المعنونة "دراسة لغوية تحليلية لل لهجة النجدية في الإتصال الإلكتروني في غرف الحوار الإلكترونية: " والتي تتمركز حول خصائص استخدام اللهجة النجدية المحلية خلال محادثة الآخرين إلكترونياً. تتمحور هذه الدراسة حول اللغة المستخدمة في غرف الحوار والتي هي لغة وسط بين اللغة المنطوقة (اللهجة النجدية) والمكتوبة (اللغة العربية الفصحى). وتهدف للإجابة عن عدة تساؤلات مهمة: هل غرف الحوار الإلكتروني تؤثر على العلاقة بين اللهجات العربية (اللهجة النجدية في هذه الحالة) واللغة العربية الفصحى في عالم الإتصال الإلكتروني الحديث؟ ماهي دينكاميكية كتابة اللهجة النجدية أثناء محادثة الآخرين إلكترونياً؟ هل الحروف والأرقام الإنجليزية تقدم غالباً أكثر مرونة للكتابة من قواعد اللغة العربية الفصحى؟ كما أنها تهدف تحديداً إلى اختبار الفرضية القائلة بأن خصائص الكتابة الإلكترونية مرتبطة بعدة عوامل كعمر مستخدم الإنترنت وجنسه ومستوى كفاءة في اللغة الإنجليزية.

في هذا الجزء من الدراسة، أود مشاركتك في توصيف اللهجة النجدية المنطوقة من خلال الاستجابة والمشاركة في استبانة خاصة بوصف اللهجات المحلية.

أود أن أفت انتباهكم بأن كافة البيانات ستكون سرية ولن تستخدم إلا لغرض البحث الأكاديمي. كما ان مشاركتكم في هذه الإستبانة هي تطوعية تماماً ولكم حرية المغادرة وعدم اكمال فقرات الإستبانة في اي وقت.

أمل أن يلقي هذا الطلب تجاوبكم وشاكرة لكم سلفاً تعاونكم الذي سوف يكون له كبير الأثر في تسهيل خطوات البحث

هذا الاستبيان يستغرق ثلاث ساعات.

مع فائق الاحترام والتقدير

الباحثة: ابتسام صالح محمد العثمان  
جامعة مانشستر- المملكة المتحدة

[Ebtسام.Alothman@postgrad.manchester.ac.uk](mailto:Ebtسام.Alothman@postgrad.manchester.ac.uk)

## Appendix C: The “IRC Language Questionnaire”

[SURVEY PREVIEW MODE] Survey - Windows Internet Explorer  
http://www.surveymonkey.com/s.aspx?PREVIEW\_MODE=DO\_NOT\_USE\_THIS\_LINK\_FOR\_COLLECTION&sm=ByW7B8r4hlyAn%2bcPrz9uMZd0Y3kj%2F7%2fmlBBEEOZ2Kk%3d

File Edit View Favorites Tools Help

# The University of Manchester

## الكتابة داخل غرف الحوار الإلكترونية

### الجزء الأول: مقدمة

أخي/أختي

يسرني دعوتكم للمشاركة في إجراء دراسة تحليلية كجزء من متطلبات أطروحة الدكتوراه الخاصة بي في علم اللغويات في كلية اللغات والثقافة جامعة مانشستر. المملكة المتحدة المعنونة "الدراسة لغوية تحليلية لإتصال الإلكتروني في غرف الحوار الإلكترونية" والتي تتمركز حول خصائص استخدام اللهجة المحلية خلال محاكاة الآخرين إلكترونياً. تتحور هذه الدراسة حول اللغة المستخدمة في غرف الحوار والتي هي لغة وسط بين اللغة المنطوقة (اللهجة النجدية) والمكتوبة (اللغة العربية الفصحى). ويهدف للإجابة عن عدة تساؤلات مهمة: هل غرف الحوار الإلكتروني تؤثر على العلاقة بين اللهجات العربية (اللهجة النجدية في هذه الحالة) واللغة العربية الفصحى في عالم الإتصال الإلكتروني الحديث؟ ماهي ديناميكية كتابة اللهجة النجدية أثناء محاكاة الآخرين إلكترونياً؟ هل الحروف والأرقام الإنجليزية تقدم غالباً أكثر مرونة للكتابة من قواعد اللغة العربية الفصحى؟ كما أنها تهدف تحديداً إلى اختبار الفرضية القائلة بأن خصائص الكتابة الإلكترونية مرتبطة بعدة عوامل كعمر مستخدم الإنترنت وجنسه ومستوى كفاءة في اللغة الإنجليزية. وفيما يلي استبيان خاصاً لاستنباط خصائص اللغة التي تستخدمها في غرف الحوار الإلكترونية المختلفة. تضم هذه الاستبيان ثلاث محاور رئيسية: البيانات الشخصية: خصائص استخدام اللهجة المنطوقة في غرف الحوار ومعتقداتك وأرائك حول استخدام هذه اللغة إلكترونياً. يطلب الجزء الثالث من الاستبيان عدداً من المعلومات الشخصية كالإسم والعمر. أود أن ألفت انتباهكم بأن كافة البيانات ستكون سرية ولن تستخدم إلا لغرض البحث الأكاديمي. فأرجو الإهتمام بتعبئة البيانات كاملة. أما الجزء الرابع والخامس من الاستبيان فيعتمد عدداً من العبارات التي تتعلق بخصائص الكتابة الإلكترونية في غرف الحوار الإلكترونية والجاهتك الشخصية تجاهها. الرجاء قرائتها بتعمين ثم اختيار الإجابة التي تميز عن رأيك علماً بأن الإجابة الصحيحة هي التي تميز عن رأيك بنقطة وأمانة. أمل أن يلقى هذا الطلب تجاوبكم ومشاركة لكم سلفاً لتجاربكم الذي سوف يكون له كبير الأثر في تسهيل خطوات البحث.

أود أن ألفت انتباهكم إلى أن مشاركتكم في هذه الاستبيان هي تطوعية تماماً ولكم حرية المغادرة وعدم إكمال فترات الاستبيان في أي وقت.

هذا الاستبيان يستغرق أقل من عشر دقائق

مع فائق الاحترام والتقدير

الباحثة: ابتسام صالح محمد العثمان  
جامعة مانشستر - المملكة المتحدة

Ebtesam.Alothman@postgrad.manchester.ac.uk

تالي

SurveyMonkey يت تشغيل بواسطة

100%

Desktop epte EN 12:41 16/05/2012

[SURVEY PREVIEW MODE] الكتابة داخل غرف الحوار الإلكترونية Survey - Windows Internet Explorer  
http://www.surveymonkey.com/s.aspx?PREVIEW\_MODE=DO\_NOT\_USE\_THIS\_LINK\_FOR\_COLLECTION&sm=ByW7BRs4hlyvAn%2bcPxc9uMZdOY3k%27%2fnIB8EOZKk%3d

File Edit View Favorites Tools Help

1824

The University of Manchester

الكتابة داخل غرف الحوار الإلكترونية

الجزء الثاني: الموافقة على المشاركة

فضلاً اقرأ الأسئلة التالية بتمعن وأقر على "نعم" في حال رغبتك بالمشاركة وتجنبة الاستيلاء

1. هل استوعبت طبيعة الدراسة الحالية وأهدافها؟

نعم  لا

2. هل تترك بأن مشاركتك في هذه الدراسة هي تطوعية وأنه تستطيع الإسهام وعدم إكمال الاستبيان في أي وقت بدون أي مجبر وبدون أي عواقب؟

نعم  لا

3. هل توافق على المشاركة في هذه الدراسة وتعبئة الاستبيانه التالية؟

نعم  لا

التالي السابق

SurveyMonkey بتسجيله بواسطة

100%

Desktop epte EN 12:42 16/05/2012

[SURVEY PREVIEW MODE] المكتبة داخل غرف الحوار الإلكترونية Survey - Windows Internet Explorer  
http://www.surveymonkey.com/s.aspx?PREVIEW\_MODE=DO\_NOT\_USE\_THIS\_LINK\_FOR\_COLLECTION&sm=ByW7BRs4hJyAn%2bcPxc9uMZd0Y3kj%2f7%2f1B68EOZKk%3d

File Edit View Favorites Tools Help

# The University of Manchester

المكتبة داخل غرف الحوار الإلكترونية

الجزء الثالث: البيانات الشخصية

4. في حال رغبتك الاطلاع على نتائج الدراسة لاحقاً، فضلاً قم بكتابة ايميلك الخاص

5. Age العمر

<18

18-25

26-35

35-45

>45

6. الجنس Gender

ذكر

أنثى

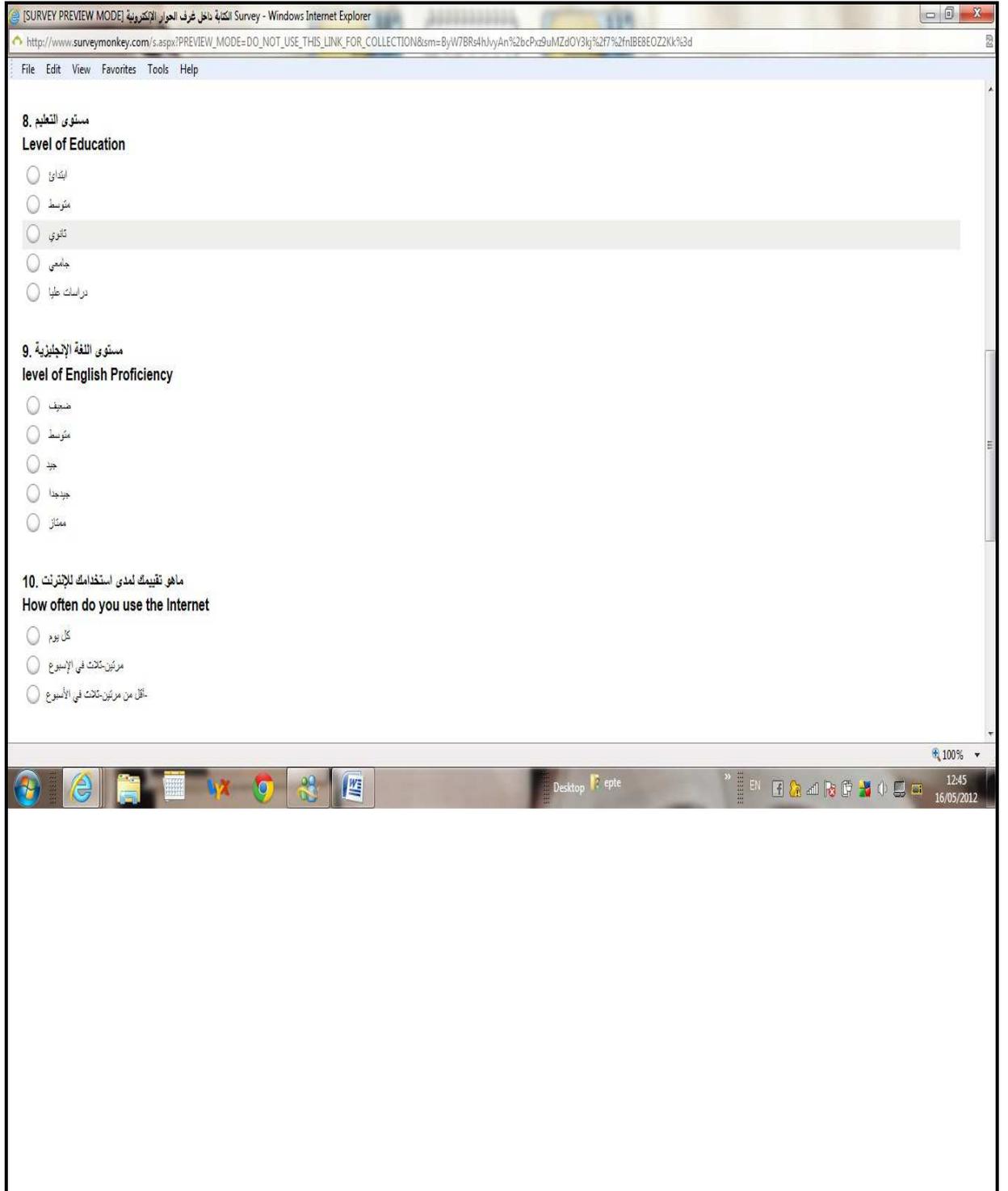
7. الجنسية

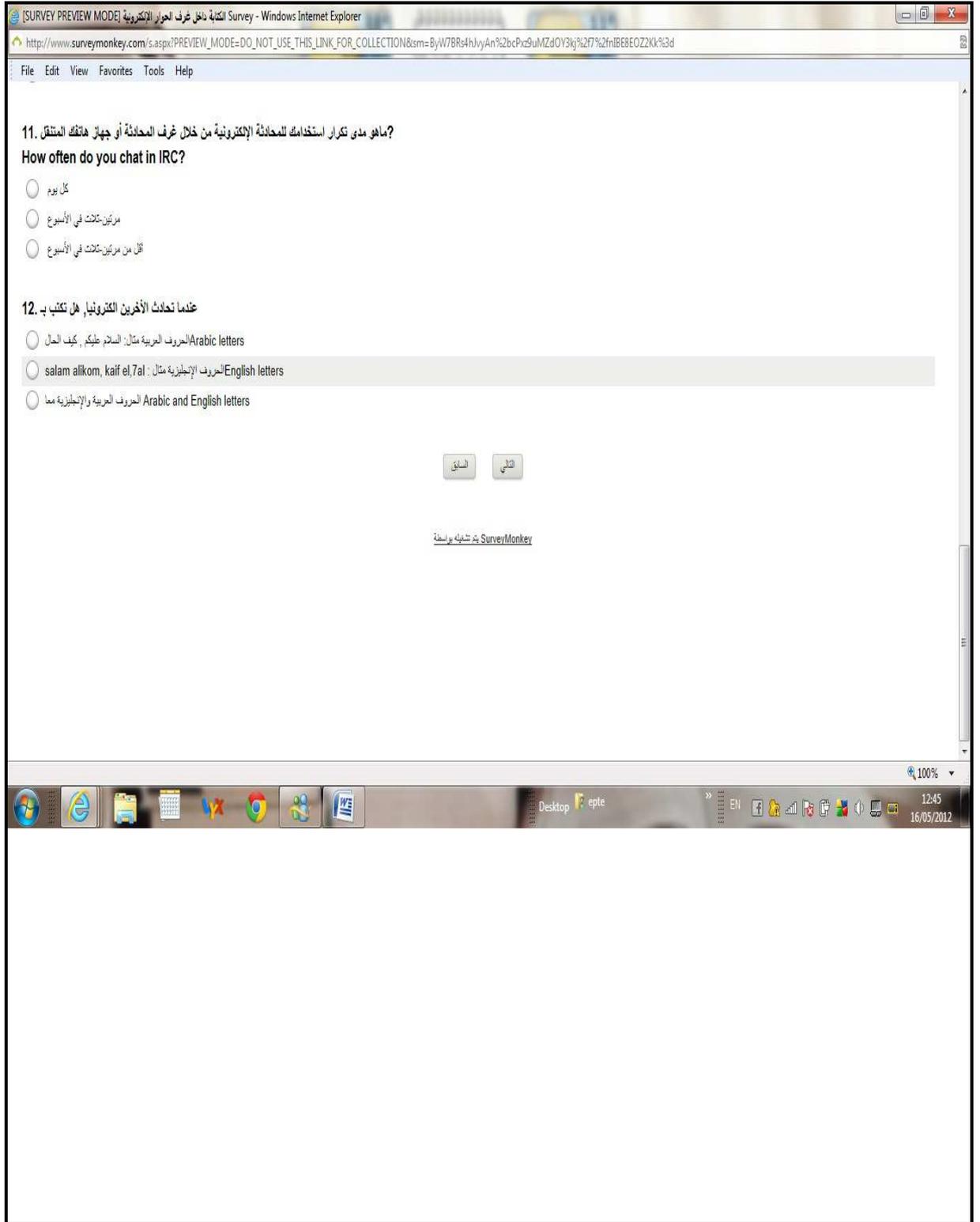
سعودي

غير سعودي

100%

Desktop epte EN 12:43 16/05/2012





[SURVEY PREVIEW MODE] الكتيبة داخل غرف الحوار الإلكترونية Survey - Windows Internet Explorer  
http://www.surveymonkey.com/s.aspx?PREVIEW\_MODE=DO\_NOT\_USE\_THIS\_LINK\_FOR\_COLLECTION&sm=ByW7BR64HvyAn%2bcPxs9uMZd0Y3kj%27%2fmlBE8EOZ2Kk%3d

File Edit View Favorites Tools Help

# The University of Manchester

## الكتيبة داخل غرف الحوار الإلكترونية

### الجزء الرابع: خيارات اللغة والكتابة

عندما تحدث الآخرين إلكترونياً فأنت تكتب عادة اللغة التي تتحدث بها والتي تسمى لهجة النجديّة في هذه الحالة وهي تختلف تماماً عن اللغة العربية الفصحى في أن اللغة العربية الفصحى ذات قواعد نحوية وإملائية ثابتة. في الأسفل لديك قائمة من خيارات اللغة والكتابة المختلفة التي تختص باستخدام لهجتك المنطوقة في محادثة الآخرين إلكترونياً. فضلاً قم بقراءة كل عبارة بعناية ومن ثمّ قم بمدى استخدامك فعلياً للخيارات المطروحة

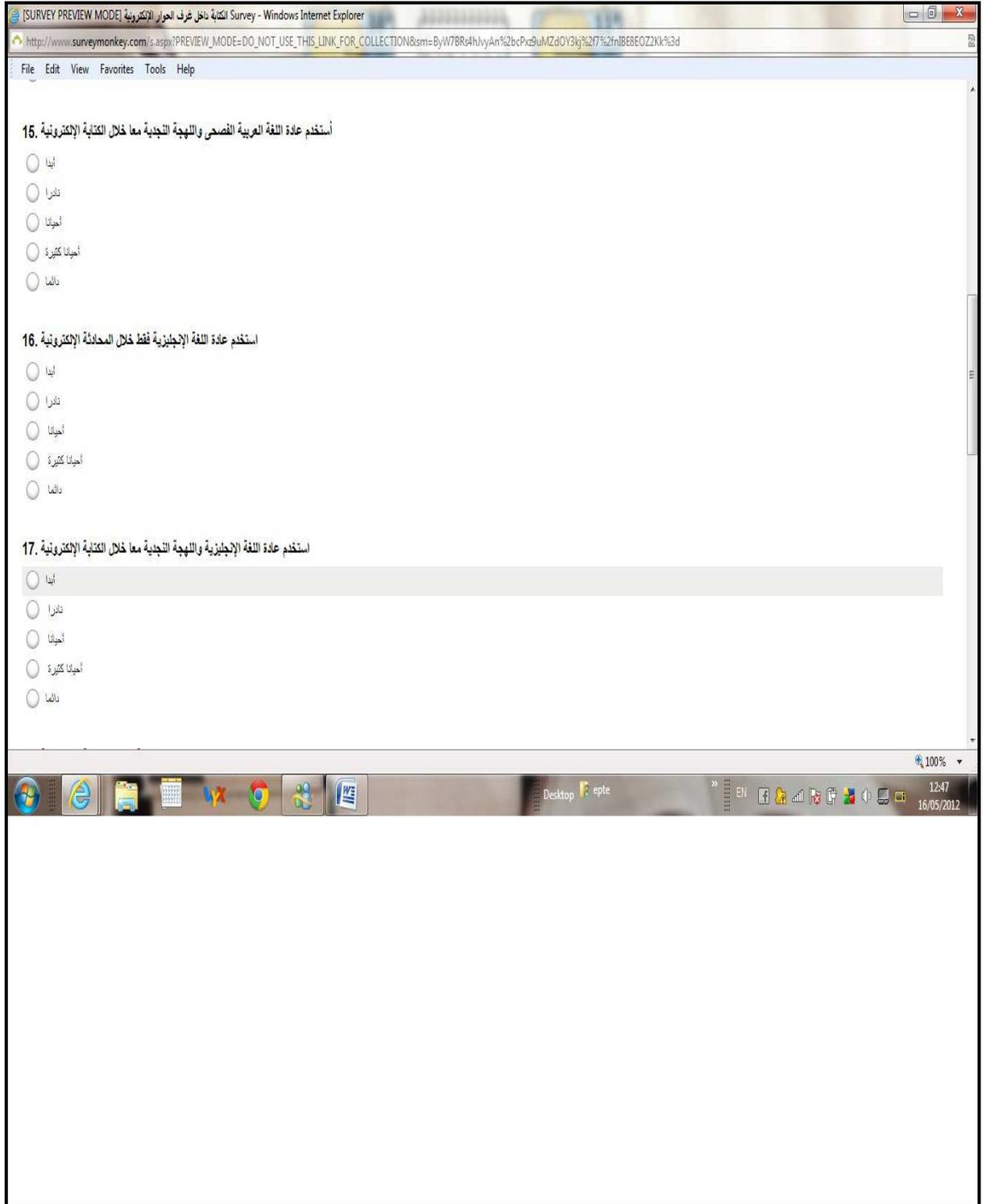
13. استخدم عادة اللغة العربية الفصحى خلال محادثة أصدقائي إلكترونياً.

أبداً  
 نادراً  
 أحياناً  
 أحياناً كثيرة  
 دائماً

14. خلال المحادثة الإلكترونية: "اللهجة النجديّة" استخدم عادة اللغة التي أتحدث بها.

أبداً  
 نادراً  
 أحياناً  
 أحياناً كثيرة  
 دائماً

100% Desktop epte EN 12:46 16/05/2012



[SURVEY PREVIEW MODE] Survey - Windows Internet Explorer  
http://www.surveymonkey.com/s.aspx?PREVIEW\_MODE=DO\_NOT\_USE\_THIS\_LINK\_FOR\_COLLECTION&sm=ByW7BR4hhvyAn%2bcPz9uMZdOY3g%2f%2fhIB8EOZZk%3d

File Edit View Favorites Tools Help  
دائماً

18. استخدم عادة كلمات إنجليزية خلال المحادثة الإلكترونية.

أبداً  
 نلحراً  
 أحياناً  
 أحياناً كثيرة  
 دائماً

19. برأيك لماذا يقوم بعض مستخدمي الإنترنت باستخدام الحروف والأرقام الإنجليزية أثناء الكتابة الإلكترونية بدلاً من استخدام الحروف العربية؟

لأنه من الأسهل الكتابة بالحروف والأرقام الإنجليزية على لوحة  
 لأنه لا تتوفر لوحة مفاتيح عربية في بعض الأحيان  
 لأنه من الصعب كتابة اللهجة المنطوقة (التجدي) باستخدام الحروف العربية

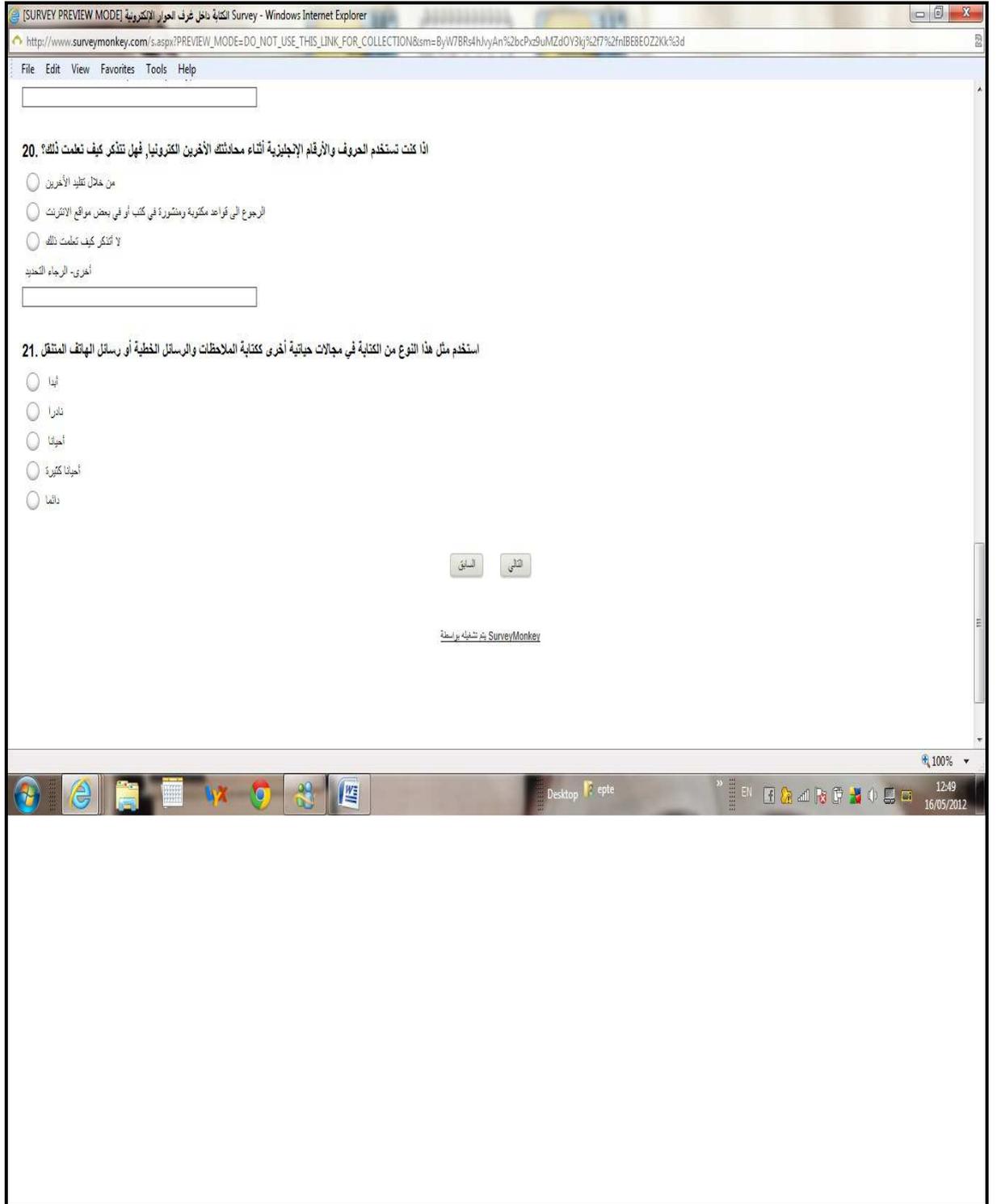
أخرى - الرجاء التحديد  
Others (Please specify)

20. إذا كنت تستخدم الحروف والأرقام الإنجليزية أثناء محادثتك الآخرين إلكترونياً، فهل تتذكر كيف تعلمت ذلك؟

من خلال تطبيق الآخرين  
 الرجوع إلى قواعد مكتوبة ومشورة في كتب أو في بعض مواقع الإنترنت  
 لا أتذكر كيف تعلمت ذلك

أخرى - الرجاء التحديد  
أخرى - الرجاء التحديد

100%  
Desktop epte EN 12:48 16/05/2012



[SURVEY PREVIEW MODE] Survey - مكتبة داخل غرف الحوار الإلكترونية - Windows Internet Explorer

http://www.surveymonkey.com/s.aspx?PREVIEW\_MODE=DO\_NOT\_USE\_THIS\_LINK\_FOR\_COLLECTION&sm=ByW7BRs4hvyAn%2bcPzduMZdOY3kj%2f7%2fmiBE8EOZ2Kk%3d

File Edit View Favorites Tools Help

# The University of Manchester

## المكتبة داخل غرف الحوار الإلكترونية

### الجزء الخامس: اتجاهات واعتقادات حول أساليب الكتابة الإلكترونية

في الأسفل لديك قائمة من المعتقدات والآراء التي تختص باستخدام لهجتك المنطوقة في المحادثة الإلكترونية. فضلاً عن بقراءة كل عبارة بعناية ومن ثم قيم مدى موافقتك أو عدمها مع كل عبارة

**22. كتابة اللهجة المنطوقة (التجديية) خلال محادثة الآخرين إلكترونياً هو إساءة للغة العربية الفصحى**

لا أتفق بشدة

لا أتفق

محايد

أتفق

أتفق بشدة

**23. أجد صعوبة في قراءة وكتابة لهجاتي التي أتحدث بها باللغتين العربية وبالحرف عريبة**

لا أتفق بشدة

لا أتفق

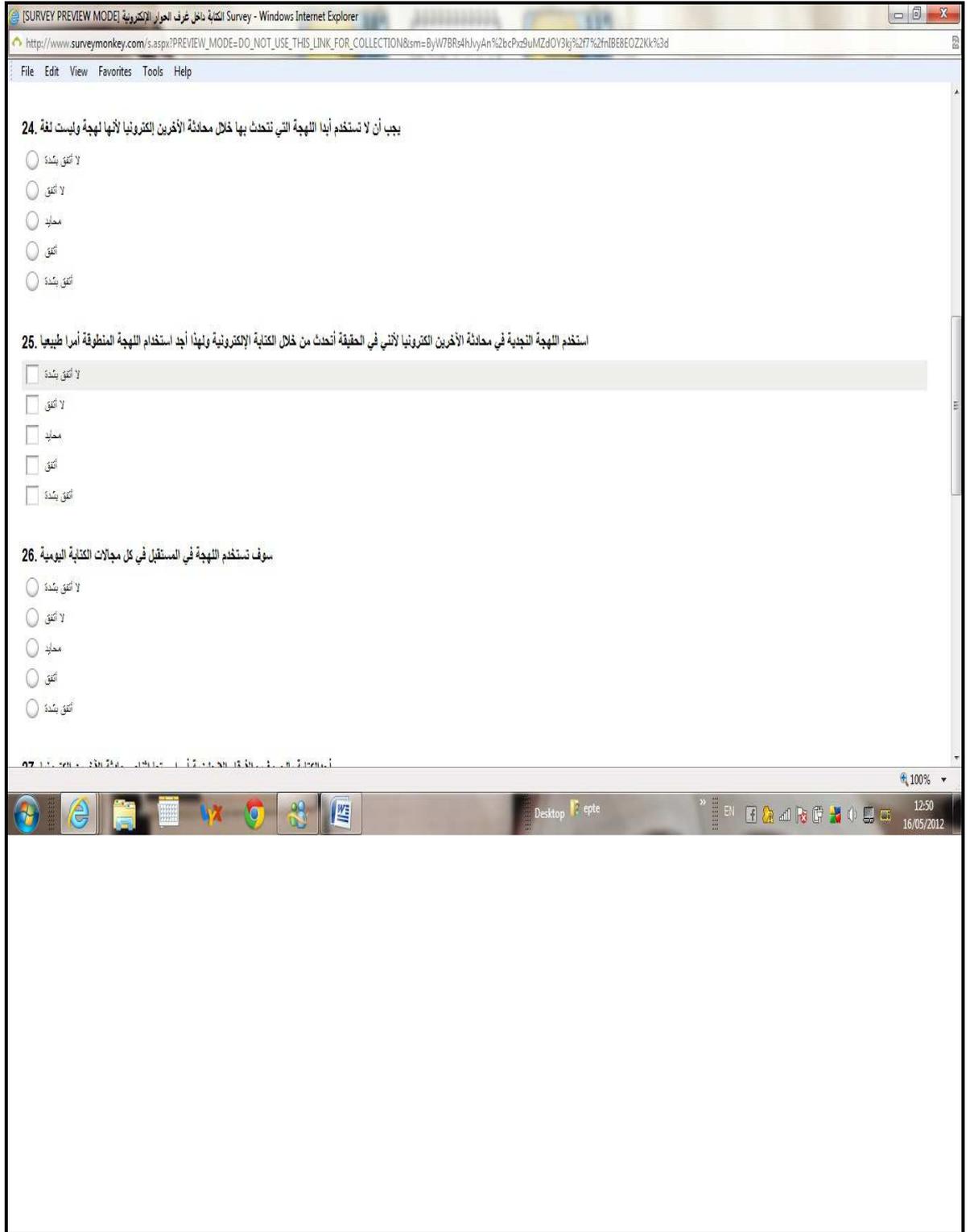
محايد

أتفق

أتفق بشدة

100%

Desktop epte EN 12:50 16/05/2012



[SURVEY PREVIEW MODE] Survey - Windows Internet Explorer  
http://www.surveymonkey.com/s.aspx?PREVIEW\_MODE=DO\_NOT\_USE\_THIS\_LINK\_FOR\_COLLECTION&sm=ByW7BRs4HjvyAn%2bcPz9uMZd0Y3k%27%2fmlB68EOZ2Kk%3d

File Edit View Favorites Tools Help

27. أجد الكتابة بالحروف والأرقام الإنجليزية أمرا ممتعا أثناء محادثة الآخرين إلكترونيا.

لا أتفق بشدة  
 لا أتفق  
 محايد  
 أتفق  
 أتفق بشدة

28. أجد الكتابة بالحروف العربية أثناء محادثة الآخرين إلكترونيا أمرا مزعجا.

لا أتفق بشدة  
 لا أتفق  
 محايد  
 أتفق  
 أتفق بشدة

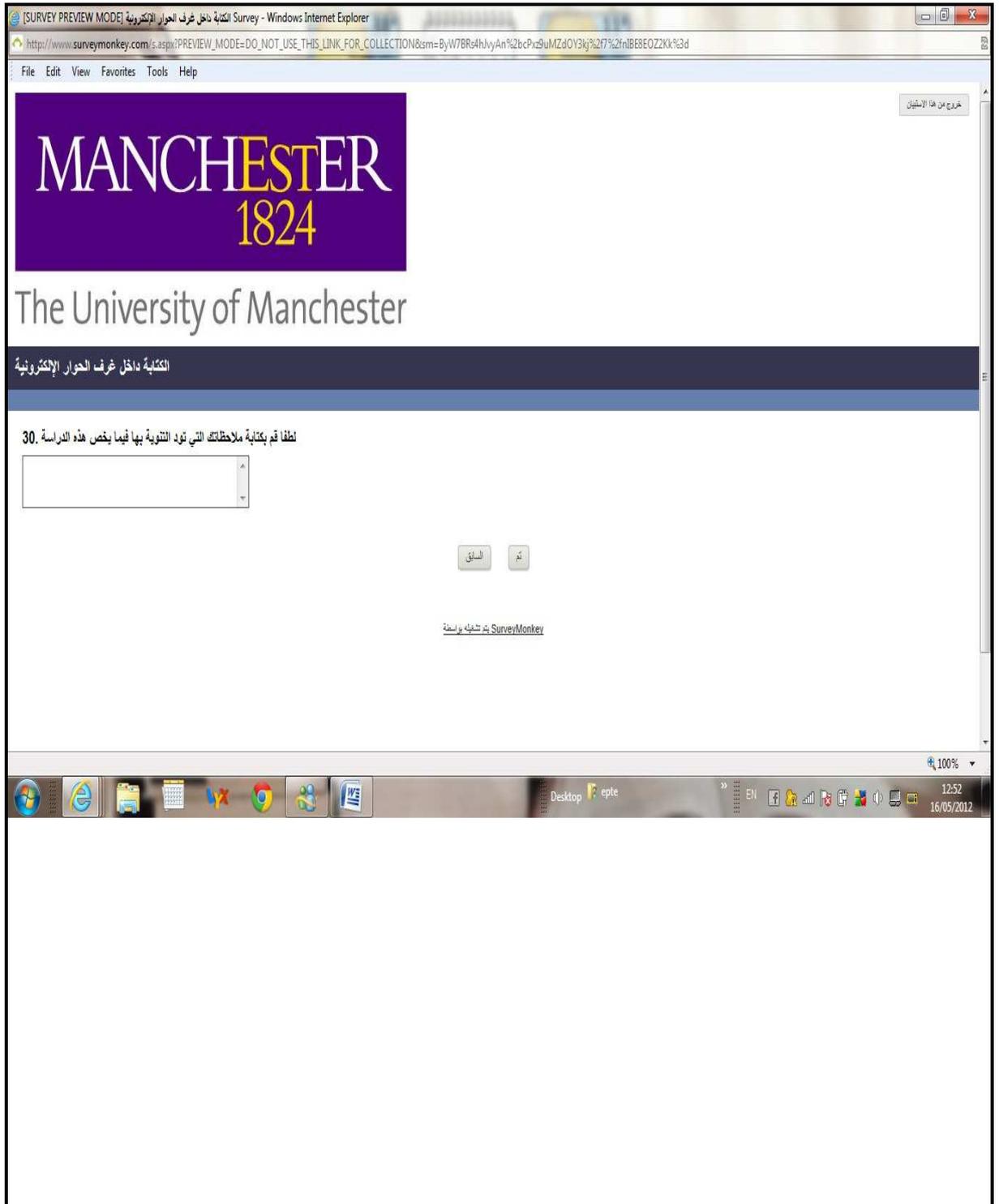
29. كتابة اللهجة النجدية المنطوقة بحروف وأرقامجليزية خطر يهدد اللغة العربية الفصحى.

لا أتفق بشدة  
 لا أتفق  
 محايد  
 أتفق  
 أتفق بشدة

السابق التالي

100%

Desktop epte 12:51 16/05/2012



## Appendix D

### The most common abbreviations in English CMC (Danet 2001)

AFK	away from keyboard
B4	before
BBL	be back later
BTDT	been there, done that
BTW	by the way
CU	see you
CUL	see you later
CYA	see ya
FAQ	frequently asked questions
FTF	face to face
FYI	for your information
GAL	get a life
IMO	in my opinion
IMHO	in my humble opinion
IRL	in real life
L8R	later
NRN	no reply necessary
OTOH	on the other hand
REHI	hi again
ROFL	rolling on the floor laughing
TIA	thanks in advance
TNX	thanks
WB	welcome back

## Appendix E

### The most common English emoticons (Sanderson 1993, Paul 2009)

Emoticons	Meaning
:-) :) :o) :] :3 :c) :> =] 8) =) :} :^)	happy face, pleasure, humour, etc
:D C:	big smiley
()	hug
:-D :D 8D xD XD =D =3 <=3 <=8	laughing, big grin
:( ;*( T_T TT_TT T.T Q.Q Q_Q ;_ ;{ ;-(	crying, sadness
:-(: (:c :< :[ :{	frown
D: D8 D; D= DX v.v	horror, disgust, sadness
;-) ;) *) ;] ;D	winking in any of its meaning
:-P :P XP :-p :p =p :-P :P :-b :b	tongue sticking out, cheeky/playful
:-O :O O_O o_o 8O OwO O-O 0_o O_o O3O o0o ;o_o; o...o 0w0	surprise, shock
c.c C.C	disapproval
:-/ :/ :\ =/ =\ :S	sceptical, annoyed, undecided, uneasy, hesitant
:	straight face disgusted, grim, no expression
:-X :X :# :#	embarrassed
O:-) O:3 O:)	angel' innocent
^o)	raised eyebrow
>:) >:) >:-)	evil
B) B-) 8) 8-)	shades
^>.>^ ^<.<^ ^>_>^ ^<_<^	sarcasm , Irony
D:< >:( D-< >:-( :-@ ;( `_' D<	angry, mad
<3 <333	love, heart, lots of hearts
</3	broken heart, no love
=^_^= =>.>= =<_<= =>.<=	happy blushing, embarrassed blushing, bashful
\o/	praise, excitement, jumping for joy
\o o/	greeting, hello, goodbye
d'-' d'_' d'-'b d'_'b	listening to music, headphones
o\o	high fiver
@ }-;-'----	rose
[:)	user is wearing a walkman
8-)	user is wearing glasses
:-@	user is screaming

