

**A STRUCTURAL ANALYSIS OF MOROCCAN ARABIC  
AND ENGLISH INTRA-SENTENTIAL CODE  
SWITCHING**

**NAJAT BENCHIBA**

Thesis submitted to the University of London in partial fulfilment of the requirements  
for the degree of Doctor of Philosophy

Department of Linguistics  
School of Oriental and African Studies  
August 2008



ProQuest Number: 10673120

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10673120

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

DEDICATION

To My Beloved Parents

## ABSTRACT

A phenomenon of language contact between different speech communities is that of code switching which is a result of language contact between speakers of diverse language(s) and/or dialect(s).

The aim of this thesis is to quantitatively and qualitatively detail the grammatical outcomes of intra-sentential code switching in natural parsing by bilingual speakers of Moroccan Arabic and English in the UK and to assess the way in which the Matrix Language Frame Model (MLF) (Myers-Scotton 1993b, 2002) is a suitable linguistic model for bilingual discourse. Such natural switching is highly regularized and syntactic features are maintained through normal grammatical constraints as will be detailed. A description of grammatical approaches to code switching is outlined with focus on one particular model, the Matrix Language Frame the concept of which was first pioneered by Joshi (1985) and elaborated upon in further detail by Myers-Scotton (1993b, 2002). I also draw upon the Minimalist model MacSwan (1999) for further analysis of inter-language parameters and language universals with regard to constraints on code switching as well as comparisons made with the Monolingual Structure Approach (Boumans, 1998).

It is not the aim of this thesis to advocate a one-size-fits-all approach to constraints on code switching as this has proved to be the Achilles heel of all theoretical approaches to code switching over the last few decades (Pfaff 1979, Poplack 1980, Di Sciullo, Muysken & Singh 1986, Bentahila & Davies 1983) but to validate and corroborate the viability of the Matrix Language Frame Model. Natural data of Moroccan Arabic and English code switched discourse collated for this thesis provide further empirical support required to test the validity of the Matrix Language Frame model well as providing a quantitative database for further research. I advocate

my own set of eleven generalizations pertaining to intra-sentential code switching and highlight a new emerging speech style amongst second and third generation speakers I have termed *Reactive Syntax* where it becomes evident that innovative speech styles and syntactic strings of utterances highlight creativity amongst these generational groups. This thesis concludes with an evaluation of the data collated together with an examination of the suitability of the Matrix Language Frame Model and suggestions for further research.

## CONTENTS

<i>Abstract</i> .....	<i>iv</i>
<i>List of Maps and Tables</i> .....	<i>x</i>
<i>Acknowledgements</i> .....	<i>xii</i>
<i>Symbols and Abbreviations</i> .....	<i>xiii</i>
<i>The Individual Moroccan Arabic Sounds</i> .....	<i>xiv</i>
CHAPTER ONE – OUTLINE OF CODE SWITCHING .....	1
1.1 Language Choice .....	2
1.2 Types of Code Switching.....	4
1.3 Early Analysis of Code Switching .....	7
1.4 Minimalist Approaches to Code Switching .....	10
1.5 Code Switching, Interference and Bilingual Considerations .....	11
1.6 Code Switching and Borrowing.....	13
1.7 Code Switching and Bilingual Proficiency.....	18
1.8 Conclusion .....	21
CHAPTER TWO – SYNTACTIC THEORY AND APPROACHES TO CODE SWITCHING .....	23
2.1 Language Specific Constraints.....	24
2.1.1 Timm’s Linear Model .....	25
2.1.2 The Conjunction Constraint .....	27
2.1.3 The Constraints Theory.....	30
2.1.4 The Free Morpheme Constraint .....	30
2.1.5 The Equivalence Constraint .....	36
2.1.6 Woolford’s Constraint .....	39
2.2 Asymmetry .....	42
2.2.1 Matrix and Embedded Approaches .....	46
2.2.2 Joshi’s Constraint .....	46
2.2.3 Doron’s Constraint.....	47
2.2.4 The Matrix Language Frame Model .....	49
2.3 Government Approaches .....	55
2.3.1 Di Sciullo, Muysken & Singh’s Constraint.....	56
2.3.2 The Functional Head Constraint .....	60
2.4 Insertional Approach.....	66
2.4.1 Bentahila and Davies’ Constraint.....	66
2.5 Monolingual Constraints in Code-Switched Discourse .....	70
2.5.1 Null Theory.....	70
2.6 Minimalist Approaches.....	73
2.6.1 The Minimalist Program – MacSwan .....	74
2.7 Conclusion .....	83

CHAPTER THREE – MATRIX LANGUAGE FRAME MODEL .....	86
3.1 Defining the Matrix Language.....	87
3.1.1 Criticisms of the Matrix Language .....	89
3.1.2 Early Accounts of the Matrix Language .....	98
3.1.3 Levelt’s Speech Production Model .....	94
3.2 Matrix Language Approaches .....	97
3.2.1 Sridhar & Sridhar’s Dual Structure Principle.....	98
3.2.2 Joshi’s Asymmetry Model .....	98
3.2.3 Counter-Examples to Joshi’s Model .....	100
3.3 Matrix Language Frame Model .....	101
3.3.1 Structure of the MLF.....	102
3.3.2 Four Principles of the MLF .....	104
3.3.3 Matrix Language – Embedded Language Distinction.....	107
3.3.4 The Content Morpheme – System Morpheme Hierarchy .....	109
3.3.5 The 4-M Model .....	112
3.3.6 Possessive Marker <i>Dyal</i> and the 4-M Model .....	115
3.3.7 The Abstract Level Model .....	120
3.4 MLF Composite Structures and <i>Reactive Syntax</i> .....	123
3.4.1 <i>Reactive Syntax</i> and Single Word Switches .....	126
3.5 Boumans (1998) and the Monolingual Structure Approach (MSA) .....	130
3.5.1 The MSA and the Matrix Language.....	131
3.5.2 The MSA and Constituent Insertion.....	133
3.5.3 The MSA and the Finite Clause .....	135
3.6 Lexical Insertions.....	137
3.6.1 VPs as Lexical Insertions .....	142
3.7 Conclusion .....	145
 CHAPTER FOUR – MOROCCAN ARABIC GRAMMAR .....	 147
4.1 Moroccan Arabic – An Historical Outline.....	148
4.2 Moroccan Arabic Morphology and Basic Word Categories.....	151
4.2.1 Moroccan Arabic and Verbal Derivations.....	153
4.2.2 Moroccan Arabic Personal Pronouns .....	155
4.3 Moroccan Arabic Root Types .....	157
4.3.1 Moroccan Arabic Stems and Word Derivation .....	159
4.3.2 Moroccan Arabic Verbs and Inflections .....	160
4.4 Moroccan Arabic Syntax - VSO or SVO.....	163
4.4.1 Moroccan Arabic and English Word Order .....	166
4.5 Moroccan Arabic Nouns - Definite and Indefinite Articles .....	168
4.5.1 Masculine and Feminine Moroccan Arabic Nouns .....	171
4.5.2 Moroccan Arabic Nouns and Adjectival Agreement .....	173
4.6 Relative Clauses.....	176

4.7 Negation.....	180
4.7.1 Moroccan Arabic Negation and Code Switching .....	183
4.8 Conclusion.....	185
CHAPTER FIVE – THE UK MOROCCAN DATA CORPUS .....	188
5.1 History of Migration from Morocco to the UK.....	189
5.2 Data Corpus .....	196
5.3 Research Design – Uses of Quantitative and Qualitative Methods .....	198
5.3.1 Transcriptions.....	199
5.3.2 Initial Data Analysis Procedures .....	200
5.4 Primary Data.....	201
5.4.1 Informants’ Phase One.....	203
5.4.2 Informants’ Language Proficiency.....	208
5.5 Informants’ Questionnaire Results.....	210
5.5.1 Results of Phase One.....	215
5.5.2 Results of Phase Two.....	218
5.6 Conclusion .....	220
CHAPTER SIX - GRAMMATICAL CATEGORIES OF CODE SWITCHING FROM ENGLISH INTO MOROCCAN ARABIC .....	221
6.1 English Lexical Insertion in Moroccan Arabic Clauses .....	223
6.1.1 The Indefinite Zero Article .....	227
6.2 Grammatical Gender .....	230
6.2.1 Grammatical Gender Contexts.....	232
6.2.2 Grammatical Gender and the use of Diminutives. ....	239
6.2.3 English Diminutives.....	241
6.3 English Verbs and the Periphrastic ‘Do’ or ‘ <i>Dar</i> ’ + Construction .....	245
6.4 Moroccan Arabic and English Numerals and Quantification.....	250
6.4.1 Moroccan Arabic Cardinals from Three to Ten .....	252
6.4.2 Moroccan Arabic Cardinals from Eleven Onwards .....	256
6.4.3 Moroccan Arabic / English Quantifiers.....	258
6.5 Moroccan Arabic /English Demonstratives – Possible Switch Sites .....	263
6.5.1 Moroccan Arabic Demonstratives.....	263
6.6 A Morphological Analysis of English Lexical Insertions (Embedded Islands).....	270
6.6.1 English Verb Insertion Within a Moroccan Arabic Frame ( <i>Pseudo-Verbs</i> ) .....	273
6.6.2 Evidence for Embedded Verbal Insertions.....	271
6.6.3 Evidence for Embedded Verbal Insertions in The Imperfective .....	276
6.7 Conclusion .....	278
CHAPTER SEVEN – CODE SWITCHING OF CONSITUENTS - THE MLF AS A SUITABLE SYNTACTIC VEHICLE .....	281

7.1 Moroccan Arabic Prepositional Insertion .....	282
7.1.2 Moroccan Arabic / English Adverbial Insertions.....	289
7.1.3 Moroccan Arabic Inserted Conjunctions .....	296
7.2 Moroccan Arabic Discourse Marker Insertions .....	302
7.2.1 Discourse Markers: uses, Sequencing and Turn-Taking .....	304
7.2.2 Functionality Amongst <i>Darja</i> Particles.....	308
7.2.3 Discourse Markers within an Asymmetrical Framework.....	311
7.3 Moroccan Arabic Pronoun Doubling in Intra-Sentential Discourse .....	313
7.3.1 Moroccan Arabic Pronoun Doubling within the MLF .....	317
7.4 Moroccan Arabic Embedded Conditional Clauses .....	322
7.4.1 Non-Switchable Nominal Constituents .....	326
7.4.2 Embedded Elements of Cultural Constituents.....	330
7.5 Conclusion .....	332

CHAPTER EIGHT - EVALUATION OF THE MATRIX LANGUAGE MODEL AND  
GENERALIZATIONS.....

8.1 Evaluation of Main Generalization 1 .....	335
8.1.1 Asymmetry and Basic Premise of Generalization 1.....	338
8.1.2 Generalizations Evaluated.....	342
8.2 Classic Versus Composite Code Switching .....	350
8.3 Matrix Language Turnover .....	351
8.4 Problematic Data for the MLF Model.....	355
8.4.1 Further Problematic Areas .....	362
8.5 Research Questions Revisited .....	365
8.6 Conclusion .....	366

CONCLUSION .....	367
BIBLIOGRAPHY .....	372
APPENDIX 1 .....	401
APPENDIX 2 .....	403
APPENDIX 3 .....	404
APPENDIX 4 .....	405
APPENDIX 5 .....	408

## LIST OF MAPS, TABLES AND FIGURES

### MAPS

Figure 1	Map of Arabic-Speaking Peoples.....	149
Figure 2	Ethnolinguistic Groups in Morocco.....	151
Figure 3	Main Zones of Migration .....	194

### TABLES

Figure 2.1	Minimalist Framework.....	77
Figure 2.2	The Minimalist Program .....	78
Figure 3.1	Levelt's Speech Model.....	93
Figure 3.2	Morpheme Classification Myers-Scotton .....	113
Figure 3.3	Production Process Lemma Activation .....	121
Figure 3.4	Table of Embedded Verbal Inflections .....	144
Figure 4.1	Moroccan Arabic Verbal Stems .....	152
Figure 4.2	Moroccan Arabic Independent Pronouns .....	156
Figure 4.3	Moroccan Arabic Dependent Clitic Pronouns.....	156
Figure 4.4	The Inflection of the Perfect Tense.....	160
Figure 4.5	The Inflection of the Imperfect Tense.....	161
Figure 4.6	Moroccan Arabic Masculine and Feminine Nouns.....	171
Figure 4.7	Moroccan Arabic Feminine Nouns with Unmarked <i>-a</i> .....	172
Figure 4.8	Moroccan Arabic Masculine Nouns with Marked <i>-a</i> .....	173
Figure 4.9	Moroccan Arabic Negative Pronouns .....	182
Figure 5.1	Countries of Moroccan Migration.....	190
Figure 5.2	Evolution of Migrant Stocks .....	191
Figure 5.3	Estimates of Moroccan Migrant Stocks .....	193
Figure 5.5	List of Informants in Phase One of Guided Interviews.....	205
Figure 5.6	List of Informants in Phase One (not recorded).....	207
Figure 5.7	Questionnaire: Place of Birth .....	211
Figure 5.8	Questionnaire: Level of Education .....	211
Figure 5.9	Questionnaire: Language Situation at Home with Parents.....	212
Figure 5.10	Questionnaire: Language Situation at Home with Siblings .....	212
Figure 5.11	Questionnaire: Languages Used Outside the Home.....	212
Figure 5.12	Questionnaire: Which Languages Preferred .....	212
Figure 5.13	Questionnaire: In Which Languages are you More Confident.....	212
Figure 5.14	Questionnaire: Do you Usually Use Both Languages.....	213
Figure 5.15	Questionnaire: Do you Switch from Moroccan Arabic to English .....	213
Figure 5.16	Questionnaire: Why do you Usually Switch .....	213
Figure 5.17	Questionnaire: From Which Language do you Switch .....	213
Figure 5.18	Questionnaire: Why do you Usually Think People Switch.....	213
Figure 5.19	Questionnaire: Are you For or Against People Switching .....	214
Figure 5.20	Questionnaire: Classify the Following Languages.....	214
Figure 5.21	Questionnaire: Which Language do you Class as Most Important .....	214
Figure 5.22	Questionnaire: Would you Prefer to be Taught in Moroccan Arabic or English .....	214
Figure 5.23	Questionnaire: Which Language do you See as Important for the Future .....	214
Figure 6.4	Moroccan Arabic Diminutives.....	241
Figure 6.5	Further Moroccan Arabic Diminutives .....	241
Figure 6.6	Moroccan Arabic Feminine Diminutives.....	242
Figure 6.7	French Origin Diminutives .....	242
Figure 6.8	Moroccan Arabic Diminutives (Second Root Repetition) .....	244
Figure 6.9	Modified Moroccan Arabic Diminutives .....	245
Figure 6.10	Moroccan Arabic Cardinals .....	254
Figure 6.11	Moroccan Arabic Cardinals II Onwards.....	258

Figure 6.12	Moroccan Arabic Quantifiers.....	261
Figure 6.13	Moroccan Arabic Demonstratives.....	266
Figure 6.14	Perfect Tense English Verb 'hit'.....	276
Figure 6.15	Imperfect Tense Verb 'copy'.....	278
Figure 7.1	Moroccan Arabic Prepositions.....	285
Figure 7.2	Moroccan Arabic Adverbs.....	292
Figure 7.3	Moroccan Arabic Classification of Adverbs.....	298
Figure 7.4	Moroccan Arabic Conjunctions.....	299
Figure 7.5	Moroccan Arabic Discourse Markers.....	308
Figure 7.10	Moroccan Arabic Cultural Constituents.....	331
Figure 7.11	Moroccan Arabic Cultural Verbs.....	332
Figure 8.1	Contributions to Constraints on Code Switching.....	346
Figure 8.2	Catalogue of Contributions to Constraints on Code Switching.....	348
Figure 8.3	Generalizations Overview.....	350
Figure 8.4	Research Questions Summary.....	369

## FIGURES

Figure 1.1	Moroccan Arabic Consonants.....	<i>xiv</i>
Figure 6.1	Moroccan Arabic Syntax.....	224
Figure 6.2	English Syntax.....	224
Figure 6.3	Moroccan Arabic Syntax.....	226
Figure 7.6	Chomsky.....	318
Figure 7.7	Moroccan Arabic Syntax.....	318
Figure 7.8	Moroccan Arabic Syntax Interpretation.....	319
Figure 7.9	Moroccan Arabic / French Syntax.....	323

## ACKNOWLEDGEMENTS

This thesis would not have been possible had it not been for the generous scholarship provided by Ghazi Al Magadmi to whom I am eternally grateful.

I owe a wealth of thanks to my supervisors: Lutz Marten and Bruce Ingham. Lutz, for academic advice which has helped shape this thesis – what an intellect! I am also grateful to Bruce Ingham whose insight and support have been inspirational as well as his expert knowledge of Arabic Linguistics and Dialectology - not to mention the excellent tea and cakes which have helped many a meeting!

I am indebted to the University of London, in particular to SOAS and Senate House for their fieldwork awards which have enabled me to collate data in the UK and Morocco. I am also grateful to the University of Maryland, in particular Jonathan Owens for their generous bursary to present at the Arabic Symposium, Maryland which enabled me to sharpen the premise of this thesis. I also wish to thank AIDA for their award to present at the conference in Vienna and also IGALA for their award to present in Valencia and New Zealand. I owe thanks to West Wiltshire District Council for their educational bursary which provided the necessary recording equipment.

I owe thanks to the library staff at SOAS, in particular Yelena Shlyuger. I also extend my thanks to the SOAS IT Department, in particular to Jahan for his time and assistance.

I am also grateful to Carol Myers-Scotton whose comments at the beginning of this thesis provided invaluable insights.

No linguistic research is possible without data and in this regard I extend my thanks to all my informants, as well as Al Hasaniya and to all the wonderful Moroccans!

On a penultimate note, I wish to thank my friends and family for believing in me and for their unconditional patience. To my wonderful parents, your daughter is back!

Last but not least, I thank my wonderful Juho for being with me on the last leg of this journey – we can now begin our lives.

## SYMBOLS AND ABBREVIATIONS

<p>* Ungrammatical or Unacceptable</p> <p>Adj Adjective</p> <p>Adv Adverb</p> <p>AGR Agreement</p> <p>AGRg Gender Agreement</p> <p>AGRn Number Agreement</p> <p>AGRp Agreement Phrase</p> <p>AGRs Agreement Subject</p> <p>ASP Aspect</p> <p>Comp Complementizer</p> <p>COP Copula</p> <p>CP Complementizer Phrase</p> <p>CS Code Switching</p> <p>Det Determiner</p> <p>DP Determiner Phrase</p> <p>DS Data Set</p> <p>EL Embedded Language</p> <p>Eng English</p> <p>FC Functional Category</p> <p>Fem Feminine</p> <p>FH Functional Head</p> <p>FHC Functional Head Constraint</p> <p>FPC Functional Parameter Constraint</p> <p>Fr French</p> <p>GB Government and Binding</p>	<p>Gen Genitive</p> <p>HP Head Parameter</p> <p>I/INFL Inflection</p> <p>IND Indefinite article</p> <p>IP Inflectional phrase</p> <p>MA Moroccan Arabic</p> <p>Masc Masculine</p> <p>ML Matrix Language</p> <p>MLF Matrix Language Frame</p> <p>MSA Monolingual Standard Approach</p> <p>N Noun</p> <p>NEG Negation</p> <p>NEGP Negation Phrase</p> <p>Nom Nominative</p> <p>NPO Noun Phrase</p> <p>P Preposition</p> <p>Per Perfective</p> <p>Pl Plural</p> <p>Pre Present</p> <p>REF Reflexive</p> <p>SA Standard Arabic</p> <p>Subj Subjunctive</p> <p>TNS Tense</p> <p>TNSP Tense Phrase</p> <p>UG Universal Grammar</p>
--	--

In each example, Moroccan Arabic and its corresponding symbol in the citation (e.g. MA or Eng) will be *italicized*; with the English un-italicized to differentiate between the switches:

- [1] *mša tāht el bed u me ġedi -š yəxruj*  
 Went 3SG under DEF bed and NEG1 go NEG2 leave  
 'He went under the bed and isn't going to come out'
- [2] one time *y-dir ši hejja* and another time *ydir welu*  
 One time do 3SG some thing and another time do 3SG nothing  
 'One time he does something and another time he does nothing'

## INDIVIDUAL MOROCCAN ARABIC SOUNDS

1. Consonants: Moroccan Arabic has thirty-one consonants. Figure 1.1 below lists them according to their articulatory positions:

*Figure 1.1: Moroccan Arabic consonants*

	Bilabial	Labiodentals	Apical	Palatal	Velar	Uvular	Pharyngeal	Glottal
Voiceless Stops			t ṭ		k	q		ʔ
Voiced Stops	b ḅ		d ḏ		g			
Voiceless Fricatives		f	s ṣ	ʃ		x	ħ h	
Voiced Fricatives			z	j		ġ	ʕ	
Nasals	m ṃ		n					
Lateral			l ḷ					
Trill / Tap			r ṛ					
Semi-Vowels	w			y				

(Harrell 2004: 3)

2. As per Boumans (1998), in Moroccan Arabic five vocalic phonemes are distinguished: the full vowels /a u i / and the unstable, often evanescent ‘short’ vowels /e/ and /ü /. The phonetic realization of the vowels is heavily dependent on the adjacent consonants (1998: 406).
3. The informants for this thesis were predominantly from the Western or Eastern parts of Morocco, with dialects according to their specific region. The majority of the recordings incorporated dialects of the *ġreba* ‘Westerners’ or the *Ujeda* ‘those from the Oujda region.’ Consonant and vowel quality and variations are covered in Moroccan Arabic Sounds in *Figure 1.1* above.

## TRANSLITERATION AND TRANSCRIPTION

Phonetic translation is shown below in [square brackets] using the IPA. Moroccan Arabic data presented throughout this thesis is given in *italics* with the corresponding English language in intra-sentential code switching given in normal typeface and font. Moroccan Arabic data is transliterated in accordance with the following scheme with the exceptions of the following pharyngeals ħ and ʕ which are presented as such in the data.

### CONSONANTS

Note: Emphatics are noted by a subscript dot (e.g. ṡ, ṭ)

<i>Transliteration</i>	IPA
<i>θ</i>	[θ]
<i>ð</i>	[ð]
<i>ʃ</i>	[ʃ]
<i>q</i>	[q]
<i>x</i>	[x]
<i>ħ</i>	[ħ]
<i>ʕ</i>	[ʕ]

### VOWELS

Transliteration	IPA
ā	[a:]
ē	[e:]
ī	[i:]
ō	[o:]
ū	[u:]
ä	[æ]

## CHAPTER ONE

### OUTLINE OF CODE SWITCHING

Approaches to code switching (CS<sup>1</sup>) over the last few decades have usually taken one of two stances; that of a grammatical perspective which focuses on syntactic grammatical relations and that of a sociolinguistic perspective which largely deals with the role social factors have on speakers' relations and other socially-related phenomena. This thesis concerns itself with the former, analysing structural aspects of code switching, the aim of which is to determine the syntactic and morphological characteristics of theories posited in linguistic analysis and applying this to data collated, namely Moroccan Arabic/English code switched discourse.<sup>2</sup> This thesis thus intends to make a prominent contribution to the field of Arabic linguistics, sociolinguistics and discourse analysis. The main premise of this research is to analyse and qualify theoretical approaches to code switching and to analyse which are the most suitable vehicle for quantifying typologically dissimilar languages in bilingual clauses. The most notable of approaches to code switching is the Matrix Language Frame Model (Myers-Scotton 1993b, 2002) and the focus of this thesis is to

---

<sup>1</sup> There are several variations in spelling the term 'code switching' and these are: code-switching, Code Switching and code switching. I use the variety 'code switching' with a space, no hyphen and no capitalization throughout this thesis.

<sup>2</sup> A sociolinguistic analysis of Moroccan Arabic/English code switching is beyond the scope of this thesis due to time and space constraints. I refer the reader to Blom and Gumperz (1972) for their work on social factors which affect bilingual discourse, based on their ethno-linguistic study in Norway where they state that speakers' code choices are: "Patterned and predictable on the basis of certain features of the local social system" (1972: 424-425). See also Gumperz & Hernández Chavéz (1975) based on their research on Spanish and English code switching where they describe how code switching is a behavioural strategy reflecting notions of ethnic identity and confidentiality. In terms of social motivations for code switching and a micro-analysis of social factors associated with contact linguistics, is the major work undertaken by Myers-Scotton (cf. Myers-Scotton & Ury 1975; Myers-Scotton 1993b, 2002). Myers-Scotton (1993b) bases her Markedness Model on that of Grice's (1975) Co-operation Principle. Heller (1998) analyses the strategic use of code switching and conversation management based on Canadian French/English data. Gardner-Chloros (1995) in her analysis of Alsatian French/German data discusses the marked and unmarked choices employed by bilingual speakers in Strasbourg.

examine its linguistic viability and whether this model is suitable for the Moroccan Arabic and English intra-sentential code-switched data.

Contact linguistics and studies in bilingualism in general have witnessed a growing interest in code switching (Gumperz 1982, Li Wei 1994, Alfonzetti 1992, 1998, Poplack 1980, Auer 1984 and Myers-Scotton 1993b). This is described as the use of more than one language in the same conversation by speakers of various competencies. Although code switching has been an increasingly attractive area of research (e.g. Poplack 1980, Sankoff & Poplack 1981, Joshi 1985, Di Sciullo & Williams 1987, Belazi et al. 1994, Halmari 1997 inter alia), it remains one of the least defined manifestations of language contact. This is because the study of the phenomenon of CS borrows heavily from a variety of linguistic disciplines, each of which has contributed to it from a different perspective.<sup>3</sup> It is agreed upon by most linguists that code switching involves the alternation of at least two languages in any one bilingual conversation with linguists describing such linguistic phenomenon in an inter-sentential or intra-sentential fashion. Gumperz (1982) defines code switching as the juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or sub-systems.<sup>4</sup>

### **1.1 Language Choice**

Language choice is purely at the disposal of the switcher in the sense that fluent bilinguals do not necessarily switch to fill lexical gaps but rather *choose* to use one variety over another. This is evident in bilingual or multilingual Moroccans who not only have a High and Low variety of Arabic,<sup>5</sup> namely Standard Arabic and the more

---

<sup>3</sup> Aabi (1999: 4).

<sup>4</sup> Gumperz (1982: 59).

<sup>5</sup> See Ferguson (1959) for further analysis on Low and High varieties of Arabic and an overview of diglossia.

vernacular variety of Moroccan Arabic, but also French, Spanish and English, with Berber varieties also at their disposal. This is further corroborated by Romaine (1995) who states that in general, it would not be correct to say that speakers code-mix or switch to fill lexical gaps, at least not in the case of fluent bilinguals. Although it is popularly believed by bilingual speakers themselves that they mix or borrow because they do not know the term in one language or another, it is often the case that switching occurs most often for items which people know and use in both languages. The bilingual just has a wider choice – at least when he or she is speaking with bilingual speakers. In effect, the entire second language system is at the disposal of the switcher.<sup>6</sup> This is no more evident than amongst bilingual Moroccans where alternating between two varieties of Arabic is rather the norm than the exception:

- (1) *jib li -l drāri ma yaklāw*  
 Bring to DEF kids something eat 3PL  
 ‘Bring the kids something to eat’
- (2) *jib li -l bəz ma yaklāw*  
 Bring to DEF kids something eat 3PL  
 ‘Bring the kids something to eat’

The lexical items *drāri* and *bəz* above both mean ‘children’ but it is the choice of the speaker which item he or she employs.<sup>7</sup> As Labov (1972) states:

It is common for a language to give many alternate ways of saying the same thing. Some words like car and automobile seem to have the same referents, others have two pronunciations like working and workin’. There are syntactic options such as ‘Who is he talking to?’ versus ‘To Whom is he talking?’ or ‘It’s easy for him to talk’ versus ‘For him to talk is easy’<sup>8</sup> (1972: 188).

<sup>6</sup> Muysken (1995: 143).

<sup>7</sup> *Drāri* is used in Western Morocco, the dialect of which is considered a High variety in diglossic terms. *Bəz* is used in Eastern Morocco, a dialect which is considered androcentric and ‘very rural,’ namely a Low variety and less prestigious.

<sup>8</sup> This is not a case of style shifting but an example of code switched utterances. Romaine (1995: 170) further cites that style shifting accomplishes for the monolingual what code switching does for the bilingual. A choice between forms of one language, e.g. lexis, phonology etc. can convey the same kinds of social meanings as a choice between languages. Therefore, the varieties of Moroccan Arabic and Standard Arabic would not fall into the category of style shifting as the two are not mutually intelligible (cf. Nortier 1990) nor are they options open to the monolingual Moroccan (this is particularly apparent in rural Morocco where a high percentage of the Moroccan population is

Motivations for switching are numerous, one of which is in-group vs. out-group dynamics, in showing solidarity, implanting speaker identity, religion, topic and other socially motivated factors. Jorgensen (1998) states that: “Code switching into the minority language may be a tool to express solidarity” (1998: 239). However, amongst fluent Moroccan Arabic/French bilinguals for example, it is nigh on impossible to detect a ‘minority’ language where language alternation is simply through speaker’s choice. In this sense Moroccan Arabic in the UK would be the minority language on a macro level as English is the majority language. This is a growing problem for first generation Moroccans as language is inextricably linked to culture and identity not to mention religious practices, and language attrition is a fear for most Moroccans.<sup>9</sup>

## 1.2 Types of Code Switching

CS may also refer to switches in dialect, register and diglossia with High and Low varieties (Ferguson 1959).<sup>10</sup> The prototypical case of code switching sketched above

---

monolingual with a command of only smatterings of singularly added lexical items into the monolingual Moroccan Arabic lexicon.

<sup>9</sup> See also Jorgensen (1998: 238) and Zentella (1981) on language, power and minorities.

<sup>10</sup> Ferguson’s (1959) definition of diglossia and conversational code switching provided an innovative approach to describing language in society and bilingualism in general. He states that diglossia is a relatively stable language situation in which, in addition to the primary dialects of the language (which may include a standard or regional standards), there is a very divergent, highly codified (often grammatically more complex) superposed variety, the vehicle of a large and respected body of written literature, either of an earlier period or in another speech community, which is learned largely by formal education and is used for most written and formal spoken purposes but is not used by any section of the community for ordinary conversation (1959: 336). This also highlights conversational code switching, also known as ‘metaphoric’ and ‘situational’ as per Blom & Gumperz’s (1972) “Social Meaning in Linguistic Structures” (see also Myers-Scotton 1993b, Rampton 1995, Benson 2001). See also Gumperz & Hymes (1972) for more ethnographic analyses of language contact and behaviour where a more social approach to language contact is described in two Norwegian language varieties. They state that “choice amongst these [phonological, morphological and lexical] variables is always restricted by sociolinguistic selection constraints such that if, for instance, a person selects a standard morphological variant in one part of an utterance, this first choice also implies selection of pronunciation variables tending toward the standard end of the scale (1972: 416). They continue “the most reasonable assumption is that the linguistic separateness between dialect and standard is conditioned by social factors (1972: 417). Blom & Gumperz’s (1972) innovative analysis provided a bedrock for further research and study on sociolinguistic approaches to language contact and code switching as commented upon by Heller (1998a) who states that they [Blom & Gumperz 1972]

represents the alternational type: one in which a return after the switch into the previous language is not predictable. There is another type of code-switching where this is not the case and which may be called insertional. In this type of switching, a content word (noun, verb, rarely adjective/adverb) is inserted into a surrounding passage in the other language. As in alternational switching, participants show an orientation towards the 'other-languageness' of the insertion, either by deriving some particular interactional meaning from it, or by relating it to the speakers' (momentary) incompetence in the established language-of-interaction. In both cases, prosodic cues (extra emphasis, preceding pause) and verbal markers (meta-linguistic comments, hesitation) may serve to underline the juxtaposition and turn it into a locally noticeable phenomenon. Note that the insertion may be morpho-syntactically fully integrated or it may carry over grammatical elements into the receiving language.<sup>11</sup>

Recent studies of code switching have normally taken one of two stances; a grammatical perspective which focuses on structural aspects of code switching with an emphasis on syntactic relations and morphological characteristics of bilingual utterances or a sociolinguistic approach which is more concerned with the role of social factors, social motivations for switching, speakers' roles, competency and intentions. However, Auer (1984) posits that a balanced approach to code switching is necessary to focus on syntactic as well as sociolinguistic aspects: The prototypical

---

proposed a basic type of code switching, situational code switching, which is rooted in a social separation of activities (and associated role relationships), each of which is conventionally linked to the use of one of the languages or varieties in the community repertoire (1998a: 5). Descriptions of CS as a series of contextualisation cues ensued and have been written on quite considerably by linguists favouring a more sociolinguistic approach as opposed to that of the purely syntactic. Gumperz (1982) gives a general definition of 'contextualisation cues': Constellations of surface features of message form are the means by which speakers signal and listeners interpret what the activity is, how semantic content is to be understood and how each sentence relates to what proceeds or follows. These features are referred to as contextualisation cues (1982: 131). This then inspired other linguists to list a series of functions enabling switches (McClure 1988; Nishimura 1997; Zentella 1997; Romaine 1995). A notable work on socio-cultural aspects of code switching and a social psychological approach was that of Myers-Scotton (1983, 1993b, 1998) in describing certain psycho-social factors in why speakers favour one language variety over another. Such social approaches are beyond the scope of this thesis due to space as structural and syntactic relations of code switching are analyzed.

<sup>11</sup> Auer (1998: 1).

case of (discourse-related) code-switching seems to have the following main characteristics (a) It occurs in a sociolinguistic context in which speakers orient themselves towards a preference for one language at a time i.e. it is usually possible to identify the language-of-interaction which is valid at a given moment, and until code-switching occurs; (b) through its departure from this established language-of-interaction, code-switching signals 'otherness' of the upcoming contextual frame and thereby achieves a change of 'footing' where the precise interpretation of this new footing needs to be 'filled in' in each individual case, although previous episodes may also be brought to bear on the interpretation of the case at hand; (c) it seems possible to describe the mechanisms by which code-switching relates to the two codes and to the context in which it occurs in very general ways. Contexts are theoretically innumerable, of course, as are the interactional meanings of code-switching; however the ways in which these meanings are construed remain constant from one community to the next; (d) code-switching may be called a personal or group style. As a group style, its use may be subject to normative constraints valid within a speech community; however, it certainly is not a variety in its own right; (e) most code-switches occur at major syntactic and prosodic boundaries (at clause or sentence level). Since switching serves to contextualize certain linguistic activities, the utterance units affected by the switch must be large enough to constitute such an activity. For this reason, code-switching does not provide much interesting data for syntactic research; (f) although code-switching bilinguals may be highly proficient in both languages, balanced proficiency is by no means a prerequisite. Indeed, code-switching is possible with a very limited knowledge of the 'other' language.<sup>12</sup>

Constraints-based approaches to code switching have blossomed over the last thirty years with an emphasis on posting universal constraints on code switching with

---

<sup>12</sup> Auer (1998: 4).

constant re-evaluations and upgrades of theories. Code switching terminology, however, over the years has been less well-defined in describing language contact phenomena.<sup>13</sup> In fact, it was previously described as a random and haphazard consequence of language contact (Labov 1972, Weinreich 1968) but this has since been revised to a phenomenon with highly regularised grammatical constraints and well-structured syntactic relations. Heath (1989) describes code switching as being a pattern of textual production in which a speaker alternates between continuous utterance segments in one language  $L_x$  and another language  $L_y$  with abrupt and clear cut switching points, often at phrasal or clausal boundaries (1989: 23). This is common in early attempts at describing code switching which see it uttered with clear-cut and well defined syntactic boundaries. However, in this thesis I hope to show that this is not always the case, most notably amongst third generation speakers where switching takes place not only in intra-sentential segments but also mid-morphemically.

### 1.3 Early Analysis of Code Switching

Early work on code switching lent itself largely to very structured and specific switch sites where it was possible to switch from one language to another; between pronominal subjects and verbs (Timm 1975, Gumperz 1976, 1982), between

---

<sup>13</sup> Code switching as well as code mixing, language mixing, fused lects are all well-attested and documented in the literature with Auer (1998) describing code switching and fused lects as the polar extremes of the continuum and language mixing as the point between them (1998: 1). Auer (1998) continues in his description of such language terminology with code switching reserved for those cases in which the juxtaposition of two codes (languages) is perceived and interpreted as a locally meaningful event by participants. The term LM, on the other hand, will be used for those cases of the juxtaposition of two languages in which the use of two languages is meaningful (to participants) not in a local but only in a more global sense, i.e. when seen as a recurrent pattern. The transition from CS to LM is therefore above all an issue to be dealt with by interpretive sociolinguistic approaches since it is located on the level of how speakers perceive and use the 'codes' in question. Stabilized mixed varieties will be called fused lects. The transition from LM to FL is primarily an issue for grammatical research; essential ingredients of this transition are a reduction of variation and an increase of rule-governed, non-variable structural regularities (1998: 1). This thesis subscribes to the sole term 'code switching' for any juxtaposed use of two different language varieties in any single discourse irrespective of length of utterance, string or lexical insertion.

pronominal objects (Pfaff 1979) or simply listed switch sites and boundaries (Poplack 1980) within typologically similar languages which facilitated not only switch sites, but rendered constraints easy to follow and prescribe.<sup>14</sup> However, such early analysis of code switching was based on typologically similar language groups such as Spanish and English deeming the overall analysis as local solution constraints with an inductive motivation (Myers-Scotton, 1993b: 24). However, subsequent analyses employing typologically different languages have been evidenced: Finnish-English, (Poplack et al. 1987), Arabic-French, (Näit M'Barek & Sankoff 1988), Arabic-French, (Bentahila & Davies 1983), Tamil-English, (Sankoff et al. 1990), Wolof-French, (Poplack & Meechan 1995) and Standard Arabic-French-Moroccan Arabic (Aabi 1999).

Faced with numerous criticisms of such earlier approaches to CS, others came to the fore, most notably that of Joshi (1985) who built his production model on Garrett (1975). Joshi (1985) was one of the first to pioneer an asymmetrical model for bilingual discourse, highlighting two major languages at play and their respective syntactic relations is that the more dominant language, the matrix language, sets the grammatical frame and so insertions from the less dominant language, the embedded language, slot into the grammatical frame set by the ML. This model has been further advanced by Myers-Scotton (1993a) with the *Matrix Frame Model* (MLF) which distinguishes between lexical properties and functional categories or 'system morphemes'.<sup>15</sup> Further differentiation between such lexical elements and functional

---

<sup>14</sup> Poplack's (1980) linear model of the equivalent constraint will be discussed in more detail in Chapter three where a close analysis will highlight strengths and weaknesses utilising my data and that of other researchers.

<sup>15</sup> Such system morphemes were later revised and upgraded using the 4M Model (Myers-Scotton 2002) which in essence describes how morphemes are activated and accessed during speech production which is a more refined version of the original MLF model (1993a). The main argument is centred upon the notion that different lemmas underlying different types of morphemes become salient at different levels. One type of system morphemes is called 'early system' morphemes. The other type comprises two

categories highlighted permissible switch sites in CS (Azuma, 1993), Bentahila & Davies' *Sub-categorisation Constraint* (1983).

Interestingly, the cross-fertilisation of CS theories has been scant to say the least with each theory carving a niche of its own in the world of syntax and CS. Subsequent theories have been born since Joshi (1985), each with their own constraints and rule-based boundaries as will be discussed in this thesis, most notably in the next Chapter. There has been a move towards advocating general principles in governing CS as opposed to strict constraints-based approaches and this is based on such principles of monolingual grammar.<sup>16</sup> Theories based on this assumption tend to appeal to such grammatical properties as inter-constituent relationships (e.g., government, case assignment) and/or language-specific features of lexical categories (i.e., sub-categorisation of grammatical arguments, inherent morphological features; Poplack, 1980). The constrained nature of CS was further detailed in studies focusing on structural relations (Klavans 1985). CS and its relation with C-command and government were outlined and faced much criticism in Di Sciullo et al. (1986) where the underlying conclusion is that CS cannot occur where government relations hold. Empirical evidence by numerous linguists rendered this approach too severe and

---

kinds of later system morphemes ('bridge' and 'outsider system' morphemes). This will be addressed in more detail in Chapter three where the Matrix Language Frame (MLF) model is examined in detail.

<sup>16</sup> Gardner-Chloros & Edwards (2002) outline how one must firstly explain the notion of grammar before describing code switching constraints and grammatical relations: "the notion of 'grammar' needs to be submitted to a similar analysis before deciding what it means to 'seek grammatical regularities' in code-switched speech. For present purposes we can identify at least 5 different senses of the term 'grammar':

1. Prescriptive/pedagogical grammar: How you should use the language.
2. Chomskyan / Universalist Grammar: Theories about principles/constraints underlying syntax and morphology of all human languages – a concept like 'Government' is claimed to be a fundamental element in all grammars.
3. Formal / 'Parsing' grammars (Mahootian 1993): Explicit, formal statements about grammatical structure of particular languages (variants of Phrase Structure grammar).
4. Cognitive/functional/word grammars: Theories which do not suggest that there is a strict division between syntax, meaning and discourse functions.
5. Idiolectal competence: Within this notion George (1990) draws a distinction between what speakers know / believe about their grammar and how these beliefs are actually internally represented ('psychogrammar').

Grammatical studies of CS have on the whole been based on grammars in Sense 2 or Sense 3 (2002) 'Touching Base: The relevance of grammatical models to code switching,' *Actas/ProceedingsII Simposio Internacional Bilingüismo (1434)*.

limiting in its constraints and government relations. Moving away from government and onto feature checking (Chomsky 1995) led to the Functional Head Constraint (Belazi, Rubin & Toribio 1995) where feature agreement is of paramount importance prohibiting CS where a mismatch occurs.

#### 1.4 Minimalist Approaches

More recently the Minimalist approach (MacSwan 1999) based on Chomsky's (1995) standard theory restricts CS at structural sites showing cross-language differences in monolingual features (Poplack 1980).<sup>17</sup> Although a more detailed analysis of the models outlined here will be developed in Chapter Two) and their compatibility with my MA/Eng data and or any idiosyncrasies or issues highlighted, it is evident given the wealth of literature on CS and the wide-ranging counter examples that analyzing CS through purely syntactic means is an end in itself. As Gardner-Chloros & Edwards (2002) detail:

The behaviour of code-switching speakers eludes grammatical description in that it is highly variable (between and within communities and even on the part of individuals), and in that it exploits the propensity of speech – unlike writing – to avoid full, grammatical', sentences. It also leads to the development of more or less local conventions of its own, i.e. displays rule-creation mechanisms like other natural languages (2002: 1448).

This is a key argument in CS analysis as far too many linguists have adopted a 'one size fits all' constraints-based model. In accordance with Gardner-Chloros & Edwards' (2002), it should be noted that their argument is:

---

<sup>17</sup> The assumption that bilingual syntax can be explained by general principles inferred from the study of monolingual grammar has not yet been fully substantiated. While formal theories of grammar may well account for monolingual language structure, including that of monolingual fragments in CS discourse, there is no evidence to suggest that the juxtaposition of two languages be explained in the same way. Bilingual communities exhibit widely different patterns of adapting monolingual resources in their code-mixing strategies, and these are not predictable through purely linguistic considerations (Poplack 1987). The differences here then as per Poplack (1980) are that the mechanisms of monolingual and bilingual grammars are not assumed a priori to be identical (1980).

Not about whether grammar plays a role in CS, but about how best to characterize the level at which grammar operates. (2002: 1449).

Clearly then, grammatical analysis alone cannot, in my opinion account for the finer characteristics of CS alone. This thesis seeks to employ the broader characteristics of the Matrix Language Frame model (MLF) in purely descriptive terms and analyzes the discourse of bilingual Moroccan Arabic/English data through syntactic micro study. As Gardner-Chloros & Edwards (2002) summarize:

Although syntax plays an important role in CS, it cannot be assumed a priori that the constructs of syntacticians are necessarily the best means for characterizing the process of performance data such as CS. The possibility of throwing light on this question depends partly on whether or not it is right to assume that all bilinguals alternate in some meaningful way between two clearly distinguishable sets of rules – and this is a question which manifestly cannot be decided by grammatical analysis alone (2002: 1449).

### **1.5 Code Switching, Interference and Bilingual Considerations**

An ongoing problem in studies on CS has been that of the borrowing versus code switching distinction. Interference, on the other hand has not played a major role in the debate between code switching and borrowing, being termed a speaker-specific deviation from the language being spoken due to the influence of the other deactivated language.<sup>18</sup> The basic premise here is that there should be a strictly monolingual context in which only one language must be operational. If traces from the other language which is supposed to be deactivated appear, interference will take place.<sup>19</sup> Interference also differs from borrowing mainly in that the former is contingent and systematic and the latter is collective and systematic.<sup>20</sup> Namely, interference is not a coded, systematically structured feature but is rather idiosyncratic in its nature. By contrast, borrowing is a property of the speech community rather than

---

<sup>18</sup> Grosjean (1995: 262).

<sup>19</sup> Aabi (1999: 4).

<sup>20</sup> Mackey (1968: 569).

the individual. It is not sporadic and is highly regularised. While interference can touch all levels of language,<sup>21</sup> phonological, syntactic, lexical, semantic and/or pragmatic, borrowing is often restricted to the lexical level wherein a word or short expression is taken from another language.<sup>22</sup> There are in essence two stances with regard to interference: positive transfer and negative transfer where the former normally facilitates the learning of L2 via L1. Negative transfer as evidenced in some migrant communities and adult second language learners is where previous language learning hinders the acquisition of L2.<sup>23</sup> CS on the other hand involves a bilingual context which assumes the activation of two languages although they may have different levels of operationality in the sense that one language is less activated than the other but never completely deactivated (cf. Myers-Scotton 1993, Grosjean 1995).

From a global societal perspective, of course, most of the world's speech communities use more than one language and are therefore multilingual rather than homogenous.<sup>24</sup> A leading factor of language change and shift is that of ever increasing migratory patterns, most notably rendering generations of bilingual speakers.<sup>25</sup> Such language shift is a change from the habitual use of one language to that of another (Weinreich 1967:68). Two manners in viewing bilingualism in relation to language shift are those outlined by Lambert (1975) namely additive and subtractive bilingualism. Additive is when speakers learn an L2 in addition to their L1 for work, or any other reason. Morocco is an example of such multilingualism where

---

<sup>21</sup> This is by no means an exhaustive account of interference. For an in-depth analysis see Romaine (1995) and Grosjean (1995).

<sup>22</sup> Ibid Grosjean (1995: 213).

<sup>23</sup> Ibid (1995: 52-55) for a further illustration of how the notion of 'static' and 'dynamic' is postulated for the concept of interference. Static interference is a reflection of permanent traces from one language onto the other such as permanent accent and the meaning extensions of particular words. Dynamic interference is associated with ephemeral influence from the other language not being spoken such as accidental slips on the stress pattern of a word and momentary use of a syntactic structure. In either case, Grosjean notes that it is easier to study interference within a monolingual rather than bilingual mode.

<sup>24</sup> Romaine (1995: 8).

<sup>25</sup> For further details on factors which contribute to societal multilingualism and code switching see Sridhar (1996: 48).

multifarious languages are accommodated and spoken widely in both rural and urban domains including Arabic, French, Spanish and different Berber varieties<sup>26</sup> as well as English with the growth of its global influence. Subtractive bilingualism is becoming increasingly common by third generation speakers whereby their L1 is replaced by a dominant L2 variety.

### 1.6 Code Switching and Borrowing

Distinguishing between CS and borrowing is by no means a facile task. However, certain linguists (e.g. Gumperz 1982, Poplack, Wheeler & Westwood 1987) highlight the use of phonological and/or morpho-syntactic integration as a key feature in its differentiation. Borrowing then is thought to involve phonological and/or morphological integration of elements or structures from lexicon into another as with the Moroccan Arabic *bisklit* for ‘bicycle’, *tumubīl* for ‘automobile/car’ and *zalamīt* for the French ‘*les allumettes*’ (matches) to name but a few. Phonological assimilation is considered as the chief, if not the only determining feature of borrowing (Halmari 1997:173). The following is an example of such integration:

- (3) Oli *putter* klöntti  
Was butter lump  
‘That was a lump of butter’  
(Finnish/*English*, Halmari 1997:47)

*Putter* above shows its almost complete assimilation from English into the phonological system of Finnish.<sup>27</sup> This is a usual pattern of almost total phonological integration from one system into the lexicon of another. However, this is not the only distinguishing feature of borrowed versus code switched items as outlined by Myers-

---

<sup>26</sup> The linguistic situation of Morocco will be discussed in more detail in Chapter Five. For a more detailed account see Sadiqi (2003), (Ennaji, 2005).

<sup>27</sup> For more on Finnish/English code switching, see Halmari (1997).

Scotton (1995) where she explains that while most established forms may well be phonologically integrated to the ML, by no means all borrowed forms show such integration (1995: 601). In addition, the total phonological integration of borrowing implies the complete un-assimilation of CS.<sup>28</sup>

However, the CS /borrowing debate is less than clear cut and phonological integration is not the only overriding feature of its explanation. Further morpho-syntactic features have been put forward as a determining factor in distinguishing CS from borrowing (e.g. Poplack 1980) in identifying CS as a completely separate language phenomenon. There has been a wave of supporting evidence in exemplifying CS from borrowing (Pfaff 1979, Bentahila & Davies 1983, Gardner-Chloros 1987, Heath 1989, Bassiouney 2006). The example below uttered by a first generation Moroccan in the UK portrays classic integration of the borrowed item ‘biscuit’:

- (4) *jib li ya wahəd al biscuita*  
 Bring to me one DEF biscuit AGR fem  
 ‘Bring me a biscuit’

The above example shows the Matrix Language (ML henceforth) to be Moroccan Arabic with the Embedded Language (EL) English supplying the borrowed and integrated lexical item ‘biscuit’ which has undergone complete phonological and syntactic integration into the MA. The final *-a* indicates the feminine marker usually reserved for inanimate objects in MA.<sup>29</sup> This is further highlighted in Heath (1989):

- (5) *šrīt wahəd l ananasa*  
 I-bought one DEF pineapple-FEM  
 ‘I bought a pineapple’  
 (MA/French, Heath 1989).

<sup>28</sup> Aabi (1999: 7).

<sup>29</sup> This is discussed in more detail in Chapters Three, Five and Six where an innovative differentiation between syntactic and semantic object agreement is analysed in intra-sentential strings.

As discussed by Aabi (1999) the EL lexeme *ananasa* has undergone full integration at the phonological, morphological and syntactic levels. At the phonological level, the sounds /n/ and /s/ have been pharyngealised into /ŋ/ and /s/, pharyngealization being a property of MA but not French. At the morphological level, integration caused the inflection of the French lexeme with the Moroccan Arabic feminine morpheme /-a/. At the syntactic level, the form *ananasa* has been integrated into the Moroccan Arabic system thus producing a type of structure [which] would be ungrammatical in French. Note here that in stark contrast to Moroccan Arabic, French syntax allows nouns such as *ananas* to be preceded by strictly one definite or indefinite article. Moroccan Arabic nouns, on the other hand, can be modified by a double determiner (numeral + definite article) to indicate definiteness (Aabi, 1999: 8). As suggested by Bentahila & Davies (1991) not all morpho-syntactically integrated items are borrowings but are classic cases of CS:

- (6) qeli            “tu pisses” w    kayn msākīn ta-ymši-w        l- la    toilette  
 he-told-me “you pee” and was    poor<sub>plr</sub> go-AGR PL        to DEF toilet
- beš        y-pissi-w  
 so that pee-3PL  
 ‘He asked to pee and there were poor others going to the toilets for a pee’  
 (MA/French, Bentahila & Davies 1991: 384).

This is further corroborated in my own data where English lexical items are fully integrated both morphologically and syntactically into the MA matrix frame:

- (7) šhāl            tə-complain,    wa ana ma    na-ħmæl        -š    a    nisa    illi  
 How much TNS complain and I    NEG<sub>1</sub> tolerate 1SG NEG<sub>2</sub> DEF women REL  
 y-complain-iw bezzaf  
 complain 3PL a lot  
 ‘How she complains and I can’t stand women that complain’

The above example, although I am cognizant of the fact that there are fully integrated forms of the English ‘complain,’ is a classic case of MA CS because borrowing is a

process that is open to monolinguals as well as bilinguals, whereas the meaning or communicative function of the alternation in (6) is hardly recognisable to the Moroccan Arabic monolingual (Aabi 1999: 9) although contextualisation cues may make this intelligible nonetheless.<sup>30</sup> It should further be noted that switches such as (7) above are not isolated idiosyncratic incidences but rather are quite common, particularly amongst third generation MA/Eng speakers. This will be elaborated upon further in subsequent Chapters.

The determining factor then, on employing such integrated forms, or classic CS, is speaker audience and competency. This type of ‘borrowing’ is referred to as ‘nonce’ borrowing (Poplack et al. 1998) although I reserve the right to reject the category of ‘nonce’ borrowing in the same way as outlined by Myers-Scotton (1992: 32), who argues that it is too extraneous and does not add any explanatory value to the study of CS<sup>31</sup>. Hypothetical examples are a good way of corroborating or vindicating analyses and drawing conclusive illustrations on how CS and/or borrowing is processed. The following shows that the grammatical linking unit, if *beš* in MA, can only be followed by a finite clause:

(8) \*nqra šwiya beš *rèussir* à l'examen  
 Work 1SG a bit so that to succeed at the exam  
 ‘I work a bit harder to pass my exams’

(9) nqra šwiya beš **n-rèussir** à l'examen  
 Work 1SG a bit so that succeed 1SG at the exam  
 ‘I work a bit harder to pass my exams’

(MA/French, Bentahila & Davies, 1983: 323)

<sup>30</sup> The debate over lexical insertions in recent years has become oriented around two prominent poles. On the one hand, Poplack and her associates (e.g. Poplack, 1988; Poplack, Wheeler & Westwood, 1990, Poplack & Meecham, 1998) have shown in a number of studies that lexical L2 insertions are often fully integrated into the grammar of the matrix language. Such insertions are interpreted as borrowings. Myers-Scotton (1997, 2002) has also described the grammatical integration of lexical insertions, but considers the phenomenon in terms of a theory of code-switching. Both writers have also noted that so-called bare forms which lack expected matrix language grammatical properties are anomalous with respect to these explanations (Owens 2005: 23).

<sup>31</sup> See Bentahila & Davies’ (1983) study on French and Moroccan Arabic code switching where they give examples of such switches (1983: 322-4) and attribute them to the need for sub-categorization rules. In accordance with the Matrix frame, the grammar must adhere to sub-categorization rules of either French or Moroccan Arabic and anything else is rejected.

As outlined above, (8) is deemed ungrammatical as it violates sub-categorization constraints in MA. However, in (9) above, *beš* introduces a finite clause and satisfies this constraint. Further examples follow:

(10) \**qul li ya beš go ʔamma bi nafsi wa na-* talk to him  
 Tell to me so that go there with myself and talk 1SG to him  
*bi nafsi*  
 with myself  
 ‘Tell me so I go there myself and talk to him myself’

(11) *qul li ya beš n-go ʔamma bi nafsi wa na-* talk to him  
 Tell me so that go 1SG there with myself and talk 1SG to him  
*bi nafsi*  
 with myself  
 ‘Tell me so I go there myself and talk to him myself’

Again, (10) above is ungrammatical as it does not follow sub-categorization requirements for MA which stipulate that *beš* must be followed by a finite clause whereas (11) does since *beš* introduces the finite clause as required and hence is grammatical. Therefore, we can conclude and concur with Myers-Scotton (1993a: 163) that code switching and borrowing undergo largely the same morpho-syntactic procedures of the matrix language during language production. Therefore, it transpires that the motivation for distinguishing them in order to assess models of morpho-syntactic constraints seems to evaporate, at least for content morphemes. I conclude that syntactic constraints which govern CS in general must therefore govern borrowing in largely the same format particularly as they are “part of the same development continuum, not unrelated phenomena” (Myers-Scotton 1993a: 163). This thesis does not seek to provide a detailed and exhaustive account of borrowing as it is, largely due to time and space, beyond the scope of this research but serves to highlight certain morpho-syntactic categories and better illustrate my intra-sentential code-switched data.

## 1.7 CS and Bilingual Proficiency

Bilingual speakers who code switch have been analysed at best as either proficient in their linguistic abilities or at worst as having a poor command of their own native tongue and are generally viewed negatively.<sup>32</sup> Poplack (1981) found that complexity of intra-sentential code switching required that the speaker has a sophisticated knowledge of the grammars of both languages, as well as knowledge of how those grammars map onto one another. Less proficient bilinguals favoured single-word and tag switches, while more proficient bilinguals code switched at the phrase and clause level as well. Code switching in itself implies at least *some* degree of competence and proficiency in two languages or more. In bilingual and multilingual societies, code switching should be seen as the norm. As Gumperz (1982) points out:

Speakers communicate fluently, maintaining an even flow of talk. No hesitation pauses, changes in sentence rhythm, pitch level or intonation contour mark the shift in code. There is nothing in the exchange as a whole to indicate that speakers don't understand each other. Apart from the alternation itself, the passages have all the earmarks of ordinary conversation in a single language (1982: 60).

Bilingual discourse should be analyzed in terms of where the switches take place in order to fully assess a speaker's language ability.<sup>33</sup> Inter-sentential switches which take place in between clauses and are the most facile of switches as they set out clear-cut switch demarcation lines:

- (12) *Meši bəzzef.* As I said, *qlil ši fin ta-lqa* a nice man  
Not a lot. As I said, little thing where find 2SG a nice man  
'Not a lot. As I said, it's rare you find a nice man'

---

<sup>32</sup> Stevens (1983) found that local attitude to code switching in Tunisia greatly favoured Arabic French. This is largely due to tradition, religion and culture.

<sup>33</sup> See also Levelt (1989) for bilingual language production model based Levelt's model which is unique in that it attempts to integrate independent, automatic modules into a complete 'speaking system'. He has five models involved in language production (conceptualizer, lexicon, formulator, monitor system and articulator). De Bot (1992) was the first to postulate a bilingual language production model based on Levelt's (1989) model for monolinguals. The subsystems hypothesis assumes that each language system constitutes a subsystem in which both languages would be stored together (Paradis, 1981).

- (13) *ruḥ ʔamma u min tə-wsəl*, get Dad on the phone.  
Go there and when arrive 2SG, get Dad on the phone  
'Go there and when you arrive, get Dad on the phone'

(12) and (13) above split the bilingual discourse into two neat clauses, or Complementizer Phrases (CPs)<sup>34</sup> with the Moroccan Arabic clause syntactically juxtaposed to the English clause. With regard to intra-sentential code switching, in terms of linguistic competence, Lipski (1982) describes intra-sentential switching as the most highly-developed of language use. However, one does not have to be fluent in two or more languages to be able to code switch. Many English speakers employ token Arabic phrases in their everyday speech, invoking religious sentiments, "Bismillah" 'In the name of Allah', "Allah yarḥamu" 'May he rest in peace', "Al ḥamdullilah," 'Praise be to Allah', and employing traditional banner statements during Islamic festivals such as "ʔīd Mubārak" 'Happy Eid.' Also, these token statements are commonly used during everyday activities: "Bi sāḥḥa" 'with health' is said to someone when buying new clothes or after having a bath or shower and "Yarḥamak Allāh" 'God bless you' (after sneezing). In this vein, it can be said that speakers do not need to be fully-fledged bilingual speakers and are actually able to 'get away' with being a *pseudo* bilingual if the right things are said at the right time. This is no more evident than amongst some third generation Moroccan Arabic/English speakers when engaging with first generation speakers, or grandparents back in Morocco. With some third generation speakers born and raised in the UK, their language of 'comfort' is English with Moroccan Arabic reserved for family discussions, bi-cultural settings and the like. As detailed by Poplack (2004):

In many bilingual communities, speakers conventionally make use of both languages with the same interlocutors, in the same domains, and within the same conversational topic (Poplack 2004: 10).

---

<sup>34</sup> The unit of analysis is CP as per Myers-Scotton (1993a) to be discussed in more detail in Chapter Three under the rubric of the Matrix Language Frame (MLF) model.

This then is also very much the reserve of language choice where speakers, irrespective of fluency select the moment at which they wish to insert lexical items from one language into the matrix frame of another. As Fishman (1972) states: “Habitual language choice in multilingual speech communities is far from being a random matter of momentary inclination” (1972: 437).<sup>35</sup> Linguistic proficiency is not based solely on capability but also upon other language factors which influence language choice such as audience, subject, competency, background and education.

Moroccan Arabic/English speakers then are able to negotiate their use of language which largely depends on the factors listed above, but most importantly, whether or not they are proficient enough to use two languages in a productive manner, or if not, to use one language productively, and the other receptively. Culture also plays a large part in the behaviour of ethnic minorities the world over as culture is in essence the bedrock of language preservation. Furthermore, as Duran (1994) states:

Viewing and constructing the world from one cultural point of view may appear to be more normative and refined and therefore more conventionally accepted. The same constructs can be viewed however, from two or more world views in a rich bilingual/multicultural environment. In this case, one language might help the other, and sometimes, both together may create a new idea, image, thought, behaviour, outlook, organization, and adaptation and thus move culture to new adaptive places in the dynamics of cross-cultural life (1994: 75).<sup>36</sup>

---

<sup>35</sup> See also Myers-Scotton (1980, 1983, 1993b) where the Markedness Model sets out maxims to account for code switching behaviour. These maxims are: (i) the Unmarked Choice, (ii) the Marked Choice, and (iii) the Exploratory Choice. A detailed analysis of this is beyond the scope of this work.

<sup>36</sup> See Corder (1981) on the notion of Interlanguage and second language acquisition. Corder in his visual model of Interlanguage describes the use of two or more languages and the effect they have on the speaker. In this analysis of the effects of interlanguage are set out in three parts: (i) that Language A and Language B have their unique results unto themselves, (ii) that each language will always in some way be tied to the other, (iii) that a third way of using language cognitively and communicatively will be unique and specific to the developed bilingual; metalinguistic ability, cognitive flexibility, conceptual elaboration, divergent thinking, bilateralism, (right/left brain hemispheric function), cognitive pluralism, code switching, code-mixing, and other forms of language pattern not common to monolingualism.

## 1.8 Conclusion

I have, in this chapter, attempted to outline introductory elements to CS. I have also sketched a brief account of borrowing and interference and their relation to CS, an exhaustive account of which is neither possible nor relevant. Following Myers-Scotton (1993b), I would claim that CS and borrowing are constrained in the same way. The second Chapter deals with an analysis of the main approaches to code switching, commencing with the linear models, followed by other models which incorporate Government and Binding as their basic premise. In addition to these, I consider more insertional approaches and examine the main theories and syntactic approaches to intra-sentential code switching in general and their applicability to Moroccan Arabic/English data collated for this thesis. The principal approaches to such linguistic exchanges are linear, Government and Binding, asymmetric, constraint-based, non-constraint-based and the minimalist approach.

In Chapter Three, I examine Myers-Scotton's Matrix Language Frame model as this forms the basic structure for Moroccan Arabic/English data examined in this thesis. Furthermore, the MLF model is examined in detail in outlining various insertional approaches, the content – system morpheme dichotomy and various extended sub-models, namely that of the 4-M model and the Abstract Language model. Chapter Four discusses grammatical categories of Moroccan Arabic grammar. These are an essential aspect of this thesis which highlights the syntax specific to this language enabling a better understanding of the matrix language – embedded language dynamic in which certain structures are possible in a code-switched domain and others are not. It is important to include a grammatical sketch of Moroccan Arabic as English insertions and inflectional affixations are made clearer when the basic construct is analysed. This facilitates the study of intra-sentential code switching in this domain.

Chapter Five commences with a discussion on Moroccan migratory patterns over the last few decades. This is then followed by a focus on the Moroccan Arabic data corpus, methodology and research design in using quantitative and qualitative methods. Transcription methods and a full analysis of the statistics collated from the data are examined in full. Chapter Six analyses certain grammatical categories in code switching from English to Moroccan Arabic under the premise of the MLF the validity of which is corroborated and verified. The treatment of Moroccan Arabic and English code switching within the Matrix Language Frame model is further examined and validated in Chapter Seven. In this regard, the main purpose of this chapter is the analysis of whether embedded Moroccan Arabic nouns and other grammatical categories adhere to the main applications and principles of the MLF model (Myers-Scotton 2002) and whether this is in essence a viable and suitable vehicle overall for the linguistic analysis of Moroccan Arabic and English code switching. The final Chapter summarises the findings of the research and concludes with an analysis of the MLF model in the light of all the data presented in the thesis and also in comparison with other models as discussed in Chapter Two and throughout the thesis. The eleven generalizations which I have made throughout the course of research are examined and concluding remarks drawn.

## CHAPTER TWO

### SYNTACTIC THEORY AND APPROACHES TO CODE SWITCHING

This chapter seeks to examine the main theories and syntactic approaches to intra-sentential code switching in general and their applicability to Moroccan Arabic/English data collated for this thesis. The principal approaches to such linguistic exchanges are linear, Government and Binding, asymmetric models, constraints-based, non-constraints-based and the Minimalist approach. In essence, code switching analysis over the years has focused on finding the most suitable and all-encompassing solution to grammatical constraints governing intra-sentential CS and is divided amongst linguists who think there are constraints (Belazi, Rubin, & Toribio 1994, Bentahila & Davies 1983, Di Sciullo, Muysken & Singh 1986, Myers-Scotton 1993b, Poplack 1980, Sankoff & Poplack 1981, Pfaff 1979, Santorini & Mahootian 1995) and those who claim that there are no universal, purely syntactically driven constraints on CS thus setting it apart from monolingual speech applications (Bokamba 1988, Clyne 1987, MacSwan 1999). Early attempts at explaining CS and its rules and regularities initially commenced with lists of where syntactic switches may or may not occur and these are discussed at the beginning of this chapter. Timm (1975) began with such a structured approach outlining the main switch sites, positing five constraints (using data from typologically similar languages; English and Spanish),<sup>37</sup> namely that switching does not occur within NPs containing nouns modifying adjectives, between finite verbs and their infinitival complements, and between pronominal subjects and their verbs. An attempt by Wentz & McClure (1976)

---

<sup>37</sup> Employing the typologically similar languages English and Spanish was as Myers-Scotton points out a “local solution” with regards to intra-sentential CS (1993a: 24). Linguists during this period and shortly after began to move away from the more descriptive approach and towards more theoretical models as syntacticians working independently within the “Extended Standard Theory” of the 1970s enumerated constraints on transformations, constraints on phrase structure and constraints on surface structure (Newmeyer, 1986).

and Pfaff (1979) to refine this approach was again modelled on typologically similar languages and was faced with much criticism as what was previously thought to be disallowable switch sites in intra-sentential English and Spanish CS, was later identified as allowable. However, these early approaches served to highlight the salient point that CS is not a random and haphazard phenomenon but is in fact rule-governed and stylistic and this is analysed in this chapter together with the most prominent morpho-syntactic approaches to CS, namely linear, asymmetric, government and the Minimalist framework. The most prominent are Sankoff & Poplack's (1981) Two x Constraints theory, the Matrix and Embedded language theory (Joshi 1985, Myers-Scotton 1992) and finally the Government Principle (Muysken et al. 1986). The salient questions are in which way do these models apply to languages? Do such constraints apply to only typologically similar languages? Are they universally acceptable? I will detail the concepts of the models and how they fit into my Moroccan Arabic/English data but will focus more on the Matrix Language theory as this is the current model I am adopting of all three variables as my research inherently points to an overriding asymmetrical model where syntactically of the two (or more) languages involved in bilingual discourse, a grammatical 'leader' and embedded 'led' dynamic emerge within the syntactic frame.

## **2.1 Language Specific Constraints**

Bilingual discourse analysis and research into the syntax of code switched utterances has gained momentum in the last few decades. This in turn opened the floodgates to the postulation of an array of language-specific constraints in CS with researchers attempting to procure a constraint that is empirically sound and viable for most languages. The last couple of decades witnessed an emergence of more structurally

defined models to account for CS where its occurrence has been shown to be governed not only by extra-linguistic factors (social and situational) but notably intra-linguistic (structural). The following sections examine such approaches with counter-examples highlighted where necessary given the Moroccan Arabic and English code-switched data collated. We begin with the early linear approaches.

### **2.1.1 Timm's Linear Model**

Amongst the first to define CS on a more linear approach was Timm (1975) who described how there were certain syntactic environments where code switching could not occur where certain grammatical rules are valid for certain languages. In so doing, Timm posited five constraints on his data of Spanish/English CS. Switching was found not to occur; a) between pronominal subjects or objects and finite verbs; b) between finite verbs and their infinitive complements; c) between auxiliaries and main verbs; d) between negation and the negated verb; or e) in certain noun phrases containing an adjective (Timm, 1975: 477-80).

Earlier research typified CS as random and haphazard with no scope for rule-governed domains or bilingual speaker control.<sup>38</sup> This was firmly attested by Weinreich (1953):

The ideal bilingual switches from one language to the other according to appropriate changes in the speech situation (interlocutors, topics, etc.), but not in an unchanged speech situation, and certainly not within a single sentence (1953: 73)

Labov (1971) concluded that:

---

<sup>38</sup> One of the forerunners in CS research was the publication of research by Jan Blom & John Gumperz (1972) where they dealt with CS between dialects of Norwegian. Another notable study was that of Haitian Creole and French (Stewart 1968) although this study does not concern itself with diglossic CS championed by Ferguson (1959).

No-one has been able to show that such rapid alternation is governed by any systematic rules or constraints and we must therefore describe it as the irregular mixture of two distinct systems (1971: 457).

The prevailing view at the time was that language contact studies largely considered CS as a by-product of the imperfect bilingual, where an inability to maintain a language in one code was a disadvantage. On the whole, attitudes towards CS were generally seen as negative (Romaine 1995) where people in Morocco who code-switched were seen as 'still colonized' (Bentahila & Davies 1983). CS is also defined as a pattern of textual production in which a speaker alternates between continuous utterance segments in one language  $L_x$  and another  $L_y$  with abrupt and clear-cut switching points, often at phrasal or clausal boundaries (Heath 1989: 23). We later witness how 'clear-cut' is often not the case particularly in new emerging forms of syntax of distinct typologies.

The revolution of Chomsky's Universal Grammar led a charge for researchers to posit a similar universality to CS with several authors claiming universal validity for any data in any language pair. The last couple of decades witnessed an emergence of more structurally defined models to account for CS where its occurrence has been shown to be governed not only by extra-linguistic factors (social and situational) but notably intra-linguistic (structural). The most prominent are Sankoff & Poplack's (1981) Constraints theory, the Matrix and Embedded Language Theory (Joshi 1985, Myers-Scotton 1992) and finally the Government Principle (Muysken et al. 1986) to be discussed.<sup>39</sup> We continue with language-specific constraints.

---

<sup>39</sup> In contrast to the linear approaches posited (Pandit 1990) where switching takes place at a surface structure level there has been a contrastive governmental approach where the structure rules are hierarchical. This was to a certain degree within the theory of Chomsky's (1981) government and binding (GB) theory. Here the analytical thrust focused on dependency rather than equivalence so that a switch cannot occur between two constituents if they are lexically dependent on each other as outlined in Di Sciullo, Muysken & Singh (1986): Government Constraints (GC) theory:  
(a) If X governs Y, ...Xq ... Yq....(Di Sciullo et al., 1986: 5). In other words, "Whenever constituent X governs Y both constituents must be drawn from the same language" (Appel & Muysken, 1987: 365). This traditional formula of the X-bar theory is stated, "the lexical governor and the highest lexical

## 2.1.2 The Conjunction Constraint

Kachru (1978) in his analysis of Hindi and English CS found that English conjunctions cannot appear alone in Hindi-English CS<sup>40</sup> and thus employed the notion of The Conjunction Constraint:

*The Conjunction Constraint (Kachru 1978: 39):*

*In code-mixing of South Asian languages the English conjunctions and, or etc., are not used to conjoin non-English NPs or VPs.*

Kachru (1978) highlights an ungrammatical example which shows the ungrammaticality of the code switched item as it conjoins two mixed language clause:

- (15) \*maī usko akhbār deta but diyā nahī  
I to-him newspaper would give but gave NEG  
'I would have given him the paper, but I didn't'  
(Hindi-English, Kachru 1978: 39)

At first glance, the fact that this constraint is language-specific in that it only pertains to Hindi-English combinations appears to limit the empirical scope of this theory. Kachru (1982) later admitted that there is not yet a study which gives a typology of constraints which is empirically valid across all languages. Also, other linguists have empirically tested the validity with this constraint on different language pairs and found this not to hold true (Bentahila & Davies 1983, Bokamba 1989). It also does not hold true of my own data in the below English to MA to English switch:

---

element of the governed maximal projection need to be in the same language" (1986: 119). This then translates that typical cases would be that of case assignment or subcategorization where within a maximal projection, no switch is permissible. In its bare form, this constraint was a target for many counter examples as complements of verbs and appositions are amongst the most common switch types (Boumans 1998). Switching between functional categories such as Negation and Determiner are considered ill-formed. However, this constraint would be valid if it were language and complementizer specific:

He's a wəld lə hrām  
boy DEF badness  
'He's a son of a bitch'

As it happens, both governed and ungoverned, as well as governing content words and constituents can be 'switched' (inserted) so that, while it will be possible to describe codeswitching in GB terminology, the government principle itself does not seem to constrain switching (Boumans, 1998: 22).

- (16) I told him *bəssaħ* he said no  
           but  
       'I told him but he said no'

And also, in the other direction, MA to English to MA respectively:

- (17) *meši* *Əamma* but *Əamma*  
       NEG there but there  
       'Not there but there'

The example as cited by Bentahila & Davies (1983) on further introspective analysis seems a little odd:

- (18) ana t-anxarj had -ši kul -u et tan-dir l ma  
       I take 1SG DEM thing all of it (MASC) and do 1SG DEF water  
       'I take all of this stuff out and add water'  
           (Moroccan Arabic/*French*, Bentahila & Davies, 1983: 310)

However, as Clyne (1987) details:

It is a matter of doubt whether the notion of grammaticality can be applied at all to data as variable as that of code-switching. 1987: 744).

It would seem then that most linguists are in a vulnerable situation in attempting to 'strait-jacket' code switching into using a uniform technique in describing its grammatical structure and constraints. Gardner-Chloros & Edwards (2002) cite the reason for this as *variability* (2002: 1435) which is why CS data poses problems for grammatical description (2002). Gardner-Chloros & Edwards (2002) continue with the second reason as being that of general grammar which is essentially a description of properly formed sentences, as found principally in the written language. It is at least questionable whether CS discourse can be meaningfully analysed in terms of syntactic categories such as 'noun', 'noun phrase', 'clause' etc (2002: 1435). Auer (1998) also describes the difficulty in applying a grammatical

analysis with regards to CS as it will only be able to account at best for some of the patterns in the data.<sup>41</sup>

A further reason in trying to syntactically dissect CS is cited as that of certain lexical let-outs as detailed in Gardner-Chloros & Edwards (2002) who describe how code switchers take advantage of various 'let-outs' to avoid the straitjacket of grammatical rules. One example is the use of a type of CS variously described as ragged (Hasselmo, 1972), paratactic (Muysken, 1995), disjointed (2002: 1436).<sup>42</sup> However, according to Bhatia & Ritchie (1996), note that the term "constraint" is used in two very different senses in the code switching literature, one descriptive and the other theoretical. In the descriptive sense, when we speak of constraints on code switching, we mean only that some code-switched constructions are well-formed and others are ill-formed as shown in (19) below (Belazi, Rubin & Toribio, 1994; cf. Timm, 1975):

(19a) The students *habían visto la película italiana*  
The students had seen the Italian movie

(19b) \*The student had *visto la película italiana*  
The student had seen the Italian movie

Bhatia & Ritchie (1996) continue to describe how although the word order is the same in both instances [above], the switch in (19b) is judged to be ill-formed. This fact shows that code switching behaviour, is constrained or rule-governed, but does not in itself tell us what the nature of the underlying rule system it (1996: 258). Before we begin with an historical account of the first grammatical models in CS theory, it should be noted that the notion of "constraint" as described by Bhatia & Ritchie (1996)

---

<sup>41</sup> Auer (1998: 3).

<sup>42</sup> A novel approach to analyzing CS is required in order to circumvent such negative and disparaging accounts of CS. This will be evidenced later in this thesis in my analysis of second and third generation Moroccan Arabic/English speakers who employ a new form of CS mainly in employing new verbal combinations: I was *kūl*-ing 'I was eat-ing' where the lexical item *kūl* 'to eat' is paired with the English gerund to derive a new morphological variety. I have termed this '*Reactive Syntax*'.

is that which applies to a system of linguistic rules or to the form of a representation, and attempts to capture a range of linguistic facts (1996: 258). Among the first to quantify the meaning of “constraint” and apply such a rule-governed approach is Poplack (1980) utilising an equivalence principle model.<sup>43</sup>

### **2.1.3 The Constraints Theory**

This was one of the principal and most influential attempts at applying a syntactic constraint on CS by Sankoff & Poplack (1981) based on word order equivalence.<sup>44</sup>

The two main constraints were that of the Free Morpheme Constraint and its sister the Equivalence Constraint.

### **2.1.4 The Free Morpheme Constraint**

The FMC predicts that codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme unless the lexical form is phonologically integrated into the language of the bound morpheme:

*The Free Morpheme Constraint (FMC) (Poplack: 1980: 585-586)*

*Codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme.*

A similar constraint was also posited by Sankoff & Poplack (1981) with an additional provision that the free morpheme is not phonologically assimilated:

*The Free Morpheme Constraint (Sankoff & Poplack 1981: 5)*

---

<sup>43</sup> Although the equivalence model is identified with Poplack et al. (1980), similar work, or even its origins can be traced back to that of Lipski (1978) who states that: “Those portions falling after the switch must be essentially identical syntactically” (1978: 258). Also, following in the same vein, Pfaff (1979) states that: “Surface structures common to both languages are favoured for switches” (1979: 314)

<sup>44</sup> Other researchers Lipski (1978) and Pfaff (1979) also suggested quasi-similar constraints based on such a linear word ordering but Poplack’s (1980-81) was the first to associate a more principled-based approach in such a maverick fashion.

*A switch may not occur between a bound morpheme and a lexical item unless the latter has been phonologically integrated into the language of the bound morpheme.*

Therefore, this constraint predicts that *flipeando* ‘flipping’ is a possible form.<sup>45</sup> The lexical English form ‘flip’ is integrated in the phonology of Spanish and together with the Spanish progressive suffix *-eando* is allowed. However, *eateando* would not be possible as ‘eat’ is not phonologically integrated into the Spanish. Halmari (1997) cites examples in her Finnish/English data where phonologically unintegrated English nouns are combined with Finnish bound morphemes such as *libraryin* and also *lunchboxiin* (1997: 76); a clear counter-example to the many accounts against employing the linear model.

Furthermore, this equivalent constraints theory suits the proposed Spanish/English data well as they are naturally in tune in terms of their shared word order, and the same categories of verb, noun and adjective.<sup>46</sup> Although this constraint has been generally acknowledged and empirically validated by numerous researchers (Bentahila & Davies 1983, Berk-Sligson 1986, Clyne 1987), due to the overriding universal claim of such a constraint, there have also been counter examples in languages which are not as typologically similar. Bentahila & Davies (1983) cite counter-evidence in their Moroccan Arabic/French corpus, as does Berk-Seligson (1986) in her Spanish/Hebrew examples, further counter-evidence is conveyed in the Italian/English data from Belazi, Rubin & Toribio (1994) and Farsi-English

---

<sup>45</sup> Sankoff & Poplack (1981: 565).

<sup>46</sup> However, even within similar typological languages there have also been counter-examples as shown by Silva-Corvalán (1980) in Mexican-American bilinguals where the equivalence constraint is violated.

(Mahootian 1993).<sup>47</sup> My own examples from my Moroccan Arabic/English corpus also show counter-evidence:

(20) He's *sally*-ing  
prayer  
'He's praying'

(21) She's *salh*-ing the floor  
sweep  
'She sweeping the floor'

(22) No thanks, I'm *sawm*-ing  
fast  
'No thanks, I'm fasting'

The Free Morpheme Constraint (FMC) Sankoff and Poplack (1981) based on word order equivalence predicts then that codes may be switched after any constituent in discourse provided that constituent is not a bound morpheme. An exception is unless the lexical form is phonologically integrated into the language of the bound morpheme. This constraints theory suits the proposed Spanish/English data well as they are naturally in tune in terms of their shared word order, and the same categories of verb, noun, adjective etc<sup>48</sup>. However, due to their overriding universal claim of such a constraint, there have been counter examples in languages which are not as typologically similar. Bentahila & Davies (1983) cite counter-evidence in their Moroccan Arabic/French corpus, as does Berk-Seligson (1986) in her Spanish/Hebrew examples. My own examples from my MA/English corpus also show counter-evidence in its information structure amongst second and third generation Moroccan Arabic/English speakers:

---

<sup>47</sup> Chan (2003) relays how "One way to save the Free Morpheme Constraint is to say that the alleged violations are in fact not genuine cases of code-switching. Instead, they are cases of *borrowing* or *nonce borrowing* (e.g. Poplack & Meechan 1995, Poplack, Wheeler & Westwood 1989)" (Chan 2003: 11).

<sup>48</sup> However, even within similar typological languages there have also been counter-examples as evidenced by Silva-Corvalán (1994) in Mexican-American bilinguals where the equivalence constraint is violated.



morphemes from the other language. Any items that are possible are assigned to the well-worn category of borrowing:

- (26) *le patron une fois t-y licencié xaddem*  
The boss one time TNS-AGR license working  
'The boss makes one of his employees redundant'  
(Moroccan Arabic/French, Aabi, 1999: 30)

In the example above, *licencié*, in this constraint as it is an intra morphemic switch is deemed a borrowing under the description of the Free Morpheme Principle as this is not considered a true code switched variety. In this manner then, *licencié* is not a violation to the Free Morpheme Principle. Also, Poplack et al. (1995) can account for the counter-examples to the Free Morpheme Principle by the notion of nonce borrowing.

Such MA/Eng examples as set out above in its information structure conveys that the speakers are British-born Moroccan bilinguals as migrant Moroccans, thus far, have not been recorded using such juxtaposed stems and gerund affixes. British-born MA speakers use such constructions in roughly 80% of their discourse in a bilingual string and only with speakers of similar cultural backgrounds. In my earlier research, such structures were thought to be only when using cultural verbs alone such as *šalli* 'pray' and *šawm* 'fast'. However, more recent data collection portrays how this is an increasingly generation-specific phenomenon. It appears that any verb can be used therefore signalling that a single culture is not the overriding feature but that of bilingualism. Verb such as *šabbān* 'to do the laundry' and *šarb* 'to drink' are common with the onus being on verbs of action as they are mainly used with the present progressive tense.<sup>51</sup> Depending on speaker audience, the MA verb + Eng gerund construction i.e. 'she's *šabbning*' is favoured over and above an MA verb + MA affix in this generation grouping as it is becoming more and more common and socially

---

<sup>51</sup> Recordings show that this grammatical pairing in this generation demographic employs use of the progressive and past simple only. There has been no evidence to date of any other tense pairing.

acceptable. This is a trend that will continue to permeate other verb types and be used more and more often. Also, phonologically, MA emphatics, pharyngeals and uvulars are all maintained in the original MA formation followed by the gerund and never modified or anglicised in any way. The information structure here in essence is a reactive syntax of a distinct generation group.

Wentz & McClure (1976) also modified this constraint:

No words with morphology from both languages can exist without first having the stem integrated into the language of the suffix phonologically and semantically (1976: 245).

In Berk-Seligson's (1986) study of Spanish/Hebrew CS corpus she cites only two instances of a violation of the FMC in her corpus one of which is illustrated below:

(27) *Ze lo maanyén otí akonséžas haéla.*  
That not interesting to me those folktales  
'Those folktales are not interesting to me'  
(Spanish/Hebrew, Berk-Seligson, 1986)

In (27) the noun 'akonséžas' consists of the Hebrew bound definite article *ha-* usually realized without the initial aspirated consonant in colloquial speech, and the Judaeo-Spanish noun 'konséžas.' There is no Hebrew pronunciation evident in the realization of the noun, the result being a Judaeo-Spanish free morpheme attached to an Hebrew bound morpheme: a violation then of the FMC.<sup>52</sup> It then appears that lack of universal validity of such a constraint proved to be problematic for a groundbreaking

---

<sup>52</sup> However, Berk-Seligson (1986) later concludes that the FMC constraint is in fact 'robust' due to the absence of clause-switching among Spanish/Hebrew bilinguals in contrast to Spanish/English bilinguals who switch with great frequency. The contrasting example posed by Spanish/Hebrew is accounted for by the presence of bound morphemes that introduce a variety of clauses in Hebrew, and the free morpheme status of comparable markers in Spanish (1986: 355).

principled approach.<sup>53</sup> So much so that it was acknowledged by Sankoff & Poplack (1981) that:

It is not quite clear how the Free Morpheme Constraint might operate in a situation involving English and some highly inflected or agglutinative language, nor what might be the scope of the Equivalence Constraint for languages with highly different word orders (1981: 7).

### 2.1.5 The Equivalence Constraint

This constraint predicts that CS will occur at points where the surface structures of the two languages map onto each other neatly and concisely.<sup>54</sup>

#### *The Equivalence Constraint (EC)*

*Code-switches will tend to occur at points in discourse where juxtaposition of L1 and L2 elements does not violate a syntactic rule of either language, i.e. at points around which the surface structures of the two languages map onto each other (Poplack 1980: 586).*

A later modified version was set out in Sankoff & Poplack (1981):

#### *The Equivalent Constraint (Sankoff and Poplack 1981: 5-6)*

*The order of sentence constituents immediately adjacent to and on both sides of the switch point must be grammatical with respect to both languages involved simultaneously. This requires some specification: the local co-grammaticality or equivalence of the two languages in the vicinity of the switch holds as long as the order of any two sentence elements, one before the switch point and one after the switch point, is not excluded in either language.*

A more diaphanous interpretation is where surface structures common to both languages are favoured for switches (Pfaff 1979: 314). Others have elaborated on the EC in finer detail as outlined by Lipski (1978):

---

<sup>53</sup> The majority of violations listed emanate from agglutinative languages where intraword switches are numerous, see Levelt (1989) and Hankamer (1989) on Turkish.

<sup>54</sup> Again the second of these two constraints was totally borne out of the Spanish/English corpus but as we see lends itself in essence to a more coincidental relativity of the similarity of the structures of the languages involved.

Whereas, the portion of the code-switched utterance that falls before the codeswitch may indeed contain syntactically divergent elements, those portions falling after the switch must be identical syntactically (1978: 258)

Continuing in the same vein as the linear-ordering approach, Sridhar & Sridhar (1980) postulated their own constraint based on their corpus of Kannada/English:

The internal structure of the guest constituent [EL constituent] need not conform to the constituent structure rules of the host language [ML], so long as its placement in the host sentence obeys the rules of the host language. (1980: 412).

Such a linear quantitative based constraint has leant itself to a few counter-examples, again from the Moroccan Arabic/French data as outlined by Bentahila & Davies (1983). In Arabic, adjectives follow their nouns as in 'big man' is *rajul* (N) *kbir* (A) and the same ordering can be said of French bar some notable adjectival exceptions:

(28) j'ai vu un ancient *tilmid dyalli*  
I saw an old student of mine  
'I saw an old student of mine'  
(French/Moroccan Arabic, Bentahila & Davies, 1983: 319)

Numerous other examples are cited Berk-Seligson (1986) on Spanish/Hebrew data, Bokamba (1988) on Lingala/French and Swahili/English and Myers-Scotton (1988b) on Swahili/English. It is also worth considering whether the constraints outlined above as the FMC and the EC were largely formulated on naturally occurring data or isolated cases. This is an important distinction as lexical items or switches in isolation or taken out of context for micro analysis will garner misleading results on which to base universally valid formulae. I will have to be aware of this when collecting my data prior to later analysis. Also, are typological factors important when considering such constraints given the diversity of Moroccan Arabic/English? It would appear that such factors should have little or no bearing on whether the languages are

agglutinative or not as a principle should hold irrespective of type.<sup>55</sup> Both constraints are arguably devoid of any mention of asymmetry which is central to most if not all analyses of CS where there are at least two languages at play. This has been noted by Doron (1953) in suggesting a non-linear approach should perhaps be addressed in her data of Spanish/English corpus where she states:

I am suggesting that what block switches such as [\*seven *chiquitas* houses] 'seven small houses' are not considerations about differences of the order of constituents but considerations about agreement. The fact that word-order is not the only thing that distinguishes grammars of different languages seems to be neglected (1953: 50).

Surface order then as described in CS grammatical rules proved to be a problem for typologically different languages, or in essence, languages other than English and Spanish. In fact, Poplack (1980) states that: "A switch is inhibited from occurring within a constituent generated by a rule from one language which is not shared by another." (1980: 586). This is clearly not the case given all the counter evidence (e.g. Bentahila & Davies 1983;<sup>56</sup> Sankoff & Maineville 1986;<sup>57</sup> Berk-Seligson 1986; Myers-Scotton 1993a). The equivalence constraint was also described using such phrase structure rules by Woolford (1986) whose model was based on the premise that the grammars of the two languages operated in the same way during CS

---

<sup>55</sup> There are further counter-examples from non-agglutinative and inflectional languages (e.g. Petersen 1988 on English/Danish).

<sup>56</sup> Some of the examples cited by Bentahila & Davies (1983) were hypothetical where they constructed sentences and tested their grammatical viability on Moroccan Arabic/French informants. Whether this is truly valid or not has been commented on by some linguists.

<sup>57</sup> Sankoff & Maineville (1986) further modified this linear constraint as they found the original too restrictive syntactically. This revision shows how the CS constraint should take place between phrasal boundaries and not between phrases. Using phrase structure trees, they restricted the equivalence requirements to the two sister constituents which are immediate descendents of the same node. The below shows how the constraints apply only to the pairs (x, y), (y, z), (e, f) and (j, k) but allows switches between f and g (A and B stand for two different languages) (Aabi, 1999: 25).

as well as monolingual speech with each generating its own phrase structure rules. This is to be discussed in the following section.

### 2.1.6 Woolford's Constraint

As conveyed below, when grammatical rules are common to both languages as for example, English and Spanish, it is impossible to cite the source of the grammar of the constraint:

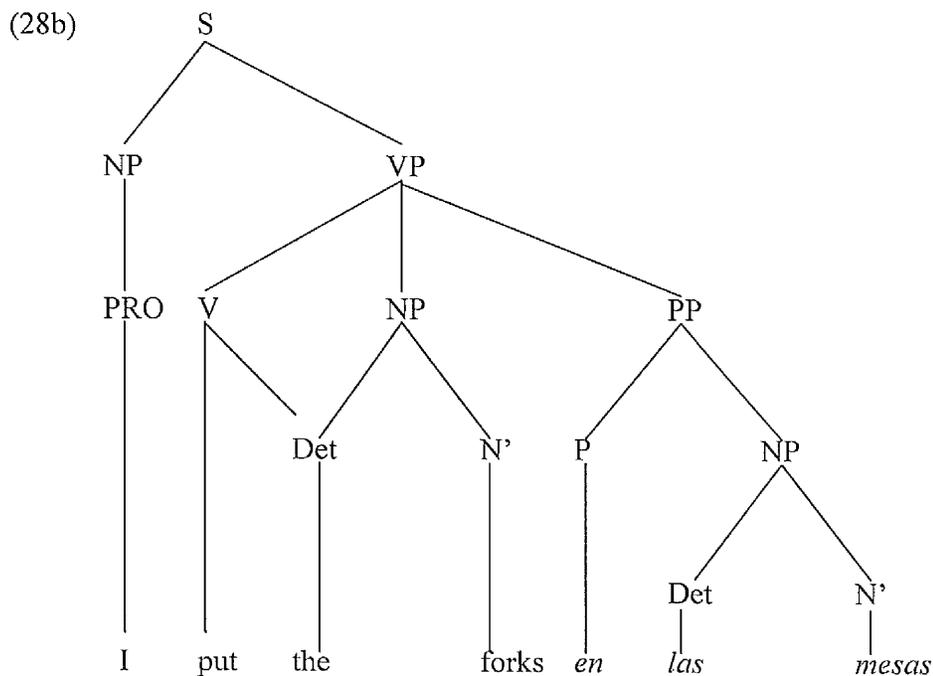
Woolford's (1983) constraint is thus:

*Woolford (1983) Word and Constituent constraint:*

- (a) *Word constraint: There can be no code-switching within a word,*
- (b) *Constituent constraint: There can be no code-switching within a constituent in which the deep structure word order is different in the two monolingual grammars.*

(28a) *I put the forks en las mesas*  
          on the tables  
      'I put the forks on the tables'  
      (Spanish/*English*, Woolford, 1983: 525)

Woolford (1983) based her model on the assumption that employment of two languages in any bilingual utterance operates in the same way during code switched discourse as that of monolingual utterances. The below illustrates this ordering:



(English/Spanish, Woolford, 1983: 525)

Employing such a linear model, Woolford (1983) states that the commonality of the above structure is not dissimilar in either Spanish or English and hence its terminal nodes can be filled by either language. This model then clearly works well with Spanish and English data in CS grammatical constructions as they are mutually complimentary and adhere to linear ordering. As with other linear models, Woolford's (1983) model was caught up against numerous counter examples, Myers-Scotton (1993a); Halmari (1997); Bentahila & Davies (1983). These unequivocal counter examples showed the non-linearity of code switching using Woolford's (1983) model; one of which is that of Bentahila & Davies (1983) who cite a conflicting structure in the position of *bint* the following:

- (29) *C'est une pauvre bint*  
 It is INDEF poor girl  
 'She is a poor girl'  
 (MA/French, Bentahila & Davies 1983: 319)

Empirical evidence above shows how the Moroccan Arabic noun *bint* must be post-modified as to that as pre-modified as shown above. MA syntax simply does not allow for this as does not the following example:

- (30) \*It's a *dār* nice  
house  
'It's a nice house'

However, the below as uttered by third generation Moroccan Arabic/English bilingual speakers allows the following construction, in keeping with normative English syntax:

- (31) It's a nice *dār*  
house  
'It's a nice house'

Evidently, this constraint was too restrictive and Sridhar & Sridhar (1980) sought to water down the acute restrictiveness by stating:

The internal structure of the guest constituent need not conform to the constituent structure rules of the host language, so long as its placement in the host sentence obeys the rules of the host language (Sridhar and Sridhar 1980: 411).

In effect, they restrict the constraint to the point where the switch of a constituent begins and not its whole internal (phrase) structure (Aabi, 1999: 29).<sup>58</sup> However, this constraint also faced numerous counter examples, most notably in the Moroccan Arabic/French empirical data of Bentahila & Davies (1983). The different versions of the equivalence constraint are similar in essence both in terms of the methodological foundation on which they base their claim and their failure to account for the large number of counter-examples cited in the CS literature. They are locational in perspective in the sense that constraints are defined in terms of linear order between

---

<sup>58</sup> This then would allow the code switched utterance of [29] above, thus not violating the syntax or juxtaposed placement of the noun *bint* in the Moroccan Arabic example.

the two languages at the location of the switch irrespective of which is the host (ML) or guest (EL). This linear perception of syntax is theoretically anticipated to fail to account for CS constraints be it inter-phrasal or intra-phrasal. The empirical evidence from several languages (cf. Pandit 1990, Nortier 1990, Myers-Scotton 1993a) including Moroccan Arabic/French and Moroccan Arabic/Standard does not attest the validity of the equivalence constraint in all its forms (Aabi: 30).

## 2.2 Asymmetry

Historically, much of the debate on CS grammatical constraints still evolves around the question of whether the interaction of a pair of languages is organised according to the internal rules of one or both grammatical systems. Should the switch be formed according to the rules of one grammatical system, ML, the guest EL constituents will simply be inserted into the frame of the ML? The resulting switch will therefore be insertional and the identification of the ML will be obvious and automatic. On the other hand, if the switch is formed according to the rules of two grammatical systems, the configuration of the CS will be alternational in the sense that the constituents of each language preserve the frame specific to their system (Aabi 1999: 13). The insertional hypothesis and the asymmetrical model in general was the touchstone of Joshi's (1985) asymmetric which asserts that CS is insertional and claims that:

Speakers and hearers generally agree on which language the mixed sentence is 'coming from'. We can call this language the matrix language and the other language the embedded language (1985: 190-1).

It must be noted that Joshi's (1985) asymmetrical model assumes that CS is unidirectional where switching is allowed from the EL into the ML and not vice versa. We will see how this is clearly not the case particularly amongst third generation Moroccan Arabic/English speakers when I further describe the innovative notion of

*reactive syntax* where complex intra-sentential structures are switched. Why Joshi (1985) would think that CS is unidirectional is perhaps due to the data at his disposal.

In analysing Joshi's model, Myers-Scotton (1993a) states that:

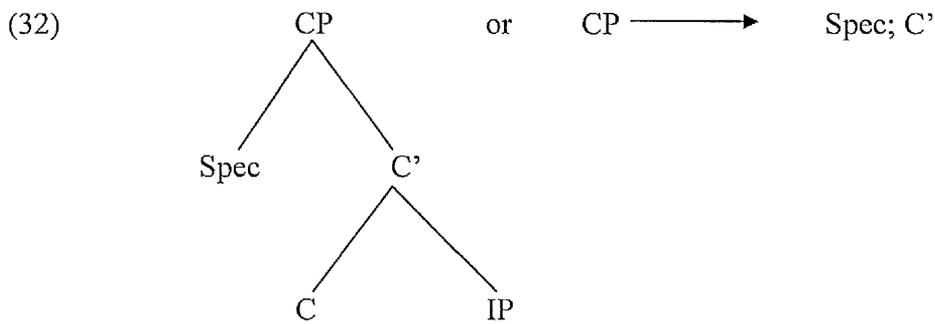
First there is no switching of categories at all in the MLF model, but rather a switching of procedures from those of the ML to the EL. This happens only when EL islands are formed. Second, there is no obvious motivation to restrict the direction of the inhibition and activation process (1993a: 36).

This will be discussed in more detail in Chapter Three where a full analysis of the Matrix Language Frame (MLF) model is considered. With regards as to whether CS is insertional or alternational, the argument is still less clear and in my opinion of little relevance as it is not only difficult to draw a line where insertional code switching begins and alternational ends, but the onus should be on an agreed definition of CS rather than inconsequential.

Within the insertional/alternational framework, the search for grammatical constraints on CS in most current studies (Poplack & Meechan 1995, Myers-Scotton 1995) is confined to a maximal unit of analysis. In syntactic theory, grammatical rules and principles are generally set within the sentence boundaries, i.e. grammar is not subject to any constraints beyond the sentence level. In much the same way, syntactic analysis of CS necessitates the determination of the maximal unit of analysis within which constraints on CS should be defined.<sup>59</sup> Taking Myers-Scotton & Jake's (1995) unit of CS analysis in intra-sentential CS, I also adopt the Complementizer Phrase (CP) as the maximal projection and frame in analysing Moroccan Arabic and English CS. As per Aabi (1999: 15) the following is a tree representation of the basic CP premise (identical to the Government and Binding (GB) model):

---

<sup>59</sup> Aabi (1999: 14).



Jake (1994: 279-81) described the CP using Moroccan Arabic/French data showing grammaticality or lack of with regards to Spec positioning:

(33) [CP moi [IP e [VP dxlt]]]  
           I                    dxlt]] n  
       'I went in'                    went in

(34) \*[CP [moi] [ana] [IP e [dxlt]]]  
           I            I            went in  
       'Iwent in'

According to Jake, the grammaticality of (33) derives from the fact that the Spec position of CP is empty, hence the insertion of the French pronoun. By contrast, the same position in (34) is filled by the Moroccan Arabic *ana* and the insertion of the French pronoun *moi* caused a double filling of the same slot which violates the basic syntactic principles of Moroccan Arabic (Aabi 1999: 15).

Switching boundaries take different structural junctures occurring within as well as beyond sentence boundaries. Inter-sentential boundaries with clear-cut and well-defined switch points are separated by individual Complementizer Phrases (CPs):<sup>60</sup>

(35) Tell him Mum. *Ma bge -š y-šannəl*  
       Tell him Mum. NEG1 want NEG2 listen 3SG  
       'Tell him Mum. He doesn't want to listen'

<sup>60</sup> Myers-Scotton & Jake (1995) adopted the Complementizer Phrase (CP) as the unit of analysis in CS research. They describe the CP as "a syntactic structure expressing the predicate argument structure of a clause" (1995: 982).

Intra-sentential (co-ordination) sentences normally portray a greater level of fluency:

- (36) *šūf mazyen u* come back  
Look well and come back  
'Look properly and come back'

Intra-sentential CS mid-clause, mid-morpheme normally uttered with such alacrity that it is a sign of a true bilingual with grammatical constraints maintained at all times:

- (37) *iwa ena li stupid li mšū ʔamma wa hit- it u*  
Well I REL stupid REL went 1SG there and hit PAST him  
'Well I'm stupid for going there and hitting him'

Intra-sentential which requires a greater command or fluency in both languages as speakers apply the rules of syntax which govern language A into that of language B mid-thought or mid sentence, and consequently may be avoided by all but the most fluent of bilingual speakers (Lipski 1985). One of the most famous examples of CS switch sites is that of Poplack (1980) which did little in actual fact to convey real evidence of intra-sentential switches as the switch, in her research occurs at a convenient break between two clauses:

- (38) Sometimes I'll start a sentence in English *y terminó en español*  
Sometimes I'll start a sentence in English and finish it in Spanish  
'Sometimes I'll start a sentence in English and finish it in Spanish'  
(English/Spanish, Poplack, 1980)

Myers-Scotton (1995) notes also how the above example (Poplack 1980) does little to inform us of real intra-sentential CS as the switch site occurs either side of two independent clauses:

Regardless of whether the conjunction relating them belongs to one language or the other, CS between the two clauses involves the alternation of two monolingual grammatical frames which makes the distinction between ML and EL in this case inappropriate (1995: 982-3).

Clearly, inter-sentential analyses to CS is bereft of any real syntactic relations as they occur independently and would naturally satisfy any well-formedness conditions in each language. Other approaches have gained prominence over the last few decades, namely the asymmetric applications.

### **2.2.1 Matrix and Embedded Approaches**

Several researchers have developed distinctions that include a matrix and embedded language dynamic. These include Wentz (1976), Sridhar & Sridhar (1980) Joshi (1981, 1982, 1985), Pandit (1990) and Petersen (1988) and Azuma (1993). As Myers-Scotton states (2002), the languages involved in an asymmetric dynamic are known as the Matrix Language and the other language the Embedded Language. The same terminology is used throughout this thesis. We commence with Joshi's (1981, 1983) constraint which paved the way for other researchers.

### **2.2.2 Joshi's Constraint**

Joshi (1981, 1983) was one of the first to document the notion of a classic asymmetry with respect to the degree of participation of the languages involved 'speakers and hearers generally agree on which language the mixed sentence is 'coming from.'

Joshi's constraint is as follows:

*Joshi's (1983) constraint:*

- (a) Assymetry constraint: Constituents can switch from the matrix language to the embedded language, but not vice versa,*
- (b) Closed class constituents: Closed class items cannot be switched.*

Inevitably, as with all constraints and syntactic restrictions on code switching, numerous counter-examples are empirically evident and cited in the literature:

- (39) Where are they, *los* language things?  
          DET  
      ‘Where are they, the language things?’  
      (English/*Spanish*, Poplack, 1981: 175)

The above shows how a closed class item, or system morpheme as is later known (Myers-Scotton, 1993a) is in language A, in this case, Spanish when it should not be permissible under the closed class constituent constraint. Later work which built on Joshi’s (1983) model includes Doron (1981), Klavans (1985), Nishimura (1986) and Myers-Scotton (1990s – to be discussed below) were all influenced by Joshi’s initial developed ideas.

### 2.2.3 Doron’s Constraint

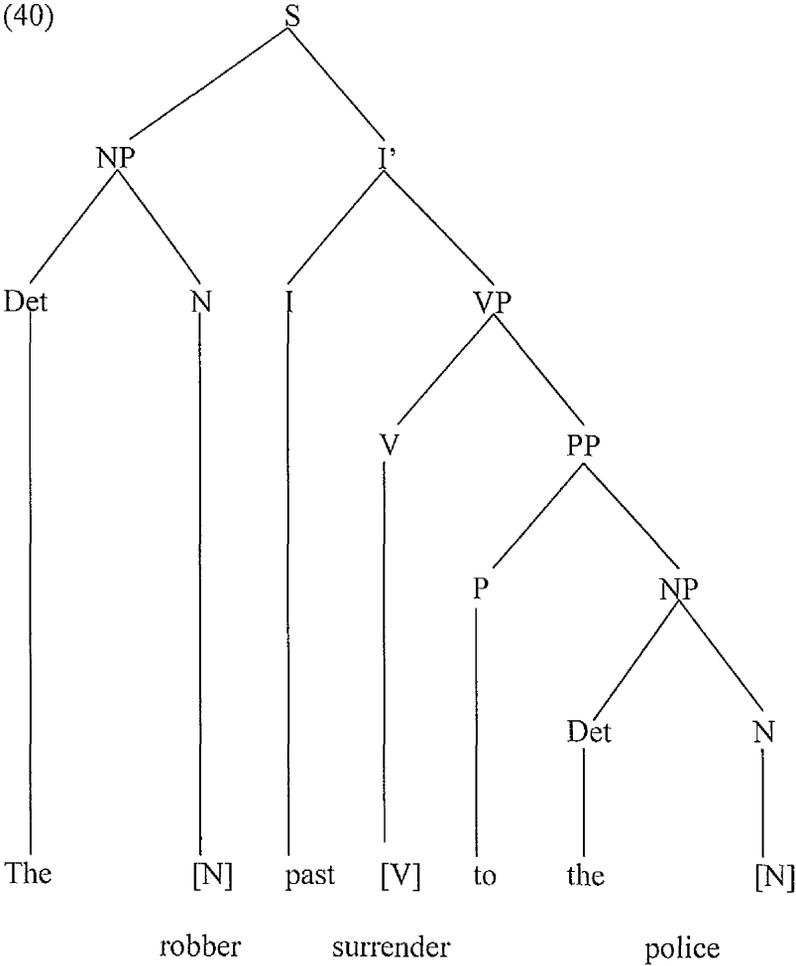
Doron (1981) for example incorporated the basic premise of Joshi (1985) but further added two additional constraints:

*Doron’s (1981) constraint:*

- (a) *Agreement constraint: Lexical categories which must be marked for agreement cannot be inserted into a position unspecified for agreement.*
- (b) *Case marker constraint: Case markers, including prepositions, from one language cannot be mixed with noun phrases from another language.*

The direction of switching is argued as being asymmetrical as: “Switching a category of the matrix grammar to a category of the embedded grammar is permitted, but not vice versa” (1981: 192). Conceptually, then this suggests that the constituents are formulated in the mental lexicon in the matrix language and then follows on to the embedded language.

Azuma (1993) made the main assumption that CS involves the process of a monolingual string where sentence processing involves a stage of frame building as well as a stage of content insertion. The first stage of frame building consists of accessing and retrieving closed class categories, and then the content insertion stage proceeds by inserting open class items set at the processing stage according to the mapping imposed by closed class categories (Aabi 1990: 34). The following illustration outlines the main premise of the closed class category versus open class item dichotomy:



Using the above as an example, closed class items are those which can only be chosen from the matrix language, i.e. verbs etc. and open class items, or content words, can come from the embedded form. The below further illustrates this:

- (41) *ḡti-ha al car beš ti-jib shopping beletti*  
Give to her DEF car so that bring 3SG shopping later  
'Give her the car so she can go shopping later'

The above shows how the content words 'car' and shopping' are in the embedded language, English in this case, whilst the closed class items,<sup>61</sup> verbs for example come for the matrix language, Moroccan Arabic. This model was successful despite some counter-examples (Nishimura 1997 in Japanese data) and thus was further modified by Myers-Scotton (1993a) in the Matrix Language Frame model.

This makes intuitive sense as we use the language we are most at ease using and draw from the embedded language for nouns, verbs, adjectives, noun phrases, collocations etc. However, I think it is too narrow a constraint to limit the *direction* of the switch from ML to EL only as will be discussed in the next Chapter. This then leads us onto the very nature of asymmetry as outlined by Myers-Scotton's (1993b) Matrix Language Frame model.<sup>62</sup> This model is described as more successful in its claim for universality than its predecessors (Aabi 1999: 36).

#### 2.2.4 The Matrix Language Frame Model

Inspired by Joshi's (1985) notion of asymmetry of a matrix language and embedded language in intrasentential CS, Myers-Scotton (1993b) has detailed and further modified the model to replace a more linear and insertional constraint with that of a more abstract universally valid hypothesis.<sup>63</sup> The premise of the following MLF is based on Myers-Scotton's (2002) Uniform Structure Principle:

---

<sup>61</sup> Closed class items were later referred to as system morphemes in Myers-Scotton (1993a) which as she states must come from the matrix language and further revised in the 4M model (2002) to be discussed in detail in Chapter Three.

<sup>62</sup> This has since been developed as a comprehensive 4M model that fine-tunes the earlier content versus system morpheme dichotomy (Myers-Scotton & Jake 2000a, 2000b, 2001).

<sup>63</sup> Myers-Scotton uses speech error data from Garrett (1975, 1988, 1990) to corroborate her data and highlights the use of closed class and open class items differentiating between system and content morphemes.



*Uniform Structure Principle:*

*A given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for this type must be observed whenever the constituent appears. In bilingual speech, the structures of the ML are always preferred, but some embedded structures are allowed if the ML clause structure is observed.*

This model of the Uniform Structure Principle where there is always only one and an identifiable ML providing the grammatical frame and an EL where its grammar slots neatly and concisely into the ML in mixed clauses follows three main principles:

*The Morpheme Order Principle:*

*In ML + EL constituents of singly occurring EL lexemes and any number of ML morphemes, surface morpheme order (reflecting surface syntactic relations) will be that of the ML.*

This attests that the ML is the dominant language of both codes and determines the ordering of the constituents.<sup>64</sup> The MLF suggests that there is a ML-EL hierarchy where the ML is the dominant language. ‘Leader,’ and sets the grammatical frame and the EL ‘led’ language slots neatly and grammatically into it as in below:

(42) *I bġit waħəd al* big car for *Ġid milād i*  
I want one DET birthday – POSS  
‘I want a big car for my birthday’

(43) *Šuft – u ambaraħ fi l* car park *wa Ġtān i* his number  
I saw him yesterday in DET car park and gave me his number  
‘I saw him yesterday in the car park and he gave me his number’

---

<sup>64</sup> It should be noted that ML assignment is dynamic in that the ML itself can change. The identity of the ML can change either synchronically or diachronically. Synchronically, a change within the same conversation is possible; an extreme case would be a change within the same sentence. Diachronically, a change may occur when the socio-political factors in the community promote some type of shift to an L2 (Myers-Scotton 1993b: 70).

The main frame set by semantic and morpho-syntactic procedures enables the assignment of one ML to one language at any one time, in the case above, Moroccan Arabic. Both languages are 'on' at all times during bilingual usage the difference being at the point of activation. Interestingly, intuitive tests of informants have been able to 'detect' which was the main language in language surveys. Only well-formed utterances are able to be formulated with a generalization that non-well-formed utterances will not be given, unless socio-pragmatic factors dictate otherwise. For example, new fashionable phrases which are perhaps grammatically incorrect but have become socially accepted norms.

*The System Morpheme Principle (SMC)*

*In ML + EL constituents, all system morphemes which have grammatical relations external to their head constituent (i.e. which participate in the sentence's thematic role grid) will come from the ML.*

The above model dictates that function morphemes can only be drawn from the ML. System Morphemes are morphemes with [+quantification] so morphemes which are neither thematic role-receivers nor assigners i.e. function words and inflectional affixes. However, this appears to be a problem in Myers-Scotton's model as not all system morphemes will come from the ML and in fact, the data she used to base her findings, was actually monolingual data hence the observation that all system morphemes, or closed class items must come from the ML. Aabi (1999) gives an example which highlights the employment of a system morpheme in Arabic:

- (44) *Exactement, il faut qu e l- commission soit établie au niveau national et regional.*  
Precisely, it needs that the commission be established at level national and regional  
'Precisely, the commission needs to be established at the national and regional level'

(Moroccan Arabic/French, Aabi, 1999: 37)

As Aabi (1999) goes on to describe, the ML is clearly French since all elements in the utterance (TNS, AGR, Comp, N, Adv, P), except for the determiner *l-* are French. On its own, the Moroccan Arabic Det cannot function as an EL island, and therefore (44) above is clear evidence that the system morpheme principle does not hold universally (1999: 37). In the MLF model, content morphemes are constrained (though not in the strict sense as system morphemes) and are described as morphemes with [-quantification] which are thematic role receivers or assigners i.e. noun and verb stems (1993a).

In cases where a category is realised as a system morpheme in one language and as a content morpheme in the other, CS is expected not to happen. Myers-Scotton illustrates her proposal with the case of pronouns which can be realised as free pronouns in one language, hence content morphemes, and as clitics in the other language, hence system morphemes (1993a: 121). For instance, in Moroccan Arabic/French conversations, a clitic subject pronoun cannot replace a free topic pronoun, nor can the topic pronoun replace the clitic counterpart.

(45) \**je*    *ġādi*  
 [I-clitic go]

(46) \**ana*    *vais*  
 I-topic go  
 (Moroccan Arabic/French, Bentahila & Davies, 1983: 312).

Further examples in my data of such constructions include the following:

(47) \**ana*    go  
 I-clitic go  
 'I go'

(48) I        *bġit*  
 I-topic want  
 'I want'

The above examples then further cement Myers-Scotton's (1993a) analysis where it is impossible (where there is no doubt) to substitute a clitic for a non-clitic morpheme.

Therefore, the ML supplies system morphemes and the EL can only provide content morphemes. An example from Moroccan Arabic/English illustrates their claim:

- (49) *qult lu waš reserve it- na al blayas*  
Say 1PST – to him – did reserve – PST suffix 1PL DEF place PL  
'I said to him whether he reserved our seats'

This example clearly illustrates how the system morphemes, inflections, pronouns, etc are provided by the ML, Moroccan Arabic and the added content morphemes in the EL are in English/French.

*The Blocking Hypothesis: (Myers-Scotton, 1993a: 83-120):*

*In ML + EL constituents, a blocking filter blocks any EL content morpheme which is not congruent with the ML with respect to three levels of abstraction regarding subcategorization.*

This hypothesis further restricts the morpho-syntactic role of the EL by allowing only certain embedded language content morphemes to occur in mixed constituents. In addition there are EL islands (structurally dependent, usually collocations) which are constituents with both the system morpheme and content morpheme from the EL and ML islands with all morphemes from the ML. EL islands, according to Myers-Scotton (1993a: 137) must show internal structural dependency relations (like AGR between French Determiner and its complement) and must be composed of two lexemes/morphemes in a hierarchical relation.<sup>65</sup> Therefore there are three possible types of constituents in intra-sentential CS: ML + EL, ML islands and EL islands both of which must be well-formed according to their respective grammars:

---

<sup>65</sup> Myers-Scotton concedes that the EL island hypothesis is the least well described in her model and represents the potential Achilles' heel (1993b: 137) of the MLF model. As described by Aabi (1999): "We find that Myers-Scotton does not specify explicitly what motivates EL islands. All we know is that they violate the ML hypothesis. Their exclusion is not independently motivated, and seems like a circular argument that is used to explain away violations to ML hypothesis. There is no indication as to when EL islands must (not) occur, i.e. we do not know whether they can be constrained as can their single constituent counterparts" (1999: 37).



both languages 'on' and at times, both grammars are accessed and over compensate unnecessarily. All in all, given the numerous counter-examples to the MLF model, Myers-Scotton's (1993a) syntactic analysis of CS is as Aabi (1999) describes: "A powerful account of CS constraints" (1999: 42).<sup>66</sup> This is in lieu of making a grammatical mistake. After Poplack's (1980) linear model, and Myers-Scotton's (1993a) asymmetrical model, a third group mainly based on Chomsky's (1981) generative grammar appeared in CS literature with emphasis mainly on the Government and Binding paradigm (Disciullo, Muysken & Singh, 1986; Halmari, 1997), the Functional Head Constraint (Belazi, Rubin, & Toribio), the Null Hypothesis (Mahootian, 1993) and the Minimalist approach (MacSwan, 1999). The assumption is that heads will require specific syntactic elements in their environment which reciprocally derive their properties from their heads (Muysken 1995: 85). The following section outlines the main models and constraints.

### **2.3 Government Approaches**

In contrast to the linear approaches (Pandit 1990) posited where switching takes place at a surface structure level there has been a contrastive governmental approach where the structure rules are hierarchical. This was to a certain degree within the theory of Chomsky's (1981) government and binding (GB) theory.<sup>67</sup> The government constraint to CS is motivated by the assumption behind X-bar theory, namely that syntactic constituents are endocentric in the sense that their properties derive from those of their heads (cf. Muysken 1995: 185).

---

<sup>66</sup> cf. also Halmari, (1997: 90) for further comments on the MLF model.

<sup>67</sup> Woodford's (1984) is a straight dependency theory – which is actually a re-hash of the Equivalence Constraint - based on the government and binding formulae.

### 2.3.1 Di Sciullo, Muysken & Singh's Constraint

In this constraint, the analytical thrust focused on dependency rather than equivalence so that a switch cannot occur between two constituents if they are lexically dependent on each other as outlined in Di Sciullo, Muysken & Singh's (1986) Government Constraints (GC) theory:

*Di Sciullo, Muysken & Singh's (1986) Government Constraint:*

*If X governs Y, ...Xq ...Yq....(Di Sciullo et al., 1986: 5)*

In other words as suggested states by Muysken (1987): "Whenever constituent X governs Y, both constituents must be drawn from the same language." (Muysken, 1987: 365), or "the lexical governor and the highest lexical element of the governed maximal projection need to be in the same language" (1986: 119). Some possible sequences as listed by Di Sciullo, Muysken & Singh (1986) with indices assigned to constituents are as follows (1986: 12):

*The Government Constraint: Some switching contexts (Di Sciullo, Muysken & Singh 1986: 12):*

- a) NP<sub>q</sub> VP<sub>p</sub>
- b) AUX<sub>q</sub> VP<sub>p</sub>
- c) V<sub>q</sub> DET<sub>q</sub> N<sub>p</sub>
- d) P<sub>q</sub> DET<sub>q</sub> N<sub>p</sub>
- e) NP<sub>q</sub> COPULA<sub>q</sub> App
- f) V<sub>q</sub> QP<sub>q</sub> Ap
- g) V<sub>q</sub> COMP<sub>q</sub> Sp
- h) V<sub>q</sub> CONJ<sub>q</sub> S'<sub>p</sub>

This translates into a claim that typical cases would be those of case assignment or sub-categorization where within a maximal projection, no switch is permissible. As per Di Sciullo, Muysken & Singh (1986: 8), the principle of language indexing requires that (i) the governor assigns the language index to the governee, and that (ii) the head of the governee must carry the language index. Illustrating their constraint with the VP domain, they claim that complementisers, determiners, prepositions and quantifiers within the VP must come from the same language as the governing verb. This constraint works well for Hindi and English which is the data they based their research on. The following illustrates the *Sub-categorisation Constraint*:

- (52) I told him that *rām bahut bimār hai*  
 I told him that Ram very sick AUX  
 ‘I told him that Ram is very sick’  
 (English/*Hindi*, Di Sciullo, Muysken & Singh 1986, 17 (43a))

The above example in compliance with the Sub-categorisation Constraint shows how the verb “*told*” is in the same language, English, the heads of the constituents including the relative clause marker “*that*” and this is in line with the government ordering. The below taken from a non-interview setting in my data shows a typical relative clause structure with the language of the relative marker continues for the remainder of the complementizer phrase:

- (53) *La, huwwa gel li* that he hasn’t got any  
 No, he said to me that he hasn’t got any  
 ‘No, he told me that he hasn’t got any’

Syntactically, *gel-li* governs ‘that’ but as we can see in the above example, the first part of the clause is in MA and the relative marker is in English.

The example below illustrates the categories which must bear the same language index within the VP according to the Sub-categorisation Constraint:

- (54) a: I saw *that* he left → V COMP  
 b: I saw *the* man → V DET  
 c: I went to Rome → VP  
 d: I went *very quickly* → VQ

In its bare form, this constraint was a target for many counter examples (cf. Pandit 1990, Myers-Scotton 1993a, Belazi, Rubin & Toribio 1994, Romaine 1995) as complements of verbs and appositions are amongst the most common switch types (Boumans 1998)<sup>68</sup> and such a government model faces challenges from typologically dissimilar languages. The problem with this approach is that there are many counter-claims that unravel it. Romaine (1989) in her Panjabi/English corpus shows how switching between V and its NP constituent is possible:

- (55) *parents* te depend Honda  
       on           be/become AUX  
       ‘It depends on the parents’  
       (Panjabi/English; Romaine, 1989: 124)

Therefore, the permissibility of switching between a V and NP is to the chagrin of this constraint. A further modification was later posited by Di Sciullo, Muysken & Singh (1990) as follows:

*Government Constraint (Di Sciullo, Muysken & Singh (1986):*

*\*[X<sup>p</sup> Y<sup>q</sup>] where X L-marks Y, and p and q are language indices.*

L-marking is a restricted notion of lexical government as defined by as:

<sup>68</sup> This definition is not particularly clear as it prevents switches between V and Det of NP, but Det is dominated by NP, a maximal projection which intervenes. Therefore, the whole domain of government was too sweeping as it included the whole maximal projection.

*Lexical Government Chomsky (1986: 15):*

*X L-marks Y if and only if X is a lexical category that theta governs Y.*

The revised Di Sciullo, Muysken & Singh (1990) model with the implementation of L-marking then made permissible switches for example, between Determiner and their NP or between verbs and locational verbs which were not catered for in the earlier version of government (Muysken 1995: 186-8). However, even this revised model was deemed too strong and counter-examples were numerous cited in the CS literature as below from Nortier illustrates (1990):

- (56) ana ka-ndir *intercultureel werk*  
I do 1SG intercultural work  
'I am doing intercultural work'  
(Moroccan Arabic/*Dutch*, Nortier, 1990: 134).

In this approach, code switching is only possible where a governed element presides over the governee and includes a 'language carrier' whose language index is identical with the language index of the governor. They cite a case when a Language q (Lq) governor such as a transitive verb, governs a Language p (Lp) complement (noun) and the latter is accompanied by an Lq marker can override any potential fallout and acts as a safety net so that the below is possible:

- (57) ha ricevuto il *diplôme*  
Have- receive DEF diploma  
'She has received the diploma'  
(Italian/*French* (Di Sciullo, Muysken & Singh, 1986: 13)

A further modification to the government models analysed above of Di Sciullo, Muysken & Singh (1986) and Halmari (1997) was that of Belazi, Rubin & Toribio (1994).<sup>69</sup>

---

<sup>69</sup> This follows on from Abney (1987).

### 2.3.2 The Functional Head Constraint

Belazi, Rubin & Toribio (1994) posited a model which clearly distinguishes between lexical heads (V, N, P) and functional heads (Q, NEG, model INFL). As Aabi (1999) details, they argue that switching between a lexical head and its complement occurs unconstrained in contrast to the restriction on switching between a functional head and its complement (1999: 45). They formulate the above as follows:

*The Functional Head Constraint (Belazio, Rubin & Toribio 1994: 228):*

*The language features of the complement f-selected by a functional head, like all other relevant features, must match the corresponding features of that functional head.*

In turn, five functional heads are listed each with a complement:

Functional heads and their complements (Belazi, Rubin & Toribio 1994:

- a) Complementizer (i.e. C) and IP
- b) Inflection (i.e. I) and VP
- c) Determiner (i.e. D) and NP
- d) Quantifier (i.e. Q) and NP
- e) Negations (i.e. NEG) and VP

This constraint works within language-specific data, in this case, Spanish and English as this was the data set employed for the Constraint. The below gives a few examples of permissible switches which fully support the Functional Head Constraint (Belazi, Rubin & Toribio 1994):<sup>70</sup>

---

<sup>70</sup> cf. Chan 2003: 35.

(58) *No switching between C and IP*<sup>71</sup>

El profesor dijo [*c that*] [*ip the student had received an A*]  
The professor said that the student had received an A  
'The professor said that the student had received an A'  
(Spanish/English, Belazi, Rubin & Toribio 1994: 224, (10b))

(59) *No switching between I and VP*

Je [I serai] [*vp parti fi -l* *É\$ra*]  
I will be gone at DEF ten  
'I will be gone at ten o'clock'  
(French-Tunisian Arabic, Belazi, Rubin & Toribio, 1994: 230)

(60) *No switching between Q and NP*

q Poco] [*np estudiantes*] *finished the exams*  
Few students finished the exams  
'Few students finished the exams'  
(Spanish-English, Belazi, Rubin & Toribio, 1994: 229, 18a)

The Functional Head Constraint then is perhaps too astringent in its application and rigid in its switch sites which are a common problem of all constraints-based theories in intra-sentential code switching as researchers uniformly validate new syntactic theories and constraints on a specific language set. The below examples from Mahootian and Santorini (1996) give numerous counter-examples to the Functional Head Constraint (1994):<sup>72</sup>

(61) *Switching between I and VP*

No parce que hanno *donné des cours*  
No because have given of the lectures  
'No, because they have given lectures'  
(Italian/French, Di Sciullo, Muysken & Singh, 1986: 15: (37a), as  
quoted in Mahootian & Santorini, 1996: 466 (5a))

<sup>71</sup> Halmari (1997: 94) reported that the Spanish-English bilinguals she consulted disagreed with the judgements in Belazi, Rubin & Toribio (1994) and accepted code-switched sentences in which the Spanish complementizer "*que*" takes an English clause as complement, violating the Functional Head Constraint. The Spanish-English bilinguals MacSwan (1997) consulted also disagree with claims of Belazi et al. (1994) (Chan, 2003: 38). For more discussion see Chan (2003).

<sup>72</sup> Chan (2003) describes how in addition to the data given in showing sound empirical evidence against the Functional Head Constraint (1994), there are other data sets (including Teng 1993) where there are switches between NEG in Cantonese and English code switched data, see Chan, (2003: 37) for further examples.

- (62) *Switching between Q and NP*  
I'll take some *naemaek*  
I'll take some salt  
'I'll take some salt'  
(English/*Farsi*, Mahootian & Santorini, 1996: 466, (6a))

Bhatt (1995) sets out numerous counter-examples to the Constraint, and interestingly, this example is from Spanish and English:<sup>73</sup>

- (63) *Switching between D and NP*  
El *flight* que sale de Chicago  
The flight which leaves from Chicago  
'The flight which leaves from Chicago'  
(Spanish-*English*, Pfaff 1979: 305, (41), as quoted in Bhatt, 1995: 2, 5a)

Counter-examples are evident when various data and language sets are used and this is the case across all theories and constraints of intrasentential code switching.<sup>74</sup>

There are then a multitude of criticisms based on valid<sup>75</sup> empirical data. As Chan (2003) comments:

In a nutshell, the validity of the Functional Head Constraint as a universal is seriously in doubt in view of the counter-examples. Notice that these counter-examples come from a lot of language-pairs, far more than the ones which Belazi et al. (1994) observed in formulating the Functional Head Constraint (Chan 2003: 37).

The very fact that numerous constraints-based models have been criticised and/or disproved prompted Muysken (1995) to review the major approaches to code switching whereby he acknowledges the main problems and consequences of posited constraints and concludes that in essence, no one specific model can account for all

---

<sup>73</sup> See Chan (2003: 37) for further examples.

<sup>74</sup> In highlighting the weakness of The Functional Head Constraint put forward by Belazi et al. (1994), Halmari (1993, 1997) also cites examples from Finnish and English where there are many examples of intra-sentential switches between determiners and nouns (1997: 91-94).

<sup>75</sup> By true, this means non-hypothetical examples of code switched utterances as these are not, in my opinion, valid switches nor should they contribute as empirical evidence. Bentahila & Davies (1983) based much of their French/Moroccan Arabic data on hypothetical examples which they in turn tested out on one bilingual French/Moroccan Arabic speaking family. One therefore should question how valid this analysis is. This will be discussed in more detail in Chapter Six.

data that has been cited in the relevant literature on intrasentential code switching.<sup>76</sup> Also, it must be remembered that Poplack's constraints, The Equivalent constraint and the Free Morpheme Principle (1980) as well as Myers-Scotton's Matrix Language Frame model (1993a) are revolutionary in the field of contact linguistics and code switch studies in general. To simply dismiss these is too severe an approach to take. Thus Muysken (1995) in an attempt to reconcile the "No Constraints" Paradigm and the "Constraints" Paradigm adopts a new approach where code switching is described as neither random nor governed by any one particular set of syntactic constraints. The below lists four possible switch scenarios:

Contexts for code-switching (Muysken 1995: 196)

- a) Switching is possible when there is no tight relation (e.g. of government) holding between two elements, so called paratactic switching;
- b) Switching is possible under equivalence;
- c) Switching is possible when the switched element is morphologically encapsulated, shielded off by a functional element from the matrix language;
- d) Switching is possible when at the point of the switch a word could belong to either language, the case of homophonous diamorph (e.g. *in* in English, German or Dutch).

In essence then, Muysken (1995) has sought to bring together the main theories and constraints of intrasentential code switching, these are: *The Government Constraint* (Di Sciullo, Muysken & Singh 1986), *The Equivalence Constraint* (Poplack 1980, Sankoff & Poplack 1981), The Matrix Language Frame (MLF) Model

---

<sup>76</sup> See also Bokamba (1989) and Clyne (1987) who deem the "Constraint" Paradigm as an incorrect approach to code switching. This is a blanket conclusion as this in essence would imply that code switched utterances are random processes bereft of syntactic principles and parameters which do not adhere to grammatical regulations.

(Myers-Scotton 1992, 1993, 1995) and finally Clyne's concept of *Triggering* (1987). Muysken's (1995) approach is based on that of neutrality where the premise is that bilingual speakers, in code-switched discourse, neutralize any potential grammatical conflicts and hence are able to string together bilingual discourse. Chan (2003) in describing Muysken's (1995) model of "Neutrality" observes:

There is an underlying principle which unifies these patterns: They are guided by the principle of neutrality. According to Muysken (1995), a language is a self-contained system and code-switching poses potential threats to this system. Code-switching is in theory impossible. When code-switching does appear, bilingual speakers adopt strategies which neutralize the potential conflicts between two language systems brought about by code-switching (2003: 1).

Muysken's (1995) approach is a change to previous approaches to code switching in that it is the first real attempt in reconciling major works on the syntactic analysis of bilingual discourse.<sup>77</sup> As Chan (2003) also outlines in his study of Cantonese and English code switching:

Functional categories and lexical categories exhibit different behaviour in code-switching, which further justifies such a distinction of lexical items in natural languages (2003: 1).

He describes in detail that with regard to word order between lexical categories and their code-switched complements, it is suggested by Chan (2003) that:

Lexical heads do not always determine the order of their complements. Nonetheless, functional heads always determine the word order of their code-switched complements (2003: 1-2).

Further, the difference between functional categories and lexical categories is explained in terms of their differential processing in production, based on the

---

<sup>77</sup> Muysken later developed a new proposal (1997) based on his non-unitary (1995) model discussed above, with the added emphasis that patterns of code switching are reliant on different processes, namely, *Alternation*, *Insertion* and *Congruent Lexicalization*. This will be discussed in more detail in Chapter Six – although a full analysis is beyond the scope of this thesis.

monolingual production model of Levelt (1989). This is in line with recent proposals that production processes are essentially the same in both monolinguals and bilinguals (de Bot 1992) and the spirit of (1), namely, code-switching and monolingual utterances are governed by the same constraints or principles. Not only are they governed by the same syntax, they are governed by the same production as well:

Code-switching and pure languages are governed by the same set of constraints and principles not only in syntax but also in production (Chan, 2003: 2).

Although this was an improvement on previous models, further counter-examples ensued (English-Farsi, Mahootian & Santorini (1996), Italian-French, Di Sciullo, Muysken & Singh (1986) where switches were incorrectly disallowed between Det and NP and Belazi, Rubin & Toribio (1994) admit that not all lexical categories are free to switch. In fact, lexical parallelisms as illustrated by them below caused them to rethink their original model:

(64) a: *J'ai une voiture mizyena*

b: \* $\xi$ ndi karhba belle  
I-have car nice  
'I have a nice car'

(Tunisian Arabic/French, Belazi, Rubin & Toribio, 1994: 232)

The new constraint is as follows:

(Belazi, Rubin & Toribio 1994: 232):

*The Word Grammar Integrity Corollary:*

*A word of language X, with Gx must obey grammar Gx.*

Under this Integrity Corollary, the above example (64a) is permissible as adnominal adjectives occur post-nominally in both Tunisian Arabic and French. However, (64b)

is disallowed as *belle* in French must occur pronominally and therefore the derivation fails on that count. Again, counter-examples were presented based on sound empirical evidence to be discussed in more detail in Chapter Six. It is evident that the Functional Head Constraint was liable to too many counter-examples except in seemingly well-matched data, particularly with the typologically similar languages of Spanish and English. This gave further rise and impetus for researchers to approach intra-sentential code switching from a different angle, namely to predict that such code switching is not bound to any syntactic specific constraints, but in essence, intra-sentential code switching adheres to the grammar of monolingual constraints. Such an approach was the insertional method.

## **2.4 Insertional Approach**

In counteracting the linear and government approaches, the insertional models were seen to offer a viable model for bilingual discourse. As Boumans (1998) discusses:

Neither the linear word order constraints such as Timm's (1975) five constraints and Poplack's Equivalence Constraint, nor the subcategorization or government-based constraints make a principled distinction between the roles of the two languages involved in code switching (1975: 24).

Such an insertional constraint originates from the work on Moroccan Arabic and French by Bentahila & Davies (1983).

### **2.4.1 Bentahila & Davies' Constraint**

Bentahila & Davies (1983) in an analysis of Moroccan Arabic and French intra-sentential code switching posited the below constraint as follows:

*The Subcategorization Constraint (Bentahila & Davies 1983: 329)*

*All items must be used in such a way as to satisfy the (language-particular) subcategorization restrictions imposed on them.*

Firstly, the Arabic determiners “*hað*”, “*ðak*” and “*waħəð*” take N’ (i.e. N) complements – a DET N constituent – as subcategorized in Arabic. In code-switching, the subcategorization restrictions of these determiners are satisfied by taking a French N’ complement (Chan 2003: 29). They give the following examples:

- (65) *waħəð une cousine*  
 One INDEF cousin F  
 ‘One cousin’  
 (Moroccan Arabic/*French*, Bentahila & Davies, 1983: 317, 71)
- (66) \**waħəð professeur*  
 INDEF teacher  
 ‘A teacher’  
 (Moroccan Arabic/*French*, Bentahila & Davies, 1983: 321, 93)

It must be noted however, that Bentahila & Davies (1983) cited many *hypothetical* examples during their research and the above is one hypothetical example. Research based mainly on hypothetical examples should, in my opinion, be questioned. In contrast however, during my data, there is a third generation speaker whose bilingual discourse invalidates the above Constraint:

- (67) I saw *waħəð* man in the park who kept staring. What a loser!  
 I saw INDEF man in the park who kept staring. What a loser!  
 ‘I saw a man in the park who kept staring. What a loser!’

Pandit (1990: 43) also provides another GB approach in his proposal: “Code switching must not violate the grammar of the head of the maximal projection within which it takes place.” Romaine (1989) as a concluding remark on this constraint observed:

Data such as these [code-mixing data] have no bearing on abstract principles such as government [...] because code switching sites are properties of S-structure, which are not base-generated and therefore not determined by X-bar theory. (1989: 145).

Therefore, both governed and ungoverned content words and constituents can be switched, so that while it may be possible to describe CS in GB terminology, the government principle itself does not seem to constrain CS (Boumans 1998: 22). At times in order to explain away counter-claims in data, borrowing is seen as the likely scapegoat. Other counter examples ensued between verbs and direct objects as illustrated below in Dutch/Moroccan Arabic:

(68) *anaka-dir* intercultureel werk –  
‘I am doing intercultural work’,  
*Moroccan Arabic/Dutch*, Nortier 1990

Also between direct and indirect objects in ‘*jib li-ya* een glas water or so’ – ‘Bring me a glass of water or so’, and between copula-type verbs and their predicates: ‘*wellit huisman*’ – ‘I became a houseman.’<sup>78</sup> Therefore, the government model, even as theoretically motivated it may be, is too stringently narrow a constraint with too many counter-examples. This is more evident in my data. Gardner-Chloros & Edwards (2004) claim the government models which are based on purely syntactic relations are too abstract.

Myers-Scotton also found the base generative approaches too abstract and argues that the Chomskyan generative models focus on phrase structure as the source of constraints and cannot account for single item insertion citing these as “*purely syntactic*” and “*too close to the surface*” (2002: 162). Halmari (1997) reinterpreted the first form of government in her approach to describing constraints on CS. She states how elements may come from a different language than the language of the governor provided that the maximal projection includes a language index ( $L_q$  carrier) which matches the language index of the governor (Halmari 1997: 100-1). As described by Aabi (1999: 44), Halmari (1997) proposes that case marking and

---

<sup>78</sup> See Halmari (1997) for further counter examples found in her Finnish/English data.

agreement properties can also reflect a government relation and, therefore, are able to carry the language index of the governor. In other words the governed lexical head does not have to bear the same  $L_q$  as the governor if its (the governee) case marker and/or AGR inflection does:

- (69) Mää oon koulu-ssa joskus pelannu *basketball-ia*  
I have sometimes played basketball  
'I have sometimes played basketball'  
(Finnish/*English*, Halmari 1997: 117)

Halmari (1997) explains that the switch between the Finnish verb *pelannu* (governor) and the English N 'basketball' (governee) which apparently have different  $L_q$  is due to the partitive morphemes *-ia*, which bears the same language index as the governor. Halmari adds: "This [partitive] morpheme must function as the  $L_q$  carrier which will satisfy the government constraint on code switching" (1997: 118). MacSwan (1999) in both empirical and theoretical criticism of this government model asks how:

Why government, in particular, should be related to code switching, since the relation is presumed to be an operation of UG that is invariant cross-linguistically. (1999: 6).

As her predecessors, Halmari's (1997) model is not without its counter-examples:

- (70) *The police officers have seen un ladrón*  
'The police officers have seen INDEF thief'  
(Spanish/*English*, Belazi, Rubin & Toribio, 1994: 230)

The example above then shows no evidence of the  $L_q$  of the governing verb in the governed maximal projection.

## 2.5 Monolingual Constraints in Code-switched Discourse

In short, it was claimed that nothing constrains intra-sentential code switching apart from that which constrains monolingual language use (Mahootian 1995, Chan 2003).<sup>79</sup> This was an approach which was developed after several other theories and constraints had become criticized based on a wealth of criticisms and the counter-examples to each theory, sought to render obsolete the notion of ‘well-formedness’ in keeping with the concept of Universal Grammar. In this vein, within a more generative framework, a proposal was put forth by Santorini & Mahootian (1995).

### 2.5.1 Null Theory

A constraint based on monolingual-based application is that of Santorini and Mahootian (1995) and their *Null Theory* which focuses on word order in intra-sentential code switching:

*The Null Theory (Santorini & Mahootian, 1995: 29)*

*The language of a head determines the phrase structure position of its complements in code-switching just as in monolingual contexts.*

The above theory is in essence opposed to constraints-based approaches to code switching in its entirety. This is an interesting concept as previous approaches have been fraught with counter-examples. This was later slightly modified:

---

<sup>79</sup> Chan (2003) assumes a Minimalist framework based on the Principles and Parameters approach whereas Mahootian (1993, 1995) is based on the Tree Adjoining Grammar (TAG) concept of Joshi (1985). As described by Chan (2003) TAG occurs when a speaker accesses a lexical entry, he or she accesses not only a lexical item (i.e. the word form) and its category (e.g. *noun or verb*), but also a minimal tree containing the projection of a category and empty slots for its syntactic dependencies. For instance, if the speaker accesses the English words a minimal tree is retrieved in which there is a V node (with the verb “ate” under it) together with two empty DP positions, one for the subject and the other for the object. Granted that V is the head, the position of its complement (i.e. the object DP) is already fixed in the post-verbal position according to the grammar of English. Derivation will proceed in such a way that the empty DP slots are replaced or “substituted” in TAG’s terms, by other DP trees (i.e. the subject and the object) containing lexical items, yielding, for instance, “*Oliver ate a sandwich*” (Chan 2003: 66).

*Mahootian & Santorini (1996):*

*Heads determine the syntactic properties of their complements in code switching and monolingual contexts alike.*

The Tree Adjacency grammar model (Mahootian 1993, Mahootian & Santorini 1996) relies on general principles of phrase structure rather than on specific code switching constraints. However, no such approach has been provided thus far to fully account for code switching data, irrespective of language type, neither has a theory been proffered which fully incorporates the machinery of Universal Grammar and can apply itself as Theory X to suit all code switched discourse for all languages and this theory is also no exception. In fact, Mahootian (1995) when formulating the Null Theory states:

We do not yet know enough about the relation between frequency distributions of specific grammatical patterns in monolingual speech data and properties of grammar to handle frequency in bilingual data with any assurance. (1995: 185).

Inevitably, well-sourced counter-examples ensued (Timm 1975), MacSwan (1999):

(71) \**El* no wants to go  
He not want to go  
'He doesn't want to go'  
(*Spanish/English*, Timm, 1975)

(72) \* He doesn't *quiere ir*  
He doesn't want to go  
'He doesn't want to go'  
(*Spanish/English*, Timm, 1975)

Then above show how (71) and (72) are considered ill-formed and counter the theory posited by Mahootian & Santorini (1986). Gardner-Chloros & Edwards (2004) state that no principles proposed to date account for all the facts, and it seems unlikely that

'grammar', as conventionally conceived, can provide definite answers. I conclude that rather than seeking universal, predictive grammatical rules, research on CS should focus on the variability of bilingual grammars.<sup>80</sup>

The generative approach of Mahootian (1993) is deeply-seated in theory and as described by Gardner-Chloros & Edwards (2004). Mahootian & Chan are alike in proposing that constraints on CS grammar operate at the level of phrase structure, and specifically that constraints affect the ways in which heads of phrases select their complements. Mahootian is concerned with the 'surface' ordering of constituents, and shows how the different phrase structure rules of English (an SVO language), and Farsi (SOV), determine that certain potential switches will not occur (Gardner-Chloros & Edwards (2004: 103). The basic premise of the Null Theory is based on general syntactic relations focusing mainly on the positioning of heads and complements and completely rejects the notion of any 'third grammar'. In this manner, the Null Theory adopts the GB Theory of directionality and word order is maintained:

$$(73) \quad X' \longrightarrow X \text{ YP}$$

(X precedes YP if head-initial; X follows YP if head-final).

However, it is at odds with GB as the latter states that branching directionality is not encoded in the head, but Null Theory proposes that such directionality is encoded. In stating that there are no specific mechanisms purely aimed at code switching, Mahootian (1993) describes how the two lexicons in bilingual speech, with their individual phrase structures, remain distinct. The Null Theory is further advanced in

---

<sup>80</sup> However, Poplack's (1980) equivalent constraint and that of Myers-Scotton's (1993a) Matrix Language Frame Model revolutionised our approach to code switching. These cannot be simply ruled out due to counter examples as submitted by empirical evidence. It is true that no one model is free from criticism or can claim universal validity, however, Myers-Scotton's MLF model is a useful vehicle (see the next Chapter for a closer analysis).

Mahootian and Santorini (1996) who propose that a head determines the phrase structure position, syntactic category and feature content of its complement, i.e. its sub-categorisation features (1996: 472). Furthermore, they state that a verb (V) (lexical head) dictates the position of its complement, allowing the switch below (in 74a) between a V-complement language and a complement-V language, but does not allow it in (74b):

(74a) You'll buy *xune-ye jaedid*  
'You'll buy a new house'  
(*Farsi/English*, Mahootian, 1993: 152)

(74b) \* You'll *xune-ye jaedid* buy  
You'll buy a new house  
(*Farsi/English*, Mahootian & Santorini, 1996: 472)

Therefore, directionality is a parameter with language X largely determining the order of the subsequent complement (YP). However, this theory is not without its criticisms as the assumption that word order is a lexical property as determined by heads is not unanimously accepted.<sup>81</sup> With this in mind, a more minimalist framework (MacSwan 1999, Chomsky 1995) was put forward and it is dealt with in the following section.

## 2.6 Minimalist Approaches

Within the framework of generative grammar, the Principles and Parameters construct provided a concise alternative to language-specific rules as posited in early generative grammar approaches. Such principles and parameters are deemed to be based on language universals and form the basis of our inherent language faculty. Hornstein, Nunes & Grohmann (2005) in fact describe the Principles and Parameters framework as: "The most fully worked out version of a Principles and Parameters approach to

---

<sup>81</sup> Chan (1999: 67).

UG” (2005: 19). Chomsky (1993) describes how a theory of grammar should be as simple and minimal as possible and should not require more than the basic elements.

### 2.6.1 Minimalist Program – MacSwan

MacSwan (1999) in his thesis on a minimalist approach to code switching,<sup>82</sup> describes how all syntactic operations and principles are relevant in defining the class of well-formed code-switching constructions.<sup>83</sup> He based his research (1999) on the Minimalist Program (MP) to explain phrasal code switching. MacSwan (1999) shows how bilinguals have separate lexicons for each language spoken together with separate phonological systems and addresses the question of how the mind represents two languages in simultaneous bilingualism concluding that evidence from code switching suggests bilinguals have discrete and separate lexicons for the languages they speak, each with its own internal principles of word formation, as well as separate phonological systems. MacSwan (1999) in his research on English and Nahuatl code switching defines a bilingual speaker as:

An individual who alternately uses two or more languages at or below sentential boundaries, and who has had continual, sustained exposure and practice in these languages since infancy (1999: 22).<sup>84</sup>

He also posits a structure which avoids the concept of a ‘third grammar’ and instead employs a monolingual apparatus in the most minimal of frames.<sup>85</sup> It is also the basic

---

<sup>82</sup> cf. Chomsky (1995) on which the Minimalist model is based and to which ‘minimalist syntax’ refers.

<sup>83</sup> MacSwan (1999: 147).

<sup>84</sup> This is rather a bold statement given the fact that many bilinguals are not only productively bilingual, but also receptively and this then would rule out the sense of fluency applied by MacSwan (1999). Furthermore, many bilinguals do not have exposure “from infancy” but become fluent as a child. My second generation Moroccans came to the UK at the average age of 10 with only Arabic as their native tongue (some may have had Berber, French and Spanish also) and so learned English by their peers and through schooling.

<sup>85</sup> MacSwan (2005) believes that: “principles of language design urge us to begin with the simplest assumptions, namely, that there is no difference in the way language is represented in the mind/brain of a bilingual and of a monolingual; we should admit additional mechanisms or design assumptions which specifically apply to bilinguals only when compelled to do so by the evidence” (2005: 277). This then is the basic premise and concept of the Minimalist model. This is based on his (1999) analysis of code switching where: The numeration is constructed by the application of *Select* to either lexicon; the

premise of this highly-theory driven approach that specific constraints to code switching are best avoided where assumptions would in essence as stated by MacSwan (1999):

Favour accounts of code switching which make use of independently motivated principles of grammar over those which posit rules, principles or other constructs specific to it (1999: 146).

The idea that there are no CS-specific constraints is also present in Woolford (1983) and Mahootian (1993), amongst others, as described earlier in this chapter and the approach advocated by MacSwan (1999) suggests that CS data may be explained by relying upon independently motivated principles of grammar, with no CS-specific mechanisms required. Bentahila & Davies, although not using a Minimalist framework also detail how in their Arabic/French code switched data, the syntactic characteristics are not governed by “ad-hoc constraints” or surface structure equivalence (1983: 328). MacSwan (1999) describes the Minimalist approach in that it is the central, leading aim of Chomsky’s (1995) minimalist program to eliminate all mechanisms that are not necessary and essential on conceptual grounds alone; thus, only the minimal theoretical assumptions may be made to account for linguistic data, privileging more simple accounts over complex ones. These assumptions would naturally favour accounts of code switching which make use of independently motivated principles of grammar over those which posit rules, principles or other constructs specific to it. The idea that no code switching-specific mechanisms may be admitted is also consistent with views expressed in most current work on code switching and I shall pursue this idea here as well. In general terms, this research

---

derivation proceeds as in the monolingual case, with Merge and Move building and rearranging structure to the extent uninterpretable features may be checked, with no CS-specific mechanisms permitted (1999: 148). The strategy adopted by MacSwan is to locate language-specific conflicts in the feature specifications of functional categories in order to explain the code switching data (1999: 156).

program may be stated in the below constraint where the minimal code switching apparatus is assumed (1999: 174). In MacSwan's (1999) study of intra-sentential English/Nahautl code switching, he goes on to describe a key aspect of Minimalism:

*Constraint on CS: MacSwan (1999)*

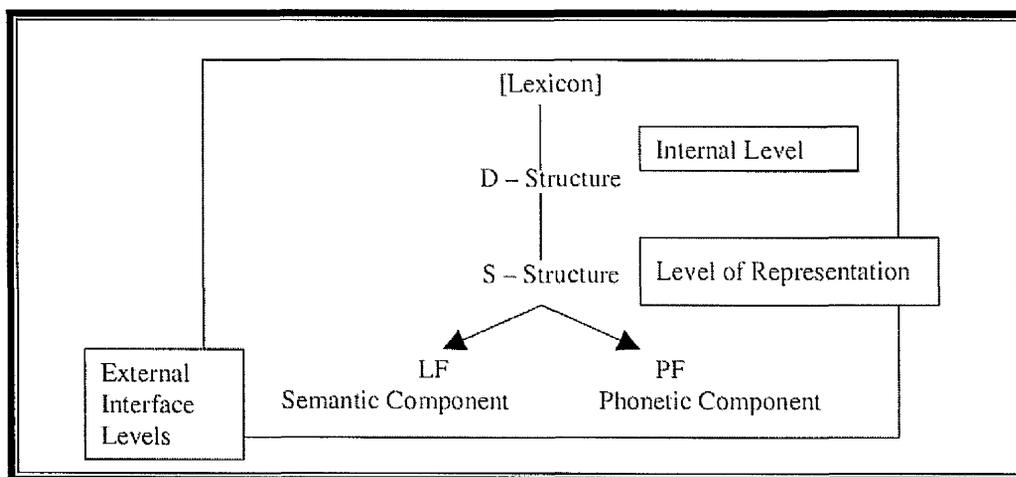
*Nothing constrains code switching apart from the requirements of the mixed grammars.*

MacSwan (1999) describes how the above constraint does not imply that there are no unacceptable code-switched sentences. The MacSwan (1999) Constraint is used in its technical sense in syntactic theory, entailing that there are no statements, rules or principles of grammar which refer to code switching. A bit more concretely, his constraint entails that we ignore differences between the identities of particular languages for the purpose of linguistic theory. The language faculty (and associated learning principles) is a generating function which selects a particular language  $L_x$  or  $L_y$  (... $L_n$ ), given input from  $L_x$  or  $L_y$  (... $L_n$ ). Thus, the value of  $L$ , determined by the language faculty crucially may not be a construct in linguistic theory; its value derived, *determined* by the theory of grammar (and associated learning principles). Hence, while distinctions like "Spanish," "French," "English" and "Berber" are meaningful for the many interesting questions of language use, they do not enter into the apparatus of syntactic theory, and should play no role in an account of code switching. Clearly, however, there are language-particular requirements; in the minimalist program these are taken to be represented in morphology.

An explanation of grammaticality in code-switched sentences must therefore appeal to mechanisms motivated to account for grammaticality in monolingual sentences, or appeal to conflicts in the requirements of mixed the languages (that is, conflicts in their parametric settings), or to other factors independently motivated for

linguistic theory.<sup>86</sup> Our conception of such conflicts is very much determined by our conception of the organization of the grammar (MacSwan, 1999: 146). Within the Minimalist framework, feature checking is of paramount importance as a mechanism for triggering movement:

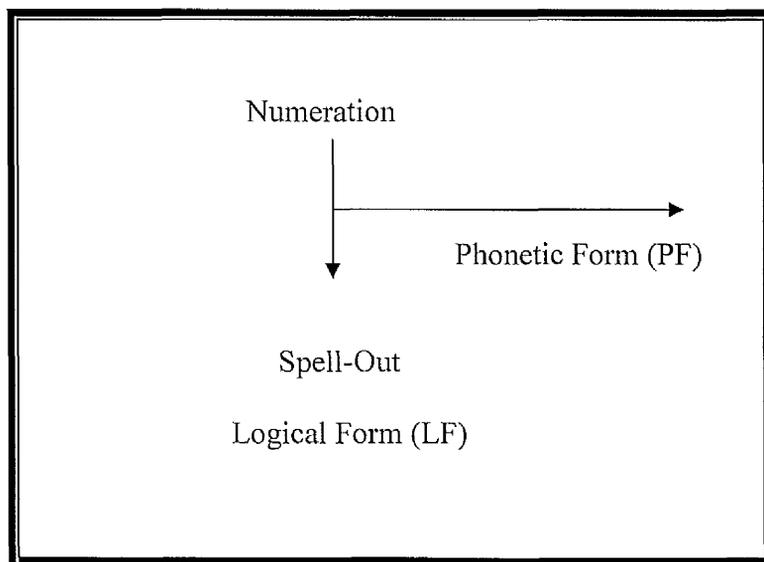
*Figure 2.1: Minimalist Framework. The linguistic form with two interface levels: the external level containing both the Logical Form (LF) and the Phonetic form (PF) and the internal level containing both D-Structure and S-Structure*



However, under the Minimalist Program, the internal level is questioned as language according to this approach, is mainly concerned with sound and meaning, this rendering both D-Structures and S-Structures as peripheral. This then leads to the following diagram under MP:

*Figure 2.2: The Minimalist Program has only the below components*

<sup>86</sup> The implication here then is that distinctions between languages do not feature in syntactic theory and should play no role in an account of code switching (MacSwan 1999: 146).



As noted by MacSwan (2000), a very important feature of the Program is that all learning is lexical and all parameters are micro-parameters associated with individual lexical items (2000: 44). However, if we assume that the computational system is invariant across languages, and that parameters are part of the lexicon which the computational system uses to build up larger structures, then the question of which particular language system is in use is answered straightforwardly. Each lexical item introduces features into the derivation, and these features must be checked. Languages differ with respect to their feature matrices, as set by experience. The language faculty need pay no attention to the socio-political identity of words (our associations of *tree* with “English” or of *árbol* with “Spanish”). It only knows that these lexical items have features which enter into the derivation, and that these features must be checked; when features mismatch, or when uninterpretable features cannot be checked, the derivation crashes, whether the set of lexical items is associated with one particular language or two (or more).

Thus in the minimalist program, a *conflict in language-specific requirements* is just a conflict involving lexical features, and the interface of distinct “languages” is trivially solved. However, as Chomsky (1995) emphasizes, the nature of the syntactic

rule system responsible for mapping  $N \rightarrow \lambda$  is “radically different” from the system which takes  $N \rightarrow \pi$ . We assume no linguistic variation in the syntactic computation; the same operations apply to lexically-encoded features to derive observable differences between particular languages. However, unlike syntax, PF rules vary cross-linguistically, and have different orders (or rankings) with respect to one another – orders which also vary cross-linguistically (Bromberger and Halle, 1989). Thus, as MacSwan (1999) concludes for reasons having to do with the structure of the PF computation, switching at PF may indeed be impossible (1999: 177). MacSwan (199b) details the following theory as integral to the overall Minimalist Program:<sup>87</sup>

*PF Disjunction Theorem (MacSwan 1999)*

- (i) *The PF component consists of rules/constraints which must be (partially) ordered/ranked with respect to each other, and these orders/rankings vary cross-linguistically,*
- (ii) *Code switching entails the union of at least two (lexically encoded) grammars,*
- (iii) *Ordering relations are not preserved under union,*
- (iv) *Therefore, code switching within a PF component is not possible.*

---

<sup>87</sup> MacSwan (2000) states the following as advantages of a Minimalist approach to code switching: (a) because linguistic differences are encoded in particular lexical items, the grammatical contribution of each language in a code switched sentence can be clearly identified, (b) because the syntactic component of the computational system ( $C_{HL}$ ) may be assumed to be invariant cross-linguistically, no “control structure” or “third grammar” is required to mediate between contradictory requirements, (c) because Minimalism focuses on minimal assumptions (allowing only those suppositions which correspond to “virtual conceptual necessity”), it is a natural framework in which to take seriously the view that there are no code-switching-specific constraints. This forces us to examine the data more rigorously, and may often lead to new insights in bilingualism and the theory of grammar, (d) because the Minimalist Program is motivated by many theoretical and empirical considerations in the context of monolingual data (Chomsky 1995), pursuing a Minimalist approach to code switching allows us to remain consistent with current work in syntactic theory as it relates to monolingual language, and (e) because the phonological component of the computations system ( $C_{HL}$ ) is assumed to be different in nature from the syntactic component, and because rules/constraints of the phonological system are ordered/ranked with respect to each other, we may disallow code switching in phonology but still permit it in syntax in a natural way (2000: 51).

In light of the above Theorem, MacSwan (2000) states how the above principles allow for greater freedom of previously deemed ungrammatical utterances (see below Poplack 1981):

- (75) \* Juan está *eat*-iendo  
Juan is eat GER  
'Juan is eating'  
(*English/Spanish*, Poplack 1981)

But with the PF Disjunction Theorem, the below is possible as the English lexical stem 'park' is already entered into the Spanish phonology and so no ill-formed construction ensues:

- (76) Juan está *parqu* -eando su coche  
Juan is park GER POSS car  
'Juan is parking his car'  
(*English/Spanish*, MacSwan 2000)

However, even after employing the PF Disjunction Theorem, it still does not account for the wealth of Moroccan Arabic/ English data which shows high levels of English preposition affixing to Moroccan Arabic stems:

- (77) *I'm forever mesh* -ing *the tabla* and *hiya me ta-dir welu*  
I'm forever wipe GER the table and she NEG do 3SG nothing  
'I'm forever wiping the table and she does nothing'

In (77) above, the verb 'to wipe' *mes-h* is certainly not phonologically integrated into English yet this is a recurrent morphological construct. How does Minimalism account for this? In simple terms, it does not and it is to the chagrin of this theory. Other examples of non-phonologically integrated switches are as follows, all uttered by second and third generation Moroccan/Arabic speakers:

- (78) Dad's *ḍaah* -ing *the houli for al Eid*  
Dad's slaughter-GER the sheep for DEF Eid  
'Dad's slaughtering the sheep for Eid'

MacSwan, in elaboration of the Minimalist Program, details how Poplack's (1981) Free Morpheme constraint is violated<sup>88</sup> with the examples he gives but only goes so far as to comment on intra-sentential bilingual discourse where the morpheme in question is already phonologically integrated into the lexicon of the host language. Myers-Scotton (2002) in her discussion of MacSwan's Theorem, as outlined above, states that:

The PF component consists of rules or constraints that must be (partially) ordered/ranked with respect to each other, and these orders can vary cross-linguistically (1999: 45).

Myers-Scotton also states how "*phonological systems cannot be mixed*" and that for this reason code switching at PF produces "unpronounceable" elements which violate F1 (full interpretation). Myers-Scotton (2002) adds that this "Does not stand up to scrutiny" (2002: 159). In light of this, Myers-Scotton (2002) states that this line of assumptions and arguments means that any switches involving bound morphemes from the Matrix language (affixes) in the same word as an Embedded Language content morpheme are not permitted, making the PF Disjunction Theorem very reminiscent of Poplack's Free Morpheme constraint (1980) which also disallows intra-word switches (2002: 1590).

However, adopting a non-constraints-based approach to code-switching is certainly attractive and as Pinker (2003) suggests:

The Minimalist Program appears to be parsimonious and elegant, eschewing the baroque mechanisms and principles that emerged in previous

---

<sup>88</sup> MacSwan (2000) although showing evidence of counter-examples of Poplack's (1981) Free Morpheme constraint, later states that: "For now, I will assume that the Free Morpheme constraint is descriptively correct, taking it to bar word-internal switches where such switches violate the integrity of legitimate X<sup>0</sup>- level elements" (2000: 46).

incarnations of generative grammar such as the Extended Standard Theory and Government-Binding Theory (Pinker 2003: 20).

However, counter examples and criticisms were lodged based on new empirical evidence (Myers-Scotton 2000, 2002). Jake & Myers-Scotton (2005) suggest that MacSwan's attempt to show how the Minimalist Program can explain CS on its own fails. Theoretically, while either of the participating languages in CS could frame the bilingual CP, only one, the ML, does. That is, recognizing the construct of the ML as the source of the morpho-syntactic frame of each bilingual clause showing CS is necessary. This is a valid point as in any given bilingual string, one language will always provide the grammatical frame and why this is refuted in preference for a non-constraints approach is an oft too-laboured point.<sup>89</sup> Another key criticism of the MacSwan's Minimalist Program is its being based largely on phrasal code switching. In Myers-Scotton's (2002) she states that:

Like most minimalist approaches, his (MacSwan) rules out singly occurring lexemes as code switching (from the Embedded Language under the MLF model). He does this in two ways. First, any Embedded Language form that is inflected with Matrix language morphemes is simply considered a borrowing. Second, any Embedded language form without Matrix Language inflections is a borrowing if its grammatical features (such as phi-features within DP/NP and agreement on Infl/TP) differ from those of the monolingual frame of the other language (2002: 159).

This then would explain why the Moroccan Arabic/English intra-sentential data in (77) and (78) above cannot be fully explained under the Minimalist Program as it largely concerns itself with phrasal clause switches as opposed to intra-morphemic switches. Furthermore, Myers-Scotton (2002) states how MacSwan then argues when there are alien singly occurring forms in a seemingly bilingual constituent, the constituent is simply monolingual, not bilingual. That is, these alien forms are borrowings. One can see why MacSwan would like to rule out these forms because he

---

<sup>89</sup> The matrix language will be discussed in more detail in the next Chapter.

can then say that any features these forms have that are different from those of the base language are not relevant to the derivation. Thus any features mismatches are irrelevant and this explains why a derivation does not crash (2002: 159).

## **2.7 Conclusion**

This chapter has analysed the major theoretical and syntactic approaches to intra-sentential code switching in bilingual discourse over the last few decades. I have highlighted the most salient syntactic approaches. Firstly, the linear approach in analysing grammatical perspectives of intra-sentential code switching pioneered by Poplack (1980, 1981) and Poplack & Sankoff (1981) who detailed linguistic constraints which govern the interaction of two language systems. This took the form of the Equivalence Constraint where codes are switched at points where the surface structures map onto each other. This was then followed by the Free Morpheme Constraint where a switch may occur at any point in the discourse at which it is possible to make a surface constituent cut and still retain a free morpheme, the main premise being that code switches are permissible intra-morphemically so long as the basic requirements of each language are respected and maintained. However, this constraint met with numerous counter-examples, including that of Moroccan Arabic/English data as presented above. It must be noted however, that this constraint was ground-breaking and a major influence on subsequent syntactic approaches to code switching.

This was then followed by the government model as outlined by Di Sciullo, Muysken & Singh (1986). This is described by MacSwan (2000) as having the virtue that this constraint refers to an independently motivated principle of grammar (government), while the other proposals considered have not (2000: 39). However, as described, the government constraint does not stand up to empirical evidence as

outlined in the main body of this chapter due to overriding government relations. It has been further noted that the government relation is not necessary in syntactic theory (Chomsky 1995). A more recent proposal was that of Mahootian & Santorini (1993) with their Tree Adjoining Grammar and that of Null Theory which focuses on the complement relation in phrase structure within bilingual discourse. The main premise of this theory is that the language of the head determines the phrase structure position of its complements in code switching just as in monolingual contexts. However, this was later modified with their constraint (1996) which focused on less specific properties of syntactic heads. These theories met with countless counter-examples mainly due to the over-specific nature of their phrase structure approach. The Minimalist model although attractive in its simplicity bases itself on phrase structure to the chagrin of bound morpheme switches and other intra-morphemic switches. However, a monolingual-based approach to CS is respected for its simplicity and overall application to most switched data but fails on account of the data presented above.

I also presented a seminal approach as put forward by Joshi (1985) and Myers-Scotton (1993a) in the asymmetric models. In employing a structured framework in describing intra-sentential code switching, a matrix frame model was put forward, whereby the matrix language provides the grammatical frame enabling the syntactic lodging of the embedded language variety. This revolutionary model, The Matrix Language Frame model provides a solid way in which to describe natural bilingual discourse as inevitably, one of the two languages will provide the grammatical frame. Although this model has met with criticisms (MacSwan 1999, 2000, 2002), it cannot be denied that this model is far more successful in an all-encompassing analysis of code switched data irrespective of language typology. It is

with this in mind that we turn to the next Chapter which provides a full analysis of the Matrix Language Frame model and its application to Moroccan Arabic/English data.

## CHAPTER THREE

### THE MATRIX LANGUAGE FRAME MODEL

After having analyzed the major syntactic theories and grammatical approaches to intra-sentential code switching over the last thirty years, this chapter focuses on the asymmetric model first pioneered by Joshi (1985) who applied Garrett's (1975) speech error study analysis to his model and formulated a matrix and embedded model.<sup>90</sup> This was then later modified and developed (Azuma 1991, 1993) and Myers-Scotton (1993a, 1999, 2000 and 2002) as previously detailed in Chapter Two. Myers-Scotton (1993a) examined a Swahili/English corpus consisting of recorded conversations in Nairobi and proposed the Matrix Language Frame model which forms the focus of this chapter. This is a major work in a non-linear perspective, but quite different in its approach from those based on the generative syntax models as described earlier. Since the MLF was first proposed (Myers-Scotton, 1993a), there have been a number of modifications and it is currently one of the most influential models to account for intra-sentential CS and on close analysis, this model is the most encompassing of all theoretical approaches to code switching as it largely covers all data irrespective of language typology.

In this manner, it has to be understood, that *descriptively*, the MLF suits all data in the sense that there will always be one of two languages (or more) that provides the grammatical frame, the matrix language, with the other language, the embedded variety assuming that frame and adhering to it. This chapter commences

---

<sup>90</sup> Another contribution to the asymmetric approach to intra-sentential code switching was that of Hasselmo (1972) in a series of publications which examined language use amongst American Swedes. This model, the Ordered Selection describes how English morphemes can be inserted into Swedish discourse. Also, Bautista (1975) designed a model of the bilingual speaker using Tagalog and English data taken from radio broadcasts. She describes the insertion of lexical items, constituents and clauses from one language, L1 into a frame set by L2. Both of these approaches have received little exposure in the literature and the matrix approaches were born independently of both Haselmo's (1972) and Bautista's (1975) contributions.

with a definition of what constitutes a matrix language, and an overview of the most prominent asymmetric models (Joshi 1985), Azuma (1991, 1993) and Myers-Scotton (1993a, 1999, 2000 and 2002), the later of which is presented in a micro analysis of the MLF model where the hypothesis outlined in this asymmetric model is examined, and the validity of the model is evaluated using different language types as well as Moroccan/English data. The MLF model is examined in detail in outlining various insertional approaches, the content – system morpheme dichotomy and various extended sub-models, namely that of the 4-M model and the Abstract Language model. The final part of this chapter is dedicated to Boumans (1998) Monolingual Structure Approach with a discussion of its similarity to the MLF model and its definition of a matrix language. A further analysis will follow in Chapters Six and Seven based on challenges to the MLF model and other counter-examples in the literature. We begin then with a definition of the notion of a matrix language.

### **3.1 Defining the Matrix Language**

The concept of a Matrix language first gained momentum with the seminal research by Joshi (1985) who first (based on Garrett's 1975 speech error study) fully explored the asymmetry between a Matrix and Embedded variety. Later definitions ensued which were then later modified as and when more evidence or further research came to light. Myers-Scotton (1993b) in her analysis of Swahili and English code switching, defined the Matrix languages as: "The language of more morphemes in interaction types including intrasentential codeswitching" (1993b: 68). This is clearly problematic since it is difficult to clearly and concisely count which and how many morphemes this would comprise.<sup>91</sup>

---

<sup>91</sup> Boumans (1998) correctly points to the problematic nature of this early analysis of what constitutes a matrix language. He cites how: "This definition is problematic since it leads to discussions about what discourse samples would be valid and which would count as morphemes. Also it does not really avoid

Such a concept of a Matrix Language stems from the assertion that a grammatical structure containing elements from two languages can be attributed to that grammar of one of these languages (the ML), rather than to the grammar of both languages, to the overlap of both grammars or to a third 'codeswitching' grammar (Boumans 1998: 61). It would appear sensible then to attribute a grammatical frame to one language in bilingual discourse and from that then establish the embedded variety. This is the very essence of a matrix – embedded asymmetry. Other linguists' definitions have also sought to show how the ML is identified by its role in structuring the clause and/or frame Klavans (1985), Nishimura (1986), Schmidt (1986) and Pandit (1986). A more apt and precise description proffered by Myers-Scotton & Jake (1995), irrespective of the number of morphemes in a given complementizer phrase, is that as described:

*Matrix Language definition (Myers-Scotton 1995: 983):*

*The ML is the language projecting the morphosyntactic frame for the entire CP which shows intrasentential CS.*

Indeed, it is this definition that I apply in this thesis (however see my own definition below in Generalization One) to examine and analyse CS and Moroccan Arabic

---

the supposed circularity because in languages where system morphemes normally outnumber content morphemes, it is precisely the structural role of the ML which leads to the correct outcome. (Conversely, if the (matrix) language is a language in which content morphemes normally outnumber system morphemes, this method would yield the wrong result if all the content morphemes in a stretch of discourse were embedded!)” (1998: 38). This is an interesting analysis but given Myers-Scotton later modified the definition of a matrix language (2000, 2002), it renders this perspective obsolete. In fact, Myers-Scotton (2002) states that this early claim of identifying the matrix language by the number of morphemes was “later abandoned” and that “Unfortunately, some researchers trying to apply the MLF model still read only the 1993 version of the model.” Furthermore, she states that “as early as 1995, the claim was modified (Myers-Scotton and Jake 1995: 984) and does not appear in publications after 1993.” Also she states that “the reason for abandoning that claim are twofold: First, even though the language that is the source of the grammatical frame (as specified in the Morpheme Order and System Morpheme Principles) often supplies more morphemes in a bilingual CP, this is not always the case. Second, as stated, the criterion was to apply a ‘discourse sample; but exactly what would constitute such a sample is ambiguous” (2002: 61-62).

English data. The below highlights the Matrix Language frame, with Moroccan Arabic providing the grammatical frame:

(79) *i fit- it ha al barah*  
Only fit – PAST it FEM DEF yesterday  
'I just fitted it yesterday'

The verb 'fit' above as the embedded language variety clearly adheres to the syntactic frame as stipulated by the Moroccan Arabic, with the past simple suffix tagged onto the end. Clearly then, Myers-Scotton's definition (post 1993b) of a Matrix Language is not only justified but correct in both its definition and application and this is justified in most data sets as inevitable, amongst a bilingual intra-sentential string, one of the two (or more) participating languages will have to provide the grammatical frame.

### 3.1.1 Criticisms of the Matrix Language

However, the definition of a Base language or Matrix language has been criticised in the literature in analysis of grammatical constraints on CS (Gardner-Chloros, 2004: 117). Nortier (1990: 158) explains that the matrix language is about individual sentences and the base language is about a whole conversation. Klavans (1993) proposes that the inflection of the finite verb is the key to defining a base language. Myers-Scotton originally proposed a 'morpheme count' as a criterion for the definition of the ML (1993: 117) but she abandoned this criterion later, e.g. in the afterword added in the 1997 reprint of her earlier work (1993a); see also Myers-Scotton & Jake (1995). She also mentions that ML is different from "dominant language" in the psycholinguistic literature and unmarked choice in the sociolinguistic literature (Myers-Scotton, 1997) as:

Dominant language refers to the language in which the speaker is most proficient and unmarked choice is a label for the variety considered most appropriate (and therefore typically most frequent) in a specific interaction type in a specific community (1997: 268).

For her, there are two principles that define Matrix Language. One is that the ML is the language which determines the morpheme order and the system morphemes. The other principle (in the 4M model) is that the ‘outsider late system morphemes’ construct the morpho-syntactic frame and thus establish that language as the ML.

One of the overriding questions (and criticisms) aimed at the MLF model, is that of identifying the matrix language. As Myers-Scotton (2002) later describes:

It is defined as by the role it plays in the Matrix Language – Embedded Language hierarchy, realized in the Morpheme Order Principle and the System Morpheme Principle. Second, the definition of the Matrix Language is not circular, as some have suggested. To begin at the beginning – a basic premise of the MLF model is that the languages referred to as Matrix Language and Embedded Language do not participate equally in structuring intra-CP codeswitching. This unequal participation is referred to as the Matrix Language – Embedded Language hierarchy, and the Matrix Language is the label identifying the language with the larger structural role (2002: 59).<sup>92</sup>

Myers-Scotton has been criticised for implying that the matrix language in a bilingual string ‘changes often’ due to her description of the matrix language as a dynamic construct. Myers-Scotton lists the following points which are implicit in the MLF model:

1. One variety is consistently the single source of the frame of bilingual CPs; thus the source of the Matrix Language does not change within any single bilingual CP.

---

<sup>92</sup> Myers-Scotton (2002) continues to state that: “But which language is the Matrix Language? The MLF model provides the two principles as tests of the premise of unequal participation and as a way to identify the Matrix Language. If the terms of the principles, morpheme order and one type of system morpheme, both are satisfied by one and the same language, then the Matrix Language can be identified as that language. Further, the basic theoretical notion that there is a Matrix Language - Embedded Language hierarchy is supported, because the two languages do not both satisfy the roles of the Matrix Language contained in the principles” (2002: 59).

2. Theoretically, the source of the Matrix Language may change in a conversation (but not within a CP) as topics or some participants change; however, even such changes are rare or non-existent in most corpora.
3. When what I label compromise strategies (bare forms, Embedded Language islands, etc) occur within a bilingual CP, the Matrix Language for the entire CP does not change.
4. As a pragmatic strategy, structures that are marked for the Matrix Language frame (e.g. marked word order) are allowed. Yet this is not intended to be an escape hatch, allowing many counterexamples to be explained as 'marked'.
5. At most, the Matrix Language shows only minor, infrequent instances of convergence toward structures in the Embedded Language as long as the bilingual situation remains relatively stable.
6. Some communities with near-balanced bilinguals provide a different pattern from the prevalence of bilingual CPs. The MLF model still applies, but it is relevant to less data because the number of monolingual CPs increases and of bilingual CPs decreases.<sup>93</sup>
7. For such bilinguals, the dominant pattern may be alternation between monolingual CPs in each of their languages. That is, the Matrix Language still does not change within a bilingual CP; however, the Matrix Language may change within a conversation (e.g. second generation Turks in Tilburg, the Netherlands, cf. Backus (1996)).

---

<sup>93</sup> See section below on speech styles of different generation of Moroccan Arabic and English speakers and my concept of *Reactive Syntax*. Code switching is not a static theory and this is the downfall of most theoretical approaches to it as speakers today are using more and more innovative combinations and structures which quickly unravel the premises of most theories. The basic MLF model with asymmetry between one matrix language and one embedded variety is the only model which suits wide-ranging data.

Myers-Scotton describes (1-3) above as testable hypotheses and (4-7) as assumptions (2002: 112), disallowing a change within a single CP is not accepted given dynamic structures and combinations evident in the data, most notable Moroccan Arabic and English intra-sentential data. Second and third generation bilingual speakers merge Moroccan Arabic and English mid-morpheme and certainly within a CP. This is perhaps covered in (7) above where “The Matrix Language may change within a conversation,” but that on a syntactic level does not explain the reality of the situation. I can therefore make the following generalization:

*Generalization 1:*

*The Leader (of two or more languages) is that which contributes word order in the CP frame where the Led (embedded) variety adheres to the morpho-syntactic frame provided by the Leader.*

The above generalization I have formulated makes no other specifications, does not specify the number of morphemes required and allows for an array of suffixations and code-switched varieties. This is important as it makes allowances for new speech styles in different generations in languages of typological dissimilarity.

Terminology also contributes to definitions of the matrix language where the very essence of the term matrix language refers to how it organizes and expresses the grammatical relations in the sentence by means of inflection, function words and word order.<sup>94</sup>

### **3.1.2 Early Accounts of the Matrix Language Frame Model**

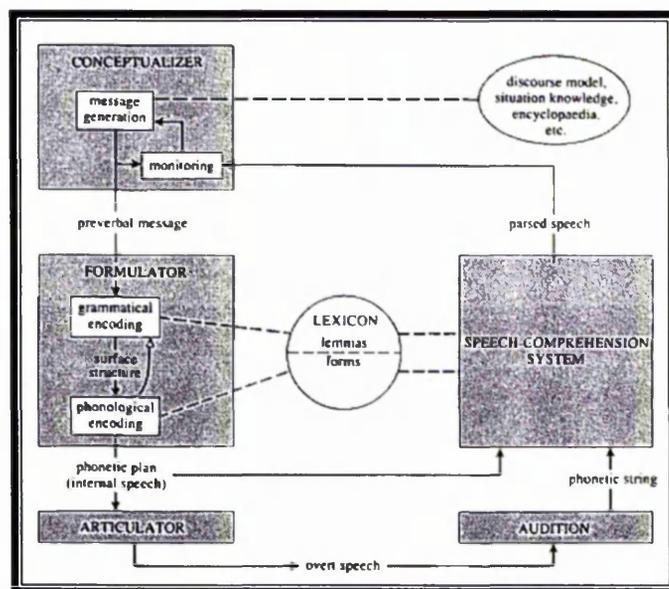
Early accounts of one language providing a dominant grammatical frame over and above the other in bilingual speech have been distinguished by several researchers

---

<sup>94</sup> cf. Boumans (1998: 34).

(Sridhar & Sridhar 1980; Joshi 1981, 1983, 1985; Pandit 1986 and Petersen 1988). Research by Joshi (1981) provided fertile ground for asymmetric approaches to code switching (Doron, 1983; Klavans, 1985; Nishimura, 1986) as well as Myers-Scotton in the 1990s. Myers-Scotton (1992) based her MLF model on Levelt (1989)<sup>95</sup> and Garrett's (1990) models of speech production as a basic outline for the MLF model as presented in the early presentation of this model (1992). This model, like that of Levelt assumes that the lexical items in the mental lexicon consist of two parts; a lemma part containing semantic and morpho-syntactic information, and a form part (lexeme) containing phonological information. In order to fully appreciate the development of the MLF model, it is important to note and examine the bilingual speech model as posited by Levelt (1989):

*Figure 3.1: Levelt's (1989) Speech Production Model*



<sup>95</sup> De Bot (1992) also based his bilingual speech production model on Levelt's (1989) monolingual speech model. In de Bot's model however, the type of organization between the monolingual and bilingual speech production models differ. At the levels of the lexicon and the formulator, their organization is what is characterized in the subsystems hypothesis (separates the conceptual from the linguistic system). At the level of the articulator, the two languages are part of an extended linguistic system and there is no systematic distinction between the two. De Bot assumes that this level-based organization between the two language systems is necessary to account for the observed phenomenon of a foreign accent. He conceives a foreign accent as representing the functional influence of the L1 on L2 production, thus an extended system is required to account for this. A full analysis is beyond the scope of this thesis due to time and space constraints. For further insight see (De Bot 1992, Paradis 1987).

### 3.1.3 Levelt's Speech Production Model

Levelt's (1989) speech model provides a good premise for language production analysis and code switching in particular.<sup>96</sup> His unilingual language model shows that languages can be utilised in single or mixed discourse and this is a useful analysis in code switched discourse (cf. Nortier 1989 for further corroboration). De Bot (1992) gives a good account of the above and provides a breakdown of the main processing accounts in Levelt's (1989) model, namely: Conceptualizer, Formulator and Articulator:

De Bot's analysis of Levelt's (1989) Speech Production Model (1992: 4):

- a. A conceptualizer: this is where the selection and ordering of relevant information takes place and where the intentions the speaker wishes to realize are adapted in such a way that they can be converted into language. The output of this component is so-called 'pre-verbal messages', in other words, messages which contain all the necessary information to convert meaning into language, but which are not themselves linguistic. In the planning of pre-verbal messages, two stages can be distinguished: macro-planning and micro-planning. Macro-planning involves the elaboration of communicative goals / intentions and the retrieval of the information needed to express these goals, while micro-planning is the "speaker's

---

<sup>96</sup> This model served as a templatc frame of reference for language production in general. De Bot (1992) describes: "The model aims at describing the normal, spontaneous language production of adults. It is a 'steady-state' model, and not a language learning model, and it hardly says anything about language perception. The model is not concerned with reading and writing and it is not aimed at the explanation of language disorders of a central or peripheral nature. A distinction is made between declarative knowledge, which includes encyclopaedic knowledge (conceptual and lexical knowledge in particular) and situational discourse knowledge, and procedural knowledge which is relevant to the processing of declarative knowledge. Procedural knowledge forms part of the different processing components. A final characteristic is that the same lexicon is used for production and perception" (De Bot 1992: 3).

elaboration of a communicative intention by selecting the information whose expression may realize the communicative goals” (Levelt 1989: 5).

- b. A Formulator where the pre-verbal message is converted into a speech plan (phonetic plan) by selecting the right words or lexical units and applying grammatical and phonological rules. In the lemma, the lexical entry’s meaning and syntax are represented, while morphological and phonological properties are represented in the lexeme. In production, lexical items are activated by matching the meaning part of the lemma with the semantic information in the pre-verbal message. The lemma information of a lexical item concerns both conceptual specifications of its use, such as pragmatic and stylistic conditions, and (morpho)-syntactic information, including the lemma’s syntactic category and its grammatical functions, as well as information that is needed for its syntactic encoding (tense, mood, pitch, aspect, case and pitch accent).
- c. An articulator which converts the speech plan into actual speech. The output from the Formulator is processed and temporarily stored in such a way that the phonetic plan can be fed back to the speech-comprehension system and the speech can be produced at normal speed.

Aspects of lexical storage and the Formulator have been well-documented and analysed (Grosjean 1986, Paradis 1978) and there are two main approaches in examining the storage and neurological compartmentalisation of bilingual lexicons. These are as follows (De Bot 1992):

- 1 There is a separate formulator and a separate lexicon for each language. This solves the problem of having to separate the two systems. It will cost some

storage capacity, but it is economical because there is no need to have a system that controls the co-ordination and separation of the two languages. It is, however, unclear how the two languages can be used simultaneously, during code-switching for example.

- 2 There is one large system which stores all the information, linguistically labelled in some way, about all the different languages. The problem which results from this solution is that it does not explain how the systems are separated in bilinguals without causing apparent problems.

In analyzing code switching, lexical storage and data systems in general are of paramount importance, but as yet, there has been little progress on a convincing conclusion on how bilingual data is stored, accessed and processed. However, as De Bot (1989) states:

When we take into account research which has been done on storage and retrieval of lexical and syntactic information by bilinguals, we could imagine a probable solution somewhere between these two extremes. Elements/knowledge of the two languages may be represented and stored separately for each language or in a shared system depending on a number of factors. The most important of these seem to be the linguistic distance between the two languages and the level of proficiency in the languages involved (De Bot 1992: 9).<sup>97</sup>

The MLF then incorporates the processing components of the Levelt (1989) model but is not in itself a language production model (Myers-Scotton 2002: 23).

---

<sup>97</sup> See also Paradis (1978) for an analysis of typologically similar and dissimilar languages where it is stated that there is some correlation between languages that are closely related (i.e. French is more closely related to Dutch than to Moroccan Arabic). Analysis of 'coherence between linguistic distance is summed up by Paradis (1978) as thus: "According to such a view cerebral representation of bilingualism would be on a language pair-specific continuum, ranging from a bi- or multi register unilingualism to a bilingualism involving two related languages" (Paradis 1978: 16). How does this then relate to a bilingual Moroccan Arabic / English language speaker? De Bot (1992) further explores how language proficiency is an obvious factor where separate or jointly stored knowledge about the two languages is concerned. A person who knows a few words and sentences in a foreign language will not have a separate system for this. The first-language system is flexible enough to add an additional register to those already in existence (De Bot 1992: 9). This viewpoint is accepted and validated given the informants recorded during data collection sessions where individuals with more receptive bilingual skills would not have a separate linguistic system as those that were productively bilingual.

Furthermore, Myers-Scotton (2002) states that this (Levelt's 1989) psycholinguistic model for language production also motivates the emphasis the MLF model places on the lexicon as source of grammatical projections connecting intentions with surface forms. This means that in its discussion of what code switching involves, the MLF model suggests that the nature of language production is relevant (2002: 14). Given that most asymmetric models assume the premise of the Levelt's (1989) model, namely, the information processing components as posited in the above diagram, we turn to early accounts of code switching in the matrix hierarchy.

### **3.2 Matrix Language Approaches**

Born independently of the models by Bautista and Hasselmo during the 1970s and 1980s, the matrix language approach offered a regulated and organized way in which to qualify intra-sentential code switching. Grammatical relations during inter-sentential and intra-sentential discourse are schematically organized between content words and function words also referred to as content morphemes and system morphemes.

As discussed previously, formal grammatical approaches to intra-sentential code switching have been developed in one of two frameworks, namely (i) the two grammatical systems are kept separate and a highly-developed switching mechanism is created enabling the speaker to switch between the two (or more) languages, (ii) the grammatical systems of the two (or more) languages are merged producing a third grammar with modifications and amendments of syntactic rules. Joshi's (1983) model then maintains two separate monolingual context-free phrase structure grammars where a switching mechanism is proposed to control code-switching between their corresponding constituents.

### 3.2.1 Sridhar & Sridhar's Dual Structure Principle

Work carried out by Sridhar & Sridhar (1980) in their analysis of Hindi and English code switching termed the matrix and embedded structure as 'host language' and 'guest language' respectively and this complimented their Dual Structure Principle as mentioned in Chapter Two and repeated here:

*Dual Structure Principle - Sridhar & Sridhar (1980: 412):*

*The internal structure of the guest constituent need not conform to the constituent structure rules of the host language, so long as its placement in the host language obeys the rules of the host language.*

Pandit (1986) referred to the matrix language as that of the 'governing' language (1986: 36). In Wentz (1977), the language of the sentence is namely the one in which the determiner and main verb are produced. Klavans (1985) however defines the matrix language as that which bears the inflection, in other words, the ML is determined by the language of the main verb. Interestingly, Lahlou (1991) cites the matrix language as the speaker's mother tongue which is a gross over-simplification as although speakers may be fluent in any language, on a micro complementizer scale, the matrix language which sets the grammatical frame varies from one 'frame' to another.

It is with this definition now clearly structured that we focus on the development of this important model, commencing with early accounts and varieties of asymmetrical models as set out.

### 3.2.2 Joshi's Asymmetry Model

As touched upon in Chapter Two, Joshi proposed two major constraints on the switching mechanism as outlined in Chapter Two and repeated below:

*Joshi's code-switching constraints (1983)*

- (i) *Asymmetry constraint: constituents can switch from the matrix language to the embedded language, but not vice versa,*
- (ii) *Closed class constraint: closed class items cannot be switched.*

Joshi's (1985) asymmetric model dictates that closed class items such as determiners, quantifiers, prepositions, possessives, Aux, Tense, and helping verbs, however, cannot be switched. However, this was met with counter-examples where closed class items were switched:

- (80) It goes without saying I think *que* ['that'] along with the picketing we are doing a boycott.  
(English/Spanish, Pfaff, 1979: 314)

The closed class item in [80] above is the relative clause marker '*que*' which is switched in this intra-sentential clause where usually, the closed class item, or relative marker in this case should be in English or the whole CP phrase is in Spanish from the beginning.<sup>98</sup> Other counter-examples also show how some closed class items can be switched:

- (81) El dentista agarraba *off* y se iba fishing.  
'The dentist would take off and go fishing'  
(Spanish/English, Pfaff, 1979: 254)
- (82) inta *hang* -ha up.  
You *it* FEM  
'You hang *it* up'  
(English/Arabic, Mohamed, 1983)

The above counter-examples highlight how stringent syntactic constraints on code-switched utterances delineate from naturally code-switched data and it is this main aspect of syntactic models and constraints which are to the chagrin and detriment of

---

<sup>98</sup> See also Prince & Pintzuk (1984) for an in-depth analysis of Joshi (1983) and data on Yiddish code-switching.

most approaches. This further serves to highlight the lack of theoretical application in describing second and third generation Moroccan Arabic / English speakers who use a more intricate form of code-switching, referred to as *Reactive Syntax* where speakers use innovative and highly-intricate strings in bilingual discourse (to be discussed below). It is *Reactive Syntax* as current and innovative ways of new speech styles, particularly amongst second and third generation of Moroccan Arabic / English speakers are constantly evolving and the syntax is then 'reactive' as opposed to static. In so doing, these utterances do not 'fit' into any of the models and constraints thus presented except for the MLF and its matrix – embedded hierarchy.

### 3.2.3 Counter-Examples to Joshi's Model

The below examples show further examples of third generation Moroccan Arabic speakers which certainly counter the model by Joshi (1983) and highlight an emerging code switching variety where 'closed' class items and/or system morphemes are switched which not only belie most models and theories:

- (83) Cut *it ha wa* put *it ha al hih*  
Cut PAST it FEM and put PAST it FEM DEF there  
'I cut it and put it there'

The clitic pronoun object marker '*ha*' in [83] above, should, by all accounts, linear, government, minimalist and asymmetric be in Moroccan Arabic as the matrix language is in Moroccan Arabic. However, the boundaries of syntax have been superseded and this gives rise to new emerging forms. It must be noted that these are not speech errors nor are they ad hoc statements but are in fact an accepted form of speaking. Further examples are given below:

(84) *rmiy- ah ʔla al ʔabla*  
Throw it MASC on DEF table  
'Throw it on the table'

(85) He shouted at her *bʔel* someone in a rage  
He shouted at her like someone in a rage  
'He shouted at her like someone in rage'

In (84) above the definite article is according to Joshi's (1983) model a closed class item and should according to this framework also adhere to the language of the matrix clause, namely Moroccan Arabic and the same in (85) where the preposition should also be in Moroccan Arabic.

However, the model is successful in that it paved the way for other researchers to examine and adopt the matrix / embedded asymmetrical hierarchy and apply it to intra-sentential code switching. With this concept of asymmetry, the distinction between closed and open class items was redefined and further developed in Azuma's (1990, 1993) model. A departure from Garrett's (1988) speech production model and largely influence by Joshi (1983), Azuma (1993) formulated the Frame Content Hypothesis whereby "closed-class items are accessed and retrieved" (1993: 1072). This model is similar to that of Myers-Scotton (1993) but certainly less encompassing and elaborate. The next section details the MLF in full from *Duelling Languages* (1993b) to Myers-Scotton's later work with revisions and amendments to the earlier model (2002).

### 3.3 Matrix Language Frame Model

Myers-Scotton corroborated and assumed not only the terminology, although slightly modified,<sup>99</sup> but built on the asymmetric matrix model as pioneered by Joshi (1983,

---

<sup>99</sup> In the matrix language – embedded language hierarchy, Myers-Scotton replaced Joshi's (1983, 1985) closed class items with 'system morphemes' and open class items with 'content morphemes', namely lexical versus function morphemes.

1985) and proposed a new model with its own syntactic rules and statements. Initially, Myers-Scotton's work was influenced by psycho-linguistic research and speech production models and it is within this framework that the MLF is born rather than the more formal grammar theories as have been previously proposed. In her earlier work, Myers-Scotton formulated the MLF as "the model [which] calls on a suprasyntactic level to motivate its constraints" (1993b: 82) and later described it as:

The Matrix Language Frame Model specifically was designed to explain structural configurations found in codeswitching, specifically classic codeswitching. (2002: 10).

### 3.3.1 Structure of the MLF

The MLF can be structurally divided into two types, that of inter-sentential and intra-sentential CS. As discussed previously, inter-sentential CS has mainly been approached in the field of sociolinguistics given the straightforward nature of its syntax. In this manner, grammatical constraints and considerations are not applied to inter-sentential code switching. However, intra-sentential code switching, which is the primary focus of the MLF model, is solely studied from a grammatical perspective given the intricate juxtaposition and collision of two separate languages, each with its own syntax, lexicon and phrase structure rules. Inter-sentential CS involves monolingual (a) + monolingual (b) utterances, however, intra-sentential CS involves bilingual utterances within the same unit of discourse. Using Myers-Scotton's (2002) unit of analysis I take the CP (maximal projection) as the frame of discourse as this is more concise and appropriate in micro-analysis of intra-sentential code switching.<sup>100</sup>

---

<sup>100</sup> Myers-Scotton (2002) goes on to state in 'championing' the need for a CP that: "In contrast with sentence or clause, its status is clear. A CP is the highest unit projected by lexical elements. It can be defined unambiguously in terms of phrase structure as a complementizer or an element in Specifier (Spec) position followed by an IP (cf. Myers-Scotton & Jake 1995: 982). Second, CP is a unit used by many statisticians, no matter what model they espouse, because of this status, there is no reason to assume its use here implies preference for any syntactic theory. Third, because CPs can contain null elements in Comp (Complementizer) position or elsewhere, using the CP as the unit of analysis avoids

Furthermore, Myers-Scotton (2002) states that employment of the CP is more appropriate than a sentence or clause as:

Grammatical constraints on codeswitching only become potentially interesting within a sentence. That is, if one sentence is in language X and the next in language Y, the grammars of the two languages are hardly in contact. However a moment's thought tells you that even within a sentence, the grammars may not be in contact. The problem using the sentence as a reference point is that what is called intra-sentential codeswitching (i.e. a sentence showing morphemes from more than one language) can contain many different structural configurations (2002: 55).

The example below outlines inter-sentential code switching in Moroccan Arabic and English:

(86) \_\_\_\_A\_\_\_\_. CP \_\_\_\_B\_\_\_\_.

*Bgit heða.* And that as well  
I want this And that as well  
'I want this. And that as well'

The example below outlines intra-sentential code switching:

(87) CP \_\_\_\_A + B\_\_\_\_.CP

*Jib li a newspaper wa het u* there  
Bring to me DEF newspaper and put it (MASC) there  
'Bring me the newspaper and put it there'

The MLF model has been created to clearly define grammatical outcomes of intra-sentential CS and of all models previously created or submitted, it is the MLF model which as Muysken (1995) declares as "Must be closer to the truth" (1995: 188), the truth being the most suitable in describing intra-sentential code switching within the Matrix Language and Embedded Language hierarchy. A further description of the MLF is in *Contact Linguistics* where previous work is explained and revisions to previous models made. Myers-Scotton (2002) explains the four general premises in the orientation of the MLF.

---

problems regarding the status of constituents with nulls" (2002: 55). These reasons are borne in mind and taken into consideration for this thesis.

### 3.3.2 Four principles of the MLF

In the MLF (Myers-Scotton 2002), four general principles guide the overall approach to all contact phenomena, not just code switching. (Of course, the parts of these principles that are not specific to bilingual speech also apply to monolingual data).

These are outlined below:

#### Four Principles of the MLF (Myers-Scotton 2002)

- 1 The Matrix language Principle. There is always an analyzable or resolvable frame structuring the morphosyntax of any CP. This frame is called the Matrix Language. In bilingual speech, the participating languages never participate equally as the source of this Matrix Language.
- 2 The Uniform Structure Principle. A given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for this constituent type must be observed whenever the constituent appears. In bilingual speech, the structures of the Matrix Language are always preferred, but some Embedded Language structures are allowed if certain conditions are met.
- 3 The Asymmetry principle for Bilingual Frames. Bilingual speech is characterised by asymmetry in terms of the participation of the languages concerned. In what I now call classic codeswitching, only one of the participating languages is the source of the Matrix Language. In other contact phenomena (such as composite codeswitching), the Matrix Language may be a composite of abstract features from more

than one language, but asymmetry still marks the contributory roles of the participating languages. This asymmetry is evidence of the universal drive in language to achieve uniformity in the structural frame of any variety, to avoid meaningless variation-although, this outcome never entirely exists in any language. Still, the drive is always there, and in bilingual speech it is especially expressed as part of the movement toward the morphosyntactic dominance of one variety in the frame.

- 4 The Morpheme-Sorting Principle. All morphemes are not equal. This is an example of asymmetry of a different type. That is, at the abstract level of linguistic competence and production, there are different types of morphemes. In bilingual speech, the outcome of these abstract differences is that all morphemes from the participating varieties do not have equal possibilities of occurrence (2002: 10).

In analysis of Moroccan Arabic and English data, the MLF is found to be the most suitable and all-encompassing model in linguistic research to date in explaining intra-sentential code switching in that it will always cover the interface between two (or more) languages in discourse analysis and contact linguistics in general. The below examples show typical structures which pertain to the MLF model and its basic theoretical concept:

(88) *al yawm ʕnd- hum open day beš na- sajjəl*  
DEF today have 3PL open day so that record 1PL  
'Today they have an open day so I can register'

In (88) above, the matrix language is quite clearly Moroccan Arabic with English providing content morphemes, namely a noun phrase which is inserted into the

syntactic frame of the MLF. See (89) and (90) below which clearly set out the matrix language in Moroccan Arabic, thus providing the grammatical frame. This then is in line with Generalization 1 outlined above:

(89) that's why *xallit - hum fi l catalogue hetta tiji*  
That's why I left them in DEF catalogue until come 2SG  
'That's why I left them in the catalogue until you come'

(90) *iwa, drive i ha al garage u xalli ha Øamma*  
well, drive it FEM DEF garage and leave it FEM there  
'Well drive it to the garage and leave it there'

We can see with examples of data listed thus far, how the MLF accounts for different structures in bilingual discourse. Furthermore, this model largely seems to be supported by a wealth of data from typologically different code-switched language pairs and as described by Boumans (1998): "There are significant advantages in the matrix language or insertional approach" (1998: 46). Although the MLF is clearly a model, it is lexically based and as Myers-Scotton says:

This means that it emphasizes the abstract procedures directed by lemmas in the mental lexicon. Some of these procedures necessarily refer to phrase structure, but also to the role of oppositions elsewhere at more abstract levels. Admittedly, to say a model is lexically driven seems to mean different things to different researchers. To me, what is most relevant to the discussion in this volume are the following points: as already noted, lemmas in the mental lexicon underlie surface-level lexical elements. The lemmas contain lexical rules and these rules contain all the necessary information to realize surface constructions. This means that a specific lemma entry contains (i) the morphological information that is associated with a surface-level content morpheme, (ii) syntactic properties (a subcategorization frame)<sup>101</sup> of that morpheme, and (iii) a semantic and a pragmatic representation. Each type of information within a lemma forms the input for a particular type of formal operation; thus, in some sense lemmas are compounds of operations (cf. e.g. Hoekstra, van der Hulst, and Moortgat 1980, Aronoff 2000). Thus the MLF model is not primarily a phrase structure model (i.e. not a syntactic model). This does not mean that the MLF model is not potentially compatible with

---

<sup>101</sup> Subcategorisation is defined as the "Specification, usually by a set of rules of what kind of items may represent a class or category in a particular environment, eg when a transitive verb (but not an intransitive verb) must occur together with a noun, object etc" R. R. K. Hartmann & F. C. Stork (1972: 223-224).

most contemporary syntactic models. Certainly, it seems to be very compatible with many of the views of Jackendoff (1997) on the relationship of the lexicon to syntactic and phonological components (2002: 14).

An example taken from Boumans (1998) in describing the matrix language asymmetry in Moroccan Arabic and French clearly shows which language is the matrix proffering the grammatical frame:

- (91) *les restaurants mehlul-in hetta l l- wahed – a d – lil*  
DEF-PL restaurant open- PL until to DEF one-F of DEF night  
'The restaurants are open until one in the morning'  
(Moroccan Arabic / *French*, Slaoui, 1986)

Furthermore, in highlighting bilingual intra-sentential discourse, it is evident that one language will always provide the grammatical frame, the matrix language and this fact cannot be refuted where the embedded variety is left as the secondary language, a closer inspection of which is detailed below.

### 3.3.3 Matrix Language – Embedded Language Distinction

The MLF distinguishes between the distribution of the two languages and renders them asymmetric in intra-sentential CS. As stated by Myers-Scotton in an outline of the Matrix Language – Embedded Language asymmetry, these are in fact the key oppositions and their syntactic relation is essential in analysing the grammatical framework in bilingual code switched varieties. See below aspects of the ML / EL asymmetry:<sup>102</sup>

ML / EL asymmetry:

---

<sup>102</sup>In Chan's (2003) Cantonese / English code switching analysis, he describes how Myers-Scotton & Jake (1997) state that the EL morpheme must be congruent with its ML counterpart in terms of predicate-argument structure, conceptual-lexical structure and morphological realization patterns (2003: 153).

- 1 The participating languages in codeswitching do not contribute equally. The language marking the larger contribution is called the Matrix Language and the other language is called the Embedded Language. Within the terms of the MLF Model, 'contributing more' does not mean *more* morphemes, although this is often the case. Rather, contributing *more* means *more abstract structure and structure of a certain type*. Specifically, the Matrix Language-Embedded Language opposition is most salient in regard to mixed constituents. (Mixed constituents are those with morphemes from two or more languages or, as we will see when phenomena other than classic codeswitching are discussed, abstract structure from two or more languages.)
  
- 2 The importance of recognizing the abstract structure behind surface phrase structures is largely encapsulated in the second opposition, that between content morphemes and system morphemes.<sup>103</sup> Content morphemes are the main elements conveying semantic and pragmatic aspects of messages, and system morphemes largely indicate relations between the content morphemes. Clearly, then, these two types of morpheme perform different functions in language in general, monolingual or bilingual. But because system morphemes are related to constituent structure, where they will come from which of the participating languages may contribute content

---

<sup>103</sup> The use of the terms 'content' and 'system' morphemes was employed by Myers-Scotton (1993b) in lieu of the classically employed 'closed class' and 'open class' distinction as used by most psycholinguists is described by Myers-Scotton "because 'content' is an easily understood term, it causes few problems; it is superior to 'thematic' just because it is more semantically transparent, even though it is true that all content morphemes are thematic. The term 'system' morpheme' requires more discussion. First, the term comes from Bolinger (1968), who used this as a cover term for both inflectional morphemes and function words. Second, my reason for using 'system morpheme' is not to be a contrarian in a world that habitually uses the terms 'open' and 'closed' class elements or 'thematic' and 'functional' elements. Rather, 'system morpheme' is used because it identifies a class of morpheme more precisely than either of the other widely used terms, 'closed class' or 'functional element' (2002: 71).

morphemes to bilingual CPs, but not all can contribute critical system morphemes. This is the domain of the Matrix Language<sup>104</sup> (2002: 16).

The above shows how the more dominant of the two languages in an intra-sentential string will always provide the grammatical frame and the distribution is asymmetrical. The MLF is structurally based such that EL units are inserted into the ML frame with feature checking and an adherence to Universal Grammar and the well-formedness condition maintained. The below example shows Moroccan Arabic providing the grammatical frame as the ML with English items inserted:

(92) *qli- ha u eat-i -ha*  
Fry it FEM and eat 2SG it FEM  
'Fry it and eat it'

The above structure is quite a common format amongst Moroccan Arabic English bilingual speakers, with either English or Moroccan Arabic inserted into the Matrix frame.

### 3.3.4 The Content Morpheme – System Morpheme Hierarchy

Identification of the content morpheme – system morpheme<sup>105</sup> distribution is essential for the MLF model as they are important in constructing grammatical frames on which the Matrix Language assumes as the dominant structure. The content–system

---

<sup>104</sup> Myers-Scotton (2002) describes how the Matrix Language-Embedded Language opposition refers to linguistic competence-in the sense that, psycho-linguistically, the bilingual's two or more languages do not achieve equal activation in bilingual speech. Decisions (largely unconscious) made at the pre-linguistic conceptual level result in one language dominating (the Matrix Language sets the grammatical frame of speech). The less dominant language (the Embedded Language) participates largely by supplying lexical elements that are integrated into that frame. The content-system morpheme opposition refers to how lexical elements are organized in the mental lexicon and differentially accessed in the language production process. This affects how they participate in bilingual CPs as well (2002: 16).

<sup>105</sup> This is one of the major aspects of Myers-Scotton's (1993a) model taken from *Duelling Languages* and has since been revised in (2002) with the 4-M model and this then takes on more importance. As outlined by Myers-Scotton (2002) in analysis of the revised 4-M model, the new model accomplishes two things: (i) it provides more precise explanations for what occurs in classic codeswitching, and (ii) with extensions, it offers a different approach from what is generally found in the contact literature to explain the form of other contact phenomena (2002: 69).

morpheme distinction is motivated by the way the two types of morphemes pattern according to frame-building properties. An established tenet of syntactic theory is that different morphemes relate differently to the thematic grid of an utterance; this is independent motivation for the content-system morpheme distinction (Myers-Scotton 2002: 29). The initial distinction between the content morpheme and system morpheme opposition as per Myers-Scotton (1993) is thus:

*Content Morphemes – System Morphemes (1993: 99-101):*

- a) [+Quantification] – System morphemes are quantificational, e.g. quantifiers, determiners, and possessive adjectives,
- b) [-Thematic Role Assigner] – System morphemes do not assign thematic roles,
- c) [-Thematic Role Receiver] – System morphemes do not assign themselves roles.

Content morphemes on the other hand e.g. nouns, verbs, adjectives and some prepositions assign and/ or receive thematic roles as they are non-quantificational.<sup>106</sup>

This opposition gave rise to the early work (1993) and the Morpheme Order Principle:

---

<sup>106</sup> Apart from these categories, most of the system morphemes are what are assumed to be functional categories – determiners, quantifiers, modal verbs and inflectional affixes (Chan 2003: 152). Chan (2003) goes on to cite counter examples in terms of specifying content morphemes and system morphemes against the two above principles “In one case, what are classified as EL content morphemes, for example, pronouns, do not always appear in mixed constituents”. As mentioned previously, major criticisms of the MLF were based on earlier work (1993) and not revised models and structures (post 2002). In dealing with this ‘problem’ of mixed constituents, the 4-M model was introduced, which clearly and concisely deals with the micro nature of the content morpheme – system morpheme principle. This is described in the following section. Boumans (1998) describes Myers-Scotton’s (1993) content-system morpheme principle as “more complicated.” However, this is clearly not the case given that it consistently and precisely gives a clear and well-thought out model of functional versus non-functional categories. Boumans (1998) goes on to criticise the model in stating that: “Differentiating system and content morphemes in a principled way that is valid for all languages is an extremely complicated task. I will only mention some major objections to the criteria advanced by Myers-Scotton. The criteria must not be manipulated such as to make certain word classes fit into the right category. Firstly, the argument that pronouns and descriptive adjectives are potential thematic case receivers since they are dominated by the category NP (1993b: 126) raises the question as to whether thematic roles are assigned to constituents or content morphemes. Elsewhere, Myers-Scotton argues that content morphemes are distinguished from system morphemes because the former are

*Morpheme Order Principle (1993: 83)*

*In ML + EL constituents consisting of singly-occurring EL lexemes and any number of ML morphemes, surface morpheme order will be that of the ML.*

The below examples outline this principle:

(93) *ana na-šri i* quality  
I buy buy 1SG only quality  
'I buy only quality'

(94) *haðu really ʕjbūni*  
These really like me  
'I really like these'

(95) *la wsaɣəd šrit set. U li kanu ʕndi fi l* bedroom dirt hum  
DEF pillows I bought set. And REL were I have in DEF bedroom, I put them  
*Θamma*  
there  
'The pillows that I bought, a set. And those that I have in the bedroom, I put them there'

The content – system morpheme opposition is clear and in examples (93), (94) and (95), the grammatical frame is provided by the Moroccan Arabic, which, socio-linguistically suits empirical evidence as the speaker was born and raised in Oujda, Morocco and came to the UK in 1971 with Moroccan Arabic as her preferred language of communication. Interestingly, the above conversation was recorded between the Moroccan-born mother and British-born daughter, both of whom are fluent bilinguals. Given the above principle, we can see how the opposition clearly

---

potential thematic role assigners of receivers (1993b: 100). If one maintains that thematic roles are assigned to content morphemes rather than to constituents, then it is not clear why descriptive adjectives are thematic role-receivers. Secondly while free form pronouns are thematic role receivers and thus content morphemes, clitic pronouns are system morphemes because these are actually "agreement particles which are co-indexed with a null NP head" (1993b: 126). Myers-Scotton (2002) in her defence states: "Admittedly, for some linguists, the status of some lexemes in relation to this dichotomy is an open issue; that is, there is not a consensus on the thematic status of some elements, such as adjectives. However, there is general agreement that all nouns receive thematic roles and most verbs (but not the copula) and most prepositions assign thematic roles. Thus the status of the most central elements bearing content in the CP is clear" (2002: 70-71). Once again, criticism is always of Myers-Scotton's earlier work and not the later revised editions / models.

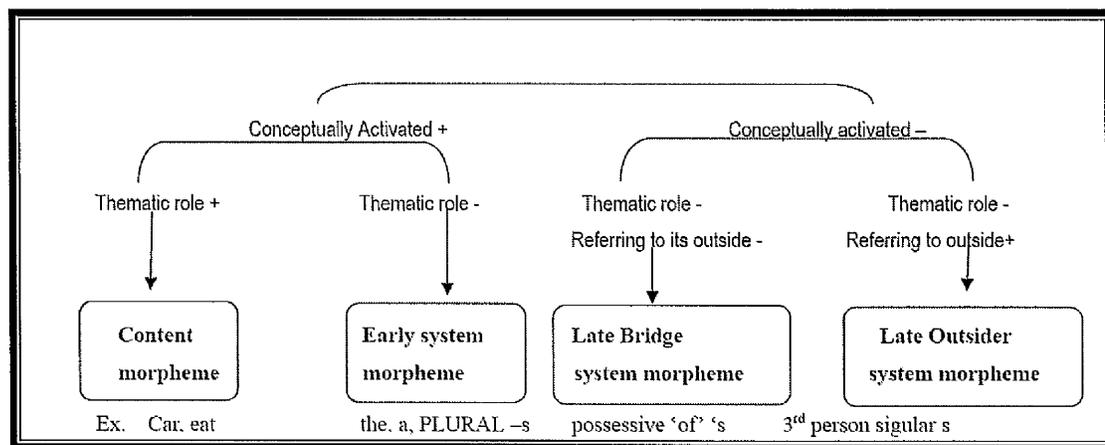
suits the Moroccan Arabic / English data. However, Myers-Scotton made further revisions to this model and this refined model is known as the 4-M model (2002).

### **3.3.5 The 4-M Model**

The MLF model not only accounts for a very wide range of code switching examples, but can explain most of them under the Matrix Language – Embedded Language and the content – system morpheme oppositions. But these oppositions do not explain them all on their own. An extended and revised model, namely the 4-M model which supports the MLF, is an extension of the earlier content-system morpheme opposition and was formulated in order to explain a wider range of code-switching data. This model results from a collaboration with Jan Jake (cf. Myers-Scotton & Jake 2000a, 2000b 2001). The model was revised so that it can now not only explain code switching as a concept but also offers plausible explanations for distributions in a wide range of other data in different languages. In essence, the 4-M model takes the content-system morpheme opposition in the MLF model and breaks down the class of system morpheme into three types. The MLF model had offered a formal way of distinguishing content and system morphemes (content morphemes participate in the thematic grid of an utterance by either assigning or receiving thematic roles, but system morphemes do not. Prototypical content morphemes that receive thematic roles are nouns (e.g. in the thematic role of Agent, nouns are often mapped onto the grammatical relation of Subject; in the role of Patient or Theme, they are often mapped onto the grammatical relation of Internal Argument or Direct Object.) Prototypical content morphemes that assign thematic roles are most verbs and some prepositions. In contrast, system morphemes neither assign nor receive thematic roles.

Most function words and inflections are system morphemes.<sup>107</sup> The below diagram outlines the key aspects of the 4-M model:

Figure 3.2: Morpheme Classification (Myers-Scotton 2002: 73)



In this revised extended model of Myers-Scotton’s earlier work (1993b), the above diagram stipulates that:

1. Content morphemes: content morphemes assign / receive a thematic role and are activated at the lemma level. They are directly selected according to the speaker’s intention.
2. Early system morphemes: If a system morpheme is activated at the lemma level, it is an early system morpheme. Although they do not have a thematic role, they contribute to the mapping of the conceptual structure to the lemma-like content morphemes. Myers-Scotton and Jane (2000: 96) define early system morphemes as “always realized without going outside

<sup>107</sup> Myers-Scotton states that the different types of morpheme under the 4-M model are differently accessed in the abstract levels of the production process. Specifically, content morphemes and early system morphemes are accessed at the level of the mental lexicon, but late system morphemes do not become salient until the level of the Formulator. The 4-M model implicates a model of surface distributions of morpheme type that is based on abstract competence (2002: 17).

of the maximal projection of the content morpheme that elects them” and further that “their form depends on the content morpheme with which they occur”.

3. Late system morphemes: Late system morphemes neither assign or receive thematic roles nor are they activated at the lemma level. They are activated at the formulator level when the lemma sends directions to construct a grammatical constituent. Late system morphemes are further categorized as two-bridges or outsiders.
4. Late bridge system morphemes: Like early system morphemes, bridge system morphemes depend on information inside the maximal projection in which they occur. Unlike early morphemes, they do not contribute to conceptual structures. They integrate content morphemes into a larger constituent, e.g. the possessive markers “of” and “s” where they link two nouns within a noun phrase. This is also used to characterise the Moroccan Arabic possessive marker “*dya*l” which otherwise would prove problematic for the 4-M model.
5. Late outsider system morphemes: Outsider morphemes differ from bridge morphemes in that they “depend on grammatical information outside of their own maximal projection” (2000: 100). They are then structurally assigned at the positional / surface level. An example would be the 3<sup>rd</sup> person singular –s as a late outsider morpheme.

Furthermore, Myers-Scotton (2002) states the 4-M model has four types of morphemes: The Differential Access Hypothesis of the 4-M model is that the four types of morpheme are related in different ways to the production process. First,

content morphemes are the only morphemes whose lemmas link them directly to speakers' intentions. Speakers' intentions activate language-specific semantic/pragmatic feature bundles that underlie the conceptual information that content morphemes will convey. In turn, these bundles point to lemmas in the mental lexicon. The lemmas underlying content morphemes are directly selected and their content is salient at the level of the mental lexicon. Second, the lemmas underlying one type of system morpheme also become salient at this level; these are called early system morphemes because of their early saliency. Their lemmas are activated when the lemmas supporting content morphemes point to them. These indirectly selected lemmas further realize the conceptual content of the semantic/pragmatic feature bundles. Third, the other two types of system morpheme become salient when their lemmas are activated at the level of the Formulator. They are called late system morphemes because the hypothesis is that their saliency is delayed until the Formulator level. They are activated by the directions sent to the Formulator by the lemmas underlying content and early system morphemes; these are directions to build larger linguistic units such as CPs and IPs.

### 3.3.6 Possessive Marker *Dyal* and the 4-M Model

In Moroccan Arabic / English code switching, the late bridge system morphemes are crucial in clearly identifying the use of the frequency markers. One such marker is that of the Moroccan Arabic possessive marker "*dyal*":

(93) *haða dyal* that man  
 This POSS that man  
 'This is that man's'

The above construction identifies *dyal* in the 4-M model as a late bridge system morpheme as these bridges build structures within a maximal projection as they are

structurally assigned to do so by the well-formedness requirements of the relevant constituent (Myers-Scotton 2002: 91). In the above example [93], *dyal* bridges the demonstrative *hada* with the relative marker and NP. Well-formedness is indeed a pre-requisite in that *dyal* cannot be used in an ad hoc manner nor can it be randomly assigned. The following examples highlight this:

(94) \**dyal* me  
POSS me  
'Mine'

(95) \**dyal* his  
POSS his  
'His'

(96) \**dyal* her  
POSS her  
'Hers'

Therefore, it can be seen that *Dyal* as a possessive marker can only be used in Moroccan Arabic and not in combination with English or any other language in direct suffixation:

(97) *dyal* i  
POSS me  
'Mine'

(98) *dyal* u  
POSS his  
'His'

(99) *dyal* ha  
POSS her  
'Hers'

Why, therefore, given the above data can *dyal* only be used in Moroccan Arabic? Myers-Scotton (2002) states that: "I have no ready answer to this question" (2002: 91). The following example given by Bentahila & Davies (1998) in criticism of the System

Morpheme Principle shows how *djal* which is a system morpheme, should only come from the matrix language, but in this case, comes from the embedded language:<sup>108</sup>

- (101) ....de quel degre de connaissance *djal* la personne  
....on which degree of knowledge of the person  
(Moroccan Arabic / French; Bentahila & Davies, 1998: 38)

In short, Myers-Scotton (2002) cannot answer why the possessive marker *djal* is in Arabic, which in the above example is the embedded language, and not in French, which is the matrix language. We can only assert then that *djal* is a ‘wildcard’ and can be inserted almost anywhere and will violate most principles, except of course, as noted in examples [96] to [101] when it is suffixed where the suffix can only be in Arabic. This then allows us to make the following generalization:

*Generalization 2:*

*Possessive marker “djal”:*

*“Djal” as a possessive marker can only take a direct Moroccan Arabic suffix in Moroccan Arabic.*

The following examples highlight this:

- (102) It is clearly *djal* -*hum* so give it back  
It is clearly POSS their so give it back  
‘It is clearly theirs so give it back’

---

<sup>108</sup> Myers-Scotton (2002) in answering her critics and in response to examples [97] to [99] above states “It is true, as Bentahila & Davies point out, that I identify *djal* as a system morpheme in *Duelling Languages* (2002: 106) and do so in such a way that I imply that *djal* is the type of system morpheme relevant to the System Morpheme Principle. I admit I am guilty of this implication; either I was confused myself or my wording was an oversight. The 4-M model makes the status of morphemes such as *djal* very clear: it is a bridge late system morpheme, not an outsider. Therefore, its presence (ostensibly from the Embedded Language in [102] above does not violate the principle. Although bridge late system morphemes pattern with outsider late system morphemes as [-conceptually activated], bridges are then differentiated from late outsiders by another opposition, [+/-look outside maximal projection] Bridges build structure within a maximal projection; they are structurally assigned to do so by the well-formedness requirements of the relevant constituent. But they do not look outside their immediate maximal projection for their form; they are not co-indexed with an element outside that maximal projection, as are late outsider system morphemes. This is a crucial difference. In this case, *djal* functions very much like English of or French ‘*de*’ in joining together two NPs” (2002: 91).

In (102) above, *dyal* can only be suffixed with Moroccan Arabic following the language of *dyal*. Otherwise, the whole possessive clause would have to be in English 'theirs'. However, another circumstance would be to have an NP + bridge + NP associative construction (2002: 91):

- (103) If it's *dyal* that girl, then you have to give it back  
If it's POSS that girl, then you have to give it back  
'If it's the girl's, then you have to give it back'

The presence then of late system morpheme, 'bridges', depends on the maximal projection in which they occur. This is a code-switching mechanism and is well-used by bilingual speakers to 'get round' morphemes which can only take a direct suffixation in Moroccan Arabic. They integrate elements in a constituent when the well-formedness conditions for those constituents call for them. Syntactically, another example of this is the marker *ɛnd* 'with/on,' etc which follows the same grammatical bearings as that of *dyal* above. The preposition *ɛnd* cannot be used with any other language in direct suffixation and can only take Moroccan Arabic:

- (104) \**ɛnd* I  
Have me  
'I have'

- (105) \**ɛnd* he  
Have he  
'He has'

- (106) \**ɛnd* she  
Have she  
'She has'

Instead, *ɛnd* can only take a Moroccan Arabic suffix:

- (107) *ɛnd* - i  
Have me  
'I have'

- (108)  $\xi$  nd - u  
 Have he  
 'He has'
- (109)  $\xi$  nd - ha  
 Have she  
 'She has'

Therefore, we can make the following generalization:

*Generalization 3:*

*Suffixation and  $\xi$ nd*

*" $\xi$ nd" as a prepositional stem can only take a direct Moroccan Arabic suffix in Moroccan Arabic.*

Therefore, in light of the above, the second type of late system morphemes, 'outsiders' look outside their immediate maximal projection for information about their form (in this case Moroccan Arabic suffixes). Both types of late system morpheme are structurally assigned in contrast with both content morphemes and early system morphemes, which are conceptually activated (2002: 17). This revised model<sup>109</sup> is certainly more refined and content-specific than that of Myers-Scotton's earlier work (1993b) with the 4-M model detailing the activation stage in the mental lexicon and the formulator as well as clearly and concisely further distinguishing between content morphemes and system morphemes. This is further highlighted in the Abstract Level model which is a supporting paradigm of the MLF.

---

<sup>109</sup> Myers-Scotton (2002) states that the 4-M model adds precision to the MLF model, but also its implications seem to explain configurations in data well beyond the scope of the MLF model. First, because of these wider applications, the 4-M model connects a theory of grammar with language production and processing in some more general ways than the MLF model (2002: 85). Clearly then, the 4-M model provides a more comprehensive approach to code switching enabling it to encompass most, if not all, data. Of course, there will always be exceptions to any rule of code switching but this is largely due to the innovative nature of language use and its evolution.

### 3.3.7 The Abstract Level Model

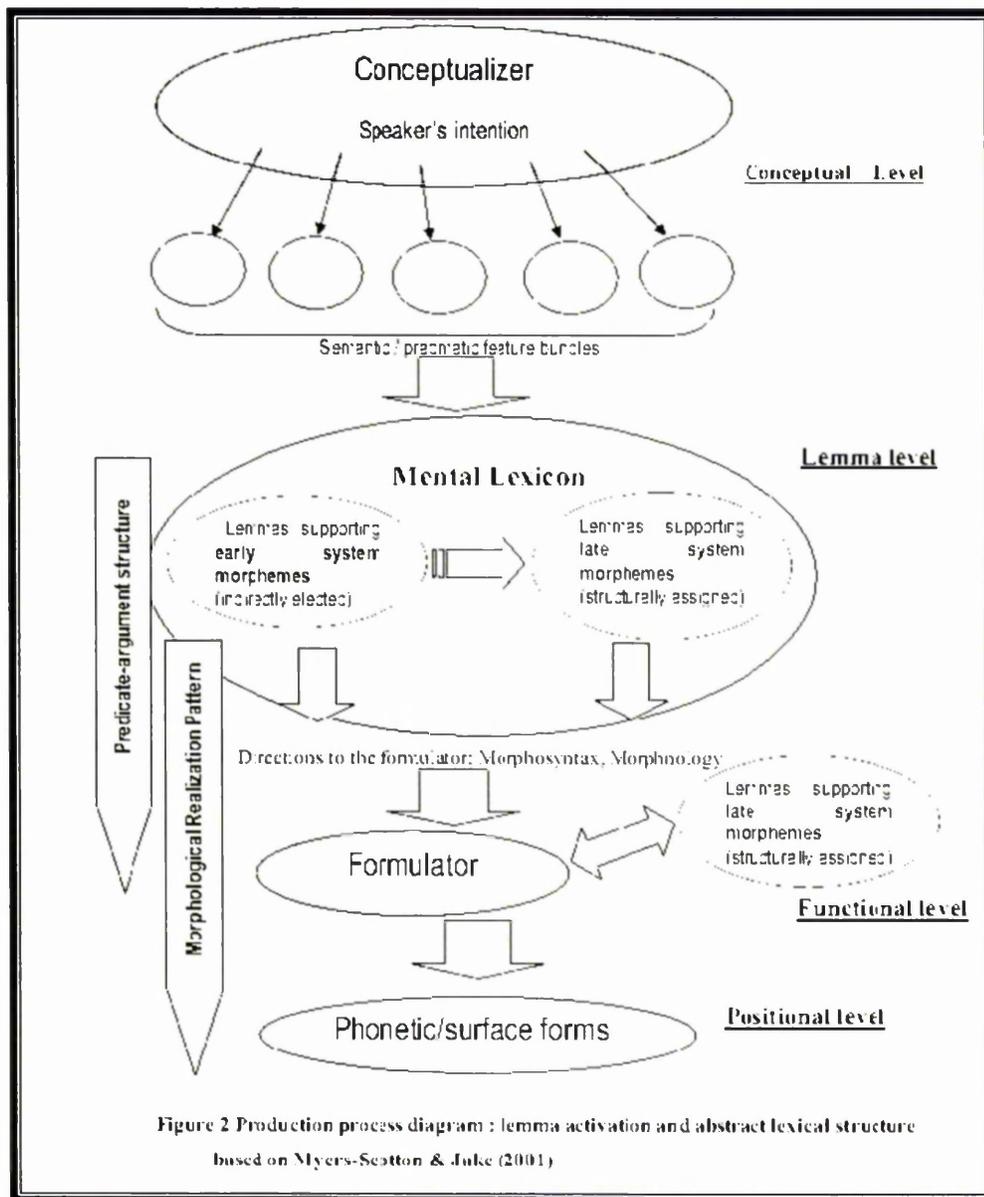
Although the MLF model involves a sound syntactic theory in explaining bi-lingual intra-sentential code switching, it cannot on its own, explain the wide-ranging amount of data in the world's languages. As a result, the MLF is supported by the 4-M model as described above and in addition to this, Myers-Scotton (2002) and in collaboration with Jake (2000a, 2000b, 2001) also added the Abstract Level model which aims to encompass as much data as possible. This model grew out of earlier research carried out by Myers-Scotton.<sup>110</sup> The major premise of the Abstract Level model is that all lemmas in the mental lexicon include three levels of abstract lexical structure. The three levels contain all the grammatical information necessary for the surface realization of a lexical entry. The levels refer to (i) lexical-conceptual structure, (ii) predicate-argument structure, and (iii) morphological realization patterns<sup>111</sup> (2002: 19). Therefore, this third contributing model to the MLF in essence describes as Myers-Scotton states: "what will count as sufficient congruence in CS so that certain constructions are possible for certain language pairs" and also "how it provides a principled explanation for the nature of the abstract morpho-syntactic frame that structures bilingual clauses" (2002: 19). The diagram below shows the production process and highlights the relation between the 4-M model and Abstract Lexical structures:

---

<sup>110</sup> Outlines of the Abstract model were introduced in Myers-Scotton & Jake (1995) and made more detailed by Myers-Scotton and Jake (2000b).

<sup>111</sup> However the crux of the issue is in deciphering what sufficient congruence means. This notion has not yet been adequately refined. Very definitely, *sufficient* does not mean complete congruence because, of course, content morphemes across languages are rarely *completely* congruent.

Figure 3.3: Production process diagram: lemma activation and abstract lexical structure based on Myers-Scotton & Jake (2001)



The Abstract Level model then, in supporting the MLF, premises that language production is made through three abstract levels of the lemma and these are as follows:

1. Lexical conceptual structure: At this level, psycholinguistic and sociolinguistic intention in the conceptualizer activates “language specific semantic/pragmatic feature bundles” (Myers-Scotton & Jake, 2001) between the conceptualizer and the mental lexicon.

2. Predicate-argument Structure: At the next level, thematic structure is mapped onto grammatical relations e.g. Agent, Subject, Beneficiary, Indirect Object.
3. Morphological realization pattern: At the third level, grammatical relations are realized on the surface e.g. word order, agreement morphology. This completes the construction of the output form ready for input into the production processes. When ELs appear in an ML frame, their congruence with ML counterparts must be checked at the three levels of abstract lexical structure in the mental lexicon.

The importance of this model is that it is a support mechanism for the MLF and it covers most data sets and typologically dissimilar language pairings in intra-sentential code switching. In this sense:

The net result is that a bilingual clause can be structured by levels from more than one language and combined with abstract structure from another language, resulting in a composite. (2002: 22).<sup>112</sup>

This is a good model in explaining third generation Moroccan Arabic / English speakers and their innovative and varied intra-sentential code switched varieties.

Myers-Scotton continues to state that:

Instead of having a matrix language for a bilingual clause that is largely isomorphic with a single language, in some contact phenomena, the Matrix Language itself is a composite” and she defines the composite matrix language as “an abstract frame composed of grammatical projections from more than one variety. It can result when speakers do not have full access to the desired matrix Language, or when there is competition between languages for the role of Matrix Language (reflecting socio-political competitions that affect Matrix Language selection at the conceptual level) (2002: 22).

---

<sup>112</sup> Myers-Scotton (2002) goes on to state that: “Such results have gone under a variety of labels, such as ‘transference’ and ‘interference’ and even ‘creolization’. However, labelling phenomena does not explain them” (2002: 22).

The Abstract Level model then captures data and bilingual structures which are in essence ‘compromise’ strategies. The below example from Turkish / Dutch data (cited in Myers-Scotton 2002) illustrates a clear example of compromise strategies at play:

(110) bunlar herkes kendi *prijs* söyli-yor  
‘These everyone self price say-PROG 3S  
(Turkish / Dutch, Backus, 1990: 109)

As described, example (110) includes a bare form (*prijs* ‘price’) from Dutch, the Embedded Language, in Turkish / Dutch code switching in Tilburg, the Netherlands. That is *prijs* occurs without the case suffix that would make it well-formed in the language of the Matrix Language, Turkish. (To be well-formed in Turkish, the Dutch noun should have a possessive third person singular suffix because of the reflexive that precedes it. Also, because of the possessive, which makes it specific / definite it needs and accusative suffix too. Thus *prij-i-ni* is expected) (Myers-Scotton 2002: 98). We can take this one step further and introduce the notion of a *Reactive Syntax* whereby innovative varieties of bilingual data is uttered by second and third generation of speakers; in this case, Moroccan Arabic / English speakers. This is an important aspect of this thesis as seemingly experimental aspects of code switched data in natural discourse are becoming more and more common, most noticeably amongst third generation speakers. The composite and innovative nature of discourse in both second and third generation speakers is very interesting as it allows us to witness new syntactic varieties not present in first generation Moroccan Arabic / English speakers.

### 3.4 MLF Composite Structures and *Reactive Syntax*

Myers-Scotton (2002) Myers-Scotton (2002) distinguishes two types of intra-sentential CS: classic code-switching and composite CS. In classic CS, only one of the

participating languages is the source of the morphosyntactic structure of the bilingual clause, whereas the morphosyntactic structure consists of two languages in composite code-switching (2002: 8).

In defining the new concept of ‘composite code switching, Myers-Scotton characterizes it as:

A phenomenon with morphemes from two languages within a bilingual CP, and with the abstract morphosyntactic frame derived from more than one source language. Composite codeswitching occurs in such phenomena as language attrition and shift.<sup>113</sup> It occurs when speakers – because of psycholinguistic or socio-political factors – do not have full access to the morphosyntactic frame of the participating language that is the desired source of the Matrix Language. Or possibly, the target of a target Matrix Language is not clear to the speakers themselves. The result is that a composite Matrix Language frames the resulting bilingual CP. Thus, in effect, composite codeswitching necessarily entails convergence (2002: 105).

The examples below portray innovative combinations of Moroccan Arabic and English syntax where speakers merge two constructions but maintain Universal Grammar, and well-formedness is not compromised. Contrary to Myers-Scotton above, this is not an example of language attrition nor is it language shift of sorts; it is a new way of speaking amongst third generation Moroccan Arabic speakers in the UK.

The examples below show some of the many structures and combinations possible:

- (111) *jib -- i        maξak        basket -a*  
Bring POSS with POSS basket FEM SG  
‘Bring with you a basket’

In (111) above, the speaker is in fact merging two structures, that of Moroccan Arabic and English. The MA version would be *silla* (basket) which ends with a feminine marker (inanimate singular) and it is evident that the speaker is applying the same construct to that of the English counterpart ‘basket’. Further examples follow:

---

<sup>113</sup> Language shift is defined as “A change from one habitual use of one language to that of another” (Weinreich 1967: 68).

- (112) There are too many *rajul -z* at this party  
 There are too many man PLURAL at this party  
 ‘There are too many men at this party’

In (112) above uttered by a third generation Moroccan Arabic / English speaker, the combination is rather striking as the speaker is using the plural form of *rajul* (man) which is *rajel* (men) and adding an English plural marker –z to further highlight the plurality of the noun ‘men’. This is the very essence of the *Reactive Syntax* I am putting forward as an innovative concept as it has not been discussed in any of the literature, and certainly not amongst British Moroccan speakers. We cannot term this a consequence of language attrition as the speaker in the above example (112) is bilingual, nor can we term it interference as the speaker is evidently emphasising the number of men at the party. This data can be captured in Myers-Scotton’s (2002) Abstract Language Model explained above. A further example of the *Reactive Syntax* concept is in (113) below:

- (113) those *klēb-z* are always getting on my nerves, *wlād la hrām*  
 Those dog PLURAL are always getting on my nerves, sons DEF badness  
 ‘Those dogs (men) are always getting on my nerves, sons of bitches’

As in (112) above, (113) again uses a ‘double’ plural marker to re-iterate and re-emphasise the plurality of the subject. It is now not a question of asking why such constructions are used, but rather to explain them. The English plural suffix marker –s/z receives almost total morpho-syntactic integration and as Myers-Scotton (2002) describes “this is evidence that there is sufficient congruence at all levels [for this to occur].” (2002: 98).<sup>114</sup> The most common of switches in Moroccan Arabic / English

---

<sup>114</sup> Undoubtedly, at the conceptual level, speakers, whether the result is monolingual or bilingual speech, map their intentions onto language. They make a number of decisions, largely unconsciously. They consider the socio-political and psycholinguistic possibilities of this mapping and its consequences. For monolingual speakers, this means they consider their dialectal and stylistic choices.

code switching and in *Reactive Syntax* are that of singly inserted lexical items, some of which are cultural terms, expressions, or general nouns in either direction: Moroccan Arabic to English or English to Moroccan Arabic.

### 3.4.1 *Reactive Syntax and Single Word Switches*

In essence, by far the largest group of switches concerns the insertion of single words in one language in sentences in the other language.<sup>115</sup> In line with Myers-Scotton's (1993b, 2002) MLF model, these belong to the category of content morphemes as they express semantic and pragmatic aspects and can both assign and receive thematic roles.<sup>116</sup> They follow the Morpheme order Principle as outlined in Chapter Two and revisited below:

#### *The Morpheme Order Principle*

*In ML + EL constituents of singly occurring EL lexemes and any number of ML morphemes, surface morpheme order (reflecting surface syntactic relations) will be that of the ML.*

Furthermore, these insertions can be classified in several ways: they can be classified according to the word classes to which they belong and they can be classified according to their syntactic category (Nortier 1990: 140).<sup>117</sup> In Tuc's (2003) well-written account of Vietnamese and English code switching, he gives the following examples of single noun switches (which account for more than half of his data):

---

For example, they consider the potential effect of speaking casually (e.g. *hey, dude*) versus much more formally (e.g. *excuse me, sir*) (Myers-Scotton 2002: 23).

<sup>115</sup> I do not distinguish between borrowings and code switched singly occurring lexical items as they are part of the same mechanism. Furthermore, Myers-Scotton's (1993b, 2002) MLF and Uniform Structure Principle offer parsimonious accounts that do more than incorporate single lexical insertions.

<sup>116</sup> A further analysis of single word switches as well as other data appears in Chapter six.

<sup>117</sup> In Nortier's (1990) analysis of Dutch and Moroccan Arabic code switching, 71% of data collated were of single noun switches. Also, in Tuc's (2003) analysis of Vietnamese and English code switching, 50.61% of the recorded switches were of single noun insertions (where names of persons, places, streets and buildings were excluded). This large quantum of switched nouns conforms with other major research on bilingual studies (Poplack 1981, Berk-Seligson 1986, Trefferse-Daller 1991).

(114) Nó là *comedy*  
It is comedy  
'It is a comedy'  
(*English / Vietnamese*, Tuc, 2003: 57)<sup>118</sup>

(115) Có hai *group*  
Have two group  
'There are two groups'  
(*English / Vietnamese*, Tuc, 2003: 57)

The above nouns 'comedy' and 'group' are in the singular form as Vietnamese nouns only take the singular form. Another example is cited in Nortier (1990):

(116) ġadi ykun ġnd-ha                    *straf*  
Will it is have POSS FEM punishment  
'She will be punished'  
(*Moroccan Arabic / Dutch*, Nortier, 1990: 141)

In the above example of Nortier's (1990) data, single word switches account for the majority of switched data. In the Moroccan Arabic / English data presented in this thesis, we can safely assert that the majority of the switched data are that of single inserted lexemes. However, in the vein of *Reactive Syntax*, the difference that researchers have not pointed out is the difference between singly inserted lexical items by speakers of different generations. The composite switches and items switched in the first generation for example, are of a generic nature and the 'expected' choice. However, with the second but particularly the third generation inserted nouns in innovative combinations and in clauses where English is the matrix language occur where the inserted single noun is in Arabic:

---

<sup>118</sup> Tuc (2003) states: "Like a general typological characteristic of East and Southeast Asian languages such as Chinese, Thai, Khmer, Hmong, the salient feature of Vietnamese nouns is that they do not in themselves contain any notion of number or count. As the Vietnamese noun has no obligatory marking of singular or plural, and simply has the property of transnumerality, it is invariant in form. The nouns themselves remain the same regardless of whether they are singular or plural. In this respect, they are all singular" (2003: 56-57).

(117) He's such a *ħmār*  
He's such a donkey  
'He's such a donkey'

Amongst first generation speakers, this would be uttered in (a) uniformly in Moroccan Arabic or (b) uniformly in English but not as a single noun insertion as shown below:

(118) *huwwa ħmār*  
He's a donkey  
'He's a donkey'

Or:

(119) He's such a donkey

Example (117) above is important as it sets out novel way of speaking, *Reactive Syntax*, amongst a certain generation of Moroccan Arabic speakers. It must be noted however, that this is not an issue regarding fluency or language shift as the singly inserted items are all known in the counter-language, Moroccan Arabic. It is simply a question of language choice and insertion. As Myers-Scotton (2002) states:

If they [speakers] do engage in codeswitching or other forms of contact language, they will have to select-again generally unconsciously - a Matrix language to provide morphosyntactic structure for bilingual speech (2002: 23).

British-born Moroccan Arabic speakers choose when and when they use this composite *Reactive Syntax* in bilingual discourse. Their speaker audience is of course important as they have to be in surroundings where there is at least one other bilingual speaker. This form of speech does not affect intelligibility as first generation speakers are more than comfortable and understand this discourse. Ziamari (2007) describes how "informants use amongst themselves a different, more developed version" (2007: 276) in analysing code switching amongst younger generations of speakers and that "they [speakers] practice [sic] linguistic innovations including lexical creations and

prosodic markers which are peculiar to their group, thus excluding adult groups. We can clearly see how the MLF model in its basic form caters for this *Reactive Syntax* and a wide range of data. The following examples convey more of second and third generation speakers:

(120) He gave her a *saqla* and then she left  
He gave her a slap and then she left  
'He slapped her and then she left'

Compare (120) above uttered by a third generation speaker with that of a first generation speaker below:

(121) *sqal -ha u mšēt*  
Slap PAST her and left FEM  
'He slapped her and she left'

There would not be any single insertion as in (120) above in analysis of first generation Moroccan Arabic speakers who, after analysis and introspection, favour uniform Moroccan Arabic or uniform English CPs or longer clauses. We can safely predict therefore that:

*Generalization 4:*

*Amongst singly occurring Moroccan Arabic lexical insertions where the Leader is English and provides the morpho-syntactic frame; these will mainly be from second and third generation Moroccan Arabic speakers.<sup>119</sup>*

We can observe then the insertion of non culturally-specific terms is an innovation in code-switched analysis particularly as Moroccan Arabic insertions in non-Arabic languages is less common than insertions into Moroccan Arabic

---

<sup>119</sup> By single lexical items, it disregards words of religious or cultural affiliation as there is no English counterpart and so the Arabic must be used. These are words such as *inshaAllah* 'God willing' and the like.

(according to Dutch / Moroccan Arabic data collated by Boumans 1998).<sup>120</sup> Boumans gives the following examples:

- (122) *ik had hizb sabbih helemaal geleerd*  
I had hizb praise entirely learnt  
'I had memorized all of hizb sabbih'  
(Dutch / Moroccan Arabic, Boumans, 1998: 302)

In (122) above, *hizb* refers to a part of the Qur'an and of course there is no Dutch equivalent so we expect this to be the natural choice for the speaker. We see that Dutch is the matrix language and in line with the Uniform Structure Principle, the contributing MA noun does little to detract from the Dutch grammatical frame. It would have been useful if researchers of Moroccan Arabic code switching had approached this research in terms of comparing and contrasting different generations of speakers (Bentahila & Davies 1986, Boumans 1998, Nortier 1990). Boumans (1998) offers his own approach to intra-sentential code switching and is critical of Myers-Scotton's (1993b, 2002) MLF model. This is analysed in the next section commencing with his Monolingual Structure Approach (1998).

### 3.5 Boumans (1998) and the Monolingual Structure Approach (MSA)

In analysing Dutch and Moroccan Arabic intra-sentential code switching, Boumans (1998) presents an insertional model, namely the Monolingual Structure Approach. This in essence is not a new approach nor it is a new theoretical model. Rather, it

---

<sup>120</sup> In fact, Boumans (1998) states that: "The insertion of MA nouns, adjectives and nominal constituents in Dutch matrix structures yields a very different picture. In the first place, MA insertions are far less numerous. For this reason it is not possible to investigate the distribution of MA nouns in Dutch matrices in detail. Yet with respect to the available examples we can observe some striking qualitative differences between Dutch and MA embedded elements. MA insertions turn out to be more associated with certain identifiable factors, namely culturally specific vocabulary, repetition in discourse, and topicality" (1998: 301). It would have been useful if Boumans had analysed the MA insertions in terms of differences in Moroccan generations and produced a qualitative report. Boumans (1998) goes on to imply how the MA insertions are only of culturally or Islamic-specific terms which are all nouns or fixed expressions. I can agree that the majority are nouns, but as we have seen and will see in Chapter six, the majority are by no means culture-specific and the focus here is on an innovative generation of bilingual speech.

builds on previous insertional models and theories and actually bases itself on the MLF model (1993b). This is attested by Boumans (1996) in his own words:

I have simplified and reformulated the model (the MLF) to make it serve as a simple format for the description of actual attested data (1996: 46).

The main premise in his study in 1998 is that “the Monolingual Structure Approach assumes that each matrix structure originates in the grammar of only one language” (1998: 89). This is explored in full below followed by counter-examples to the MLF model, explanations and conclusions.

### **3.5.1 The MSA and the Matrix Language**

The MSA is not inherently a new concept nor can it claim to be. Boumans in attempting to differentiate it from other models and approaches to intra-sentential code switching defines the MSA in the following terms: “Each matrix structure is assumed to originate in the grammar of only one language, the Matrix language” (1998: 61). The main focus then of this model is that of the matrix language and this is an essential part of this model in that the internal make-up of the matrix structure must be entirely attributable to the ML grammar (1998: 89). In defining the matrix language which is a key aspect of this model, Boumans states that:

The concept of a Matrix language stems from the assertion that a grammatical structure containing elements from two languages can be attributed to the grammar of one of these languages (the ML), rather than to the grammar of both languages, to the overlap of both grammars or to a third ‘codeswitching’ grammar (1998: 61).<sup>121</sup>

---

<sup>121</sup> Boumans (1998) goes on to describe how “languages in contact with each other are seldom equal in status. In many cases one of the languages can be characterised as the bilingual community’s ‘own’ language, whereas the other language in use is imposed by an economically and/or culturally dominant speech community. For this broad generalisation I use the terms Community Language and Superimposed Language. The unequal status of the languages involved is reflected in the patterns of intra-sentential CS” (1998: 62). This broad definition of two different sets of language types is too demarcated, as language situations are seldom as clearly set out. In data collated for this thesis, and other data presented throughout, it is clear that bilingual speakers can change languages mid morpheme and not only dependent upon social factors such as in-group or out-group associations.

Boumans describes the matrix language as follows: “The matrix language is a device to describe grammatical structures containing morphemes from more than one language” (1998: 65).<sup>122</sup> It is agreed that identification of a matrix language in mixed clauses facilitates the description of code-switched data. The following examples show Moroccan Arabic as the matrix language with French as the embedded variety:

(123) *les restaurants meħlul-in hetta l l- waħd-a d l lil*  
 DEF-PL restaurant open-PL until 11 to DEF-one of DEF-night  
 ‘The restaurants are open until one in the morning’  
 (Moroccan Arabic / *French*, Slaoui, 1986 Annexe I: 60)

(124) *ta-n-fi-q-u le matin, ka-n-šerb-u atay nigrū*  
 Wake-up-3PL DEF-M morning drink-3PL tea black  
 ‘We wake up in the morning and drink black tea’  
 (Moroccan Arabic / *French*, Slaoui, 1986 Annexe VIII: 12)

The above examples clearly show Moroccan Arabic as the matrix language, whereby in (123) it shows how it is possible to switch between the French noun ‘*les restaurants*’ and the predicative adjective ‘*meħlul-in*’ which matches the features specifications of plurality and masculine gender. This then fits well with the MSA and all major theories of code switching, linear, government and asymmetric models. The MSA in essence has two types of matrix structures proposed in this model; that of finite clauses and that of constituent insertion.

<sup>122</sup> This model is clearly based on the MLF with slight adjustments. Boumans (1998) in further explaining his concept of the matrix language that he is “concerned with two types of matrix structure, namely the finite clause and clause constituents headed by one of the major word categories Noun, Verb, Adjective, Adverb and Preposition. I will use the term matrix structure to refer to both types at the same time. The notion of the ML attributes the morphosyntactic properties of the matrix structure to the grammar of one of the contributing languages rather than to both languages. The embedded language only contributes embedded elements; that is it does not participate in creating the matrix structure” (1998: 65). This then encapsulates the meaning of the matrix language and its structure in the MSA. However, unfortunately, Boumans (1998) continues to cite out-moded definitions of Myers-Scotton’s matrix language which previously (and has since been modified and corrected) whereby she states that the matrix language is “that the ML is the language of more morphemes in interaction types including intrasentential codeswitching” (1993b: 68). Boumans (1998) states that: “This quantitative criterion will not work either” (1998: 74).

### 3.5.2 The MSA and Constituent Insertion<sup>123</sup>

In describing the asymmetrical nature of the MSA model, Boumans employs almost identical terminology to that of Myers-Scotton (1993b, 2002). Terms such as matrix language, embedded language, content words and system or function words. However, he adds slight adjustments to the MLF model:

One of the major tasks of the MSA and of matrix language approaches generally is the classification of EL material that is inserted in mixed constituents. Is it content morphemes or rather content words that are embedded? The examples just discussed show that the idea of content morpheme insertion as advanced in Myers-Scotton's MLF model is too restricted to describe the attested insertion patterns. Furthermore, embedded compound words cannot be explained by content morpheme insertion alone, unless the ML and EL share the same structures of compound words. 'Content word' instead of 'content morpheme' is a broader term that covers the attested derived and inflected forms as well as EL compounds. However, 'content word' is a rather indiscriminate expression introducing its own problems of demarcation (1998: 68-69).

This is rather a case of splitting hairs as the innovative analysis of such an asymmetric model was pioneered by Joshi (1985) and progressed by Myers-Scotton (1993b, 2002) in a grounded analysis of bilingual data.<sup>124</sup> A definition of the constituent level according to the MSA is as follows:

*Boumans (1998) constituent level*

*On the constituent level, the ML is the language to which the internal structure of the constituent as expressed by the distribution of all morphemes within the constituent can be attributed. The distribution of a morpheme concerns both*

---

<sup>123</sup> Boumans (1998) describes constituents as units in the analysis of sentences as hierarchical structures. Constituent classes, such as DET, N, NP, PP etc., are identified primarily by distributional facts. Secondary considerations include semantic coherence, prosodic evidence, facts about the distribution of pronouns. Boumans continues to state that "No criterion for constituency is entirely unproblematic and that alternative analyses of syntactic structures are possible and one can often opt for a more refined classification of constituents that takes smaller distributional differences into account" (1998: 69). This is clearly evident.

<sup>124</sup> Boumans (1998) describes content and function morphemes as complementary categories using Myers-Scotton's terminology but states that "While the terminology used here has clear affinity with Myers-Scotton's terms 'content morpheme' and 'system morpheme', remember that the latter terms have very specific meanings in Myers-Scotton's MLF model" (1998: 70).



departs from the fundamental aspects of the MLF model. The second type of constituent insertion as per the MSA is that of finite insertion.

### 3.5.2 The MSA and the Finite Clause

Based on Levelt's (1989) language production model, Boumans describes how the finite clause can itself be a constituent within another clause as is the case with complement clauses or relative clauses.<sup>125</sup> In further analysis of the matrix language, he states that in the case of the finite clause there is fortunately a suitable independent criterion: the verbal inflection, or perhaps the more precisely inflection for tense, is probably the best indicator of the ML (1998: 76) and goes on to state that the same language that provides the inflection of the tensed verb also organizes the relative order of the verb and its arguments.<sup>126</sup> The below sets out how the MSA adopts Klavans' (1985) definition of the base language on the finite clause level:

*Boumans (1998) finite clause:*

*The Matrix Language (ML) on the sentence level is the language of the inflection bearing element of the tensed verb.*

Boumans is not alone in describing the language of the ML as that of the inflection-bearing element. Indeed, several researchers prior to the MSA make a similar linguistic claim (Treffers-Daller 1994). Cited in Boumans, Moyer (1995) also points out that:

---

<sup>125</sup> Levelt (1989) describes the distinction between basic and finite clauses as "A basic clause contains one and only one main verb. There is evidence that the basic clause is the primary planning unit in speech production and on this level grammatical functions are assigned to arguments. But the finite clause "concerns the ordering aspect of grammatical encoding; it reflects which categorical procedures do the word and phrase ordering of the retrieved functional information" (Levelt 1989: 256) cited in Boumans (1998: 73).

<sup>126</sup> This has been proposed by researchers over the last few decades (Wentz 1977, Pandit 1986) and once again shows the MSA as a repetition of previous models and theories. See also Klavans (1985) who describes the inflection-bearing element of the verb as indicative of the ML.

Ouhalla's (1991) functional categories proposal does provide theoretical support for Klavans' view of INFL as an important category in code-switching. According to Ouhalla the ordering of TNS and AGR categories with respect to each other is shown to be ultimately responsible for word order variation with across languages (1995: 194-5).

Below is an example of a third generation Moroccan Arabic speaker. Notice the language of the inflection-bearing element:

(129) *He's jib- ing me some chocolates beš na- put on some weight.*  
He's bring-GER me some chocolates so that put on some weight 1SG  
'He's bringing me some chocolates so I can put on some weight'

The above example is an interesting case in point as the inflection is carried on the Moroccan Arabic verb *jib* with the verb inflected in English -ing to enable the merged MA verb plus GER '*bringing*.'<sup>127</sup> The above clearly follows English syntax yet the inflection element is in Moroccan Arabic '*jib*' and it is this fusion of morphemes from two separate languages which provides an interesting database for further research. However, Boumans states that "It must be emphasised that it is the verbal inflection that correlates with the word order, and not the lexical verb itself" (1998: 76). In order to illustrate this, Boumans gives the following example in Moroccan Arabic and Dutch where the order Verb-Subject in a declarative sentence is in accordance with Moroccan Arabic syntax and not Dutch:

(130) *ṣta-ha-ni de buurman*  
Give-3F-1SG DEF neighbour  
'The neighbour gave it to me'  
(Moroccan Arabic / Dutch, Hocine)

---

<sup>127</sup> This example contradicts Myers-Scotton's (2002) list of Matrix Language points where we have a bi-lingual CP unit combining both a Moroccan Arabic verb and an English suffix. Myers-Scotton states that: "However, there is also evidence that when speakers are nearly equal at home in both languages, almost ironically, Embedded Language islands lose their importance. Instead, switching between CPs becomes very frequent as well as switching between sentences, which of course may include more than one CP" (2002: 149). Backus (1996) in his analysis of Turkish and Dutch intra-sentential code switching makes the point that: "At some point, intra-sentential CS takes over" (1996: 334). Ziamari (2007 and also personal communication) describes such code switching as a constantly evolving phenomenon with the creation of new syntactic structures.

As discussed earlier, NP insertion is one of the easiest of insertions and is employed as a code switching mechanism and strategy by bilingual speakers. This is amply covered in the MSA. This model then as described by Myers-Scotton (2002) is a test of earlier models and in this case I am inclined to agree. The MSA makes identification of the matrix language a focal point of its model and also describes layered insertions of morphemes and how this analysis offers a solution for seemingly problematic cases of function morpheme insertion in a variety of other language pairs (1998: 80). The main point of departure between the MLF and MSA models (although the MSA is built on the MLF) is that the former goes beyond simple syntactic explanations for code switching and is an all-encompassing model which suits a wide variety of data. The MSA however proposes a purely syntactic approach to code switching focusing on constituent insertion and its morpho-syntactic placement within the matrix frame. Both models deal with the asymmetry between a matrix and an embedded variety with the MSA advocating that both languages are activated to the same degree. The MLF however states that activation is triggered at different times. Also, the content morpheme – system morpheme dichotomy which is a key concept in the MLF is not a salient part of the MSA (although Boumans uses the same terms). Boumans Moroccan Arabic and Dutch research provides a good data for further research in particular with the data presented on lexical insertion. The following section deals with such constituent insertions.

### **3.6 Lexical Insertions**

Lexical insertion<sup>128</sup> is one of the most commonly inserted constituents in code switching with verb insertion quickly gaining momentum particularly amongst

---

<sup>128</sup> One of the controversies in the study of CS is the treatment of single-item insertion. Poplack and her associates argue that “lone other-language items” insertion is ‘borrowing’ and should be distinguished from longer stretches of switches, which they define as code-switching. They propose

evolving speech styles in younger generations today. Pennington (1998b) describes 'lexical bilingualism' to describe this pattern of code-mixing and suggests that:

The knowledge of English by Hong Kong Chinese is more a matter of familiarity with a certain number of words and phrases than fluency in a second language (1998b: 9).

Such insertion is described as 'layered' insertion in code switched varieties with evidence of such layering found in Moroccan Arabic and French (Abbassi 1977, Bentahila & Davies 1983, Slaoui 1986, Nait M'Barek & Sankoff 1988, Lahlou 1991, Wernitz 1993). The following examples illustrate layered insertion with [131] providing an example of an embedded Arabic noun:

- (131) *attention il ne faut pas chang- er t- tewşil*  
Caution 3-SG-M NEG should NEG change-INF DEF receipt  
'Be careful not to change the receipt'  
(Moroccan Arabic / French, Bentahila & Davies, 1991: 383)

Insertions within insertions are also frequently uttered by near-fluent or fluent bilinguals and the structures can become quite complicated:

---

that if other language items are morphosyntactically integrated into the recipient language, it [unclear what 'it' refers to] is identified as lexical borrowing. If not, it is a case of CS. They further set a continuum of lexical borrowing. 'Established loan words' which "typically show full linguistic integration, native-language synonym displacement, and widespread diffusion, even among recipient-language monolinguals" (Poplack & Meechan, 1995: 200) are on the one end. On the other end is 'nonce borrowing' which just satisfies the criterion of morphosyntactical integration. Other researchers (Myers-Scotton, 1993; Bentahila & Davies (1983); Treffers-Daller, 1994) do not distinguish lexical borrowing and CS as different processes. Myers-Scotton (1993a) argues that "B forms and singly occurring CS forms undergo ML morphosyntactic procedures in the same way" (1993: 206). However, "the *lexical entries* (original emphasis) of CS and B forms must be different, since B forms become part of the mental lexicon of the ML, while CS forms do not" (1993: 163). She further divides lexical borrowings into cultural borrowings and core borrowings. Cultural borrowings are "words for objects and concepts new to the culture" (2002: 41). They often fill gaps in the recipient language (1993a: 206) and may appear in the monolingual speech of either bilinguals or monolinguals, or in the codeswitching of bilinguals (2002: 41). Core borrowings are "words that more or less duplicate already existing words in the? L1" (2002: 41). Myers-Scotton argues that core borrowed forms typically enter the recipient language gradually through code-switching (2002: 41), whereas cultural borrowed forms appear abruptly "for the obvious reason that they are needed to fill gaps" (1993a: 206). She proposes frequency as the criterion for distinguishing between CS and lexical borrowings. She predicts that culturally borrowed forms will show high relative frequency and core borrowed forms will show high frequency compared to CS forms.

- (132) *tu perds wahəd l- demi heure*  
You lose INDEF DEF half hour  
'You lose half an hour'  
(Moroccan Arabic / *French*, Bentahila & Davies, 1991: 383)

Boumans (1998) describes the above (in the MSA) as a French finite clause with an inserted Arabic NP (*wahəd l-demi heure*) in which a French content word is inserted (*demi-heure*) is inserted. Conceptually, in the MLF model, *demi-heure* is seen as EL collocation and many Embedded Language islands are adverbial phrases of time or place; that is, they are adjuncts. This means that they are outside the predicate-argument structure projected by the main clause verb (Myers-Scotton 2002: 141). Further Backus (199b) describes EL islands or collocations as 'chunks' as they are very formulaic in nature. The concept of layered insertion then is anathema to the MLF model due to the acute asymmetry of the ML and EL dynamic where the Embedded Language is 'on' to some extent within the bilingual CP. Muysken (2000), in analysing the juxtaposition of two languages in a single clause introduced the Adjacency Principle:

*Adjacency Principle:*

*If in a code-mixed sentence, two adjacent elements are drawn from the same language, an analysis is preferred in which at some level of representation (syntax, processing) these elements also form a unit (Muysken 2000: 61).*

The argument against the above Adjacency Principle is that not all insertions or EL islands are singly occurring units. Myers-Scotton argues that singly occurring Embedded Language forms are projected differently from the mental lexicon than are full maximal projections (e.g. NPs) (2002: 144). She cites the following reason for this conclusion:

Singly occurring Embedded Language forms require less checking and undergo less processing at all levels of abstract grammatical structure than do Embedded Language phrases (phrases that will include grammatical as well as lexical elements). Singly occurring nouns simply must be sufficiently congruent with a Matrix Language morpheme counterpart at the level of lexical-conceptual structure to occur. (If there is not a counterpart, then under the Uniform Structure Principle--even this congruence is not necessary as long as the Embedded Language nouns follow the relevant language's procedures regarding potential Matrix Language predicate-argument structure and morphological realization patterns.) In the case of verbs, congruence may be more difficult to establish with a Matrix Language counterpart, since predicate-argument structure is at issue too (2002: 144).

Why do such insertions occur if the speakers are in essence near or fluent bilinguals? At the point of spell-out, dependent upon subject matter, speaker intention, fluency and proficiency, code-switchers are in essence at liberty to lexically insert nouns or noun phrases as well as verb phrases and prepositional phrases as long as grammaticality is maintained. Further, within the rubric of the MLF, such Embedded Language islands occur because of "structural mismatches, and in some data sets they are as frequent as pragmatically motivated islands, if not more so" (2002: 146). The following example as cited by Myers-Scotton is shown below:

- (133) *hunak binikhi aktar li?annu we get in the mood*<sup>129</sup>  
 IP / IMP / speak more because / IP we get in the mood  
 'There we speak more because we get in the mood'  
 (Arabic / *English*, Okasha 1996 corpus, cited Jake & Myers-Scotton,  
 1997a: 30)

The indefinite composite article *wahad* and *ši* are very commonly inserted in bilingual clauses and these are usually followed by the definite marker 'l':

- (134) *ka-t-gul wahad l- žeww*  
 Say 2Sg INDEF / DEF atmosphere]  
 'You say it's a certain atmosphere'  
 (Moroccan Arabic, Boumans, 1998: 185)

<sup>129</sup> The original as cited in Myers-Scotton (2002) *binikhi* should actually read *binihki* – this is an error in the Arabic transliteration.

- (135) *mšī-t ʔemma wahəd l nhār wa ma laqit-u- š* so I came back  
 Went there INDEF DEF day and NEG 1 find him NEG 2 so I came back  
 ‘I went there one day and didn’t find him so I came back’

Also, the indefinite composite *šī wahəd* ‘someone’ below is an oft-used structure:

- (136) *šī wahəd ja ou ma open-i-ti – š al bāb?*  
 INDEF one came and NEG 1 open 2SG NEG 2 DEF door  
 ‘Someone came and you didn’t open the door?’

- (137) *ya-t-haka-w wahəd les histories*  
 tell 3PL INDEF DEF / PL stories  
 ‘They tell each other some fantastic stories’  
 (Algerian Arabic / French, Boumans & Caubet, 2000: 154)

In the MLF, (137) is described as a French NP preceded by Algerian Arabic *wahəd* and this is considered as an Embedded Language island under this model. *Les histories* above as a single unit is an internal Embedded Language island under a node with *wahəd* as its highest left branch).<sup>130</sup> Moroccan Arabic requires a ‘determiner complex’ before a noun when a demonstrative is used as in *wahəd* or *ʔak* ‘that’ with the complex consisting of two elements, the demonstrative and a determiner (2002: 14). Nortier also notes that “When the article is used at all it is always in Moroccan Arabic” (1991: 200-1).<sup>131</sup> Therefore, the data supports the premise of the MLF in line with the matrix language – embedded language dynamic. Embedded language islands are full constituents consisting of embedded language morphemes occurring in a bilingual CP with the embedded language island showing structural dependency relations. Minimally, there can be two content morphemes (e.g. noun and modifier) or a content morpheme and a non-derivational system morpheme (2002: 139). Certainly, NPs although frequently inserted are not the only lexical categories which are easily embedded as islands into a syntactic frame. VPs are also frequently inserted,

<sup>130</sup> Myers-Scotton (2002: 115).

<sup>131</sup> Refer to Chapter Six for a further analysis of Moroccan Arabic determiners which contrasts the use with bare forms.

particularly intra-sententially and the combinations are striking as they are interesting for further research.

### 3.6.1 VPs as Lexical Insertions

Evolution of communication and language is producing new data sets which are the focus of this thesis, in particular the way in which Moroccan Arabic and English bilingual speakers are using more novel and innovative ways of stringing discourse where embedded language verbs receive matrix language inflections with ease. The MLF has not had much exposure to such new structures, therefore we cannot apply the MLF in its totality to the structures below, but use the MLF in its basic form to describe matrix and embedded forms. Myers-Scotton states that: “Admittedly, there are also language pairs where very few Embedded Language verbs appear inflected with Matrix Language affixes” (2002: 138).<sup>132</sup> A wealth of data presented in this thesis then will add to the way in which we examine intra-sentential code switching and in particular, the juxtaposition of matrix and embedded combinations currently bereft in the literature. The examples below show English verbs with Moroccan Arabic inflections where the direction of the discourse is English verb + MA inflection:<sup>133</sup>

(138) *iwa* close – *i* – *ha*      *u* *jib-i*      *li*      *l*      *maftēh*  
Well close TNS it FEM and bring 2SG to me DEF key  
'Well close it [the door] and bring me the key'

---

<sup>132</sup> Myers-Scotton also states that there is some evidence that simple familiarity with codeswitching as a medium of communication may result in new structures developing that were not present in earlier data sets. This comment applies especially to the treatment of verbs (2002: 138).

<sup>133</sup> My data and close examination shows that the present tense is used far more than any other tense variety. Also, when the perfect tense is employed, regular verbs are only use in order to avoid irregular verbal stems and in so doing restrict odd-sounding verbs such as “*do-it ha*” ‘I did it (feminine). The only exception is with the verb ‘make’ where verbal combinations such as “*make-it ha*” I made it (feminine) are often used. This code switching strategy is a new feature (see also Caubet 1993) and will be further examined in Chapter Six. (The periphrastic ‘Do’ construction is not analysed in this Chapter as it will be covered in later Chapters).

(139) tell *-i - h* *y-zid li ya* *šwiya dyal* sugar, it's too strong!  
Tell 2SG MASC add to me little of sugar, it's too strong!  
'Tell him to add some sugar, it's too strong!'

(140) shut *-i -h* *walla na-shut-i-h* for you  
Shut 2SG it MASC or shut 1SG it MASC for you  
'Shut it or I will shut it for you'

In Nortier's (1990) Dutch and Moroccan Arabic data analysis, she cites only one example in her entire data set where a similar variety is recorded:

(141) dak *š - š* *lli t-* *bezig fi- h*  
That the thing that you IMPF busy with it  
'That thing that you are busy with'  
(*Dutch / Moroccan Arabic*, Nortier 1990: 144)

However, Nortier states that;

This is the only recorded example of simultaneous syntactic and morphological integration in my entire data collection. Although I have heard other instances of this phenomenon in non-recorded conversations between bilingual Moroccans, I will give it the status of an exception. (1990: 144).

It is certainly not the exception amongst bilingual Moroccan Arabic speakers who merge structures with any other language, French, English, Spanish or Dutch (witnessed personally). In fact, verbal stems with inflections from L2 would be a useful study for the future of code switching in terms of syntax, morphology, psycholinguistics as well as sociolinguistic analysis. Boumans (1998) also describes the paucity of such constructions in his data:

The *dar* plus infinitive construction is by far the most common way to insert Dutch verbs in MA clauses. However, Dutch infinitives and verb stems are occasionally inserted without the MA verb *dar* carrying the inflectional affixes. These insertions are not very frequent in the corpus and tend to be accompanied by pauses and hesitations (1998: 259).

Pauses and hesitations were not evident in the data and in fact, speakers use this code-switching variety with such alacrity that it can be considered an aspect of their fluency

and proficiency. Further examples of English verb + MA inflection are conveyed

below:<sup>134</sup>

Figure 3.4: Embedded verbal inflections in English verb + Moroccan Arabic inflections

	English stem +	Present Tense	Past Tense	Moroccan Arabic suffix (Masc or Fem)	Combinatorial result	Gloss
1	Shut	<i>i</i>	-	<i>i</i> (M)	Shut- <i>i-ih</i>	You (F) Shut it
2	Bring	<i>i</i>	-	<i>ih</i> (M)	Bring- <i>ih</i>	You (F) Bring it
3	Make	-	<i>it</i>	<i>ha</i> (F)	Make- <i>it-ha</i>	She made it
4	Eat	<i>i</i>		<i>ha</i> (F)	Eat- <i>i-ha</i>	You (F) Eat it

The above (without including periphrastic *Dar+* constructions) the most commonly employed verb + inflection combinations and interestingly, mainly amongst first generation of bilingual speakers.<sup>135</sup> In the case of changing the direction where verbs inserted are Moroccan Arabic stems + English inflection, this is also recorded although it appears not to be as prevalent as the data above states amongst first generation speakers:

(142) She's *seth* -ing by herself  
 She's dance GER by herself  
 'She's dancing by herself'

(143) I'm not *kūl* -ing it, it looks gross  
 I'm not eat GER it, it looks gross  
 'I'm not eating it, it looks gross'

<sup>134</sup> See Chapter Five for elicitation methods, data collection and methodology in general.

<sup>135</sup> Data also shows that this code-switching strategy is also used amongst second and third generation Moroccan Arabic / English speakers but not to the same extent.

As discussed in Chapter Two, the Moroccan Arabic + English gerund structure as shown above in (142) and (143) is a phenomenon used amongst second and third generation bi-lingual speakers only (example of reactive syntax) and not amongst the first generation. No other syntactic combinations have been recorded or witnessed thus far. Therefore, recapitulating on this section, this section shows an innovation of a new speech style with new structures and verbal stems. The data presented does not counter the MLF as the Uniform Structure Principle is maintained throughout however, a readjustment of the model will be necessary in order to incorporate evolving syntactic structures and the addition of embedded inflections.

### **3.7 Conclusion**

In this chapter I have provided a detailed and thorough analysis of Myers-Scotton's (1993b, 2002) MLF model in light of different data sets together with the principal Moroccan Arabic and English code-switched discourse. The MLF as a theoretical notion has been considered together with the integral meaning of what constitutes a matrix language which is a key concept in this asymmetrical approach. I have selected the MLF to present my data in light of the evidence presented in Chapter Two whereby other theories and approaches, through micro analysis, were deemed too stringent or had too many counter examples and did not suit the data presented in this thesis. Early accounts of asymmetric approaches have been considered (Wentz 1977; Sridhar & Sridhar 1980; Joshi 1981, 1982, 1985; Pandit 1986 and Petersen 1988) as well as an analysis of Levelt's (1989) Speech model which is an integral aspect of linguistic analysis. I analyse the structure of the MLF together with the four main principles and these are shown to be structurally sound within the data presented. This is followed by an examination a major aspect of the MLF which is that of the Matrix Language and Embedded Language dynamic and this is portrayed by the salient features of content morphemes and system morphemes, which are analysed in the two

further supporting models, namely, the 4-Model and the Abstract Level model. These two models provide a more explicit account of the MLF and provide principled explanations of the abstract morpho-syntactic frame that structures bilingual clauses in bilingual intra-sentential discourse based on the premises of the mental lexicon.

The Monolingual Structure Approach is examined and is regarded as an interesting insertional approach and this model is compared with that of the MLF. Certain generalizations are made in this chapter which will be further examined in Chapters Six and Seven. Finally, lexical insertions form the concluding part of this chapter where noun insertion is highlighted followed by innovative structures also known as *Reactive Syntax*, as are used by second and third generation speakers in natural Moroccan Arabic and English discourse. The form of Moroccan Arabic, morphology and root patterns are discussed in the following Chapter in order to gain a lucid and tangible grasp of inflectional insertions and verbal derivations in the code-switched data presented.

## CHAPTER FOUR

### MOROCCAN ARABIC GRAMMAR

This chapter deals with Moroccan Arabic grammar as a spoken oral dialect known as *Darija* and used by the inhabitants of Morocco and the Moroccan diaspora. There is no single Moroccan Arabic dialect in Morocco as a single homogenous variety but rather there are different dialects across the length and breadth of the Kingdom from Chaouen in the North, to Oujda in the East and south to the Moroccan Sahara.<sup>136</sup> There is a paucity of literature on Moroccan Arabic language with research focusing mainly on pedagogical texts (Abdel Massih 1970, Brunot 1950 and more recently Heath 1989 and Harrell 2004). A non-exhaustive account of Moroccan Arabic history will commence this chapter together with the co-existence of Berber, French and Spanish influences and integration. This is followed by a close analysis of Moroccan Arabic morphology based on the Arabic consonantal skeleton or root where word classes will be discussed.

A brief sketch of the main differences between Moroccan Arabic and Classical Arabic is outlined in order to illustrate major differences between the two varieties to better identify the placement of Moroccan Arabic as a regional dialect amongst Arabic dialects today. The first section of this chapter will be devoted to MA basic word order and syntax (Ennaji 1985) where several word orders are possible and these are discussed and evidenced in the data. Nominal sentences, nouns and noun insertion together with adjectives and their formation in monolingual MA and in code-switched intra-sentential varieties follow in section two where it will be shown that MA exhibits a frequent use of nominal sentences consisting of a subject (noun or pronoun)

---

<sup>136</sup> According to Abboud (1970) the major division of modern Arabic dialects is formed by a line running roughly from the western borders of Egypt to Chad, dividing them into Eastern and Western groups. The numerous Moroccan dialects constitute, together with similar dialects spoken in Algeria, Tunisia and Libya, the main part of the Western group. See also Heath (1989) on Moroccan Arabic dialects.

and a predicate (noun, pronoun, adjective, adverb, or prepositional phrase (Nortier 1990). Basic MA word categories are examined with the larger part of this chapter concerned with verbal clauses, relative clauses and negation. It is important to outline the grammaticalization of Moroccan Arabic as English insertions and inflectional affixation are made more transparent and lucid when basic structures are analysed and this facilitates the study of intra-sentential code switching. A more detailed analysis of Moroccan Arabic as a grammatical system follows in Chapters Six and Seven.

#### 4.1 Moroccan Arabic – An Historical Outline

The most commonly used languages in Morocco today are *Darija*, Classical Arabic (for media, religious sermons, the legal system, formal education and literature only and not as a contact language), Berber and French with Spanish used in the North (Chaouen, Tetouan, Ceuta and Melilla). MA belongs to the Semitic group of languages,<sup>137</sup> a family of languages spoken by over 300 million speakers across North Africa, the Middle East and the Horn of Africa, the most widely spoken of the Semitic languages being Arabic (150 million), Amharic (25 million), Tigrinya (6.7 million) and Hebrew (5 million).<sup>138</sup> According to Abbassi (1977), MA is the descendant of at least three Arabic dialects: (i) non-Bedouin dialects, (ii) Bedouin dialects, and (iii) Andalusian dialects (1977: 19-21). The first type was introduced in Morocco in the seventh and eighth centuries A.D by the first Arab conquerors who originated from urban centres of a conquered Middle East. Bedouin dialects were introduced by subsequent nomadic tribes, namely the Beni Hilal and Beni Salim who settled in Morocco around the beginning of the eleventh century. The Andalusian dialects were introduced by the refugees who fled Spain around the thirteenth century and these are

---

<sup>137</sup> Greenberg, J (1963) "Some universals of grammar with particular reference to the order of meaningful elements," in *Universals of Language*, pp. 73–113. Cambridge: MIT Press.

<sup>138</sup> cf. Versteegh (1997) for a comprehensive account of Semitic languages and Arabic in the Arab world.

also referred to as urban dialects since they developed in urban centres in Spain (Abbassi 1977: 21).

Figure 2: Map of Arabic-speaking peoples<sup>139</sup>



The language situation in Morocco today is fast-moving and non-static particularly given its diglossic situation<sup>140</sup> and number of koinés that have resulted from specific situations of contact between different types of MA with Berber varieties (El Aissati (1977: 26). Further, Caubet (1993) also discusses the radical changes in the language situation in Morocco in the last fifty years (post colonialisation), due mainly to factors such as rapid population growth, urban development and urban migration, mass media, education, contact with other languages mainly French (an official language) and Spanish not to mention Classical Arabic. According to Ennaji (1985), there are at least five dialects of MA today and these are spoken in Tangier, Oujda, Casablanca, Fez and Marrakech but Ennaji maintains that more can be added according to one's observation (1985: 5).<sup>141</sup> All

<sup>139</sup> [http://www.intersolinc.com/newsletters/newsletter\\_32.htm](http://www.intersolinc.com/newsletters/newsletter_32.htm)

<sup>140</sup> Ferguson, C. (1959a). 'Diglossia.' In *Word* 15: pp. 325-40.

<sup>141</sup> Heath (1989: 6-7) notes more dialects can be added if we take into account religion, namely Muslim and Jewish dialects. He states that Jewish dialects have been widely spoken by members of the Jewish community in some major cities of Morocco, like Fez and Casablanca, especially before the creation of

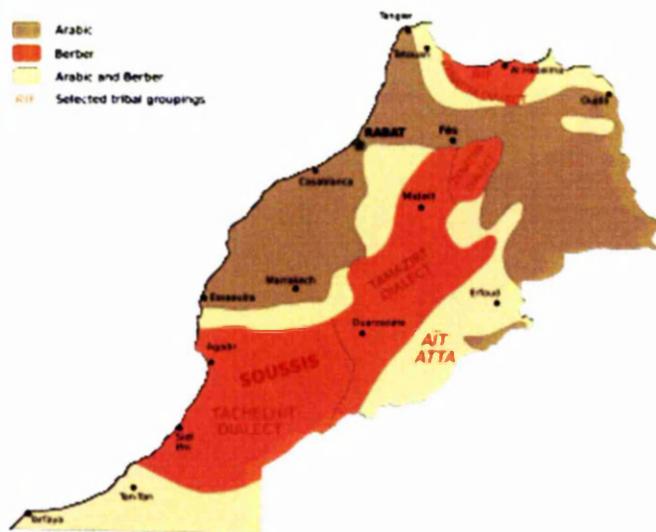
Arabic dialects in Morocco are mutually intelligible due to mass media, social interaction and migratory movement. However, there are striking differences in vocabulary which vary from region to region which can be identified at the lexical level. El Aissati (1997: 27) cites the following differences in lexical items where the term “dog” for example in Casablanca and Oujda is *kelb* whereas in Tangier it is *dyru*. Similarly, the equivalent of the verb “to look for” is *qelleb* in Casablanca and Tangier but *dewwer* in Oujda, and the word for “boy” is *weld* or *derri* in Casablanca and Oujda but *ḡayel* in Tangier. Some of the major differences in lexical categories lie in the historical connection and integration of Berber nouns and phrases, where Berber is still a major language today. Approximations have been cited for the number of speakers of Berber in Morocco today and is estimated to be used by a third of the population today (Abbassi 1977: 13), by 40% of the population (Youssi 1992) and by 50% of the population (Boukos 1995).<sup>142</sup>

---

the State of Israel which was an impetus for large numbers to emigrate to Israel. Some of the characteristics of Jewish dialects according to Heath are shown as follows. At the lexical level, these dialects can be identified by the use of the terms *ra* “to see” and *hebb* “to like” instead of *sāf* and *bḡa*, which are used in Muslim dialects. At the phonological level (although this thesis does not provide an account of Moroccan phonology due to space restrictions), Heath cites the merger of /š j/ with /s z/ as an identifying trait of Jewish dialects. This same merger however was pointed out by Heath (1989: 6) as a characteristic of the city of Meknes as a whole and not only a feature of a Moroccan Jewish dialect (El Aissati 1997: 26).

<sup>142</sup> Berber is enjoying a revival since the ascension to the throne by King Mohamed VI of Morocco who has reinstated it in school curricula. Berber is also broadcast daily on the national news, radio and used in literature.

Figure 3: Ethnolinguistic groups in Morocco<sup>143</sup>



There are three main Berber varieties in use today; Tashelhit, Tamazight and Tarifit with the first being spoken in the High Atlas, the Anti Atlas and the Sous valley; the second spoken in the Middle Atlas and part of the High Atlas and Tarifit spoken mostly in mountain and rural areas<sup>144</sup> but Moroccan Arabic is still clearly the most dominant and widely-used variety. The following sections deal with non-concatenative MA and an examination of its syntax and basic word order(s) and how this conforms with code-switching and in particular in line with the MLF model. We commence with MA morphology and general morphological notions.

#### 4.2 Moroccan Arabic Morphology and Basic Word Categories

One of the main distinguishing features of Semitic languages is their root-and-pattern morphology. The root is a semantic abstraction consisting of two, three or (less commonly) four consonants from which words are derived through the superimposition of templatic patterns (Holes 1995: 81). This consonantal skeleton

<sup>143</sup> <http://en.wikipedia.org/wiki/Morocco>

<sup>144</sup> Sadiqi, F. (1997). *Grammaire du Berbère*. L'Harmattan.

occurs in patterns with various vowels and additional, non-root consonants. The two concepts of root and pattern are fundamental to the structure of Arabic in general as well as Moroccan Arabic words.<sup>145</sup> The root usually has some fundamental kernel of meaning which is expanded or modified by the pattern (Harrell, 2004: 23). Arabic and Moroccan Arabic verbs have a complex morphology where they follow derived form patterns that stem from the consonantal skeleton, such verbs are capable of carrying a lot of semantic information depending on which Measure they stem from, vowel additions and consonant pattern type. The following example gives a basic outline of KTB, the tri-consonantal root for aspects or the concept of ‘writing’:

*Figure 4.1: Moroccan Arabic Verbal Stems*

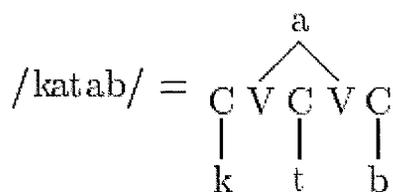
<b>Verbal Stem</b>	<b>Gloss</b>
<i>KTeB</i>	‘he wrote’
<i>KTBeT</i>	‘she wrote’
<i>meKTuB</i>	‘written’ or ‘pre-destined’
<i>KTaB</i>	‘book’
<i>KTuB</i>	‘books’
<i>maKTaB</i>	‘office’

McCarthy (1981) proposed a successful linguistic model for Arabic morphology under the framework of auto-segmental phonology (Goldsmith 1976) where a stem is represented is represented by three types of morphemes: root morphemes consist of consonants, vocalism morphemes consist of vowels, and

<sup>145</sup> Such a root-and-pattern phenomenon has become the prototype for the evaluation of the few non-linear morphological models which have emerged in recent years.

pattern morphemes are CV-skeleta; some stems include affix morphemes e.g. {*st*} in *staktab* (Measure 10) ‘to seek to write’. For example, the analysis of *katab* (Measure I) produces three morphemes: the root morpheme {*KTB*} ‘notion of writing’, the vocalism morpheme {*a*} ‘perfect – active’ and the pattern morpheme {CVCVC} – Measure I.<sup>146</sup>

1. Derivation of Measure I



Other derivations follow the same consonantal paradigm with a fixed and transparent shape rendering it easy to identify the both the Measure and semantic association from the shape of the stem. Moroccan Arabic follows the same verbal patterning.

**4.2.1 Moroccan Arabic and Verbal Derivations**

Following the Classical Arabic verbal stem, MA verbs are clearly defined by their shape and semantic content. The following Measures (Harrell 2004: 29) convey MA verbal Measures:

1. Measure I

Measure I is the most frequently encountered type of verb stem. The pattern differs for each of the different root types. For sound roots the pattern is *F*ξ*eL*:

*k**te**b*            ‘to write’

<sup>146</sup> Kiraz, G. (1994). *Computational Analyses of Arabic Morphology* (pp. 5).

*šreb* 'to drink'

*kber* 'to grow'

## 2. Measure II

All Measure II verbs have the common characteristic of a medial doubled consonant.

The pattern for strong roots, both sound and doubled, is *FeXeL*:

*beXeEd* 'to move aside, to make distant'

*beddel* 'to change'

*xeffef* 'to lighten'

## 3. Measure III

There are relatively few MA verbs which have the pattern of Measure III. The pattern for sound, double and middle-weak verbs is *FeXeL*:

*sāXef* 'to heed the advice of'

*qāda* 'to finish'

*jāweb* 'to answer'

## 4. Measures Ia, IIa(V) and IIIa(VI)

The stems of these Measures are derived by pre-fixing *tt-* or *t-* to the stems of Measures I, II and III. There is no distinction to be made for the different kinds of roots in these Measures. The single *t-* is always used in deriving Measures IIa(V) and IIIa(VI) from Measures II and III:

<i>ġetta</i>	‘to cover’	IIa(V)	<i>tġetta</i>	‘to be covered’
<i>sāfef</i>	‘to put in line	IIIa(VI)	<i>tšafef</i>	‘to line (oneself)’

Therefore, both the base form and derived forms in Moroccan Arabic follow the Measures as listed above.<sup>147</sup> The Moroccan Arabic dialect forms are not taught pedagogically, (as opposed to Classical / Standard Arabic which is taught at school) but rather the root types acquired naturally (cf. Universal Grammar) and given the Moroccan community in the UK, this is taught in the home, from parents to children who are brought up bilingually. Also, Moroccan children attend Arabic classes which take place at weekends and are now part of GCSE and ‘A’ Level curricula if chosen. Before commencing with Moroccan Arabic root types, the following section deals with Moroccan Arabic personal pronouns which come in two varieties; the first are independent personal pronouns and the second are dependent clitic pronouns.

#### 4.2.2 Moroccan Arabic Personal Pronouns

Two types of Moroccan Arabic personal pronouns have separate forms for the first, second and third persons as well as for first person singular and plural varieties. Both sets of pronouns distinguish between masculine and feminine gender in the third person singular and the independent personal pronouns have, in addition, separate masculine and feminine forms for the second person singular. The following Figure 4.2 shows the various forms MA personal pronouns can take:

*Figure 4.2: Moroccan Arabic Independent Pronouns*

<sup>147</sup> Moroccan Arabic has no verb Measure corresponding to measure IV of written Arabic. Measures VII, VIII, IX and X occur with only a very limited number of verbs. cf. Harrell (2004).

	Singular		Plural	
First Person	<i>ana</i>	'I'	<i>hna</i>	'we'
Second Person (M)	<i>nta</i>	'you' (masc)	<i>ntuma</i>	'you' (pl)
Second Person (F)	<i>nti</i>	'you' (fem)		
Third Person (M)	<i>ntiya</i>	'he'	<i>huma</i>	'they'
Third Person (F)	<i>huwwa</i>	'she'		

Figure 4.3: Moroccan Arabic Dependent Clitic Pronouns

	Singular		Plural	
First Person	<i>-i, -ya, -y, ni</i>	'me, my'	<i>-na</i>	'us, ours'
Second Person	<i>-ek, -k</i>	'you, your'	<i>-kom</i>	'you, your' (pl)
Third Person (M)	<i>-u, -h, -eh</i>	'his'	<i>-hom</i>	'them, their'
Third Person (F)	<i>-ha</i>	'hers'		

In stems to which suffixed pronouns are added, there is a phenomena of elision and inversion are encountered in stems which take the vowel endings e.g. *ktef* 'shoulder' and *ketfi* 'my shoulder' and *wekkel* 'he fed' and *wekklek* 'he fed you' (Harrell 2004: 135). Furthermore, Moroccan Arabic does not have a gender equivalent morpheme for the English 'it' as the Moroccan Arabic specifies the gender of the animate or inanimate item it refers to anaphorically. The following shows how MA always specifies for gender:

- (144) *ʕtiʔ -u al book u ma rədd -u š*  
 Gave him DEF book and NEG1 return it (MASC) NEG2  
 'I gave him the book and he didn't return it'

Example (144) above shows how book is masculine and the MA follows its masculine assignment of gender as 'it' is referred to using the MA masculine clitic pronoun 'u'. Had this been a feminine noun, 'ha' would have been employed to assign feminine gender.<sup>148</sup> Subject pronouns occur only when the verb is absent, or since Moroccan Arabic is a 0-subject language (Rizzi 1982) in cases of emphasis. Pronominal elements are expressed by the use of suffixes with clitics able to attach to verbs, nouns and prepositions (Nortier 1990). The following sets out some examples:

(145) *šeft -ha θamma*  
 'I saw her there'

(146) *xellit-ek la heder ta- learn-i lesson*  
 Left you DEF order learn 2SG lesson  
 'I left you so that you learn a lesson'

MA clitic pronouns attach themselves in the normal way even in intra-sentential code-switched sentences, and this is a mark of the fluent bilingual. Switches as in (146) above, attaching clitics and aspectual affixation, are very common amongst all generation groups and are a sign of *Reactive Syntax* as discussed previously.

### 4.3 Moroccan Arabic Root Types

The Moroccan Arabic verb is very complex which means that it is capable of carrying a great deal of semantic information. There are four main categories of verb which are categorised by the quality of the radicals (Nortier, 1990: 13). These verbs are as follows:<sup>149</sup>

1. Simple (or sound): the three radicals are different from each other as in *kteb* 'he wrote'.

<sup>148</sup> Analysis of code-switched discourse and grammatical gender follows in Chapter Six.

<sup>149</sup> Nortier, J. (1990). *Dutch-Moroccan Arabic Code Switching among the Moroccans in the Netherlands*. Fortis.

2. Hollow: the second radical  $-w-$  or  $-y-$  has become invisible; its place is taken by a vowel as in *šāf* 'he saw'.
3. Weak: the third radical, originally  $-w-$  or  $-y-$  has been replaced with a vowel as in *bga* 'he wanted'.
4. Doubled or geminate: the second and third radical are equal as in *hell* 'he opened'

Furthermore as per Harrell (2004: 24) the following summarises the different common root types in MA:

A. Trilateral

1. Strong

- (a) Sound: e.g. *KTB* of *kṭeb* 'he wrote', *XBZ* of *xubz* 'bread'
- (b) Doubled: e.g. *ŠMM* of *šemm* 'he smelled', *KLL* of 'all'

2. Weak

- (a) Middle-weak: e.g. *S(V)F* of *šāf* 'he saw', *X(V)F* of *xūf* 'fear'
- (b) Final-weak: e.g. *MR(V)* of *mra* 'he polished', *HD(V)* of *hda* 'near'

B. Quadrilateral

1. Strong: e.g. *TRJM* of *terjem* 'he translated', *TNBR* of *tenber* 'postage stamp.'
2. Weak
  - (a) Second element weak: e.g. *Ṣ(V)FT* of *ṣifet* 'he sent'
  - (b) Fourth element weak: e.g. *SQS(V)* of *ka-iseqsi* 'he asks'

C. Atypical as in *ja* 'he came'

From the verb types, stems or measures can be derived and this root-and-pattern morphology is a highly-productive and semantically-rich process in Classical Arabic which uses up to 13 Measures from the basic consonantal skeleton, five of which are most commonly used in Moroccan Arabic.<sup>150</sup>

#### 4.3.1 Moroccan Arabic Stems and Word Derivation

Given the templatic MA consonantal skeleton as described above, MA stems are easily formed by the addition of a prefix or suffix to the already established verb pattern. Word derivation from MA verbal stems follows a patterned form and its derivation is considered not too complex. Again, citing Harrell (2004), the following outlines the main affixes assigned to stems, nouns, adjectives and particles:<sup>151</sup>

- a. The personal prefixes and suffixes of the imperfect tense of the verb e.g. the *-n* of *nšūf* 'I see'.
- b. The personal suffixes of the perfect tense of the verb e.g. the *-et* of *kanet* 'she was'.
- c. The prefix *-m* of various noun, adjective and verbal derivations e.g. *msāfer* 'traveller'.
- d. The feminine *-a* suffix e.g. *kbira* 'big' (fem, sing.).
- e. The plural ending *-(a)t* e.g. *ḥkāyēt* 'stories'.
- f. The plural suffix *-in* e.g. *ferḥānīn* 'happy (pl.)'.
- g. The nisba ending *-ī* e.g. *fransāwi* 'French'.
- h. The suffixed pronoun endings e.g. the *-k* of *mḡak* 'with you'.

The above examples can be described as uniform affixation to Moroccan Arabic verbs, nouns, adjectives and pronoun stems in general and given the basic consonantal

---

<sup>150</sup> Harrell (2004: 29) cites these as Measures I, Ia, II and IIa(V). The other six, namely III, IIIa, (VI), VII, VIII, IX and X are each represented by only a limited number of verbs.

<sup>151</sup> Person, number and gender of the MA verb are indicated by pre-fixes and affixes.

template, there is a close connection between verb types and corresponding word classes by reference to their isomorphic shape and this aids learning MA as a grammar and in the learning of inflections and tenses.

#### 4.3.2 Moroccan Arabic Verbs and Inflections

In MA, there are two tenses, the perfect and the imperfect, the former recognisable by its suffixes indicating person, gender and number, and the latter, the imperfect, by its prefixes indicating person, gender and number. The following Figures 4.4 and 4.5 show the inflectional suffixes for both the perfect and imperfect tenses:

Figure 4.4: The Inflection of the Perfect Tense

	Singular		Plural	
First Person	-t	'I'	-na	'we'
Second Person	-ti	'you'	-tim, -tu	'you' (pl)
Third Person (M)	-	'he'	-u, -w	'they'
Third Person (F)	-et, -at, -t	'she'		

The above Figure shows the general characteristics of the Perfect Tense and general tense formation. In regular verbs, these are straightforward, e.g. *ħəll* 'he opened', *ħəllət* 'she opened' and *ħəllina* 'we opened'. The Moroccan Arabic verbal stem follows a uniform and systematic set of affixations insofar as verbs are easily identifiable from their verbal stem, Measures and shapes. Code-switched examples follow below highlighting typical MA Perfect tense formation:

(147) *ktəb briya u posta – ħa li-ħum*  
 He wrote a letter it(FEM) and posted it (FEM) to them  
 'He wrote a letter and posted it to them'

(148) *šəmmet a riħa u ma ġbat ħa š, iwa*  
 Smelt 3SG DEF perfume and NEG1 like it (FEM) NEG2, and so

take -t      ha      back li l      shop  
 Take it PAST (FEM)      back to DEF shop  
 'She smelt the perfume and didn't like it so she took it back to the shop'

The Imperfect Tense as shown below has a set of prefixes which are added to the Arabic stem in whichever form or Measure it takes:

Figure 4.5: The Inflection of the Imperfect Tense

	Singular		Plural	
First Person	<i>n(e)-</i>	'I'	<i>n(e)- + -u, -w</i>	'we'
Second Person (M)	<i>t(e)-</i>	'you' (masc)	<i>t(e)- + -u, -w</i>	'you' (pl)
Second Person (F)	<i>t(e)- + -i, -y</i>	'you' (fem)		
Third Person (M)	<i>i-, y(e)-</i>	'he'	<i>i-, y(e)- + -u, -w</i>	'they'
Third Person (F)	<i>t(e)-</i>	'she'		

The above Figure 4.5 outlines the general characteristics of the Imperfect Tense and general tense formation. The basic verb shape for the Imperfect Tense is comprised of a prefix *-n* to MA verbs which begin with a consonant plus vowel e.g. *nredd* 'I bring back', *nreddu* 'we bring back' etc. However, if the verb stem begins with two or more consonants, the pre-fix *ne-* is used e.g. *neqra* 'I read' and *neqrāw* 'we read' etc. Code-switched examples follow below highlighting typical MA imperfect tense formation:

(149) *ila bgiṭ ne-ḡra, uken bdiṭ haḍi* long time ago, *meṣi heṭta* now  
 COND1 wanted read 3SG, COND2 started DEM long time ago NEG until now  
 'If I wanted to read, I should have started a long time ago, not until now'

(150) *xalliha te-jmaḡ al ḥwayaj al awwal, ḡād te-bda* studying *dyaḷ*  
 Leave her collect 3SG DEF things DEF first, later start 3SG studying POSS  
*ha*  
 her

‘Leaver her to tidy things up first then she can start her studying’

- (151) *ila bġina ne-mšiw, xasna ne-hurry up ġurk*  
COND want go 3PL, we have hurry up 3PL now  
‘If we want to go, we have to hurry up no’

In the Imperfective, when progressivity or habituality is referred to in describing everyday actions, the durative particle *ta-* or *ka-* is commonly employed:<sup>152</sup>

- (152) *ka-šreb bezzaf u haġak drinking ma ġedi -š y-help ih*  
Drink 3SG a lot and DEM drinking NEG1 going NEG 2 help 3SG him  
‘He drinks a lot and that drinking is not going to help him’

- (153) *ka-yekseb li girlfriend dyalu ma ġendu š waqt*  
Writes 3SG to girlfriend POSS (MASC) NEG1 have NEG2 time  
‘He is writing to his girlfriend, he hasn’t got time now’

A particle used to denote the future is *ġadi* ‘going to,’ which is always followed by an MA verb in its Imperfect form. For most speakers, *ġadi* as an auxiliary is invariable in number and gender. As a participle, feminine is *ġadya* and the plural is *ġedyin*:

- (154) *ġadya te-mši beletti, i leave-u ha alone*  
Going to/will go 3SG later, only leave 3PL her alone  
‘She is going to go later, just leave her alone.’

- (155) *ġedyin ne-šūfiw later, u ne-find-u out a truth*  
Going to/will PL see 3PL later and find 3PL out DEF truth  
‘We are going to see later and we will find out the truth’

Moroccan Arabic then has a limited number of clearly defined shapes in both tenses; the perfect and imperfect tense with inflection affixes tagged on in a predictable pattern. In code-switched intra-sentential discourse, MA morphology, given its distinctive and typologically different word order, does not disallow code switching or

<sup>152</sup> The imperfective inflections of *ka-* and *ta-* are used from region to region, mainly in Northern and North Eastern Morocco. However, a common morphological feature in Oujda, the largest city in the East is to add the inflection *rah* ‘he is’, *raha* ‘she is’, *raħna* ‘we are’ (note that the first person plural uses the pharyngeal voiceless fricative), giving rise to *rah yokəl* ‘he is eating’ and *raha ta twi al kiswa* ‘she is folding the clothes’

code-switched forms from being uttered. The onus is on intelligibility and Well-Formedness and given that these are maintained, produces some very interesting strings of bilingual discourse. Following Myers-Scotton's (1993b) Uniform Principle, Moroccan Arabic and English word order is imperative in establishing which of the two languages will produce the grammatical frame on which the syntax hinges.

#### 4.4 Moroccan Arabic Syntax - VSO or SVO

Standard Arabic is a Verb-Subject-Object language however it has been argued that the VSO word order is changing to SVO.<sup>153</sup> In MA, several word orders are both possible and employed. According to Ennaji (1982), Moroccan Arabic has two basic word orders: an unmarked VSO order used in any context, and a marked SVO order used only under special discourse conditions (1982: 14). According to Boudali (1984: 56), both VSO and SVO word orders are possible in Moroccan Arabic without restrictions to the use of one or the other. Introspectively, and according to my corpus, both word orders are possible. The below gives two basic word orders which are both found:

(156) *jāw a nās walla ma zel?*  
Come 3PL DEF people or NEG yet?  
'Have the people come or not yet?'

(157) *a nās jāw walla ma zel?*  
DEF people come 3PL or NEG yet?  
'Have the people come or not yet?'

Nortier (1990: 32) also cites two basic word orders in her corpus of Moroccan Arabic and Dutch code switching:

SVO ordering:

---

<sup>153</sup> See also Ouhalla (1994).

- (158) *ana ma ka-nešreb -š l birra*  
 I NEG1 drink 1SG NEG2 DEF beer FEM SG  
 'I don't drink beer'  
 (Moroccan Arabic, Nortier, 1990: 32)

V(S)O ordering:

- (159) *weqɛu mašākīl*  
 Happen PL problem M PL  
 'Problems happened'  
 (Moroccan Arabic, Nortier, 1990: 32)

Harrell (2004: 160) in analysing the simple verbal MA sentence describes it as consisting of a noun or a pronoun as a subject and a verb as a predicate with the subject usually following the predicate:

- (160) *jāw đ-điāf*  
 Come PL guest PL  
 'The guests have come'

- (161) *mša huwwa*  
 Went he  
 'He went'  
 (Moroccan Arabic, Harrell, 2004: 160)

According to Harrell (2004: 159) and my own data analysis, the predicate precedes the subject when the subject is a demonstrative pronoun in an exclamation or when the predicate is an interrogative pronoun:

- (162) *škun hað l- bint?*  
 Who DEM DEF girl  
 'Who is this girl?'

- (163) *aš dak ši?*  
 What DEM thing  
 'What is that?'

(164) <sup>154</sup> smen haða!  
Butter DEM  
'This is butter!'  
(Moroccan Arabic, Harrell, 2004: 159)

Harrell also goes on to state that the most elementary kind of simple [MA] sentence consists of a single verb form with the subject pronoun signalled by the affix of the verb e.g. *klina* 'we ate'. Equally fundamental is the simple sentence which consists of a single word as the subject and a single word as the predicate e.g. *ħwayžu mwessxin* 'his clothes are dirty'. Simple sentences are not necessarily short, however, and the expansion of the subject and predicate brings about simple sentences of considerable length (2004: 160). In terms of MA complex sentences, as Harrell (2004) outlines, these contrast with MA simple sentences in that they consist of a combination of two or more simple sentences to form one larger sentence. In such a sentence, one of the constituent simple sentences stands as an independent sentence while the other simple sentence or sentences function as a subordinate part of it e.g. *xella l-xor ka-itsenna fih* 'he left the other waiting for him', where the simple sentence *xella* 'he left' has for an object the entire sentence *l-xor ka-itsenna fih* 'the other one waiting for him'. The simple sentence which functions as a subordinate part of a complex sentence may be referred to as the subordinate clause, and the simple sentence to which it is subordinated may be referred to as the main clause. Simple sentences function as subordinate clauses in complex sentences in only three ways, as nouns, as adjectives and as adverbs. Complex sentences differ as to whether the constituent clauses are joined by a linking conjunction by a pre-stated topic. (2004: 162). The following example is given:

---

<sup>154</sup> However, this statement can also be reversed: *hada smen*.

(165) *šifet* *šemm-u* *ixteb* *lu waḥəd* *le mra* *šnd š*  
 He sent uncle POSS ask for marriage to one DEF woman POSS some  
*nās*  
 people  
 'He sent his paternal uncle to some people to ask for a certain woman in  
 marriage for him'  
 (Moroccan Arabic, Harrell, 2004: 162)

In explanation of the above complex sentence, Harrell states that:

The noun *šemm-u* 'his paternal uncle' functions as the object of the verb *šifet* 'he sent' in the main clause and as the subject of the verb *ixteb lu* 'he asks in marriage for him' in the subordinate clause (2004: 162).

The above examples are then typical of Moroccan Arabic morphology of simple and complex sentences and shows how the word order can in fact be both VSO and SVO but how does MA fit in with English word order in code-switched sentences?

#### 4.4.1 Moroccan Arabic and English Word Order

As discussed in Chapters Two and Three, the sentence of the matrix language provides the grammatical frame irrespective of the individual syntax of each language. Myers-Scotton's (1997) theory in the MLF is re-presented below:

*Myers-Scotton (1997):*

*Matrix Language definition (Myers-Scotton 1995: 983, 1997)*

*The ML is the language projecting the morphosyntactic frame for the entire CP which shows intrasentential CS.*

Therefore, whether Moroccan Arabic is VSO or SVO and English strictly SVO has no bearing on an individual level, but whether MA or English syntax is adhered to does when the morphemes are code-switched intra-sententially as the morpho-syntax is

activated with the matrix assuming the grammatical frame. The below examples highlight both MA and English syntax in code-switched sentences:

MA word order + English word order + MA:

(166) *bda y-gull -i* something but I didn't catch it *geɛ*  
Start tell 3SG me something but I didn't catch it all  
'He started to tell me something but I didn't catch it all'

English word order + MA word order:

(167) I keep telling him *baṣaḥ ma bge -š* *y-listen*  
I keep telling him but NEG1 want NEG2 listen 3SG  
'I keep telling him but he doesn't want to listen'

English + MA:

(168) Are we going to get some *houli -s* for *ɛid walla xesna* *na-wait-iw*  
Are we going to get some sheep PL for Eid or we have to wait 3PL  
'Are we going to get some sheep for Eid or do we have to wait'

The combinations are quite striking, particularly (168) spoken by a third generation speaker where she has added the English plural marker 's' to the singular noun *houli* 'sheep' to mark plurality and this is a further example of *Reactive Syntax* where second and third generations of Moroccan Arabic speakers use a new and innovative speech styles which sets itself apart from that of first generation speakers. Furthermore, the verb 'wait is sandwiched between an MA aspectual marker and suffix. This is typical of the *Reactive Syntax* discussed in the previous Chapter where new combinations in intra-sentential code-switched strings are becoming more and more innovative and experimental. Nortier (1990) presents data which highlights differences in word order and direction from Dutch to MA but this is a straightforward switch and certainly does not incorporate the dynamic switching as evident in (168) above:

(169) *Dat is wel voldoende, u haða huwwa l- ferq*  
 That is about enough and that is DEF difference  
 ‘That is about enough and that is the difference’  
 (Moroccan Arabic / *Dutch*, Nortier, 1990: 94)

These examples then give not only a more transparent view of code switching mechanisms and employed by bilingual speakers but also give us a better concept of Moroccan Arabic syntax which is essential in order to fully grasp the strategies that bilingual speakers use in order to merge two or more syntaxes in natural discourse. We can see that the characteristics of the morphology of Moroccan Arabic is far from complex and regular in its pattern formation with few exceptions to grammatical rules, unlike that of Classical Arabic which is highly inflectional with numerous tenses, case assignment and the like. Following the basic Moroccan Arabic consonantal skeleton as described above, noun formation is taken from these radicals and assigned masculine or feminine gender. Its morphology is outlined in the following section.

#### 4.5 Moroccan Arabic Nouns - Definite and Indefinite Articles

The general functions of a noun in MA are similar to that of the English noun. According to Harrell (2004) seven main functions of the noun are distinguishable: (i) as the subject of a sentence e.g. *dexlet waħəd l-xadem* ‘a maid came in’; (ii) as the object of a verb e.g. *šuft rajel* ‘I saw the man’, *šra l-xubz* ‘he bought some bread’; (iii) as a predicate complement e.g. *ana nejjār* ‘I am a carpenter’; (iv) as an objective complement e.g. *semmit weldi ħli* ‘I named my son Ali’, (v) as the object of a preposition e.g. *maħa šaħbi* ‘with my friend’, *fe d-dār* ‘in the house’; (vi) as an adverb of time e.g. *bqa yumayn* ‘he stayed for two days’; and (vii) as a vocative *a ħli, fayn kunt?* ‘Ali, where have you been?’ The MA noun has five basic modifications

and these are definition,<sup>155</sup> annexion, pronominal possession, adjectival modification and enumeration (Harrell 2004: 186). Amongst the noun types listed above, certain MA nouns occur without prefixes and some never take the indefinite article and some take the definite article. Prefixless nouns are quite common and a few examples of these are as follows:

(170) *haḍa ʕājīb!*  
 DEM amazing!  
 ‘This is amazing!’

(171) *bab -ak maġrebi u m -uk glinziya*  
 Father POSS 2Sg Moroccan MASC and mother POSS 2SG English FEM  
 ‘Your father is Moroccan and your mother is English’

(172) *iwa haḍa xbar!*  
 Well DEM news!  
 ‘Well this is really something!’

The system of MA nouns and their annexion of definite and indefinite articles as well as being prefixless operates according to a standardised and highly regular template. In terms of the indefinite article, MA uses the morpheme *si* and *wahəd* to denote indefiniteness e.g. *si rajel* ‘a man’ and *si nhār* ‘some day or other’ and it can also be used in the plural form *si rjēl* ‘some men or other’ The other indefinite article *wahəd*<sup>156</sup> is normally followed by the definite article annexed to the noun unless the noun is of a category which does not take the definite article<sup>157</sup> e.g. *wahəd al wald* ‘a boy’, *wahəd al nhār* ‘one day’. The definite article – *al* or –*a* or –*l* is prefixed to the

<sup>155</sup> There are four degrees of definition of the noun. Nouns occur prefixless, prefixed with an indefinite article, prefixed with the definite article or prefixed with a demonstrative article (Harrell 2004: 186). A further analysis of MA nouns and their complements follows in Chapters Six and Seven.

<sup>156</sup> The indefinite article *wahəd* which is invariable for gender and which precedes its noun is not to be confused with the numeral adjective *wahed* (masculine) and *wahda* (feminine form) which follows its nouns e.g. *mra wahda* ‘one woman’ and not ‘a woman’ cf. Harrel (2004).

<sup>157</sup> Nouns which do not take the definite article are those with pronoun endings e.g. *weldi* ‘my son’ and *ixti* ‘my sister’. Also nouns “other than numerals which are the first term of a construct state” (Harrell 2004: 189 – 191), (e.g. a noun or an adjective followed directly by a noun or pronoun) e.g. *hamut rajalha* ‘her husband’s shop’.

noun e.g. *l-bāb* ‘the door’ *a rajel* ‘the man’, *al qaḍi* ‘the judge’ and this is a regular pattern in definiteness. However, certain nouns are definite without need for the article, in particular nouns of possession e.g. *ktābi* ‘my book’ is definite as the article cannot co-exist with a possessive marker. Also, abstract nouns and professions which are definite do not take the definite article e.g. *bennāy* ‘mason’, *semṣar* ‘broker’ and *derrī* ‘child’ are all definite by nature. Another category of MA nouns which do not require the definite article is proper names e.g. *Fez*. The last category is that of highly abstract nouns which must be memorized as single lexical items e.g. *seksu* ‘couscous’, *xizzu* ‘carrots’ and *matiṣa* ‘tomatoes’. In code-switched utterances, the data corpus shows that MA grammaticality is maintained when it is the Matrix Language and both the definite and indefinite articles were used in the correct method and form. The below gives some examples:

(173) *wahād al* day he went there to get me but it was too late.  
 INDEF DEF day he went there to get me but it was too late  
 ‘One day he went to get me but it was too late’

(174) *ama, al yawm* Friday so *kayn ši seksu?*  
 Mum DEF day Friday so there is INDEF couscous?  
 ‘Mum today is Friday so is there any couscous?’

(175) *šedd l- door mur -ak u jib al* take-away *maḡ -k*  
 Close DEF door behind you and bring DEF takea-awar with POSS  
 ‘Close the door behind you and bring the take-away with you’

In (175) above the MA indefinite NP construction shows how the definite noun *seksu* ‘couscous’ is made indefinite by the pre-fix *ši* and this is maintained even in code-switched discourse which conveys that the speaker is a bilingual and also adheres to the UG of languages in general. This gives us an important indication with regards to bilingualism and second language acquisition in general.

#### 4.5.1 Masculine and Feminine Moroccan Arabic Nouns

Moroccan Arabic nouns and adjectives have two main grammatical features and are either masculine or feminine, with nouns ending in *-a* almost always being feminine and most other nouns then defaulting to the masculine. Nouns have a fixed gender whereas adjectives can be either masculine or feminine with grammatical agreement following suit. MA nouns and adjectives can be either singular or plural with highly regular inflectional marking to identify them with a few nouns marked for dual although use of the dual in spoken MA is very limited. The following shows typical MA nouns in either their masculine or feminine forms, with *-a* marking feminine gender assignment.<sup>158</sup> This patterning is highly regularised with few exceptions:

Figure 4.6: MA masculine and feminine nouns

Masculine	Feminine	Gloss
<i>qerd</i>	<i>qerda</i>	'monkey'
<i>malik</i>	<i>malika</i>	'king, queen'
<i>xāl</i>	<i>xālla</i>	'maternal uncle / aunt'
<i>ʕfrit</i>	<i>ʕfrita</i>	'bad person'
<i>ʕemm</i>	<i>ʕemma</i>	'paternal uncle / aunt'
<i>emir</i>	<i>emira</i>	'prince / princess'
<i>jār</i>	<i>jāra</i>	neighbour

The above Figure 4.6 lists a limited number of MA nouns with identical corresponding feminine counterparts as MA also has separate nouns for male and female equivalents. A few of these are as follows: *weld* 'boy' and *bint* 'girl', *rajel*

<sup>158</sup> In general, names of towns and countries are feminine e.g. *Fez, maṣr* 'Egypt'. The exceptions to this rule must be learned as individual lexical items e.g. *r-rbaṭ* 'Rabat' is masculine. Nouns referring specifically to women, including personal names are always feminine e.g. *bint* 'girl', *xaddem* 'servant girl' (Harrell 2004: 98).

‘man’ and *mra* ‘woman’. Certain nouns are feminine by nature and not all have the marked *-a* ending. These are fixed nouns with a fixed feminine gender and this is interesting particularly in code-switched sentences as the adjectival description has to agree with an unmarked Moroccan Arabic noun. This will be described in detail below after a list of MA feminine nouns with unmarked *-a*:

Figure 4.7: MA Feminine nouns with unmarked *-a*

Feminine	Gloss
<i>blād</i>	‘country’
<i>dār</i>	‘house’
<i>nār</i>	‘fire’
<i>rūh</i>	‘spirit’
<i>lerḍ</i>	‘earth’
<i>zīt</i>	‘oil’
<i>yedd</i>	‘hand’
<i>semš</i>	‘sun’

The above MA nouns are uncompromising in that they can only ever be in the feminine and have feminine adjectival agreement.<sup>159</sup> The majority of them are short nouns with the pattern of *CCa*. Interestingly, there is a selection of MA nouns which are masculine in gender but have the feminine marker *-a* and these will be discussed below with regards to how they are assigned adjectival agreement in code-switched paradigms. The following gives a few examples of masculine nouns:

<sup>159</sup> This is in contrast with some MA nouns which can either be masculine or feminine depending on the speaker. Some examples are *bāb* ‘door’, *bīt* ‘room’ and *jāmeʿ* ‘mosque’.

Figure 4.8: MA Masculine nouns with marked *-a*

Masculine	Gloss
<i>bla</i>	'misfortune'
<i>hna</i>	'peace / tranquility'
<i>hwa</i>	'air'
<i>nqa</i>	'cleanliness'
<i>sma</i>	'sky'
<i>ʕya</i>	'fatigue'
<i>ħya</i>	'modesty / shame'

In terms of pedagogy, the above masculine items are learnt as single lexical entries and these are then memorized as there is no logic as to their masculine assignment given their feminine form. In code-switched CP phrases, how do nouns and adjectives fare in terms of gender assignment, NP agreement and grammaticality in general? The following section details MA nouns and adjectives.

#### 4.5.2 Moroccan Arabic Nouns and Adjectival Agreement

The MA noun takes one of two functions, either as a noun or as an attributive adjective. In the former function, according to Harrell (2004) adjectives are used freely as nouns, and as such they have all the general sentence functions of nouns e.g. as the object or a preposition in *mʕa le kbir* 'with the big one'. Adjectives used as nouns, however, have several characteristics which distinguish them from actual nouns. They regularly have either the definite article or a demonstrative article e.g. *l-*

*buyed* ‘the white ones’, *haḍak ṣ-ṣgīr* ‘that little one.’<sup>160</sup> Adjectives used as a noun modifier is directly placed after the noun it modifies e.g. *bint kbira* ‘a big girl’, *awlād zinnīn* ‘nice boys’ and *derri ṣgīr* ‘a small boy’ and correspond in number and gender:<sup>161</sup>

(176) *ktāb*                      *mezyan*  
 Book MASC SG nice MASC SG  
 ‘A good book’  
 (Moroccan Arabic, Nortier, 1990: 21)

(177) *ktūb*                      *mezyanīn*  
 Books MASC PL nice MASC PL  
 ‘Good books’  
 (Moroccan Arabic, Nortier, 1990: 21)

As well as nouns, adjective insertion is regularly embedded in code-switched discourse with either an MA or English internal structure dependent upon which is the matrix language and therefore dictates the grammatical frame and word order. If the matrix language is MA, adjective agreement, if an MA adjective is used, must correspond in gender and number. If the adjective is in English, then no agreement marking is necessary. The below examples highlight MA noun-adjective agreement in CS:

(178) *the bint’s too ṣgira*            for that  
 The girl’s too small FEM for that  
 ‘The girl’s too small for that’

<sup>160</sup> Harell (2004: 204).

<sup>161</sup> Ibid (2004) lists the chief patterns of MA attributive adjectives and these are as follows: (a) the attributive adjective regularly takes the definite article if the noun it modifies is preceded by a demonstrative e.g. *duk s-skariya le mxezzin* ‘those incorrigible drunkards’, (b) the attributive adjective never takes the definite article if the noun it modifies is preceded by an indefinite article e.g. *ṣi mmiha barda* ‘some cold water’, (c) an attributive adjective modifying a prefixless noun takes the definite article only if the noun has a pronoun ending e.g. *modda ṭwila* ‘a long time’, and (d) if the noun it modifies has the definite article, there is considerable variation as to whether the attributive adjective does or does not take the definite article (2004: 205).

(179) *a rajel a blondey mazinu*  
 DEF man DEF blonde MASC nice  
 'The blonde man is handsome'

The above examples are interesting, in particular example (179) as the speaker has maintained gender and number agreement in line with the MA noun *bint* which is marked for feminine and singular, and at the same time, the speaker has added the possessive marker to the noun and modified the adjective with the English adverb 'too'. In terms of predicate adjectives which are regularly used in MA, these also follow gender and number agreement occurring as either a predicate complement or an objective complement as in *kanuw ferhānīn* 'they were happy' as a predicate complement and *dahum buyed* 'he took them white in colour' as an objective complement. In intra-sentential code switching, agreement is maintained in the below example for instance where the objective complement is maintained in the MA variety;

(180) I want some of those *burquq tayibīn meši l hamdīn*  
 I want some of those plums ripe NEG DEF sour  
 'I want some of those ripe plums not the sour ones'

(181) *kanuw sad hta ja bah -um ou dah -um li l park y-have-u fun*  
 They were sad until came dad POSS and took them to DEF park have 3PL fun  
 'They were sad until their dad came and took them to the park so they can have fun'

In (181) above, the subject is plural and the masculine possessive marker *hum* 'them' is in agreement as is the aspectual marker *-u-*. This type of code-switched string is very common amongst all generation groups and conventional code switched demarcation lines in terms of where the grammars can merge are very progressive compared with other Moroccan Arabic bilingual data (Nortier 1990), Benhtahila & Davies (1983), and Boumans (1998). We can therefore establish that Moroccan

Arabic / English speakers in the UK are very innovative and naturally experimental in their use of both languages in a code-switched domain where intelligibility is always maintained together with the Well-Formedness constraint. Two areas that cannot be switched relatively easily are that of relative clauses and negation where the language of the contributing negative particle or relative clause has to be followed by the corresponding particle or clause. We commence the latter part of this chapter with an analysis of MA / English relative clauses in bilingual discourse.

#### 4.6 Relative Clauses

In MA, the complex sentence involves the bridging of two clauses or simple sentences to form a larger, or more complex sentence. According to Ennaji (1982) and Boudali (1984), *lli* 'who' and 'that', *men* 'who' and *aš* 'what' are relative pronouns capable of introducing a relative clause in MA linking another Noun Phrase. The relative marker *lli* is by far the most commonly used relative clause marker which in all its forms can be translated as 'who', 'that', and 'which':

(182) *al bint lli xu -ha mša l- engliz beš y-dir doctora*  
 DEF girl REL brother POSS went to England in order do 3SG Doctorate  
 'The girl whose brother went to England to do a PhD'

(183) *l- wald lli mša*  
 DEF boy REL went  
 'The boy who left'  
 (Moroccan Arabic, Ennaji, 1982: 51iii)

(184) *le mra maš men šuftak*  
 DEF woman with whom I saw you  
 'The woman with whom I saw you'  
 (Moroccan Arabic, Ennaji, 1982: 51i)

(185) *jbert l ktāb lli xallit fi dār haðak a nhār*

I found DEF book REL left 1SG in house DEM DEF day  
'I found the book which I left in the house that day'

In discussing relative pronouns, Boumans (1998) focuses only on those which are confined to those cases where a non-human head noun functions as the complement of a preposition in the relative clause. The prototypical MA relative clause which can be used in the aforementioned context as well as in all other cases is of a different type. The particle *lli* links the head noun to the relative clause, and the head noun recurs in the relative clause as a pronoun. No resumptive pronoun is used when the head noun is the Subject of the relative clause, and when it is the Object of the relative clause, it is often omitted. Moreover the particle *lli* is optional when the head noun is indefinite and not marked by the indefinite *ši*. The following are some typical examples of MA relative clauses:

(186) *ka-nduwwer ʕe l bint lli ba- ha mrið*  
Look 1SG on DEF girl whose father POSS is ill  
'I am looking for the girl whose father is ill'  
(Moroccan Arabic, Harrell, 2004 (164))

An alternative to the commonly-used *lli* is *aš* as cited below:

(187) *s- sfina f- aš kan-uw rākbīn*  
DEF ship in which they PAST 3PL riding PL in  
'The boat that they were riding in'  
(Moroccan Arabic, Harrell, 2004 (164))

(188) *al ʕām f- aš dxelt lengliz*  
DEF year in which I entered England  
'The year in which I came to England'

Ennaji (1982) differentiates between restrictive and appositive relative clauses where the former means that there are additional members of the class referred to by the

relativized element in the main clause which do not share the same qualifications of the subordinate clause (Nortier 1990):

- (189) *l- hulandiyin lli mɣa -ya f groep ka-yhetmu le mgerba*  
DEF Dutch MASC PL who with POSS in group disdain 3PL DEF Moroccans  
'The Dutch who are in my group disdain the Moroccans'  
(Moroccan Arabic / *Dutch*, Nortier, 1990: 37)

Appositive relative clauses often function as parenthetical comments which are marked by a specific intonation pattern that clearly sets them apart from the main sentence (Ennaji 1982):

- (190) *had r- rajel, lli ɣād xrej, saħb- i*  
DEM DEF man who just left friend POSS  
'The man who just left is my friend'  
(Moroccan Arabic, Ennaji, 1982: 122)

In addition to the above relative clause markers, (*i*)*lli* can occur in the sense of 'the one who' 'those who' and 'that which' without recourse for a pre-stated topic. In this sense, the relative clause is reductive without a need for a bi-clausal state as in the examples above:

- (191) *lli šətih, šətu*  
REL you saw, I saw  
'Whatever you saw, I saw'

- (192) *nti lli kadaba*  
You FEM who liar  
'You're the liar / the one who tells lies'

- (193) *illi bga ši problem y-het yidd -u fi l fire*  
REL wants INDEF problem put 3SG hand POSS in DEF fire  
'Whoever wants a problem, should put his hand in the fire'

In the data collected, code-switched sentences do not allow for the first part of the relative clause to be in Language *x* and the second clause in Language *y*. Also, in the literature, there are hardly any examples of the merging of two or more languages in

an intra-sentential code-switched CP. The only examples are where the first clause is in one language including the relative clause marker, and the final part of the clause in the second language:

(194) *a rajul illi gel-ek ðak ši,* is a liar  
DEF man REL told you DEM thing, is a liar  
'The man who told you that, is a liar'

(195) *awlād al yawm illi zādu hna fi* England, haven't got a clue  
Boys DEF day REL born PL here in England, haven't got a clue  
'Boys today who were born in England haven't a clue'

(196) *me bqiš naħmāl al smell dyellu*  
I can NEG1 carry 1SG DEF smell POSS his  
'I can no longer tolerate his smell'

This then allows us to make the following Generalization:

*Generalization 4*

*In intra-sentential code-switched CPs, the language of the relative clause marker can be switched.*

In light of the above generalization, the following examples are possible:

(197) *a rajul illi* went there  
DEF man REL went there  
'The man who went there'

(198) *the girl who mšet to town ma wellet-š li dār nišan*  
The girl who went to town NEG1 return NEG2 to house straight  
'The girl who went to town didn't go home straight after'

Given (194) above is possible as well as examples (194) to (196), it gives us an insight into the evolution of code switching as a switching strategy. We can safely assume and predict that in complex sentences as those listed above, these would not

be uttered by first generation Moroccan speakers as it is a unique speech style cultivated by second and third generation speakers. The syntactic structure in example (198) is particularly striking as the CP phrase commences with an English DP and relative clause marker, followed by a Moroccan Arabic past participle, and English PP phrase and concludes with a negative clause. Such high-volume switching is typical of a new generation of bilingual speakers in the UK. The negative clause markers in (198) “*ma*” and “*š*” are always in tandem and whether these can be switched is examined in the next section.

#### 4.7 Negation

In Moroccan Arabic, negation is formed by adding negative markers to verbal clauses or nominal clauses (Harrell 2004; Marçais 1977; Ouhalla 1994; Brustad 2000). Verbal negation is expressed by prefixing the particle *ma-* to the verbal stem or form and adding the suffix *-š*:

- *Ma* + MA VERB + *š*

(199) *ma gelli-š walu ɛla al* party next week  
 NEG1 say NEG2 nothing on DEF part next week  
 ‘He didn’t tell me anything about the party’

This same structure where *ma-* is the head of the NEGP and *-š* is its Specifier also applies to verb forms including all affixes:

(200) *ma ɛta hum li- -š*  
 NEG1 give 3Sg them to me NEG2  
 ‘He didn’t give them to me’

The NEG2 or ‘closing’ negative particle is always the last element of the verbal string together including any suffixes:

(201) *ma iqulu-ha- lna- š*  
 NEG1 tell it FEM to us NEG2  
 ‘They won’t / are not going to tell it to us’  
 (Moroccan Arabic, Harrell, 2004: 152).

The *ma-...š* structure is the most common configuration in Moroccan Arabic negation of verbal forms.<sup>162</sup> However, the substitution of *la-* for *ma-* in the negative imperative has a more general advisory or morally admonishing implication (Harrell 2004: 153):

(202) *la tkedbu -š!*  
 NEG1 lie 2PL NEG2  
 ‘Don’t tell lies!’  
 (Moroccan Arabic, Harrell, 2004: 152).

There are exceptions to the above structure where the MA verb form is pre-fixed with *ma-* but does not have the enclosing *-š* ending. The most common negatives (Harrell 2004: 153) are as follows:

(203) **MA negatives**

<i>walu</i>	‘nothing’	→	<i>ma rbeṭ walu</i>	‘he didn’t tie anything’
<i>(hetta) hejja</i>	‘nothing’			
<i>(hetta) ḥed</i>	‘no-one’			
<i>(hetta) waḥəd</i>	‘no-one’	→	<i>ma šet (hetta) waḥəd</i>	‘I didn’t see anyone’
<i>la...wa-la</i>	‘neither..nor’			
<i>Ḥimmer</i>	‘never’			

In non-verbal paradigms, negation is most commonly made by employing the particle *ma-ši* where this precedes the noun *ma-ši mazyān* ‘not good’, *huwwa ma-ši hna* ‘he is

<sup>162</sup> However, there are exceptions where the negated form does not require the *ma-...š* structure and examples of these are as follows:

(1) *ma Ḥndu walu* ‘he hasn’t got anything’  
 (2) *ma Ḥnd-ha flus* ‘she has no money’

These are usually in structures which contain the preposition *Ḥnd* ‘have’.

not here.’ When the sentence does not contain a finite verb but a participle or a predicatively used adjective, both negation forms can be used. According to some speakers, the use of the discontinuous form makes the negation stronger when both forms are used (Nortier 1990: 41):

(204) *ma ġāli -š*  
 NEG1 expensive NEG2  
 ‘It is not expensive’

(205) *ma-ši ġāli*  
 NEG1 expensive  
 ‘It is (absolutely) not expensive’  
 (Moroccan Arabic, Nortier, 1990: 410)

Moroccan Arabic pronouns operate in a similar manner, such that negative pronouns are used with non-verbal predicates including participles, adjectives, prepositional phrases and nouns. The following sets out MA negative pronouns:

Figure 4.9: MA negative pronouns<sup>163</sup>

Person	Singular	Plural
First Person	<i>maniš</i> ‘I am not’	<i>maħnaš</i> ‘we are not’
Second Person (M)	<i>mantaš</i> ‘you are not’	<i>mantumaš</i> ‘you re not’
Second Person (F)	<i>mantiš</i> ‘you are not’	
Third Person (M)	<i>mahuwwaš</i> ‘he is not’	<i>mahumaš</i> ‘they are not’
Third Person (F)	<i>mahiyyaš</i> ‘she is not’	

(206) *ma ni- š tərša*  
 NEG1 am NEG2 deaf  
 ‘I’m not deaf’

<sup>163</sup> Brustad K. E. (2000) *The Syntax of Spoken Arabic*. Georgetown University Press (Washington) (pp: 296).

(207) *ma huwwa -š hna*  
 NEG1 he NEG2 here  
 'He is not here'  
 (Moroccan Arabic, Harrell & Sobelman, 2004: 153)

Nortier (1990) states that:

In summary, a discontinuous negation is generally used for the finite verb. If there is no finite verb, the discontinuous negation can be used for a form that has some characteristics in common with a finite verb. In any case, this verb or verblike (sic) should be marked for number and gender, in combination with either person or tense. When the main verb contains a tense marker, the discontinuous negation is used. In all other cases *maši*, the discontinuous negation, is required (1990: 42).

Therefore, it can be concluded that in MA negation, the *ma-* particle must always be present to express negation and the secondary *-ši* is subject to distribution polarity sensitive expressions, see (206) above, and is therefore optional. How then does MA negation operate in code-switched sentences?

#### 4.7.1 Moroccan Arabic Negation and Code Switching

Given the *ma-...ši* construction as detailed above in verbal negation, and the need for *ma-* to always be present in MA negation, it is predicted that splitting the two negative particles within an intra-sentential CP and suffixing the *-š* to an English verbal form could be possible given the innovation forms of *Reactive Syntax*. Below is an example:

(208) *ma rəd li -š flus -i*  
 NEG1 return to me NEG2 money POSS  
 'He didn't give me back my money'

(209) *ma give ni -š back flus -i*  
 NEG1 return to me NEG2 back money POSS  
 'He didn't give me back my money'

Does (209) above denote some aspect of manoeuvrability in terms of adding the suffixed *-š* negative particle to an English verbal stem? It is evident then that MA / English code switching allows for the evolution of speech styles where the *ma-* is the head of the NEGP and *-š* as its Specifier attaches itself to an English verb. Further examples follow:

(210) *ma* like *ha -š* *wa lakin mazel zayed* *mašha*  
 NEG1 like her NEG2 but still continues with her  
 'He didn't like her but still carries on with her'

(211) *ma* tell *-it* *ha -š* *walu*  
 NEG1 tell PAST her NEG2 nothing  
 'I didn't tell her anything'

In terms of structure, it is clear that the matrix language is Moroccan Arabic with English as the embedded language adhering to both the syntax and negation structure of MA. We can therefore make the following generalization:

(1) *Generalization 5*

*In MA negation, the MA suffixed negative particle -š does not also have to be in MA.*

These are complex sentences in MA and in intra-sentential sentences, the combination is interesting in the way in which the negative particles are split and the suffixed *-š* attaches itself to the English verb. In terms of nominal sentences and MA negation in code-switched sentences, this is far less complex where the particle *ma-š* precedes any noun either in MA or English:

(212) *haða ma-ši* nice  
 DEM NEG nice  
 'This isn't nice'

- (213) that's *ma-ši* good enough  
That's NEG good enough  
'That's not good enough'

We can therefore state that in non verbal negation there is, in code-switched intra-sentential dialogue, some room for manoeuvrability as the *ši* can attach itself to non-MA stems and does not in the process impede intelligibility. Rather it is a new method of merging the syntaxes of two or more languages amongst second and third generations of bilingual speakers in the UK. In terms of nominal sentences, we have seen that the negative particle *ma-ši* has to be a single unit and cannot be split, however this can be followed by any noun phrase in a language other than Moroccan Arabic as in *huwwa ma-ši good enough* 'he is not good enough.' In MA negative clauses there is the negated clause *ma ʔndhaš flus* 'she hasn't any money' can be turned into a 'double negative' *ma ʔndha walu* 'she hasn't anything' where the suffixed *š* is dropped in lieu of a negative particle. In code-switched sentences, this 'double negative' patterning permeates the English and the following examples have been recorded *ma ʔndha nothing* 'she hasn't got anything / nothing', *ma ʔndu nothing* 'he hasn't got anything / nothing'. In these examples, both the MA and English syntaxes are followed as they are identical in negative syntactic behavioural patterns.

#### 4.8 Conclusion

This chapter has given a broad outline of Moroccan Arabic grammar which is an essential part of this thesis as it provides an important sketch of Moroccan Arabic syntax and its placement alongside English syntax and code-switched paradigms. In view of the relevant literature which is rather scarce (Abdel Massih 1973, Brunot 1950 and more recently Heath 1989 and Harrell 2004), I commenced this chapter with an historical overview of Moroccan Arabic and other Arabic dialects, describing the

multilingual make-up of Morocco. I then analysed Moroccan Arabic morphology in the structure of the Moroccan Arabic verb and the way in which verbs are formed using the Standard Arabic Measures and the templatic consonantal skeleton on which trilateral (and quadrilateral) verbs are composed. This is an integral part of this chapter as it highlights the way in which code-switched suffixes, for example, the English gerund *-ing* can be attached to the Moroccan Arabic stem. MA word order is analysed as being VSO or SVO (Ennaji 1982) and this further details how code-switched sentences, following Myers Scotton's (1993b) Matrix Language Frame model, must adhere to the language which provides the grammatical frame. In this manner, an analysis of MA word order aids research on inflectional affixation on both MA and English stems in their respective syntaxes.

Other word classes and grammatical categories are also addressed, namely, relative clauses where a generalization was made that in intra-sentential code-switched CPs, the language of the relative clause marker can be switched. This is an interesting discovery as prior to this thesis, on an introspective level, this was thought not to be the case. Finally, the chapter concludes with MA negation where it has been shown that MA negation is formed by adding negative particles to verbal clauses or nominal clauses (see also Harrell 2004; Marçais 1977; Ouhalla 1994). Given the analysis, it is clear the prefixed *ma-* particle must always be present to express negation and the secondary *-ši* suffix is subject to distribution sensitive expressions.

A second generalization in this chapter is made that in MA negation, the MA suffixed negative particle *-š* does not also have to be in MA. This was made transparent in light of code-switched data gathered and the examples shown above. This is an important chapter which lays the foundation for Chapters Six and Seven which will provide a further detailed analysis of code switching from MA to English and English to MA respectively where I will advocate the use of the MLF model

based on empirical data collated. The next Chapter then focuses on data collection methods, informants, the Moroccan community in the UK and research design.

## CHAPTER FIVE

### THE UK MOROCCAN DATA CORPUS

This chapter focuses on data collected for this thesis, commencing with the history of migrants from Morocco in the 1960s and the influx of migration in general with family reunification in the 1970s followed by the onset of second and third generations in the UK. The second part of this thesis is dedicated to the analysis of the range of informants with selection criteria a major factor in determining the quality of the data, proficient knowledge of Moroccan Arabic and being born in either the UK or Morocco being the main criteria for selection. As a native Moroccan Arabic speaker, it is interesting to carry out the interviews and recording of the data as it provides us not only with linguistic information but also with sociolinguistic information about the speakers (Nortier, 1990: 85). As Poplack (1980) describes:

The importance of data collecting techniques cannot be overemphasized, particularly in the study of a phenomenon such as code-switching, which cannot be directly elicited. The actual occurrence of a switch is constrained, probably more than by any other factor, by the norms or the perceived norms of the speech situation. The most important of these norms for the balanced bilingual was found (Poplack , 1978a) to be the ethnicity of the interlocutor, once other criteria (appropriateness, formality of speech situation) were met (1980: 595).

This enabled me to execute the individual structured interviews and more importantly, the informal individual and group conversations were recorded in as natural an environment as possible.

Data collection began as a pilot study at the beginning of the PhD in 2005 to pave the way for the main data collection used in this thesis from 2006 to 2007 partly in Morocco but mainly in the UK which enabled me to further access the Moroccan communities in the UK. This is followed by an analysis of the research design,

respondents' questionnaires submitted and results found as well as the structure and layouts of structured interviews, results and comments for further study. As discussed in previous Chapters, it will be evident that my new concept of *Reactive Syntax* is only attributable to second and third generations as it is amongst these generational groups that the syntax varies and is 'reactive' to their newly formed speech styles. This is discussed in the final part of this chapter.

### **5.1 History of Migration from Morocco to the UK**

International migration flows have increased in magnitude and complexity over the last few decades changing the social and economic dynamics of both migrant-sending and migrant-receiving countries, in this case, Moroccan and the UK.<sup>164</sup> Rapid post-war economic growth in northwest Europe created increasing unskilled labour shortages in sectors such as industry, mining, housing construction and agriculture from the 1950s.<sup>165</sup>

Moroccan migration to the UK began in the 1960s when British employers in industrial sectors began to recruit personnel from Morocco due to a shortage of unskilled and semi-skilled labour with Morocco evolving into one of the world's leading emigration countries. Moroccans form one of the largest and most dispersed migrant communities in Western Europe. Morocco has become a typical example of Skeldon's (1977) "labour frontier country" which is a category of upper and lower middle income countries whose modest social, economic and infrastructural

---

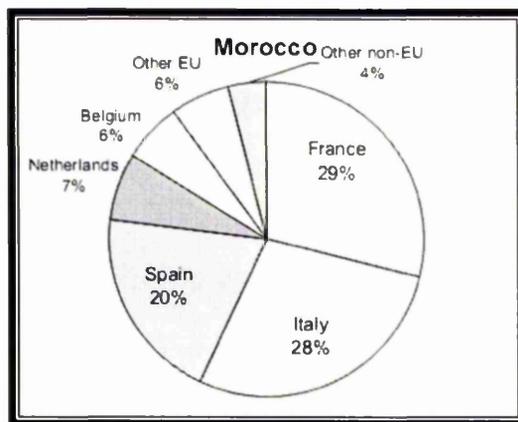
<sup>164</sup> According to Eurostat (2002), migration has affected one in two households in Morocco.

<sup>165</sup> This triggered an increasing number of 'guest workers' from poorer countries around the Mediterranean and until the 1960s, most of whom were recruited in south European countries. When this migration stagnated, attention shifted towards south Mediterranean countries and agreements on the recruitment of guest workers were signed between Morocco and the former West Germany (1963), France (1963), Belgium (1964) and the Netherlands (1969). This was an onset of spatial diversification of Moroccan migration to Europe which used to be mainly towards France. Migration then boomed particularly in 1967 to peak in 1972 (de Haas 2005). From 1991 to 2000, almost 400,000 Moroccans were granted the nationality of an EU member state, larger than any other immigrant group (SOPEMI 2003).

development has encouraged and enabled people to emigrate in large numbers (de Haas, 2005: 2).<sup>166</sup>

Arrivals to Britain were a mixture of mainly unskilled and some professional employees with the average migrant being male and in his twenties or thirties. Furthermore, connected with the young age structure, the average migrant was predominantly single and usually migrated from the family home in Morocco with the migratory trajectory being the first foray abroad. According to Eurostat 2002, the general emigration pattern of sending-countries, in this case Morocco, individual migration, primarily involving men looking for a job or education, followed gradually over time by family reunification migration and family formation migration where male migrants' economic motives dominated while for female migrants, family-related motives were more salient. Figure 5.1 below shows the main countries of migration of Moroccans:

Figure 5.1: Countries of migration in Europe<sup>167</sup>



<sup>166</sup> de Haas, Hein (2005). 'Morocco's Migration Transition: Trends, Determinants and Future Scenarios.' Global Migration Perspectives research papers series No 28. Geneva: Global Commission on International Migration.

<sup>167</sup>[http://www.eds-destatis.de/en/downloads/sif/nj\\_03\\_20.pdf](http://www.eds-destatis.de/en/downloads/sif/nj_03_20.pdf)

Between 1965 and 1972, the start of the influx of Moroccan migration to Europe, it is estimated that the number of registered Moroccans living in Europe increased tenfold from 30,000 to 300,000 to approximately 400,000 by 1975.<sup>168</sup> See Figure 5.2 below:

*Figure 5.2: Evolution of Migrant Stocks in Selected Countries in Western Europe. Sources, El Mansouri 1996 (F, NL, B, G 1968 - 1990); Basfao & Taarji 1994 (IT 1982, 1990); National Statistical Services (B and F 1998; N, G, ES, IT 1998 and 2005); López García 1999 (ES 1968-1990); IOM and Fondation Hassan II 2003 (B and F 2002)*

<b>Evolution of Migrant Stocks in Selected Countries in Western Europe (Registered Population, Regardless of Nationality, Including Second and Third Generations)</b>							
<b>Year</b>	<b>France</b>	<b>Netherlands</b>	<b>Belgium</b>	<b>Germany</b>	<b>Spain</b>	<b>Italy</b>	<b>Total</b>
1968	84,000	13,000	21,000	18,000	1,000	N/A	137,000
1972	218,000	28,000	25,000	15,000	5,000	N/A	291,000
1975	260,000	33,000	66,000	26,000	9,000	N/A	394,000
1982	431,000	93,000	110,000	43,000	26,000	1,000	704,000
1990	653,000	184,000	138,000	62,000	59,000	78,000	1,174,000
1998	728,000	242,000	155,000	98,000	200,000	195,000	1,168,000
2005	1,025,000	316,000	214,000	73,000	397,000	253,000	2,278,000

New Moroccan migratory patterns are emerging which show that between 1980 and 2004, the combined Moroccan population officially residing in Spain and Italy increased from about 20,000 to 650,000. Italy and in particular Spain have replaced France as the primary destination for new Moroccan labour migrants. An increasing proportion of independent labour migrants to Southern Europe are women who work as domestic workers, nannies, cleaners, or in agriculture and small industries. The combined effects of family-related migration, undocumented migration, and labour migration to Spain and Italy explain why Moroccan emigration has persisted in spite of increasingly restrictive immigration policies. The Moroccan migrant stock in Europe and North America has increased almost sevenfold from 300,000 in 1972, on the eve of the recruitment freeze, to at least 2.3 million around

<sup>168</sup> <http://www.migrationinformation.org/Profiles/display.cf.m?ID=339>

2000. This figure excludes undocumented Moroccan migrants, who are likely to number at least several hundreds of thousands of people.

France, given its colonial history in Morocco, is home to the largest legally residing population of people of Moroccan descent (more than 1,025,000), followed by Spain (397,000), the Netherlands (315,000), Italy (287,000), Belgium (215,000), and Germany (99,000). Smaller communities live in the Scandinavian countries (17,000), the United Kingdom (50,000), the United States (85,000), and the Canadian province of Quebec (70,000). Figure 5.3 below shows recent trends on Moroccan migratory patterns and statistics. However, given the rise in unregistered migrants, it is difficult to clearly establish an exact number of Moroccans currently residing in the UK:

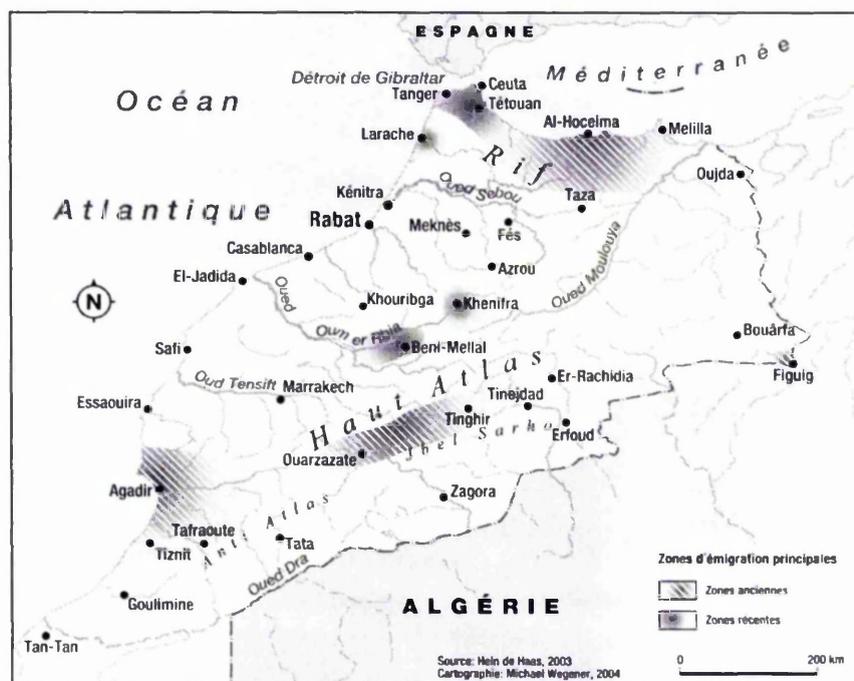
Figure 5.3: Estimates of Moroccan Migrant Stocks Worldwide

Estimates of Moroccan Migrant Stocks Worldwide <sup>169</sup>				
Country	Population registered in Moroccan consulates (2002)	Destination country statistics	Year	Source
France	1,024,766	506,000	1999	IOM 2003:217 (French census; only Moroccan nationals)
Netherlands	276,655	315,821	2005	Statistics Netherlands
Belgium	214,859	204,000	2000	IOM 2003:101
Germany	99,000	80,266	2000	IOM 2003:33
Spain	222,948	396,668	2005	Ministry for Employment and Social Affairs, Spain
Italy	287,000	253,362	2004	Istituto Nazionale di Statistica
UK	30,000	50,000	2000	Collyer 2004
Scandinavia	17,000	-	-	-
Other	13,593	-	-	-
Europe	2,185,821	1,806,117	-	-
US	85,000	21,035	2000	US Census Bureau (only Moroccan nationals)
Canada	70,000	-	-	-
North America	155,000	-	-	-
Libya	120,000	-	-	-
Algeria	63,000	-	-	-
Tunisia	16,414	-	-	-
Saudi Arabia	11,973	-	-	-
UAE	8,359	-	-	-
Other Arab countries	12,216	-	-	-
Arab countries	231,962	-	-	-
Israel	-	270,188	2005	Total immigration 1948-2003 (CBS Israel)
Sub-Saharan Africa	5,355	-	-	-
Other countries	3,959	-	-	-

<sup>169</sup> <http://www.migrationinformation.org/Profiles/display.cf.m?ID=339>

As shown above, in 2000 there were 50,000 Moroccans in the UK, and presently there are approximately 100,000.<sup>170</sup> Moroccan out-migration took hold in certain areas in Morocco and this altered the economic and social dynamic of each particular territory as well as Morocco in general. The figure below shows the major points of out-migration:

Figure 3: Main zones of migration



The main points of migration within Morocco to the UK and Europe are shown to be Oujda and the Rif region in general, pockets of the middle and High Atlas and the Souss region around Agadir. Larache is also a region where a great majority of migrants emigrate. The main regions of migration are Oujda, Larache, and Casablanca and Rabat. This is mainly due to chain migration and other factors such as marriages that take place within the same city in Morocco where British-born Moroccans return to Morocco for the holidays and meet a respective spouse and both parties return to the UK. This is a common phenomenon and has been ongoing since the start of

<sup>170</sup> Based on my conversations with the Moroccan Ambassador in London.

migration to the UK from Morocco in the 1960s, with Moroccans considered the 'preferred' choice as spouses for British-born Moroccans. Even during the current climate, this still continues today with the great majority of spouses emanating from Morocco as opposed to the UK or any other European country of citizenship. However, there are a few people that have married non-Moroccans and/or Arabs from different Arab countries although this is the exception and not the most desired as there is enormous social stigma attached to those who marry non-Moroccans as this is considered non-normative and simply 'wrong' with blame normally assigned to either 'too British an upbringing' or 'parents having gone wrong.'

Data collected for this thesis show the first generation of Moroccans arrived in the UK in the 1960s with the majority originating from three major parts of Morocco; namely, Oujda, the Larache region and Dar al Bayđa (Casablanca). The motive for migration was work-related as manual labour was being offered to Moroccan citizens together with the prospect of a better standard of living and education. Moroccan migration to the U.K. in the 20<sup>th</sup> Century can be divided into four phases (Cherti 2006, personal communication):

- The first wave, which is the most significant, started in the 1960s and was characterised by the emigration of unskilled labourers who originated mostly from the northern part of Morocco, more specifically the Jbala region (Khmiss Sahel, Beni Garfet, Beni Arouss), Larache, Tetouan, Tangiers and the surrounding areas, with a smaller community from Meknes and a larger group from Oujda.
- Family reunification then followed from the early 1970s onwards with mainly wives and young children.
- The third wave started in the 1980s with young semi-skilled professionals and entrepreneurs, mostly from Casablanca, Larache and Oujda.
- The fourth wave started in the early 1990s with the emigration of highly skilled Moroccan professionals both from Morocco itself and France. A large

majority of these most recent immigrants currently work in the finance sector in London.<sup>171</sup>

My corpus for this thesis is made up of migrant and British-born Moroccans that reside in the UK. All speakers I recorded are bilingual with variations in fluency and skill. The first part of data collection took the form of recorded structured interviews of individual informants with questionnaires of precisely targeted questions. The second part of the data collection was that of used naturally occurring conversations which were also recorded based on current events, social talk and asking about families as a main starting point for dialogue. My questions varied in two ways: beginning with questions in Arabic and then the follow up question in English and vice versa. Here generational factors, education and exposure to Arab culture have important bearings on information structure highlighting the emergence of a *Reactive Syntax*. The following section details data corpus of both Phase One and Phase Two, informants, results from questionnaires, research methodology and data collection conclusions.

## 5.2 Data Corpus

The present quantitative research was initiated as a pilot study in 2005 where random informants were selected and conversations recorded in order to establish the extent of code-switching amongst different generations of British Moroccans in the UK. Criteria for the selection of the informants for both Phase One and Phase Two of the recordings of these ethnographic interviews and free-style conversations were stringently geared so as to achieve well-balanced and well-defined corpus data. Structured interviews have been consistently used across all disciplines of the social

---

<sup>171</sup> There is a large and ever-growing Moroccan community in the South-West of England that largely goes unnoticed in surveys. In fact, they are the largest Moroccan community outside London. Data was also recorded by informants from this region.

sciences as a useful tool in eliciting targeted data from select informants. Since the 1980s, different types of interview techniques have been employed and in this research, the focused and structured interview has been chosen as precise code-switched data is sought for a full and detailed data corpus in order to analyse the quantitative material.<sup>172</sup> Patton (1987) suggests three basic approaches to qualitative types of interviews:

*i. The informal conversational interview:*

This type of interview resembles a chat, during which the informants may sometimes forget that they are being interviewed. Most of the questions asked will flow from the immediate context. Informal conversational interviews are useful for exploring interesting topic/s for investigation and are typical of 'ongoing' participant observation fieldwork.

*ii. The general interview guide approach (commonly called guided interview)*

When employing this approach for interviewing, a basic checklist is prepared to make sure that all relevant topics are covered. The interviewer is still free to explore, probe and ask questions deemed interesting to the researcher. This type of interview approach is useful for eliciting information about specific topics. Wenden (1982: 39) considers that the general interview guide approach is useful as it "allows for in-depth probing while permitting the interviewer to keep the interview within the parameters traced out by the aim of the study."

*iii. The standardised open-ended interview*

Researchers using this approach prepare a set of open-ended questions which are carefully worded and arranged for the purpose of minimising variation in

---

<sup>172</sup> For further interview methods, see Hitchcock & Hughes (1989: 79) who list nine types: structured interview, survey interview, counselling interview, diary interview, life history interview, ethnographic interview, informal/unstructured interview, and conversations.

the questions posed to the interviewees. In view of this, this method is often preferred for collecting interviewing data when two or more researchers are involved in the data collecting process. Although this method provides less flexibility for questions than the other two mentioned previously, probing is still possible, depending on the nature of the interview and the skills of the interviewers.

The following sections cite the data analysis procedures in terms of how the informants were recorded, data collection methods, research design and the manner in which the quantitative data were collated as well as transcription methods commencing in the following section with an analysis of quantitative and qualitative approaches to data collection.<sup>173</sup>

### **5.3 Research design – Uses of Quantitative and Qualitative Methods**

Almost all data used by social researchers begins in a qualitative form.<sup>174</sup> It is only after the work has been done on it, to transpose words into numbers that quantitative data come into being (Blaikie 2006: 244). Further, Turner (1994) states that:

We can regard all of the information which we acquire about the world as qualitative, and then see that under some circumstances we can use this information to create a particular kind of data, quantitative data to which the properties of number can be applied (1994: 195).

On the other hand, the differences between quantitative and qualitative data research are not as clear-cut according to Halfpenny (1996), who states that:

Quantitative data is usually produced by coding some other data, which is reduced to a number by stripping off the context and removing context from it.

---

<sup>173</sup> There has been considerable debate about the relative merits of quantitative and qualitative methods, with the protagonists invariably adopting tactics to bolster their own positions and denigrate that of the 'opposition' (Blaikie, 2006: 243 – 244). In this research both methods have been applied to better exploit the data and maximize results.

<sup>174</sup> cf. Blaikie (2006).

Later, after manipulating the numbers, they are interpreted, that is, expanded by adding content and context which enable one to see through the numerical tokens back to the social world (1996: 5).

Some research involves both quantitative and qualitative forms as is the case in this research. This research is quantitative in the sense that it collects Moroccan Arabic and English code-switched data in a survey-based manner with structured and guided interviews and through a coding process as outlined in this chapter, these data are transposed and tabulated with the results manipulated and sorted into a numerical form. This research is also qualitative in the sense that it produces percentages from the data collected and summarizes the data in a non-numerical way, in terms of which code switching theory best suits the switched data, which is the matrix language, and how the MLF model is the correct vehicle in which to interpret the data.<sup>175</sup> The logic is that combining both methods leads to a type of ‘multiple operationalism’ (Webb et al. 1966), ‘combined operations’ (Stacey 1969), ‘mixed strategies’ (Douglas 1976), ‘linking data’ (Fielding and Fielding 1986) and this mixing generates maximum results from data collected from the field.<sup>176</sup> In terms of the data collected for this research it is clear that both quantitative and qualitative methods were used to the mutual satisfaction of the primary data results produced.

### 5.3.1 Transcriptions

The transcribing process during any data collection is both time consuming and arduous even for the most organized of researchers. Before the transcription took place, different transcription systems were studied and borne in mind in order to locate the most suitable system which would enable speed of transcription as well as a

---

<sup>175</sup> See Halfpenny (1996) for further analyses of quantitative and qualitative research methods.

<sup>176</sup> For further information on mixing quantitative and qualitative research methods, refer to the concept of ‘triangulation’ which was introduced into the social sciences by Webb et al. in 1966.

clear and well-defined data bank to refer to and to draw upon for further research. I had decided prior to going to the field not to transcribe the data entirely as this would prove too time-consuming and ineffective if the data were of long stretches of monolingual utterances. Once the data were listened to and analyzed prior to transcription, I decided to use Heath's (1986) transcription system as his book *'From Code Switching to Borrowing – A case Study of Moroccan Arabic'* used a well-structured and comprehensive system for Moroccan Arabic.

Data were then transcribed using the Roman alphabet, italicizing the English utterances as and when they occurred. Given time constraints, I decided to transcribe only the code-switched discourse in each section. This proved effective as it highlighted both the inter-sentential and intra-sentential data. It also proved effective in identifying single 'island' and mono-morphemic (Myers-Scotton 1993b) switches as and when they occurred. The recorded audio material was immediately transferred to the laptop and set in Windows Media Player using Vista for maximum audio quality. This also enabled me to catalogue and store the recordings in individual files for easy access and transcribing purposes. After listening extensively to the recorded audio, a rough transcription was made and directly stored on the laptop with dates of recordings as well as names of informants. Once this was completed, the initial data analysis procedure could commence as detailed below.

### **5.3.2 Initial Data Analysis Procedures**

In terms of data collection, certain factors were borne in mind prior to execution of elicitation and were assessed at the research design stage prior to going out in the field both here in the UK and in Morocco. An intrinsic research plan was established and followed and as Blaikie (2006) cites:

Regardless of whether data are primary, secondary or tertiary, they can come from four different types of settings: natural social settings, semi-natural settings, artificial settings and from social artefacts. Research conducted in a natural social setting involves the researcher in entering an area of social activity and studying people going about their everyday lives. In a semi-natural setting, individuals are asked to report on their activities that occur in natural settings, while in an artificial setting, social activity is contrived for experimental or learning purposes. The fourth kind of social setting is in the past and involves the examination of records or traces left by individuals or groups (2006: 187).

Data collated for this research took the form of primary data in a natural social setting of individuals and small groups. Blaikie (2006) goes on to describe the different types of natural social settings as follows:

“Natural social settings involve three main levels of analysis: micro-social phenomena, meso-social phenomena and macro-social phenomena. These levels vary in scale from individuals and small groups, through organizations and communities, to institutions and large-scale social situations, such as cities and regions and multi/transnational bodies” (2006: 187-188).

In this regard, sources of data were well-scrutinized before going out into the field and the natural social setting and micro-social phenomena type were selected for collection of primary data as this allows for the recording of individuals in their natural, everyday social setting and domains as well as small social units of small groups where informants may feel more relaxed and more inclined to speak freely and naturally code-switch amongst friends and family thus rendering better results of the primary data. A further consideration at the research design stage was that of the demographics of the informants and their selection. The pre-requisites have already been discussed above on the type of informant needed. These demographic characteristics are a very important part of the research design and collated in the social survey aspect of the questionnaires distributed. The variables include gender, age, place of birth, language proficiency, background, education and the like which all have a profound effect on the quality of the primary data elicited.

#### 5.4 Primary Data

A total of 45 hours of primary audio recorded material was gathered during Phase One and Phase Two of the recording sessions. The two phases of recordings were structured so as to maximise the potential of the recordings in order to better utilise the data in different settings and in a structured and non-structured environment. This initial quantitative stage is important as it is located within the research design which includes a set of pre-determined stages, procedures and pre-tested instruments. This approach was exploited in order to allow maximum control over the data gathering and to achieve uniformity in the application of the techniques (Blaikie, 2006: 242). Once this stage was completed, the qualitative stage began with a micro analysis of all data together with a parallel analysis of the MLF model (Myers-Scotton 1993b) and how this compared with the actual data.

Phase One consisted of structured interviews with two separate interview question listings, one with questions aimed at informants born and raised in Morocco and the second aimed at informants born and raised in the UK (see Appendix 1 ) for both listings. The subjects were audio recorded, with the apparatus not clearly visible so as to minimize any disruption or shyness except for a small microphone pinned to the individual's clothing. At the start of each recording in both Phase One and Phase Two, the objective of the recording and the background to the PhD research were again relayed to the informant. Once recording commenced, each informant was asked whether they consented to the recording taking place and to then clearly state their name, age and date of birth. The recordings were transferred directly onto a laptop with the date of the recordings, the respondent's full name and interview number clearly labelled. Before each recording was transcribed, the whole interview / recording was listened through entirely to gauge not only the quality of the recording, but also to obtain a better grounding in the content and approximate level of code

switching. Each recording was then transcribed using English orthography and only where switches took place with any contextually relevant information also added where necessary. Given this set-up, after an analysis of all the data, there were no ambiguities in terms of speaker, content or data in general and this is most probably due to micro planning of the research design. Essentially, certain research points with regards to the data were borne in mind whilst (a) recording the informants, (b) during the transcribing stage and (c) focusing on the possible intra-sentential switch sites. These are listed as follows:<sup>177</sup>

1. Moroccan Arabic to English?
2. English to Moroccan Arabic?
3. Single Moroccan Arabic noun insertions (islands)?
4. Single English noun insertions (islands)?
5. Consistency of Moroccan Arabic as the matrix language?
6. Consistency of English as the matrix language?
7. Consistency of Moroccan Arabic as the embedded language?
8. Consistency of English as the embedded language?
9. Moroccan Arabic affixation to English stems?
10. English affixation to Moroccan Arabic stems?
11. Grammaticality?
12. Potential violations and/or corroboration of the MLF model?

A qualitative analysis was carried out of the data including inter-sentential as well as intra-sentential switches examining the points above as well as a quantitative analysis

---

<sup>177</sup> The above are answered in Chapter Six and Seven of this thesis with examples of the intra-sentential data explained in detail.

in order to account for frequency of switches and details of the informants' responses as per Phase One as outlined below.

#### 5.4.1 Informants Phase One

Hitchcock (1989: 79) stresses that:

Central to the interview is the issue of asking questions and this is often achieved in qualitative research through conversational encounters. (1989: 79).<sup>178</sup>

Given the above, the demographic criteria for the guided interviews were as follows:

- Age – all three generation groups must be represented
- Both genders to be as proportionately represented as possible
- Bilingual or at least a proficient command of Moroccan Arabic as well as English
- Informants from different regions of Morocco
- Only open-ended questions may be asked so as to maximize language responses

Sample questions and responses to the two interview lists as well as the questionnaires are reproduced in Appendix 1. Phase One was structured so as to elicit direct responses from each individual informant in a natural social setting. As a native Moroccan Arabic speaker and an in-group member, this was not only commenced as a quantitative project, followed by qualitative analysis, but also as an introspective investigation which enabled me to apply linguistic intuitions as well as being aware of both English and Moroccan Arabic grammars in terms of the syntaxes employed.

---

<sup>178</sup> See also Cannell, C.F. & R.L. Kahn (1968) 'Interviewing', in G. Lindzey and A. Aronson (eds), *The Handbook of Social Psychology, Vol.2, Research Methods*. New York: Addison Wesley.

In total there were 15 one-hour Phase One interviews, details of which are listed schematically in Figure 5.5 below:<sup>179</sup>

*Figure 5.5: List of informants in Phase One of guided interviews (recorded)*

Conversation #	Name	Sex (M/F)	Age	Place of Birth	Education
1	Bekhta	F	58	Morocco	High School
2	Nourredine	M	44	Morocco	University
3	Katya	F	24	Gibraltar	University
4	Saad	M	56	Morocco	University
5	Ali	M	35	Morocco	High School
6	Youssef	M	25	Morocco	University
7	Jihad	F	15	UK	High School
8	Iman	F	17	UK	High School
9	Mariam	F	14	UK	High School
10	Sara	F	13	UK	High School
11	Naima El	F	36	Morocco	High School
12	Naima H	F	36	Morocco	High School
13	Abdou	M	51	Morocco	University
14	Haja (1)	F	56	Morocco	None
15	Haja (2)	F	59	Morocco	None

As we can see, 66% of the informants in Phase One of the guided interviews (recorded) are female, with 34% male. The youngest age is 13 and the oldest is 59 with the average age of all informants 35.7 years of age. Of the informants, 66% were

<sup>179</sup> For the sake of privacy and anonymity, only first names have been used.

born in Morocco, 34% born in the UK, and one informant was born in Gibraltar and raised in the USA. Also, all three generations are covered from the first generation born in Morocco, second generation born in either Morocco or the UK and the third generation born in the UK. These interviews took place in various locations in the UK, notably the Ladbroke Grove youth centre in London where the majority of third generation interviews were conducted, with the remainder taking place in cafes, parks, restaurants and people's homes as well as during three trips to Oujda, Morocco where informants were recorded although this was deemed not as effective as the data collated in the UK as the informants spoke mainly in Moroccan Arabic and so the focus switched to respondents in the UK. Although these were structured interviews, the onus was on as natural a setting as possible and in order to facilitate this, the recording equipment was not visible during the interviews, but the microphone was attached to the informants' clothing. Each interview started with the informant stating their name, date of birth, age and place of birth. This was as a back-up should the written questionnaires in any way get lost or get mixed up, and also for the identification of the recorded material.

As a separate exercise, the same questionnaires for both British-born and Moroccan-born informants were distributed without recordings taking place. This was carried out in order to achieve a wider distribution of responses to the questionnaires, particularly if informants did not have the time to sit for a guided interview. In this manner, the questionnaire was either completed in front of the researcher, or left with the respondent and returned at a later date. A schedule of questionnaires distributed was designed so as to minimize any potential confusion of the data once returned. In total there were 25 completed non-recorded questionnaire details which are listed schematically in Figure 5.6 below:

Figure 5.6: List of informants in Phase One of questionnaires submitted (not recorded)

Conversation #	Name	Sex (M/F)	Age	Place of Birth	Education
1	Asma	F	26	UK	University
2	Samantha	F	29	UK	University
3	Iman	F	24	UK	University
4	Souad	F	47	Morocco	College
5	Yahya	M	49	Morocco	University
6	Fatima H	F	42	Morocco	College
7	Moona	F	34	Morocco	College
8	Laila	F	22	UK	College
9	Lamyae	F	30	Morocco	College
10	Hanadi	M	27	Morocco	College
11	Nizare	M	30	Morocco	College
12	Moona	F	16	UK	School
13	Si Mohamed	M	35	UK	University
14	Mariame	F	22	UK	College
15	Hisham	M	34	Morocco	High School
16	Fouad	M	34	UK	College
17	Yahya R	M	40	Morocco	College
18	Nordine	M	36	UK	College
19	Yamina	F	40	Morocco	College
20	Bechir	M	34	Morocco	High School
21	Nora	F	34	UK	College

22	Zahra	F	31	Morocco	College
23	Hooda	F	24	UK	College
24	Hooda (2)	F	22	UK	College
25	Mohamed B	M	28	UK	College

As we can see, 60% of the informants in Phase One of the guided interviews (non-recorded) are female, with 40% male. The youngest age is 16 and the oldest is 49 with the average age of all informants 31.6 years of age.<sup>180</sup> Of all the informants, 52% were born in the UK, and 48% born in Morocco. Also, all three generations are covered from the first generation born in Morocco, second generation born in either Morocco or the UK and the third generation born in the UK. All questionnaires of both the recorded and non recorded Phase One of the data were returned and processed. All in all, a fair representation of all age groups, genders and generations was incorporated in the data collection and research. The results of Phase One of both the recorded and non-recorded guided interviews questionnaires are detailed in the following section.

#### 5.4.2 Informants' Language Proficiency

All informants spoke Moroccan Arabic. The first generation were the most fluent, with second and third generations having bilingual fluency and/or being very proficient. Some informants felt their Moroccan Arabic was lacking and preferred to speak English when the questions in the interviews were in Moroccan Arabic. This is mainly due to self-consciousness as all informants are more than proficient and had no problems with any linguistic intelligibility. The majority of the first generation Moroccans born in Morocco came to the UK with no English and learnt it 'on the job'

---

<sup>180</sup> Given the preferred demographic criteria of informants during the research design stage, it was hoped that there would be as close to 50% of both genders as possible. However, later analysis shows that this was nearer to a ratio of 2:3 in favour of female respondents.

whilst at work. One informant recounts the story of how she learnt to count in English by listening to two school girls playing outside her window and skipping, shouting one, two, three, four etc... This was how the informant learned to count at age 23. Second generation Moroccans are more fluent in Moroccan Arabic than the third generation although on a case by case basis, the results vary due to exposure of Moroccan Arabic at home and on visits to Morocco. The co-existence of both languages takes place in the domestic domain from birth, with parents, more specifically mothers, speaking to their offspring in MA. It is not until children go to school that they learn English or while watching cartoons on television. However, at school, language use takes on a new dimension whereby in the linguistic continuum, the sequence of language use is such that English starts to take over as the dominant language during the primary period of language acquisition.<sup>181</sup> As Boumans (1998) states:

The respondents' competence in Moroccan Arabic and Dutch is obviously related to the amount of time spent in each country and to their age upon arrival...Generally speaking, the younger a person was when he or she emigrated, the better (s)he acquired the new language (1998: 161).

It has been suggested in the literature that the more proficient the bilingual speaker, the greater the tendency to switch, most particularly intra-sententially and at a wider variety of permissible sites (Berk-Seligson 1986, Poplack 1988, Poplack et al. 1988, Treffers-Daller 1994). It is also suggested (Poplack 2004) that those who are less proficient in one of the two languages do not eschew code switching altogether, as might be the case were code switching not the eminently social tool that it is, but rather restrict their code switching in number, type and/or discourse location according to their bilingual ability. The less proficient thus favour switch sites and

---

<sup>181</sup> See also Saffran, J. R. Aslin, R. N. & Newport, E. L. (1996) *Science* 274, 1926-1928, Pinker, S. (1994) *The Language Instinct* (Morrow, New York).

types requiring little or even no productive knowledge of the other language, such as tags, routines or frozen phrases. Bilingual proficiency is no condition for code switching. Rather, given the appropriate discourse and social circumstances, speakers who engage in the most complex type of intra-sentential code switching turn out to be the most proficient in both of the contact languages (2004: 13).<sup>182</sup> This is the norm given bilingual proficiencies, as the greater the fluency in both languages, the greater the code-switching potential and complexity of intra-sentential code-switching varieties. The results of Phase One of all questionnaires, recorded and non-recorded sessions are detailed in the following section.

### **5.5 Informants' Questionnaire Results**

In order to clearly ascertain the language attitudes to code-switching and the use of Moroccan Arabic and English in general by the informants, the investigation into the Moroccan community in the UK was conducted to lend currency and overtly marry the results of both the recorded and non-recorded sessions of Phase One and new code-switching data. The questionnaires were designed and questions selected which would best maximize informants' responses and give a better analysis of the Moroccan community in the UK and its linguistic norms.<sup>183</sup> As has been seen previously, it is predicted that the linguistic habits of the first generation are largely stabilized with new innovative and emerging forms witnessed amongst some second generation and certainly third generation informants. Linked with this is a survey of informants' language attitudes towards their own language use, linguistic repertoire, language attitude and code-switching in general. As Bentahila (1983) notes in his analysis of Moroccan French bilinguals:

---

<sup>182</sup> For more on code switching and bilingual proficiency, see Poplack, S (2004) 'Code-switching. Soziolinguistik.' In, U. Ammon, N. Dittmar, K.J. Mattheier & P. Trudgill (eds.) *An International Handbook of the Science of Language*, 2<sup>nd</sup> edition. Berlin: Walter de Gruyter.

<sup>183</sup> See Appendix for questionnaires used during Phase One of research.

The Moroccan bilingual is familiar with both French and Arabic [sic] cultures. Naturally, he will look upon the French language as one aspect of the general concept he has of “Frenchness”, and will tend to associate it with other aspects of French culture; and similarly, Arabic will be associated with other parts of the Arab cultural heritage. It would seem quite natural, then, that when a bilingual adopts one aspect of French culture, namely the French language, other aspects of this culture will also tend to become uppermost in his mind (1983: 48).

With regard to the present study, all Moroccans, whether British-born or Moroccan-born associate themselves with both British and Moroccan cultures, with first generation Moroccans associating themselves more with the deep-seated Moroccan-ness than third generation Moroccans. However, there is a trend of third generation children associating themselves more and more with the ‘bled’ or motherland with frequent trips to Morocco, a strong interest in Moroccan popular culture as well as a revival of learning written Arabic at British secondary schools. Statistically, the following demographics ensued from completed questionnaires distributed:

*Figure 5.7: Place of birth?*

Morocco	UK	Other
55%	40%	5%

*Figure 5.8: Level of education?*

School	College / University	None
22.5%	70%	7.5%

*Figure 5.9: Language situation at home with parents?*

Moroccan Arabic	English	Both
37.5%	15%	47.5%

Figure 5.10: Language situation at home with siblings?

Moroccan Arabic	English	Both	French / Spanish
35%	30%	32.5%	2.5%

Given the above, the results of the remaining questions on all questionnaires, 40 in total are as outlined below.

Figure 5.11: Languages used outside home?

Moroccan Arabic	English	Both
42.5%	42.5%	15%

Figure 5.12: Which languages do you prefer to use in general?

Moroccan Arabic	English	Both
40%	57.5%	2.5%

Figure 5.13: In which language are you more confident?

Moroccan Arabic	English	Both
32.5%	55%	12.5%

Figure 5.14: Do you usually use both languages at the same time?

Yes	No	Sometimes	Rarely
35%	5%	37.5%	22.5%

Figure 5.15: Do you switch from MA to English or English to MA?

MA to English	English to MA
57.5%	42.5%

Figure 5.16: Why do you usually switch from one language to another?

Subject	Incompetence	Audience	Family
62.5%	32.5%	2.5%	2.5%

Figure 5.17: From which language do you usually switch?

MA to English	English to MA
40%	60%

Figure 5.18: Why do you think people switch between the two languages?

Well-skilled	Incompetent	Indifferent
62.5%	15%	22.5%

Figure 5.19: Are you for or against people using two languages at the same time?

For	Against
75%	25%

Figure 5.20: Classify the following languages in order of importance?

Moroccan Arabic	English	Modern Standard Arabic
45%	45%	10%

Figure 5.21: Which language do you class as most important in your daily life?

Moroccan Arabic	English	Other
60%	27.5%	12.5%

Figure 5.22: Would you prefer to be taught in Moroccan Arabic or English?

Moroccan Arabic	English	Other
67.5%	25%	7.5%

Figure 5.23: Which language do you see as important for your future (children)?

Moroccan Arabic	English	Other
47.5%	25%	27.5%

Respondents were asked to select the language and/or response which best reflected their current usage of Moroccan Arabic and English and attitude to the respective

languages in general using the space available in the questionnaires. All results are listed as percentages, using the variable  $(100/40 [2.5\%] \times n)$ , of the total number of respondents, namely 40, those that were recorded in Phase One and those which were not recorded and filled in the questionnaires. All questionnaires handed out were returned and completed, which renders this a successful outcome.

### **5.5.1 Results of Phase One**

Given the linguistic repertoire of the informants as presented in the demographic data at the beginning of the above section, it is clear that all informants make daily use of both languages and that there is no sign of language attrition amongst second and third generation bilingual speakers. The speakers' ability to choose the appropriate variety for any particular purpose is part of their communicative competence; the choice is not random, but has been shown to be determined by aspects of the social organization of the community and the social situation where the discourse takes place (Bentahila 1983: 50). Fishman (1972) also discusses the aspect of language choice amongst bilingual communities which he defines as: "Major clusters of interaction situations that occur in particular multilingual settings" (1972: 19).

The results as presented above portray an array of conclusions in that 55% of the informants were born in Morocco with an unexpected 70% as holders of a university or college degree. This result was expected to be much lower. When asked about languages used at home with parents, 37.5% only used Moroccan Arabic and 47.5% used both languages, which shows how English has penetrated the domestic domain, whereas only 35% use Darija with siblings, the norm being English at 30% or both at 32.5%. This would largely depend on subject matter and where the conversations take place. Also, language proficiency should be considered as some speakers as shown in the results in Figure 5.13 shows that 55% are more confident in

speaking English than Moroccan Arabic and 35% in speaking Moroccan Arabic. Surprisingly, 42.5% of the respondents use MA outside of the home, and this is perhaps amongst first generation Moroccans who meet with Moroccan friends and converse in MA. In terms of language attitude, the results are quite clear and show that 57.5% prefer to use English in general and this may be due to confidence and/or language prestige, with perhaps English rated higher than MA by certain respondents in this regard. In terms of code switching, only 35% of the respondents confirmed they code-switched, with 5% citing they do not and a surprisingly high figure of 22.5% showing that they rarely code switch. This is due to the fact that bilingual speakers are sometimes unaware that they are in fact code switching with some even unfamiliar with the term until it was pointed out to them. Direction of switches shows that this is fairly even, with 57.5% switching from MA to English. This is from the more fluent bilingual speakers' use as the natural direction is to switch back to the language of fluency and proficiency. The following examples highlight directions of switches:

MA → English

(214) *š-hel min merra na-gul -u* and he doesn't want to bother  
 How from time tell 1SG him and he doesn't want to bother  
 'How many times I tell him and he doesn't want to bother'

(215) *mezel ka-mši beṣaḥ ma y-come -š back ḥta lil*  
 Still goes 3SG but NEG1 come 3SG NEG2 back until night  
 'He still goes but doesn't come back until late'

Interestingly, the data show that 62.5% switch due to subject matter with 32.5% citing language incompetence as the main factor and an expected low figure of 2.5% attributable to family and/or audience. Similarly in respect to language attitudes to code switching, 62.5% consider switching between languages as a well-skilled trait, with 15% stating it is due to language incompetence, and a high figure of 22.5%

shows indifference – this further confirms that some of the informants are not even aware of (a) the notion of code switching and (b) their own linguistic repertoire and code switching capabilities. However, 75% of those questioned were in favour of code switching citing the following selected reasons:

1. It is interesting
2. A better way of communicating
3. It enriches the content
4. Shows fluency
5. A symbol of bi-nationalism
6. Facilitates communication with other bilinguals
7. One can better express certain ideas using both languages
8. It is better to speak two languages
9. Shows good language skills

Conversely, 25% were against code switching, citing, amongst others, the following reasons:

1. One does not master one language properly
2. Language quality is very important and code switching dilutes this
3. Code switching is a bad language habit
4. It gives an impression of linguistic division
5. It can be considered inconsistent
6. It shows a failure of being able to manage any one language
7. One should be able to use individual languages competently
8. It is confusing

Although 45% of the informants consider English as most important as well as 45% Moroccan Arabic, 60% consider Moroccan Arabic as most important during their daily lives which, given that the respondents are mainly based in the UK is an unexpectedly high figure. This shows then that MA use in the domestic domain is considered very important and is frequent. In terms of future usage, 47.5% consider Moroccan Arabic important for future language use and 25% English which is quite low given the wide use of English on a global scale. It can be concluded that there is perhaps a romanticized notion of Moroccan Arabic amongst the Moroccan diaspora. Therefore, amongst Phase One of the structured research, the data show no real unexpected results or any surprising conclusions which is in line with generalizations made for the Moroccan community in the UK amongst gender, age groups and generations. Phase Two of the data collection is analyzed in the next section.

### **5.5.2 Informants: Phase Two**

Phase Two of data collection took the form of informal conversations in relaxed atmospheres so as to make as natural an environment as possible. This type of research method in essence supports Phase One as it lends itself to different code-switched styles whereby the informant is more inclined to speak at will and in a non-focused way, although I guided the conversation as and when necessary or spoke in Moroccan Arabic if too much English was being used and vice versa. Maykut & Morehouse (1994) state that the data of qualitative inquiry is most often peoples' words and actions, and thus requires methods that allow the researcher to capture language and behaviour. The most useful ways of gathering these forms of data are participant observation, in-depth interviews, group interviews, and the collection of relevant documents. Observation and interview data is collected by the researcher in the form of field notes and audio-taped interviews, which are later transcribed for use

in data analysis (1994: 46). A total of 30 hours recorded material was collated and transferred to a laptop. This material was separated and stored in different files from the guided interviews in order to facilitate data analysis procedures. It was found, as Boumans (1998) also explains in his data collection analysis that the data are heterogeneous in many ways and the variation is apparent in the sociolinguistic backgrounds of the respondents, and their speech behaviour in terms of language choice.

It is clear that first generation Moroccans prefer to speak in Moroccan Arabic as this is considered their native language and even though the great majority of first generation respondents have been in the UK for well over 35 years, they still sway towards Moroccan Arabic in natural language usage even if they start in English. In terms of second and third generation speakers, again there is a predictable result with the majority giving English as their preferred language. Also, in the recorded material, it is evident that questions posed to these two generation groups even when commencing in Moroccan Arabic, eventually returns to English as this is the language some of them feel most confident speaking or feel more familiar with in general.

It emerges then that, as Myers-Scotton states that in an analysis of second generation shifts in socio-pragmatic orientation and code switching patterns some speakers either do not have access to the grammatical structure of the variety of the matrix language or the new language is turning over to a new matrix language. That is a new composite matrix language is beginning to structure their code switching (Myers-Scotton 2002) in terms of morpho-syntactic development and this is in line with the notion of *Reactive Syntax* as discussed in previous Chapters.<sup>184</sup> This is of essential importance within the domain of bilingualism, sociolinguistics, psycholinguistics and general linguistics. As Rouchdy (2002) claims, the code

---

<sup>184</sup> See also Rouchdy (2002) with regards to language shift amongst second generation Arabic speakers in general.

switching patterns of generation two do indicate that their sociopragmatic orientation is different from their parents and it is evident that there is a composite matrix language that is being formed in bilingual intra-sentential CPs.

## **5.6 Conclusion**

This chapter then begins with an analysis of migration patterns from Morocco to Europe from the 1960s onwards with data sourced from various publications (de Haas 2005) as well as primary data. This is then followed by an outline of the classification of the data corpus of both British-born and Moroccan-born informants and the selection criteria which are essential in order to maximize and exploit a good data bank of material for this thesis as well as for further research. Approaches to research design are explained with explanations of the two methodologies adopted, namely both quantitative and qualitative methods which are mutually compatible given the data collated. The corpus is analyzed in depth with both Phase One and Phase Two of the audio recorded (and non-recorded) material tabulated and results presented and explained in full. This is an integral part of the thesis as it allows for a demographic break-down of the data and the tabulated forms facilitate referencing. On reflection of the data collected, informants' language attitudes are discussed as this contextualizes the code switched material and informs us of the micro-world of the bilingual speaker and his/her approach to both English and Moroccan Arabic in general as well as code switching as a bilingual tool. With the data classification firmly established, the following Chapters, Six and Seven analyze the code-switched data syntactically with all grammatical categories analyzed with particular emphasis on generational factors and a detailed report of the data as listed in the tables and figures presented in this chapter.

## CHAPTER SIX

### GRAMMATICAL CATEGORIES OF CODESWITCHING FROM ENGLISH INTO MOROCCAN ARABIC

Following on from Chapter Three where the most prominent syntactic approaches to code switching were examined (Joshi 1985, Azuma 1990, 1993 and Myers- Scotton 1993b, 2000, 2002), this chapter focuses on the MLF model and its application to intra-sentential code-switched data from English to Moroccan Arabic and most grammatical categories not covered in the previous chapters incorporating lexical insertions and the placement of diminutives within code-switched data and is evidenced whereby speakers employ certain discourse strategies. Further, a new approach to grammatical gender is employed where concord relies upon its semantic as opposed to phonological alignment. This is discussed in detail as there is hardly any literature on this innovative and interesting approach which I have discovered and analysed in my own data.

The periphrastic ‘do’ or *Dar+* construction is then examined with numerous examples cited as this is one of the most frequent types of verb insertion. It is in this regard that firm advocacy of the MLF is highlighted with salient examples validating the MLF model. However, its constraints and its limitations are also investigated. In this manner, the empirical validity of the model is evaluated and examined alongside other research on code-switched material in Moroccan Arabic and other European languages (Bentahila & Davies 1983, Boumans 1998 and Nortier 1990). In order to frame the data presented, it is necessary to re-capitulate the most important aspects of the MLF model,<sup>185</sup> namely the Matrix Language Principle where there is always an analyzable or resolvable frame structuring the morpho-syntax of any CP and the claim

---

<sup>185</sup> cf. Myers-Scotton (2002: 9).

that in bilingual speech, the participating languages never partake equally as the source of the ML. With the Uniform Structure Principle, a given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for constituent type must be observed whenever the constituent appears. The Asymmetry Principle is where bilingual speech is characterized by asymmetry in terms of the participation of the languages concerned, in this respect Moroccan Arabic and English. It is evident given the data presented in this thesis thus far that only one of the participating languages can provide the grammatical frame in a micro CP unit as larger stretches of discourse can switch between MA as the matrix language followed by English in a separate CP clause and back to MA as has been shown previously. As Myers-Scotton (2002) discusses, this asymmetry is evidence of the universal drive in language to achieve uniformity in the structural frame of any variety (2002: 9).

The first section then of this chapter details the grammatical categories listed above as used by respondents in their matrix language or embedded language clauses in both copula and nominal sentence constructions. The larger part of the chapter validates the application of the MLF model which is evidenced through lexical insertions, the most frequently occurring category which is integrated into MA or English clauses. Such embedding of MA affixations on English verbal stems is a common strategy amongst informants and code-switchers alike and this is discussed throughout the body of the chapter commencing with nominal constituents and adjectives recorded by informants followed by an analysis of the major grammatical categories in code-switched data from English to Moroccan Arabic. The converse direction, Moroccan Arabic to English, is discussed in the following chapter.

## 6.1 English Lexical Insertion in Moroccan Arabic Clauses

The below represents a Moroccan Arabic definite NP construction in its simplest form compared with an English NP construction:<sup>186</sup>

Figure 6.1: Moroccan Arabic Syntax

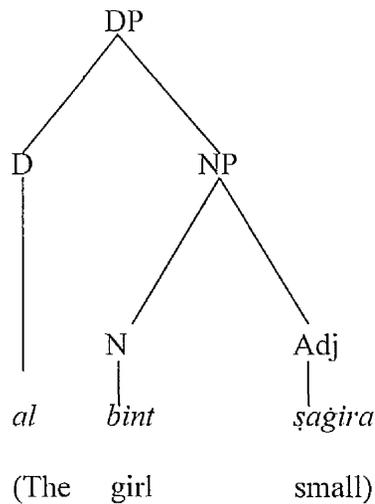
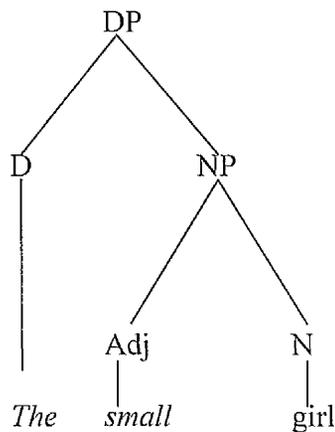


Figure 6.2: English Syntax



Therefore, given the syntax of Moroccan Arabic and English, single lexical insertions in code-switched data are the most frequently switched of all grammatical categories and this is attributed to the relative ease of their insertion (cf. Nait M'Barek & Sankoff 1988). Interestingly, single lexical insertion has no real bearing on language proficiency in the given language(s) as these are not intrinsically complex within the

<sup>186</sup> cf. Caubet 1993, vol. II Chapter III for an in-depth illustration of the use of Arabic definite articles.

grammatical structure of any given CP unit. This perhaps explains why they are the most switched category in bilingual discourse. This universal observation of single item insertion most of which are nouns has been discussed by Poplack (1980) where examples are given in Spanish and English code-switched data:

(216) But I used to eat the *bofe*, the brain. And they stopped selling it.  
(*Spanish/English*, Poplack 1980: 597)

For many researchers, single lexical insertion has raised the oft-quoted quandary of whether this is actually code switching at all or in fact lexical borrowing. This discussion is still on-going from Myers-Scotton's (2002) distinction between cultural borrowing and core borrowing and Poplack et al.'s (1988) quantitative analysis of borrowing frequencies. As discussed previously, this thesis concerns itself with the code-switched approach to code-switched lexical insertion and not with the concept of borrowing. The example below from Swahili and English data shows the single lexical insertion of customers amidst a Swahili CP phrase which is not necessarily a cultural borrowing as the noun 'customers' clearly exists in Swahili but is in fact a conscious choice of the speaker:

(217) Tuna *customers*, wengi sana kwa mpango huu  
'We have many customers in this plan'  
(*Swahili/English*, Myers-Scotton 1997: 72)

Data from Moroccan Arabic and French as examined by Nait M'Barek and Sankoff (1988) showed that constituent insertion accounted for a large number of unidirectional switches at equivalent sites and such insertions were mainly French NPs.<sup>187</sup> It can be said that single noun insertion is always 'additive' in the sense that

---

<sup>187</sup> In this documented study, there are ten times as many NP switches in all as there are switches at the equivalent site between the Arabic determiner and the French noun. Constituent insertion in research on Moroccan Arabic and Dutch was not referred to by Nortier (1989), which confirms that constituent insertion is dependent on the community rather than on the language typology itself (Poplack 2004).

affixes can be added to single nouns and not reductive as the bare form of the noun must be maintained and not reduced.

Nouns sometimes take with them certain grammatical elements, in particular the definite and indefinite articles as well as their modifiers such as adjectives (Auer 2002). The below illustrates the possible indefinite article additions in the general Moroccan Arabic syntax:

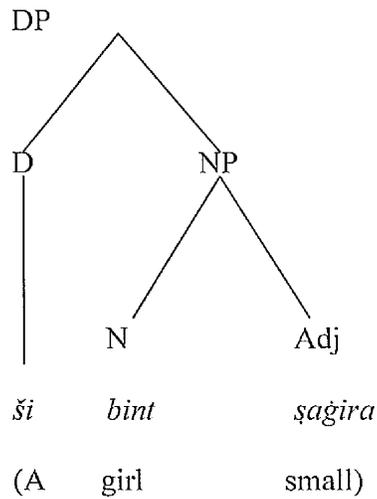


Figure 6.3: Moroccan Arabic Syntax

Figure 6.3 represents a basic Moroccan Arabic indefinite NP construction where the adjective follows the noun *bint* ‘girl’. As previously discussed in Chapter Four, Moroccan Arabic nouns are marked for gender and are usually overtly marked with the addition of a final *-a* for feminine nouns whilst most other nouns are masculine. Such nouns have fixed gender and in contrast with English nouns agree with adjectives and show plural and singular marking. English has no overt gender marking and single lexical insertions in MA intra-sentential discourse are both frequent and straightforward:

(218) *ma ʔndi -s time beš na-ħuk nif -i*  
 NEG1 have NEG2 time so that scratch 1SG nose POSS  
 ‘I don’t have the time to scratch my nose (no time for anything)’

(219) *li kəmel first, y-wash up*  
 REL finish first, wash up 3SG  
 ‘Whoever finishes first, washes up’

The insertion of English nouns in Moroccan Arabic matrix clauses is one of the most frequently switched of categories and given that the respondents are based in the UK, evidence suggests that the frequency of the direction of the switch is from Moroccan Arabic to English, most particularly amongst second and third generations. Note how the examples are in line with MLF model, where the matrix language provides the grammatical frame. On this, there is no doubt as even on a micro-CP level, the matrix language always provides the syntactic skeleton in which the embedded forms align. There is no real categorization of the types of nouns inserted as these point to ease of utterance on behalf of the speaker whereby words are inserted as and when the speaker wishes to switch. However, complex nouns such as technical words associated with the sciences, media, etc. are also switched due to lack of immediate equivalence in Moroccan Arabic. The following examples highlight this:

(220) *mšit ʔamma for check-up u ma gel u li walu*  
 Went there for check-up and NEG tell PL to me nothing  
 ‘I went there for a check-up and they didn’t tell me anything’

(221) *iwa guli l l teacher ya-aʔt ik exercise book jdid*  
 Well tell to DEF teacher give you 3SG exercise book new  
 ‘Well tell the teacher to give you a new exercise book’

In (221) above, the speaker omits the article ‘a’ when referring to ‘a check-up’ and this has been noted by other scholars with respect to loanwords from Berber and Hispanic Romance (Colin, 1945: 232; Harrell, 2004). This is referred to as the

indefinite zero article and is a frequently omitted article in intra-sentential code-switched data.

### 6.1.1 The Indefinite Zero Article

The indefinite zero article, or absence of overt marking, must not be confused with those cases where the definite prefix fails to surface. One of the contexts where the zero article is to be expected is the predicate in copula constructions.<sup>188</sup> The following examples further highlight this phenomenon:

- (222) f had 1 marħala ka-y- welli *communist*  
In DEM DEF stage become 3SG communist  
'At this stage one becomes a communist'  
(Moroccan Arabic / Dutch, Boumans, 1998: 186)

- (223) *ila bğit man, namši barra wa na-jib xemsa*  
COND want man, go 1SG outside and bring 1SG five  
'If I wanted to get a man, I just have to go out and I'll get five of them'

Boumans (1998) offers no explanation for this phenomenon as in his example, listed as (223) above. It can be said, in light of the data obtained for this research, that the majority of indefinite zero articles are uttered by either (i) first generation speakers and (ii) speakers whose first language is Moroccan Arabic and have a less proficient grasp of English.<sup>189</sup> Further, this is a direct translation of the Moroccan Arabic which does not require a definite article in the conditional sentence as evidenced in (223) above. This would then explain why this indefinite zero article is used more frequently by first generation speakers than by second and third generation speakers as statistically, the former are more proficient in MA than the latter and are more

---

<sup>188</sup> Caubet (1993 II: 260).

<sup>189</sup> I differentiate between (i) and (ii) as some first generation speakers have been in the UK for over 40 years and are clearly fluent in English. However, some are not and have spent the large majority of their time conversant in Moroccan Arabic at home, using limited English outside the domestic arena.

likely to translate grammatical constructions.<sup>190</sup> This is both an important observation and distinction. Compare example (223) above with the following second generation speaker in (224) below who replies to the mother's sentence with no sign of the indefinite zero article but rather makes full use of the English conditional clause:

(224) No Mum, if I wanted a man, *na-mši barra wa na-jib ḡšrīn!*  
No Mum, if I wanted a man, go 1SG outside and get 1SG twenty!  
'No Mum, if I wanted a man, I would go outside and get twenty!'

Therefore, it can be predicted that:

*Generalization 6:*

*Indefinite zero articles will be mainly used by first generation speakers or those with a less proficient command of an L2 variety.*

In general in Moroccan Arabic, employment of the definite article is invariable for all genders and numbers. It is assimilated completely to a following stem-initial alveolar or postalveolar coronal consonant (Nortier 1990: 22) as per the examples below:

(225) *l- bāb*  
DEF door  
'The door'

This is mirrored in intra-sentential discourse with English noun insertion:

(226) *balḡ i l- door mur -ak*  
Close 2SG DEF door behind PRON  
'Close the door behind you'

There is often a case for mirroring of MA syntactic structures by bilingual speakers reasons for which are clear as the matrix language is the MA, hence the

---

<sup>190</sup> The English article system poses real problems for Arabic speakers in general whose native language sometimes uses some definite articles in a different manner. See also Kharina (1981) and Al Fotih (2000) who describe problems in L2 acquisition of English definite articles due to language interference.

English noun insertion has to follow the grammatical frame. Definiteness in MA is also borne out through geminate consonants where it only occurs in the MA and is not ‘carried over’ into the English:

(227) *r – riħ*  
DEF wind  
‘The wind’

In (227) above, there is no possibility of code switching any of the morphemes between the definite consonant *r* and the noun itself. It can be said then that in definite DP phrases, no switching can occur when the initial consonant is a geminate. The only possible alternative would be to have a double definite article construct where there is the English definite article + the definite geminate consonant + the noun:

(228) the *r – riħ*  
DEF1 DEF2 wind  
‘The wind’

In this regard, when uttered, the geminate consonant and first definite article would be elided into the first geminate *r* of *riħ* and viewed as a basic noun without the definite article. Compare this with previous examples cited throughout this thesis where it has been shown that switching may occur after a definite article *al* car ‘the car’ and this is a frequent occurrence across all generation groups. This is an interesting discovery in the data collated for this research as it highlights certain discourse strategies that speakers employ in order to circumvent any potential syntactic conflict or difficulty. The following section further highlights other discourse strategies used by Moroccan Arabic bilingual speakers in the use of Moroccan Arabic grammatical gender in an intra-sentential code-switched domain.

## 6.2 Grammatical Gender

In this section, the use of grammatical gender is examined analyzing the switching between a modifier and a noun in both Moroccan Arabic and English code switching and how the gendered Moroccan Arabic versus ungendered English dynamic allows accommodation for a form of convergent flexibility.<sup>191</sup> Here, in MA and English bilingual utterances, if an English noun ends in an /a/ the speaker will assign an MA feminine marker /a/ in the adjoining adjective (Adj) as in *kamera kbīra* ‘big camera’. The same can be said if the English noun is devoid of an /a/ final sound as in ‘coffee’ which will then be marked for masculine in its adjectival gender in the MA as in *coffee zwīn* ‘nice coffee’. Here then, phonological shape accounts for grammatical concord and this is generally agreed upon by speakers in natural conversation<sup>192</sup>. However, sometimes such phonological gender assignment is not always as transparently defined. Sometimes MA speakers instead select a semantic frame application to nouns, irrespective of phonological shape and such semanticization is transferred across from the MA L1. Such assignment rules force the speakers in bilingual conversation to somehow ‘agree’ on noun assignment giving rise to either phonologically or semantically attributed MA adjectives and at times, there is feminine-masculine-feminine alternation until the end result is ‘levelled’ and agreed upon by both speakers. The main criterion here is absolute grammatical agreement which is essential in MA discourse and it is found that code switchers have strong grammatical intuitions of masculine or feminine-specific lexicality.

Hockett’s (1958) definition of gender is that of classes of nouns reflected in the behaviour of associated words (1958: 231). However, this thesis further analyses the physical loading of gender with phonological markings as well as semantic

---

<sup>191</sup> See Benchiba (2007).

<sup>192</sup> cf. Ingham (1994) for further explanations on number, gender and concord in Najdi Arabic.

mapping irrespective of associated words. It is only recently that grammatical gender has been considered and researched as a core linguistic category. Early work addressed the linguistic category of grammatical gender (Arndt 1970, Poplack 1982, Radford 1988) focusing on the host language semantic equivalent and links with phonological shape of host language words. Poplack's research denied that micro language variation amongst a single speech community exists and that there is a general unanimous agreement amongst speakers (1980: 25). Many approaches to grammatical assignment often attempted to develop gender assignment rules as in Tucker, Lambert & Rigault (1977) where in describing gender assignment for French, simplified it as being reflective of the natives speaker's competence in being able to trace back the noun from its terminal phone.<sup>193</sup> In a paper comparing Arabic and English speakers, Clarke, Losoff, McCracken & Still (1981) asked Arabic-speaking informants to assign genders to a series of English nouns and found assignment of the masculine gender was given to nouns which are masculine in Arabic.

In this thesis semantic mapping with concept association is a frequent occurrence amongst Arabic and English speakers, but is more transparently evident amongst fluent or first generation speakers of Arabic. It is evident that grammatical gender assignment is one of the last aspects of second language acquisition to be mastered. Whether it is *le terre* or *la terre* in French is often a stumbling block for L2 learners. It appears, then that the mastery of grammatical gender is an indicator of language fluency<sup>194</sup> and part of the native speaker's competency. Many researchers have investigated this grammatical category in L2 acquisition (Fransechina 2005) and have found that transparent features on N heads are attainable throughout one's life and not subject to a 'critical period' for learning a language. The main criteria are feature mapping, semantic or phonological assignment and their faithfulness to the

---

<sup>193</sup> Tucker, Lambert & Rigault (1977: 62).

<sup>194</sup> cf. Hawkins & Chan (1997) for their failed functional features hypothesis

usual settings and parameters of Universal Grammar. This thesis then includes an examination of grammatical gender amongst British Moroccan speakers and their faithfulness to either phonological or semantic assignment systems and their assignment choice in terms of assignment systems and the compromise default of levelling. This is analyzed in the following section in terms of gender contexts.

### 6.2.1 Grammatical Gender Contexts

In monolingual MA discourse, the noun and adjective have to agree and concord identically maps feature specifications of an adjective with the noun it modifies<sup>195</sup>. Lexically referential equivalents are also to be found in human nouns and most obviously in kinship terms, *wālid* (m) ‘father’ and its equivalent *wālidā* (f) ‘mother’, *xāl* (m) ‘maternal uncle’, *xāla* (f) ‘maternal uncle’ with other non-kinship terms being modified for the feminine with the morphological addition of *-a* such as *mumarrið* (m) ‘nurse’ and *mumarriða* (f) ‘nurse’, *usted* (m) ‘teacher’ and *ustedā* (f) ‘teacher’ therefore highlighting formulaic feature mapping:

(229)

Feminine -----	Feminine	}
Masculine -----	Masculine	
Singular -----	Singular	
	<i>kalb</i> <i>kbir</i> dog                              big noun (MASC)    Adj (MASC) SG                              SG	

Switching between a modifier in language A and a noun in language B is becoming increasingly common across all generation groups of bilingual MA/English speakers and grammatical gender switching is prevalent amongst British-born *and* Moroccan-born bilinguals irrespective of fluency. Concord in MA identically maps the above feature specifications of an adjective with the noun it modifies where feminine nouns

---

<sup>195</sup> Not only is concord mapping essential for grammatical gender but also in number as well as pronouns, prepositions and verbs.

are morphologically marked for gender and carry the feminine suffix *-a* and in contrast, masculine nouns carry a zero suffix and are devoid of any such markings. Words in MA then follow a strict morphological ordering and are either masculine or feminine:

(230) ‘aspects’ of the feminine                      ‘aspects’ of the masculine

<i>kura</i> ‘ball’	<i>fūl</i> ‘beans’
<i>xubza</i> ‘bread’	<i>tapsi</i> ‘plate’
<i>qmajja</i> ‘shirt’	<i>ktāb</i> ‘book’

(231) *a r-rajul al kbir ya-skun quddam i*  
 The man (MASC) the big (MASC) lives 3Sg in front me  
 ‘The big man lives in front of me’

(232) *al warda z- zina*  
 The flower (FEM) pretty (FEM)  
 ‘The flower is pretty’

(233) \* *assarjum ma zela mwasxa*  
 Window still dirty  
 ‘The window is still dirty’

Lack of concord as in (233) above then results in ill formedness and hence is not uttered as *sarjum* is masculine and the modifying verb and adjective should also be masculine *ma zel mwasax*. As we have seen MA, unlike English has a two gender system and MA verbs, adjectives and anaphoric pronouns always show gender agreement with MA adjectives always following the noun together with corresponding features specifications of number and gender:

MA masculine agreement

(234) *šət waħəd a rajul fi s sūq ya-biḡ karmūs*  
 See 1sg NUM INDEF man in DEF market sell 3SG figs  
 ‘I saw a man in the market selling figs’

Above, the determiner *wahəd* ‘one/a’ is masculine and so agrees with *rajul* ‘man’ which is also masculine as is the aspectual marker *ya*.<sup>196</sup> The same can be said of the below; *warda* ‘flower’ is feminine and so *mazīnt-ha* ‘beautiful’ is also feminine in line with noun-adjective concord:

MA feminine agreement

(235) *Ḥaṭani*            *warda*            *mazīnt-ha*  
 give PAST me rose FEM SG beautiful FEM SG  
 ‘He gave me a beautiful rose’

The same is valid in code-switched utterances where MA speakers cognitively determine on the basis of phonological shape the gender of an English noun (in natural discourse) and if it has a non /a/ noun-ending marker it is modified as masculine in the ML. If it the noun ends in an /a/ sound which will render it feminine in the ML. The addressee continues the gender in the same discourse. Consider the dialogue below where the phonology is ‘carried over’:

(236) Masculine agreement

Speaker 1:

*Dad Ḥaṭani*    *wahəd al*    book    *mazīnu*  
 Dad gave PAST one DEF book lovely (MASC)  
 ‘Dad gave me a lovely book’

Speaker 2:

*Baṣaḥ, fin*    *dərt ih?*            *Dyal madrāsa wella kifeš?*  
 really where put it masc of school or wha  
 ‘Really, where did you put it? Is it for school or what?’

‘Book’ above is devoid of a feminine sounding marker or an /a/ sound in the English and so is marked for masculine in the Arabic ML with a masculine pronoun (in bold).

Cognitively, book in MA *kitāb* is also masculine and the matching system here only

<sup>196</sup> Although this has a final *-a*, this should not be confused with the feminine marker *-a* as with Arabic aspectual markers, *ya* is masculine and *ta* is feminine.

shows phonological concord mapping. This is matched by Speaker 2 above who continues the same masculine agreement as a) ‘book’ in English is masculine in its shape (devoid of final ‘a’) and b) is also masculine in Arabic as in *kitab*. Here then, morphologically, phonological gender concord is matched across all generation groups:

(237) Feminine agreement

Speaker 1:

*šuft waħəd al big house, kbira wa zina*  
 I s see PAST – one DET big house, big FEM and nice FEM  
 ‘I saw a big house, it was big and beautiful’

Speaker 2:

*Liyeh, bġitu ta-šriw ha? Ġala haqaš šut waħda beautiful*  
 Why want buy 2SG it FEM because saw 1SG one FEM beautiful  
 down the road  
 down the road  
 ‘Why, do you want to buy it? Because I saw a beautiful one down the road’

Although ‘house’ is devoid of a final /a/ sound, the cognitive application is +feminine and the modifiers are +feminine in MA (in bold). This is because ‘house’ *dār* in MA is feminine and this semantic association spills over into the English. In this analysis we can therefore determine that:

*Generalization (7):*

*Phonology has a more significant bearing over semantic association in determining noun gender in English only if both the English noun and the MA noun map phonologically.*

This is the generalization for all cases of MA/English CS. It is also uni-directional in the sense that, if an English noun has an /a/ sounding marker, it will be +feminine in the MA, but, if an MA noun is either masculine or feminine, it has no bearing on the English adjective as there is no gender marking in English. This significant

phonological versus semantic negotiation has not been studied in detail at all during general CS research. However, Boumans (1998) describes how assignment of Dutch feminine gender is associated with the noun's phonological shape such that the Dutch noun *agenda* would be feminine due to the final /a/. Here his analysis is strictly phonological with no mention of the semantic application as in Generalization 7 above.

Semantic considerations, however, are part of the micro information structure of the clause as it overtly expresses whether a phonological or semantic approach has been applied and shares the speaker's own association with his private world. The process of grammaticalization involves a phonological → semantic transfer process or a semantic → phonological transfer process. This in turn should provide us with a more psycholinguistic approach to codeswitching where micro-negotiated discourse is analysed and evaluated. During my recordings I note that the sense of personal meaning of a noun is either, shared or not shared and then negotiated. At this stage there is a levelling where the final gender assignment is determined by speaker one, modified by speaker two and then fixed and agreed by both. However, we can conclude that phonological shape –a will almost always be the overriding feature in feminine gender assignment. Therefore, nouns such as 'camera', 'drama', 'cola' with clear phonological shapes will be marked for the feminine in MA and agreed by both parties or levelled as such in natural discourse.

However, identical mapping of features specifications can be overridden if we apply a semantic association to the equation. Chomsky (1995) concluded that mismatch of gender features cancels the derivation (1995: 309). However, this analysis only considered phonological aspects of such feature checking. A cognitive application may override this feature and render it feminine due corresponding word in Arabic being feminine such as in the nouns below. Contrastively, in MA/English

codeswitching, other nouns are assigned a specific feminine gender marking due to a *semantic* application alone irrespective of phonological shape:

(238)

ENG - Phonological	MA - Semantic	honey	<i>εasl</i>	door	<i>bāb</i>	house	<i>dār</i>	car	<i>tumubīl</i>
		masc	-	masc	-	masc	-	masc	-
		fem		fem		fem		fem	

Therefore such gender assignment to the English noun shows that a semantic rather than phonological adaptation has been applied by the informant at that given moment. The below highlights this notion:

Semantic Agreement

(239)

Speaker 1:

*Ama, fin dār ti al honey, šet fi l cupboard wa ma laqitā - š*  
 Mum where put you DEF saw in DEF and NEG find NEG  
 ‘Mum where did you put the honey, I looked in the cupboard and didn’t find it’

Speaker 2:

*šufi gballa, dārt ha Ōamma bi yed -i al bārah*  
 Look well put it there with hand my DEF yesterday  
 ‘Look properly, I pit it there with my own hand yesterday’

Speaker 1

*Yes, laqit ha*  
 ‘Yes, I found it’

In (239) above, Speaker 1 is British-born and assigns a masculine suffix *-a* to */laqit/* in concord with the English phonological shape of ‘honey’. However, Speaker 2, a Moroccan-born speaker, even though ‘honey’ lacks the feminine phonological marking in English, assigning a feminine MA adjective as it is feminine in the MA

and this cancels out the phonology of the English. Speaker 1 the realising the semantic compatibility and the affiliation between target and controller, agrees to this assignment rule and in turn the modifier is levelled. Therefore, our further examination leads to the generalization that:

*Generalization (8):*

*MA Semantic applications may override the morpho-phonological make-up of an English noun provided there is direct semantic concord.*

Approaching this analysis from a semantic perspective proves rather illuminating in terms of concord and generational groups. Moroccan-born migrants at times override phonological shapes and attach semantic interpretations to nouns more swiftly and accurately than British-born bilinguals. Observations show that this is the case for 95% of phonological versus semantic pairing amongst group one speakers. This could be due to fluency and pragmatic forces such as how the noun is perceived in the speaker's mind. Nevertheless, both groups switch between phonological and semantic moulding with children developing this skill very early on. Such a seeming lack of correspondence at the beginning of bilingual discourse between two or more interlocutors is quickly agreed upon with one 'winning' and accommodation is adhered to by the fellow speakers in that group. Another aspect of Moroccan Arabic grammatical gender in terms of its placement within English intra-sentential code switching is the utilization of diminutives in Moroccan Arabic, which are often employed as a discourse strategy as discussed below.

### 6.2.2 Grammatical Gender and the Use of Diminutives

Diminutives in Moroccan Arabic are rarely referred to and this is surprising given its frequent occurrence in both monolingual and bilingual discourse.<sup>197</sup> The diminutive, which is a slight variation of the original root form, in essence cites the smallness of the object in terms of its size, or is used as a term of endearment.<sup>198</sup> In Moroccan Arabic it is almost always the latter that is often used with both masculine and feminine objects or objects of affection in general. Other MA nouns which are masculine can be modified to the feminine in terms of usage are the commonly used diminutives which are normally terms of endearment, speaking to children or sweetening phrases:

(240)

Noun	Gloss	Gender	Diminutive	Gender	Gloss
<i>halūf</i>	pig	MASC	<i>ḥlilifa</i>	FEM	little piglet
<i>galb</i>	heart	MASC	<i>gliliba</i>	FEM	little heart
mine					of

Note also the chameleon-type nouns which are sometimes masculine and sometimes feminine. This shows that the use of diminutives at times, sharpens the distinction and gives absolute grammatical gender assignment. Given the hypocorism, the norm is to add a suffixation to the noun or adjective to render it a diminutive.<sup>199</sup> Harrell (2004) notes that a distinguishing characteristic of Moroccan Arabic diminutives is the fact that they are formed by the affixation of the morpheme *-i-* after the second segment of the base, after an initial cluster of two consonants followed by *i* (2004: 81).<sup>200</sup> Some

<sup>197</sup> See also Watson (2006) on the use of diminutives in Yemeni Arabic.

<sup>198</sup> Moroccan Arabic diminutives differ from the morphological formation of diminutives in Classical Arabic.

<sup>199</sup> The process of affixation is often accompanied by the labialization of the first segment of the base if it happens to be one of the dorsal consonants (k, g, x, etc) or labial consonants (l, b, m).

<sup>200</sup> Harrell states that there are several different classes of stem types, with a number of irregularities and the exact meaning is not uniformly predictable from the meaning of the derivational base (2004: 81).

examples of Moroccan Arabic diminutives follow commencing with monosyllables with trilateral roots where the diminutive particle *-iyye-* is inserted between the second and third consonants:<sup>201</sup>

Figure 6.4: Moroccan Arabic Diminutives

Base	Gloss	Diminutive
<i>kelb</i>	‘dog’	<i>kliyyeb</i>
<i>xubz</i>	‘bread’	<i>xbiyyez</i>
<i>terf</i>	‘piece’	<i>triyyef</i>

Moroccan Arabic verbs with middle-weak trilateral roots usually have the diminutive patterning *FwiyyeL* as shown in Figure 6.5 below:

Figure 6.5: Further Moroccan Arabic diminutives

Base	Gloss	Diminutive
<i>bāb</i>	‘door’	<i>bwiyyeb</i>
<i>fār</i>	‘rat’	<i>fwiyyer</i>
<i>kās</i>	‘glass’	<i>kwiyyes</i>
<i>jīb</i>	‘pocket’	<i>jwiyyeb</i>

Finally, some monosyllabic Moroccan Arabic verbs with trilateral verbs and the vowel *e* show the diminutive pattern *FʕiLa* as shown in Figure 6.6 below:

<sup>201</sup> It can be said that in disyllabic and trisyllabic forms, the same number of syllables surface in the diminutive forms. However, a full and circumscriptive analysis of Moroccan Arabic diminutives is beyond the scope of this thesis due to space and time constraints. cf. Al Ghadi (1990) gives a full account of the phonological and morphological processes involved in Arabic diminutives.

Figure 6.6: Moroccan Arabic feminine diminutives

Base	Gloss	Diminutive
<i>bint</i>	‘daughter, girl’	<i>bnita</i>
<i>rjel</i>	‘foot, leg’	<i>rjila</i>
<i>ʕsel</i>	‘honey’	<i>ʕsila</i>
<i>semš</i>	‘sun’	<i>smiša</i>
<i>ʕin</i>	‘eye’	<i>ʕwina</i>

In terms of code switching, how do Moroccan Arabic verbs fit in? Can English verbs also adopt the same diminutive forms? The following section analyzes the data collated starting with calques or loanwords in Moroccan Arabic.

### 6.2.3 English Diminutives

Given the formation of diminutives as described above, there are many English integrational forms in MA which have been assimilated and rendered ‘diminutive’ by the consonantal shape and form. A few examples elicited are as follows in Figure 6.7:

Figure 6.7: Further Moroccan Arabic French origin diminutives

Base	Gloss	Diminutive
<i>limuna</i>	‘lemon’	<i>lwimina</i>
<i>banana</i>	‘banana’	<i>bwinina</i>
<i>sandala</i>	‘sandal’	<i>snidila</i>
<i>blasa</i>	‘place’	<i>bliyssa</i>
<i>btaṭa</i>	‘potato’	<i>bṭiyya</i>

By extension, English diminutives in a Moroccan Arabic environment can be formed as long as the noun adheres to the MA monosyllabic shape. In terms of code switching, insofar as the morpheme does not contravene the MLF or Uniform Structure Principle, diminutive forms are constructed with relative ease and regularity as detailed below:

- (241) *šufi ʕla rwibitta ki mazint-ha ta-run fi l* garden  
 See PREP rabbit how pretty FEM run 3SG in DEF garden  
 ‘Look at the pretty little rabbit running in the garden’

In (241) above, the noun ‘rabbit’ is given the diminutive Moroccan Arabic pattern in line with the affixation of the morpheme *-i-* as described above. Note how the modifying adjective *mazint-ha* is given the feminine suffix *-ha* in line with the matrix language principle where the embedded forms must align themselves with the grammar of the matrix frame. In this case, the noun ‘rabbit’ is rendered feminine and therefore, the adjective must also be made feminine in line with Moroccan Arabic grammar.<sup>202</sup>

- (242) *kan labes waħəd al jwikitta brown ta-ħemeq mʕa* blue jeans...  
 Was wearing one DEF jacket DEM brown go mad 3SG with blue jeans  
 ‘He was wearing a little brown jacket with blue jeans’

The use of diminutives in Moroccan Arabic and English code-switched discourse further validates the MLF as even derived English forms strictly adhere to the MLF where the embedded morpheme tallies with the grammar provided by the syntactic frame of the matrix language. This is where the ML is the language projecting the morpho-syntactic frame for the entire CP which shows intra-sentential CS. The above example shows how the Moroccan Arabic syntax which is the ML, is

<sup>202</sup> Most diminutive nouns are made feminine to render them more affectionate and endearing and this is clearly the case with *rwibitta*. There are numerous cases where baby boys are referred to using feminine diminutive forms to optimize the endearment.

structured so that the Noun + Adjective construction is adhered to as opposed to English syntax which is strictly Adjective + Noun. Not only are MA nouns used in the diminutive form but also adjectives of colour and defect.<sup>203</sup> These are regularly used in the diminutive form and the patterning is that of F*ξ*i*ε*L showing a repetition of the second root consonant (Harrell 2004: 82). Some examples are as follows:

Figure 6.8: Further Moroccan Arabic diminutives (second root consonantal repetition)

Base	Gloss	MA Diminutive
<i>bkem</i>	‘mute, dumb’	<i>bkikem</i>
<i>ħmer</i>	‘red’	<i>ħmimer</i>
<i>bxīl</i>	‘miserly’	<i>bxixel</i>
<i>kbir</i>	‘big’	<i>kbiber</i>
<i>mliħ</i>	‘good, excellent’	<i>mlileħ</i>
<i>ṭwil</i>	‘tall’	<i>ṭwiwel</i>
<i>khel</i>	‘black’	<i>khihel</i>

The interesting point about Moroccan Arabic diminutives is the way in which bilingual speakers attach suffixes to the diminutive stem or make either modifications to the diminutive to make it sound even more affectionate or more endearing. The below are some examples of modified MA diminutive forms:

Figure 6.9: Modified Moroccan Arabic diminutives

Base	Gloss	Diminutive A	Diminutive B
<i>rajel</i>	‘man’	<i>rwijjel</i>	<i>rwiwij</i>
<i>ħuta</i>	‘fish’	<i>ħwita</i>	<i>ħwittita</i>

<sup>203</sup> cf. Harrell 2004 for more on adjectives of defect.

<i>Fouad</i>	'Fouad'	<i>Fiw</i>	<i>Fwiwiw</i>
<i>Najat</i>	'Najat'	<i>Nwijit</i>	<i>Nwijitta</i>
<i>gelb</i>	'heart'	<i>gliyyeb</i>	<i>glilib</i>
<i>bussa</i>	'kiss'	<i>bwissa</i>	<i>bwiwissa</i>

MA suffixation acts in the normal manner where possessives for example can be added as in:

Masculine suffixation to MA diminutive B stem:

(243) *huwwa glilib dyalli w na-mut eli -h*  
 He is heart DEMIN POSS and die 1Sg on him  
 'He is my little love and I am crazy about him'

Feminine suffixation to MA diminutive B stem:

(244) *hiya gliliba<sup>204</sup> dyalli w na-mut eli -ha*  
 She is heart DEMIN POSS and die 1SG on her  
 'She is my little love and I am crazy about her'

As has been shown, MA diminutives as well as the modified English MA diminutive are derived from a great number of nouns and adjectives with common distinguishing characteristics. Further, use of such diminutives is also a discourse strategy as they streamline grammatical gender assignment as the subsequent adjectival agreement must adhere to the gender of the initial diminutive and this, in bilingual discourse, is agreed upon by both speakers if the first speaker refers to a noun in the feminine by adding the feminine suffix, the second speaker will not then refer to it in the masculine form. This is a process I term 'semantic levelling', where both speakers agree upon the grammatical gender of a noun. The use of diminutives in code-switched discourse transcends generational groups and is not specific to one generation. Rather it can be concluded that all three generations make good use of

<sup>204</sup> The feminine suffixation can also be *glilibti* thereby excluding the requirement for the possessive marker 'dya'.

Moroccan Arabic diminutives as well the hybrid English forms. Given the data presented, it is evident that the Matrix language Frame Model is a suitable vehicle through which to describe the diminutive phenomenon in general monolingual data as well as that of bilingual intra-sentential data where the structural configurations are uniformly aligned with the principles and parameters of the MLF, namely that of the content morpheme – system morpheme opposition as explicated in Chapter Three and also the structural requirements of the 4-M Model as previously discussed. Another prominent discourse strategy used by all generation groups is that of the periphrastic ‘Do’ construction which is discussed in the next section.

### **6.3 English Verbs and the Periphrastic ‘Do’ or ‘*Dar*’ + Construction**

A number of researchers have discussed the use of the periphrastic ‘do’ construction in contact linguistics in various languages (Backus, 1990, 1996, Myers-Scotton and Jake 2000, Azuma 1993, Annamalai 1989, Romaine 1995). In most cases, the ‘do’ construction is employed in order to introduce an Embedded Language verb into a code switching construction and this is a frequent code switching discourse strategy employed by bilinguals of various levels of proficiency. In essence, the ‘do’ construction is a simple way in which to integrate Embedded Language verbs into a Matrix frame as long as the morpho-syntactic structure is maintained and well-formedness is not compromised in line with the Matrix Language Frame Model, its constraints and applications.

Whether operating as EL islands or as an ML, English verbs that are integrated into MA clauses are inflected for person, gender and number. This is evident from recordings logged amongst both British-born and Moroccan-born Arabic speakers. Here is one example of an English embedded verb:

(245) sawwalt-ha      waš    *reservat*                      al    *blayyəs*  
Ask-1SG PAST COND reserve 3 SG PAST DEF places  
'I asked her whether she booked the seats'

The verb 'reserve' is inflected for person and past simple. This is a very common method of integrating EL verbs into the matrix MA language. This is further solid evidence of MA being the ML in a given string of discourse. Myers-Scotton (2002) explains that when such languages participate in code switching as the Matrix Language, they do not accept Embedded Language verbs as tensed forms (i.e. with Matrix Language inflections). To solve the problem of carrying speaker intentions that an Embedded Language verb is intended to convey, they construct what has become to be called a 'do' construction<sup>205</sup> in order to compensate. This construction includes an Embedded Language verb that is the reflex of those intentions at the level of lexical-conceptual structure. However, the construction also includes the Matrix Language verb for 'do' and it is this verb that carries the inflections that the Matrix Language requires for well-formedness from a tensed verb construction. Thus the Embedded Language verb occurs without any of the inflections that are required by the Matrix Language (2002: 134). The literature thus far has offered no real explanation as to why speakers make use of this construction why there is a need to integrate embedded forms.<sup>206</sup> Boumans (1998) describes the use of the 'do' construction as "characteristic for migrant bilingualism in modern industrialized societies" (1998: 369. This is rather a bold statement as introspectively, I feel that bilingual speakers employ the periphrastic 'do' construction simply because they are able to incorporate embedded forms into a Matrix frame and this gives rise to an

---

<sup>205</sup> cf. Romaine (1995, Muysken 2000).

<sup>206</sup> Several researchers consider structural features of the 'do' construction. Jake and Myers-Scotton (2001) claim that incongruence between the tense/aspect systems of the participating languages makes the 'do' verb construction a necessary compromise strategy, but they offer little evidence. Ritchie and Bhatia (1996) offer a treatment of the 'do' construction in Hindi/English code switching within the Minimalist Program that describes the phenomenon very neatly but does not explain why it occurs in the first place. Muysken (2000) illustrates many 'do' constructions across language pairs but provides no explanation for why there are needed (Myers-Scotton 2002: 136).

increased fused variety of Moroccan Arabic and English forms. There is no reason, not even socially other than the desire to better express oneself using fused varieties such as ‘light’ verb integration in the *Dar+* construction.

This is in stark contrast with periphrastic do-constructions with the MA verb -*Dar* (imperfect *-dir*) “to put; to make; do” commonly found in Turkish *-yap-* “to make; do” in Dutch CS.<sup>207</sup> In this case, inflection is ‘avoided’ and the auxiliary ‘*Dar*’ is inflected:

(246) *dər-t tilifun al yawm*  
Do-1SG telephone DEF day  
‘I called today’

Such use of the above *Dar+* construction is partly a borrowing strategy to wholly integrate loan verbs and render them MA verbs in their affixation patterning and partly it is a personal choice and its frequency of usage confirms this observation. Cross-linguistically, it is mainly used in integrating loan verbs such as ‘telephone’ ‘fax’, and ‘text’ into the MA syntactic frame and typologically, only with a specific word class. Statistically, whether a speaker uses a periphrastic or lexical construction is a question of preference and ease of utterance amongst different generation groups. Hence the *Dar+* paradigm is used in a higher proportion of cases recorded than the inflectional construction. The same figures apply across different generation groups. It can be said, therefore, that such structural properties alone cannot predict whether either a *Dar+* construction or inflectional type will be employed<sup>208</sup>. Direct insertion in the *Dar+* construction and inflectional construction lends itself more to accommodation and integration as opposed to any language processing constraint.

---

<sup>207</sup> Backus (2000) assigns the term ‘carrier auxiliary’ to do-constructions in Turkish/Dutch CS in embedded verb types. This method of embedding foreign verbs, as described by Boumans (1998) is well-documented in Turkic, Indo-Iranian and Dravidian languages, as well as some eastern dialects of Arabic.

<sup>208</sup> Modern technological verbs are common in the use of ‘*Dar+*’ or direct insertion paradigms such *Dir email* or *email-ih* ‘email him’ which are both mutually employable.

Such English verbs as are integrated are treated as MA verbs. In the case of verb insertion it appears then that the strategy for verb formation (not which construction to use) has been carried over into the target language and this is the only patterning found. There are as yet no constraints found in my data on which type of verb can be used. Neither do levels of bilingualism, language competence or social factors have a bearing on which construction is used. The following examples highlight this flexibility which shows no trace of any syntactic or processing constraint:

(247) *cook-it ha əlbārah*  
 Cook 1SG it yesterday  
 'I cooked it yesterday'

(248) *dərt-ha ta cook əlbarah*  
 Do 1-SG it cook yesterday  
 'I cooked it yesterday'

There is no semantic difference between (247) and (248), or (249) and (250):

(249) *diri lu text*  
 Do -2-PS to him text  
 'Send him a text message'

(250) *text ih*  
 Text him  
 'Send him a text message'

These are frequently uttered constructions and the only exception is if there is a phonological difficulty in uttering a grammatical construction i.e. '*dərt lu text*' 'I sent him a text' would be preferred to '*text-it lu*' 'I sent him a text' or '*fax-it ha lhum*' as opposed to '*dərt lhum fax*' 'I faxed them' which would be the first choice. However, an observation to be made is that there is more of a propensity towards using the Dar+ construction if the noun or verb is more than one syllable in length as the inflectional construction would be too lengthy and morphologically clumsy if the word could be

truncated and this in essence balances out the number of syllables in the equation of (*Dar*+ 1,2 syllable(s)) + (N + 1,2 syllable(s)) matrix. This is the same patterning across different generation groups. Therefore in contrast to Boumans (1998) analysis, I can claim that there *is* an overriding element of choice and volition in the *Dar*+ Construction. I do not find that the semantic element of volition that is to engage intentionally in some activity disappears, along with the transitive feature of '*Dar*'. However *Dar*+ adopts the sub-categorization pattern of the embedded verb (Boumans, 1998: 225). The *Dar*+ construction is very common for certain periphrastic causatives i.e. '*daret* noise' 'she made noise' 'she was noisy.'<sup>209</sup> Also it can be improvised on an ad hoc basis with any number of MA verbs but *less* commonly with English loan verbs where MA affixes accommodate the loan verb: The following are a few examples:

(251)

*Dar* + MA verbs: *tilifun* (to telephone), *té* ( to make tea), *xubz* (to bake bread), *sinniya* (prepare the tray)

*Dar* + Eng verbs: text (send a text), fax (send a fax), coffee (make a coffee), homework (do homework)

Choice (and phonology in some cases) is therefore a salient aspect in deciding to employ an MA verb or to integrate an English verb or a *Dar* + construction. The onus is on grammaticality and well-formedness in any given MA matrix as it is not possible to predict which strategy will be used by the individual speaker. This is a striking aspect of CS in this community of bilingual speakers. Consequently, use of the MLF in describing contact linguistic and code- switched phenomenon is further

---

<sup>209</sup> El-Idrissi (1990) cites an example of '*dar* plus noun' constructions in *daret al khul fi ξini-ha* 'she put the khol in her eyes' (1990: 23). However, *kahlal ξini-ha* 'she kohl-ed her eyes' is also equally employable adding further impetus to the fact that structure alone cannot dictate usage.

supported in this section given that the only environment in which the Dar+ Construction is able to flourish and is described adequately, is that of the MLF model ,as this provides a basic asymmetrical construct which enables both the Matrix and the Embedded forms to be used. Furthermore, in these cases of the Dar+ construction Embedded Language verbs occur in code switching only as non-finite forms in this ‘do’ construction. As has been evidenced, the construction consists of a ‘do’ verb preceded by an Embedded Language infinitive (or another Embedded Language non-finite form) and any complements of the verb. As the ‘do’ verb carries any necessary Matrix Language inflections, the Embedded Language non-finite form can be considered a bare form (Myers-Scotton 2002: 161-2).<sup>210</sup>

The MLF is also the most suitable vehicle for describing the use of numbers in MA and English code switching and this is analyzed in the following section.

#### **6.4 Moroccan Arabic and English Numerals and Quantification**

Other nominal constituents in an MA / English code-switched environment are discussed in this section as it will become evident that further strategies are employed by bilingual speakers and there is a difference in the way in which first generation of speakers use and refer to numbers in CS as opposed to that of the second and third generation of speakers. In MA as in Classical Arabic there are different combinations of numbers with objects, sometimes incorporating plural forms as and when necessary. Cardinal numbers operate differently to ordinal numbers and we analyze MA numbers and quantification in general within a code-switched domain. As described by Harrell (2004) the ordinals and fractions function as ordinary nouns and adjectives, The

---

<sup>210</sup> Myers-Scotton (2002) in analyzing the ‘do’ construction also states that “a conflict of branching requirements between the Matrix Language and the Embedded Language may be behind the need for the ‘do’ construction. That is, Embedded Language verbs may be blocked from projecting predicate-argument structure because they are not congruent enough to pass the Uniform Structure Principle of the Matrix Language” (2002: 162). However, I think this is too detailed an explanation particularly due to the frequency of ‘do’ constructions across many typologically different languages.

cardinals as shown below as in *waḥād* (masc.) and *waḥda* (fem.) in essence function as adjectives:

(252) *rajel waḥād*  
Man one  
'A/one man'

(253) *mra waḥda*  
Woman one  
'A / one woman'

In a code-switched domain, the MA cardinal can only be in the MA in line with the MLF constraint on uniformity as the Matrix Language is Moroccan Arabic and so the syntax must be that of the MA. The sentences below are never possible as the MA syntax is not adhered to and it renders the basic discourse ill-formed and grossly ungrammatical:

\* (254) *rajel one*  
Man one  
'A/one man'

\* (255) *mra one*  
Woman one  
'A / one woman'

However, if the ML were English then it would be possible to have an intra-sentential utterance:

(256) *One / A rajel*  
'One / A man'

Harrell (2004) describes the other cardinals as functioning independently as plural nouns with the exception of *zuj* 'two' which occurs alternately as the first term of a construct state before a prefixless plural noun or as the first term of an analytic annexion before a plural noun prefixed with the definite article, e.g. *zuj ktūb* or *zuj d-le ktūb* 'two books'. Boumans (1998) describes the MA numeral system as "complicated" (1998: 192). However this is not the case as Standard Arabic as well as Moroccan Arabic offers a sound and formulaic template for both cardinals and numerals alike and as is detailed below, such is its systematicity, that it is maintained in intra-sentential code-switched discourse as well as in monolingual utterances.

#### 6.4.1 MA Cardinals from Three to Ten

From three to ten, the full forms of the numerals occur as the first term of an analytic annexion before a plural noun prefixed with the definite article and the short forms occur as the first term of a construct state before a prefixless noun (Harrell 2004: 206). Moroccan Arabic then uses the analytic construction with the prefixed plural noun (Boumans 1998: 193). At times, the full form of the numbers from three to ten is used and at other times a short form is used and this is largely learned and enters into the lexicon very early on in first language acquisition. The forms are listed as detailed in Figure 6.10 below:

Figure 6.10: Moroccan Arabic cardinals

Cardinal (Full Form)	Gloss	Short Form
<i>waḥəd</i> (masc)	'one'	∅
<i>waḥda</i> (fem)	'one'	∅
<i>zūj</i>	'two'	∅
<i>tlata</i>	'three'	<i>tlt</i>
<i>rbḫa</i>	'four'	<i>rbḫ</i>
<i>xmsa</i>	'five'	<i>xms</i>
<i>stta</i>	'six'	<i>stt</i>
<i>sbḫa</i>	'seven'	<i>sbḫ</i>
<i>tmnya</i>	'eight'	<i>tmny</i>
<i>tsḫud / tsḫa</i>	'nine'	<i>tsḫ</i>
<i>ḫšra</i>	'ten'	<i>ḫšr</i>

The patterning for Moroccan Arabic cardinals from three to ten is quite formulaic and straightforward given the full forms and short forms in combinations of a number and a noun:

- Number (full form) + DEF + plural noun with definite article
- Number (short form) + plural noun without definite article

Examples are as listed below in monolingual MA utterances:

Using the full form:

(257) *tsʕa d - l ktūb*  
Nine of DEF books  
'Nine books'

(258) *rbʕa d n- nāss*  
Four of DEF people  
'Four people'

Using the short form:

(259) *xəms drāhem*  
Five dirhams PL  
'Five dirhams'

In terms of intra-sentential code-switched discourse, the MA is maintained whereby the Moroccan Arabic cardinal is the most frequently inserted of the numbers as opposed to MA numbers with the plural forms of the MA nouns agreeing in phi-features, plurality and gender:

(260) there were *ʕi xemsīn waħəd ʔemma*  
There were INDEF fifty one there  
'There were roughly fifty of them there'

(261) We went for *ʕi ʕsrīn yawm li -l mağrib*  
We went for INDEF twenty day to DEF Morocco  
'We went to Morocco for twenty days or so'

In intra-sentential sentences, it is often difficult to split the cardinal number from the noun in its singular or plural form as in Moroccan Arabic, the cardinals from three to ten are in the plural form which would give rise to hypothetical utterances such as *xemsa d books* ‘five books’ which is conceivably well-formed. However, in the cardinals from eleven onwards, the noun must be in the singular form which would render a very odd-sounding result as in *Ḥsrīn girl* ‘twenty girls’. This perhaps then explains why third generation groups are now uttering sentences where the supposed singular MA noun is being entered into their lexicon as an MA plural to comply with the English syntax as opposed to the MA syntax. The following is an example of a third generation bilingual speaker who using the plural English suffix on a plural MA stem, which in itself is striking:

- (262) No, actually there were only five *rjel -z* in the club that were *zinīn*  
 No, actually there were only five men PL in the club that were nice PL  
 ‘No, actually there were only five men in the club that were nice’

The above example (262) is striking as the third generation bilingual speaker attaches an English plural marker *-z-* to the already plural noun *rjel* ‘men’ and this then renders it a double plural formation. Notice also how the adjective *zinīn* is in the masculine plural for agreement purposes. This clearly shows that the Matrix Language in this CP is English and the speaker has maintained the ‘plurality’ of the English ‘men’ and this has carried over into the syntax of the Moroccan Arabic plural noun. Psycholinguistically, this makes for a very interesting analysis as had the Matrix Language in the initial CP clause been Moroccan Arabic, given MA syntax and numeral formation as described above, we would not expect this to be the case as numbers three to ten are always in the plural in Moroccan Arabic and there would certainly be no need for an English plural suffix. Moroccan Arabic numerals then pose no problems in the average intra-sentential code-switched domain as long as the MLF Model and its

principles and parameters are maintained. The following are examples of Moroccan Arabic cardinals:

(263) Give me *xmsa* potatoes *beš na-bda la ġda*  
Give me five potatoes so that begin 1SG DEF lunch  
'Give me five potatoes so I can start preparing lunch'

(264) *š-hel šriti fi set? Meši fi hum six fnejel u stt tpasa?*  
How much bought 2P FEM in set? NEG in them six cups and six plates?  
'How much did you buy in the set? Are there not six cups and saucers?'

In both (263) and (264) above the short form of the MA cardinal formula is used, namely that of Number (short form) + plural noun without definite article following the principle of economy. In keeping with the Myers-Scotton's (1993a) Morpheme Order Principle whereby in Matrix Language and Embedded Language constituents consisting of singly occurring Embedded Language lexemes and any number of Matrix Language morphemes, surface morpheme order (reflecting surface syntactic relations) will be that of the Matrix Language.<sup>211</sup> In this regard, the terms contained in the above examples, i.e. the morpheme order and system morphemes as discussed in Chapter Three and in the Matrix Language theory itself clearly test the validity of the MLF model and test its premise as it is evident that there can only be one Matrix Language which can and does provide the system morphemes and grammatical frame in general. Further, the basic theoretical notion that there is a Matrix Language – Embedded Language which is the bedrock of the MLF Model hierarchy is further supported because the two languages, Moroccan Arabic and English do not both satisfy the roles of the Matrix Language contained in the principles (Myers-Scotton, 2002: 59).

---

<sup>211</sup> cf. Myers-Scotton (2002: 59) for further principles of the MLF Model and Matrix Language – Embedded Language asymmetry.

## 6.4.2 Moroccan Arabic Cardinals from Eleven Onwards

The cardinals in Moroccan Arabic from eleven to one hundred are used with singular prefixless nouns with the numerals from eleven through to nineteen taking the ending *-er* or *-el* when annexed to a following noun.<sup>212</sup> See Figure 6.11 for the standard format of MA cardinals from eleven to nineteen.<sup>213</sup>

Figure 6.11: Moroccan Arabic cardinals - eleven onwards

Cardinal	Suffix		Gloss
<i>ħdaš</i>	<i>-er, -el</i>	<b>+ Cardinal (Singular) Noun</b>	‘eleven’
<i>ṭnaš</i>	<i>-er, -el</i>		‘twelve’
<i>tleṭaš</i>	<i>-er, -el</i>		‘thirteen’
<i>rbeḡṭtaš</i>	<i>-er, -el</i>		‘fourteen’
<i>xemṣtaš</i>	<i>-er, -el</i>		‘fifteen’
<i>seṭṭaš</i>	<i>-er, -el</i>		‘sixteen’
<i>sbeḡṭaš</i>	<i>-er, -el</i>		‘seventeen’
<i>tmenṭaš</i>	<i>-er, -el</i>		‘eighteen’
<i>tseḡṭaš</i>	<i>-er, -el</i>		‘nineteen’

When the MA cardinals are used directly before a noun, these forms always take the *-er* or *-el* ending e.g. *sbeḡṭaš el bent* ‘seventeen girls’. Whether the speakers affixes the *-er* or *-el* ending is determined by personal choice alone as there is no difference morphologically speaking between the two varieties. ‘One hundred’ or *mya* has a combining form *myat*, as is usual with feminine nouns used as the first term of a construct (Harrell 2002). Multiples of ‘one hundred’ combine with nouns either in

<sup>212</sup> cf. Harrell (2004: 206) on numerals.

<sup>213</sup> cf. Caubet (1993) for further analysis of Moroccan Arabic cardinals.

analytic annexion or in a construct state e.g. *xems –emya de-r-rayl* ‘five hundred riyals’ (Harrell 2002: 207).

Ziamari (2007) also gives examples of French and Moroccan Arabic code-switched insertions of the MA cardinal numbers where: “*Un nom français peut être enchasé avec un quantificateur numérique fourni par la langue matrice*” (2002: 136). Ziamari gives the following example which clearly shows how the system morpheme is provided by the Matrix Language in line with the MLF model:

(265) *ka-yemšiw fuq l- ħmīr ši tnašel kilometres*  
Go 3SG PL on DEF donkeys QUANT twelve kilometres  
‘They travel by donkeys for some twelve kilometres’  
(French/Moroccan Arabic, Ziamari 2007: 136)

Ziamari states that:

Dans ces deux énoncés, il s’agit de l’insertion des substantifs: (kilomètres) dans un cadre morpho-syntaxique régi par les quantificateurs, nombres cardinaux (ṭnaš). Il s’agit également d’une construction en état réservée aux nombres cardinaux. Les [deux] exemples montrent également la valeur approximative de cette quantification numérique, marquée par [ši + nombre cardinal + l’article l- + substantive]. Donc, les règles de la langue matrice ne sont pas violée (2007: 136).

In short, the substantive element, or system morpheme must be provided by the Matrix Language and the syntactic ordering within the micro CP will always be that of the language providing the grammatical frame. Given MA as the Matrix Language, the analytic construction is maintained and the rules and regulations of the MA in its morpho-syntactic dynamic are not violated:

(266) *ka-teqra řebġa des-swayeġ d l maths*  
Learn 2SG four of hours of DEF maths  
‘You study four hours of maths’  
(Moroccan Arabic / French, Ziamari, 2007: 136)

It is evident then from the System Morpheme Principle, the Matrix Language + Embedded Language constituents show how all system morphemes which have grammatical relations external to their head constituent, and which participate in the sentence's thematic role grid, will come from the Matrix Language. As has been evidenced from the Moroccan Arabic cases cited above, there is always an analyzable or resolvable frame structuring the morpho-syntax of any given CP at any one time. In this regard and in bilingual utterances, the participating languages, in this case MA and English, never participate equally as the source of the Matrix Language. This is the overriding concept of the MLF and is fully supported by the data used in this thesis. As Myers-Scotton (2002) states in defining the Morpheme Sorting Principle, all morphemes are not equal and at the abstract level of linguistic competence and production, there are different types of morphemes. In bilingual speech, the outcome of these abstract differences is that not all the morphemes from the participating varieties (Moroccan Arabic and English) have equal possibilities of occurrence (2002: 9). In this regard, the same claim can also be verified when addressing the use of quantifiers which are discussed in the following section.

### **6.4.3 Moroccan Arabic / English Quantifiers**

Moroccan Arabic as a linguistic system contains numerous quantifiers which can be classified into a number of syntactic structure types. According to Boumans (1998) one type involves quantification words that are formally equivalent to attributive adjectives, in that they follow the quantified noun, and agree in gender and number. The other major type is the formally similar genitive construction, in which the

quantifier precedes the quantified item<sup>214</sup> (1998: 191). Harrell (2004) describes MA quantifiers as:

A small group of words which may be referred to as quantifiers which have no particular meaning of their own and serve merely as an intensification or attenuation of an expression (2004: 199).

The following are some examples of commonly expressed MA quantifiers which we will later analyze in a code-switched domain:<sup>215</sup>

Figure 6.12: Moroccan Arabic quantifiers

Quantifier	Gloss
<i>ši</i>	'some'
<i>baɣð</i>	'some'
<i>kul</i>	'all, every'
<i>kāmlīn</i>	'all'
<i>gaɛ</i>	'all'
<i>ayy</i>	'any'
<i>neṣṣ</i>	'half'
<i>gir</i>	'just'
<i>jehd</i>	'just the amount'

Harrell (2004: 199) cites some frequently used examples in monolingual MA discourse which convey the contextual meanings of the various quantifiers used:

(267)

<sup>214</sup> Boumans (1998) describes two types of genitive in Moroccan Arabic, one synthetic and one analytic. In the synthetic construction, the quantifier immediately precedes the quantified NP; in the analytic genitive one of the particles *dya*, *d-* or *naɣ* links the quantifier to the quantified NP (1998: 191).

<sup>215</sup> This list is by no means exhaustive and excludes numerical quantifiers which have already been discussed in this thesis.

<i>beḡd l-merrāt</i>	→	‘sometimes’
<i>neṣṣ maḡilqa</i>	→	‘half a spoonful’
<i>ḡir jehd ma iḡli</i>	→	‘just enough for it to boil’
<i>zehd ṣi ḡām u neṣṣ</i>	→	‘(the amount of) a year and a half’

Examples of MA quantifiers in intra-sentential code-switched environments occur quite regularly and this shows that the quantifiers are readily accessible at the point of spell-out amongst bilingual speakers:

(268) aḡla xaṭer ḡendek ṣi beḡd l... *handeling-en (die je doet)*  
 On thinking have 2SG INDEF part DEF action-PL that you do  
 ‘Because you have some actions that you do’  
 (Moroccan Arabic / Dutch, Boumans 1998: 192)

(269) *ṣi waḡad baḡi ṣi contrat beṣ y-hreb -l berra*  
 Some one wants some contract so that runs 3Sg DEF outside  
 ‘Someone wants a contract to leave the country’  
 (Moroccan Arabic / French, Ziamari (2007: 135)

(270) *waṣ kayen ṣi nouveautés?*  
 Are there some news?  
 ‘Is there any news to tell?’  
 (Moroccan Arabic / French, Ziamari (2007: 135)

The MA quantifier *ṣi* is one of the most frequently inserted quantifiers in the corpus and in general. Examples in the Moroccan Arabic / English corpus show complete integration in code-switched discourse within the MLF frame which do not impede well-formedness constraints nor do they impede intelligibility where an English substantive is directly adjoined to the MA quantifier *ṣi* +Eng/MA substantive:

(271) *ṣufi, mḡari ila bḡa ṣi money ka-naḡawn -u.. zaḡma...*  
 Look, usually COND wants some money help 1SG him...you know..  
 ‘Look usually if he needs any money, I help him, you know...’

(272) *gaḡy know-uw ṣḡul -hum, ma ḡendi men dir.*  
 All ASP know 3PL work POSS, NEG1 have REL do  
 ‘Everyone knows what he’s doing, I can’t do anything more’

Such individual lexical items are easily inserted into the Matrix frame as embedded items where the structure follows the paradigm: [QUANT + zero article + Noun]. Furthermore, English substantives can be both modified and quantified by the insertion of the MA quantifiers *kull* and *kāmlin* ‘all’ the first of which Caubet (1993) describes as “*post-posé muni d’un pronom de rappel*” and the latter as “*participe-adjectif post-posé*” (1993: 288). As evidenced earlier with lexical insertion, MA quantifier insertion within an English Matrix frame shows total assimilation to the preceding noun. The following examples highlight this:

(273) *il faut dire d -deuxième année kull -ha*  
 Must say DEF second year all it FEM  
 ‘You have to say the whole of the second year’  
 (Moroccan Arabic / French, Ziamari, 2007: 138)

(274) *Take-y hum kāmlīn maġk la ħeder y-grow-uw up šwiya*  
 Take 2SG PL all PL with you in order grow 3PL up a little  
 ‘Take them all with you in the hope that they will grow up a bit’

In the above examples, the MA quantifiers *kull* and *kāmlin* agree in number and gender with the preceding noun rendering the structural formulae:

- [I- substantive + *kull* + pronoun]
- [article + substantive + *kāmlīn*]

The above formulae are both attributable to the Matrix Language within the particular CP frame and are compatible with the principle of the MLF and thus a further validation of the MLF model.

Furthermore, individual quantifiers are then embedded into the MA grammatical frame as this is evidently the Matrix Language, but we must ask how does the syntax operate with English quantifiers in code-switched environments? It

would appear that this embedding of an English quantifier is compatible with Moroccan Arabic syntax. Similarly, the embedding of a Moroccan Arabic quantifier in an English matrix frame poses no grammatical issues:

(275) If I had some *flus* I could do it ,yeah  
'If I had some money I could do it, yeah'

(276) Any *wald le ħram* can get in these days, *wallah ma kan mizān!*  
Any son DEF badness can get in these days, by God, NEG there is measure  
'Any son of a bitch can get in these days, I swear, there are no measures in place!'

In terms of generational factors, the data suggests that examples (275) and (276) above are reserved for the non-first generation bilingual speakers, namely the second and most notably the third generations, as the first generation of bilingual speakers are more likely to continue a CP in a uniform language, namely Moroccan Arabic rather than engage in an innovative *Reactive Syntax* where youth speech styles are characterized not only by abundant lexical and structural creativity but are also motivated to mark their spatial identity.<sup>216</sup> In addition, theoretically within the MLF domain, Myers-Scotton (2002) states that any lexical item belonging to a syntactic category which involves quantification across variables is a system morpheme (1993a: 100), which neatly categorizes the quantifier in terms of syntactic placement for analysis purposes. This then matches the data presented above with [+quantification] items as the system morpheme always belongs to the Matrix Language in bilingual discourse and certainly within a CP frame. Other grammatical categories are analyzed and tested within the MLF domain in the final part of this chapter commencing with Moroccan Arabic and English demonstratives in intra-sentential paradigms.

---

<sup>216</sup> cf. Ziamari 2007: 288.

## 6.5 Moroccan Arabic /English Demonstratives – Possible Switch Sites

Moroccan Arabic demonstratives are another indicator of the manner in which lexical insertion has been fully integrated into the morpho-syntax of bilingual discourse. Conversely, English demonstratives are also easily inserted into the Moroccan Arabic frame within NP frames where adjacent embedded lexical items are mutually compatible within the particular syntactic environment. In other words, whenever an MA or English DEM is inserted within a Matrix frame, the specific syntactic environment must be one of well-formedness and grammatical compatibility. The data considered thus far in this thesis have shown how the majority of switched constituents within a CP frame are that of smaller constituents, namely adjectives, adverbs, determiners, nouns, prepositions, verbs and most frequently, nouns. This is also in line with research carried out by Bentahila and Davies (1983) on Moroccan Arabic and French intra-sentential syntactic analysis which revealed that the same types of constituents were switched in the same syntactic environments nouns constituting the largest number of switches within any given CP frame. This is also the case in the discourse analysis of other language pairs (cf. Berk-Seligson's (1986) study on Hebrew and Spanish, Poplack's (1980, 1981) study on Spanish and English). In DEM insertion, we ask whether switches between an MA DEM and an English noun are possible.

### 6.5.1 Moroccan Arabic Demonstratives

Moroccan Arabic possesses two categories of Determiners, namely nominal determiners and pronouns. There is the general 'near' demonstrative *hað*, which is invariable for number and gender and the 'far' demonstrative *(ha)ðak* which agrees in

number and gender with the noun it accompanies.<sup>217</sup> Moroccan Arabic in its use of demonstratives instantiates a resolute gender distinction in its noun / determiner system (cf. Hawkins and Franceschina 2004). The forms below are the near and far demonstratives (cf. Harrell 2004):<sup>218</sup>

Figure 6.13: Moroccan Arabic demonstratives

MA DEM	Gloss
<i>haða</i>	'this' (masc)
<i>haði</i>	'this' (fem)
<i>haðu</i>	'those'
<i>(ha)ðak</i>	'that' (masc)
<i>(ha)ðik</i>	'that' (fem)
<i>(ha)ðuk</i>	'those'

MA demonstratives always precede the noun they modify and the nouns they precede take the MA definite article unless they are of a noun type or in a grammatical construction which excludes the use of the given definite article (Harrell 2004: 144):<sup>219</sup>

(277)

- *hað l-weld* → this boy
- *(h)ðak l-weld* → that boy
- *hað l-bent* → this girl

<sup>217</sup> Speakers invariably use *ðak* or *haðak* for 'that' (masculine) and *ðik* or *haðik* for 'that' (feminine) – there is no difference except for speaker preference.

<sup>218</sup> These forms almost always have a specific noun antecedent and are often to be translated as 'this one', 'that one' or 'these ones' (Harrell 2004: 143).

<sup>219</sup> There are also constructions in which the second term is an MA pronoun, e.g. *bent men hadi?* 'Whose daughter is this?' but this is not the main thrust of this analysis and is for illustration purposes only.

- *ha(ðik) l-bent* → that girl

But:

- *hað şāħbi* → this friend (of mine)
- *hað weld əħmmi* → this cousin (of mine)

In a code-switched environment, MA Dem insertion takes the form of [DEM + article + Noun] as per the examples below:

(278) *šetti hað l- concert derna -h*  
 Saw DEM DEF concert we did it MASC  
 ‘Did you see the concert we did’  
 (Moroccan Arabic / *French*, Ziamari, 2007: 129)

(279) *hað l- game ma ya-eħajabni -š at all*  
 DEM DEF game NEG1 like 3SG NEG2 at all  
 ‘I don’t like this game at all’

In the above examples, when the Matrix Language is MA, the demonstratives follow the grammatical rules posed by the Matrix Language, namely of [DEM + article + Noun] where the English nouns are inserted into the MA morpho-syntactic frame<sup>220</sup>. Bilingual speakers, in the UG domain, invariably use the correct gender forms of the demonstrative they employ in natural bilingual discourse even if the first CP requires a masculine DEM and the second CP a feminine as per the preceding noun they modify. The following examples highlight this phenomenon:

(280) *waš haðak l- prof lli ka-yqerrick ðik l-*  
 INTER DEM (MASC) DEF teacher REL teaches 3SG DEM (FEM) DEF  
*madda?*  
 subject  
 ‘Is it that teacher who teaches you that subject?’

(281) *haðuk l- boys ma əndhum -š maħ ðik l- bint*  
 DEM (PL) DEF boys NEG1 have NEG2 with DEM (FEM) DEF girl  
 ‘Those boys don’t like that girl’

<sup>220</sup> cf. Caubet 1993 and Youssi (1992).

(282) *haða*                    problem        *dyalli meši dyallak*  
 DEM (MASC) problem      POSS NEG POSS  
 ‘This is my problem not yours’

Both (280) and (281) have an inserted non-MA noun into the MA morpho-syntactic frame, *prof* in (280) and *boys* in (281) and these, as Embedded Language islands follow the syntactic rules and regulations as set by the Matrix Language, Moroccan Arabic. Further, the preceding demonstrative selected by the speaker matches the gender and plural / singular features of the noun it modifies and this is the same whether in a monolingual or code-switched utterance. This leads to a further examination of EL islands as previously discussed but in this manner, it is clear that Moroccan demonstratives follow certain grammatical phi-features of any noun they precede and modify, either in English or Moroccan Arabic. In essence, within the DEM paradigm, EL islands are themselves full constituents consisting only of Embedded Language morphemes occurring in bilingual CPs that are otherwise framed by the Matrix Language, in this case, Moroccan Arabic. It has been shown that Embedded Language islands show structural dependency relations where minimally there can be two content morphemes as shown above with DET + Noun or Noun + modifier or a content morpheme or non-derivational system morpheme (cf. Myers-Scotton 2002: 139).<sup>221</sup> The NPs as cited above in examples (280) and (281) are in essence internal Embedded Language islands as for example in (281), ‘boys’ is in the plural form to match the plural demonstrative *haðuk* and the same in (281) here the singular form noun ‘game’, which is also devoid of any overt feminine marker (see section on grammatical gender) takes the preceding demonstrative with the exact

---

<sup>221</sup> Myers-Scotton (2002) questions whether EM islands constitute a problem for the MLF in that they represent a break in the Matrix Language Frame but remarks that EL islands do not pose a problem at all. However they do indicate that characterizing the relationship of the participating languages and their level of activation during production is more complex than simply stating that the Matrix Language is in continuous control of the morpho-syntax of the bilingual CP (2002: 139-140).

same phi features, namely singular and masculine *ḍak*. However, example (282) differs as it does not take the same form of [DEM + article + Noun] as it is of a noun type or in a grammatical construction which excludes the use of the given definite article. In this regard, the English noun is inserted into the MA Matrix frame with a zero article and after the demonstrative pronoun itself.

Myers-Scotton (2002) states that these internal Embedded Language islands take this form because they match a feature (plurality here) that is part of the abstract NP heading the maximal projection that is activated to satisfy the speaker's intentions. At the same time, plurality must appear on those elements for which plural agreement is stipulated by rule in Arabic (2002: 150). Boumans (1998) also in commenting on Embedded Language islands within a Matrix Language frame argues that:

These EL plurals, whether they occur as EL islands or within mixed constituents, typically trigger agreement where appropriate and according to the ML grammar. (1998: 36).

This also shows that bilingual speakers always have both the Matrix Language and the Embedded Languages 'on' during the course of natural discourse which enables them to access lexical items as and when needed due to either psycho-linguistic or socio-pragmatic motivations. Another viewpoint discussed by Myers-Scotton (2002) is that it is easier to use Embedded Language islands in some psycho-linguistic sense rather than to access a single occurring element access as a full collocation or a frequently occurring complex unit. This in essence further validates the need to analyze any code-switched material within the asymmetrical form of the ML – EL hierarchy as not only does it pave the way for clarity of research but more importantly, it is the most suitable form in describing and analyzing data from any typologically similar or dissimilar languages.

Does the same patterning hold true of English demonstratives in an MA Matrix frame? Given that there is no overt gender marking as such in English except in the pronoun system, English demonstratives within an intra-sentential code-switched frame are straightforward and the examples cited pose no problems of syntaxes or intelligibility. The following data set shows an example of the English demonstrative ‘some’ inserted as singly occurring lexical items within an MA Matrix frame. This goes against the claim made by Abbassi (1977) who stated that switching between a French determiner and an Arabic noun is not permissible. I am working on the assumption that as English and French are typologically similar, they should obey the same rules. Clearly they do, as the following example shows:

(283) *hetta ila some rjēl y-bother-uw ma ʔndi -š l waqt*  
 Even COND some men bother 3PL NEG1 have NEG2 DEF time  
 ‘Even if some men bother, I haven’t got the time’

The illustration above gives a typical example of single lexical insertion within an MA Matrix frame. In this instance, some is not coloured by any over gender or plural marking as this is not evident in the English language and does not need to adhere to any morpho-syntactic rules nor does it require any further analysis at this stages as it is a basic Embedded Language item. Other embedded lexical items which highlight speakers’ linguistic proficiency include lexical insertion of English as a content morpheme within in Matrix frame.

## 6.6 A Morphological Analysis of English Lexical Insertions (Embedded Language Islands)

English nominal constituents and adjectives as used by bilingual respondents have shown a high degree of usage and insertion into both nominal and verbal frames. Indeed, such is the precision and alacrity of the inserted constituents that one must

question the need and pre-requisite for language proficiency. Does one need to be fluent in both languages? Is it indeed possible to have absolute fluency in two or more languages? With regards to nominal lexical insertion, the answer is that linguistic fluency is not a pre-requisite as singly occurring lexical items are very often inserted by speakers who are fluent in one language and have a proficient command of another. In discussing Embedded Language islands, Myers-Scotton (2002) states that:

The relative presence of Embedded Language islands seems to tell us two different contradictory conclusions about the proficiency of the speakers depending on the community (2002: 148).

These are as follows:

1. When the overall prevailing pattern includes many bilingual CPs (with many mixed constituents), singly occurring forms (typically nouns) prevail. If speakers employ relatively many Embedded Language islands, they seem to be among the more proficient speakers. That is, it seems that higher language proficiency in the Embedded Language is necessary in order to feel at home producing islands.
2. However, there is also evidence that when speakers are nearly equally at home in both languages, almost ironically, Embedded Language islands lose their importance. Instead, switching between CPs becomes very frequent as well as switching between sentences, which of course may include more than one CP.

Backus (1998) however, in his analysis of Moroccan Arabic and Dutch code switching states that there is no:

EL island stage and that at some point, inter-sentential code switching takes over and the frequency of Embedded Language islands goes down again (1996; 334).

As Myers-Scotton (2002) further discusses, in classic code switching, both languages cannot participate equally, as one language must always assume the Matrix form and lead the grammatical syntactic frame. In brief, this means that while speakers must be able to produce and recognize well-formed utterances in their language or dialect, they need not have full command of the morpho-syntax of the other language. The MLF model recognizes that one language, the Matrix Language is more responsible for morpho-syntactic structure than the other Embedded language. In this manner, as previously discussed, the ML supplies essential morpho-syntactic structures for mixed constituents while the EL may supply content morphemes that are suitably inserted into this frame (2002: 25). With this in mind, we examine the frequently inserted English lexical items in Moroccan Arabic grammatical frames.

#### **6.6.1 English Verb insertion within a Moroccan Arabic frame (*Pseudo-Verbs*)**

Syntactically, English lexical insertions into an MA grammatical frame are rendered content morphemes in line with the MLF model and in mixed constituents must be framed by system morphemes. In addition to the *Dar+* Construction-type verbs discussed above, embedded English verbs are very frequently inserted in MA frames by bilingual speakers who attach MA suffixes for inflectional agreement purposes. In this manner, the English verb maintains its stem, but in intra-sentential environments it is combined with an MA inflectional affix marked for gender, singularity and plurality. In other words, it becomes a *pseudo-Moroccan Arabic verb* with a different lexical (or cultural meaning) intention. The inflectional affixes attached to the English verbal stem are, in line with the MLF, system morphemes and these must come from the Matrix Language in any given MA /English intra-sentential string. In this domain, grammatical relations must adhere to the inherent premise of the MLF Model, namely

the Morpheme Order Principle and the System Morpheme Principle as reproduced below:

1. *The Morpheme Order Principle:*

*In ML+EF constituents consisting of singly-occurring EL lexemes and any number of ML morphemes, surface morpheme order will be that of the ML.*

2. *The System Morpheme Principle:*

*In ML+EL constituents, all system morphemes which have grammatical relations external to their head constituent will come from the ML.*

*(Myers-Scotton 1993: 83).*

Therefore, the question arises, do singly occurring English verbs in Moroccan Arabic mixed clauses follow the above principles? This is examined below with data providing supporting evidence where possible. Further, the extent to which English verb insertion can be accounted for within the MLF paradigm is evaluated.

### **6.6.2 Evidence for Embedded Verbal insertions**

Non-Arabic verbs and their integration with complements in Moroccan Arabic have previously been studied due to their frequency in bilingual discourse (Abbassi 1977, Heath 1989, Caubet 1993, Nortier 1990, Boumans 1998). Switching between a finite verb and an infinitive complement or between an auxiliary and a main verb was thought not possible (Timm 1975, Lipski 1978, Abbassi 1977). However there have since been numerous counter-examples which invalidate this statement. Bentahila & Davies (1983) in their analysis of Moroccan Arabic and French data found switches between an auxiliary and a main verb occurred in natural speech (1983: 314). Further, in contrast to Boumans (1998) who states that apart from the *Dar+* Construction,

Dutch verbs inserted in MA clauses: “Are not very frequent...and tend to be accompanied by pauses and hesitations” (1998: 259).

I would claim that such insertions are indeed very common and occur frequently throughout the corpus and in all three generation groups which conveys that it is a sweeping trend and not a consequence of language shift or attrition in any way. Furthermore, given the literature and current corpus, and also in other data sets, patterns of verb insertion are frequent and very much the norm amongst bilingual speakers in the Moroccan diaspora<sup>222</sup>. Boumans (1998) observes that speakers who use the non Dar+ construction in inserting non-Arabic verbs suggest that the respondent experiences a certain need to use Dutch verbs in the context of a MA clause but lacks a productive insertion strategy (1998: 259). As has been discussed previously, speakers certainly do use certain strategies in order to integrate certain classes of lexical items, namely nouns, verbs, adjectives and other embedded clauses.

Structurally, Myers-Scotton (2002) states that:

The empirical evidence in almost all examples in code switching corpora is that Embedded Language elements form two classes, based on their opportunities to occur in mixed constituents. These classes coincide with the content-system morpheme opposition. While Embedded Language content morphemes (especially nouns) occur with relative freedom in these constituents, Embedded Language system morphemes have little or no freedom of occurrence. Specifically, under the System Morpheme Principle of the MLF model, certain types of system morpheme cannot come from the Embedded Language, but must come from the Matrix Language (2002: 72).<sup>223</sup>

We can further add Backus’s (1999, 2003) Unit Hypothesis which states:

1. *The Unit Hypothesis: Every multi-morphemic EL insertion is a unit, inserted into an ML clausal frame.*

---

<sup>222</sup> Boumans (1998) contradictorily then goes on to state that he counted 14 occurrences of Dutch verb insertion within MA clauses and that this “phenomenon is relatively frequent” (1998: 259).

<sup>223</sup> See also section in Chapter Three on the 4-M Model which is a refined and extended version of the MLF and further details the asymmetrical relation between content and system morphemes. cf. Wei (2000) whose Chinese and Japanese data fully support the 4-M Model.

Boumans (1998) also states that in complementation patterns of embedded verbs, which are functional morphemes (system morphemes), they are not commonly embedded categories, thus they are realised in the ML. In terms of argument structure on inserted English verbs within MA frames, it can be said that they assume that bilingual respondents assign an MA sub-categorization to the English verb as it takes on both the structure and syntax of Moroccan Arabic verbs. In this vein, Boumans (1998) states that we can view the embedded verb as being assigned to a class of ML verbs with a concomitant sub-categorization pattern of a particular corresponding ML verb (1998: 265). The following are typical examples of embedded French infinitives within an MA frame:

(284) *ʕad xessu y-redoubler*  
 Again he needs repeat 3SG  
 ‘Again he needs to repeat’  
 (*French / Moroccan Arabic, Bentahila & Davies, 1983: 314*)

(285) *ka-y-xessu y-dubbel zijn best doen*  
 Must 3SG double 3SG his best do  
 ‘He must double his efforts’  
 (*Dutch / Moroccan Arabic, Nortier, 1990: 167*)

Nortier (1990) refers to Non-Arabic verb insertion as a “phenomenon” but due its frequency, I would claim that it can no longer be termed a phenomenon but a natural speech act which is a by-product of bilingual discourse and very productive in the Moroccan socio-linguistic domain.

In composite as well as classic ML / English code switching where English verbs are inserted into an MA frame, both prefixes and suffixes are attached to the English stem, the imperfect being formed by the addition to the stem of a set of prefixes in the singular and the same set of prefixes plus a set of suffixes in the plural. The most frequently occurring types are to be found in both the imperfect and perfect

forms. The perfect is formed by a set of suffixes in addition to a stem. The following modified English verbs give some examples of the pseudo-verb where the formula [ENG STEM + M INFL] is maintained:

(286)

- *copīt* → ‘I copied’
- *ma lefit-haš* → ‘I didn’t leave it (fem)’
- *tryit ha* → ‘I tried it (fem)’
- *eatina* → ‘we ate’
- *fightet maξah* → ‘she fought with him’

The above constructions in the perfect tense show how the stem of the English verb is maintained and Moroccan Arabic inflectional affixes, referred to in the MLF model as system morphemes, are attached in line with the MLF. Interestingly, the last example above shows how the speaker has maintained the stem in the infinitive even when the perfect stem is an irregular verb and should be ‘fought’ + INFL. Figure 6.41 below shows the conjugation of the perfect forms of an English verb in natural code-switched discourse which highlights not only the frequency of such constructions but also their full integration in a Matrix frame.

Figure 6.14: Perfect Tense English Verb ‘Hit’ with Moroccan Arabic inflectional affixes

	<b>Fused Verb</b>	<b>Gloss</b>
First Person	<i>hit-it</i>	‘I hit’
Second Person (M)	<i>hit-it</i>	‘you hit’ (masc)
Second Person (F)	<i>hit-iti</i>	‘you hit’ (fem)
Third Person (M)	<i>hit-a</i>	‘he hit’
Third Person (F)	<i>hit-at</i>	she hit’

First Person (Pl)	<i>hit-ina</i>	'we hit'
Second Person (Pl)	<i>hit-itu</i>	'you hit' (pl)
Third Person (Pl)	<i>hit-iw</i>	'they hit'

The above shows a basic paradigm of conjugated verbs in the perfect tense with Moroccan Arabic inflectional affixes which assume the characteristics of MA verbs and obey the grammatical and syntactic rules set by the Matrix frame. In other words, the integrated English verbs assume Moroccan Arabic morphology, namely that of the Matrix language. The above conjugation is a blanket template for any English verb irrespective of verb type, phonological make-up or specification. The onus then is not on the quality of the inserted English verb but rather on the quality of the Moroccan Arabic Matrix environment so long as there is a syntactic positioning available and it conforms to the well-formedness constraints of the MLF in general. Examples are as follows:

(287) Cut *-it* *al* *lhem u* *dirt-u* *fi -l* fridge  
 Cut PAST DEF meat and put it MASC in DEF fridge  
 'I cut the meat and put it in the fridge'

(288) Put *-aha* *Θamma al* *berrah* *u* *me zel* y-complain  
 Put PAST it FEM there DEF yesterday and still complain 3SG  
 'He put it there yesterday and he is still complaining'

In the perfective form it is clear then that the inserted English verbs, together with Moroccan Arabic participles, or inflectional affixes assume the argument structure and both the syntactic and morphological patterning of the MA verb. Is this the same for the imperfective, Moroccan Arabic's second tense category?

### 6.6.3 Evidence for Embedded Verbal Insertions in the Imperfect

From the corpus collated of this thesis, it is clear that bilingual Moroccan Arabic and English speakers in the UK insert English verbs into Moroccan Arabic Matrix clauses with such ease and fluency that this warrants major attention and examination. There appear to be no impediments as to the quality and type of verb inserted other than that it should adhere to the general rules of UG and that of the Matrix Language. English verbal stems with imperfect Moroccan Arabic inflectional morphology are far more frequent than the perfective forms.<sup>224</sup> The below Figure 6.15 gives an example of the imperfect inflectional forms on a typical English stem:

Figure 6.15: Imperfect Tense English Verb 'Copy' with Moroccan Arabic inflectional affixes

	<b>Fused Verb</b>	<b>Gloss</b>
First Person	<i>n-copy</i>	'I copy'
Second Person	<i>t-copy</i>	'you copy'
Third Person (M)	<i>y-copy</i>	'he copies'
Third Person (F)	<i>t-copy</i>	'she copies'
First Person (Pl)	<i>n-copi-w</i>	'we copy'
Second Person (Pl)	<i>t-copi-w</i>	'you copy' (pl)
Third Person (Pl)	<i>y-copi-w</i>	'they copy'

Again, as with the perfect, embedded insertions in the imperfective also assume the inflectional morphology of Moroccan Arabic and the system morphemes, in line with the MLF are always in the language of the Matrix, in this case, Moroccan Arabic. The

<sup>224</sup> This may be due to the nature of the phonology as the agglutinative morphological stems extend to more and more morphemes, e.g. "carry-it-*uha*" 'I carried it' (fem).

examples below highlight the morpho-syntactic nature of the embedded imperfect forms:

(289) *iwa bğit*                      *n-show*    *him sağa ma*    *wait -š*  
Well wanted PAST 1SG show 1SG him but    NEG1 wait NEG2  
'Well I wanted to show him but he didn't wait'

(290) *beš*    *y-start -a -ha*        *xessu*        *y-jib*        *al*    *key al*    *arawal*  
So that start 3SG it (FEM) he must    bring 3SG DEF key DEF first  
'So that he can start it, he has to being the key'

This provides further supporting evidence for the MLF model and is a fundamental characteristic of embedded verbs.<sup>225</sup> Ziamari (2007) in her analysis of Moroccan Arabic and French code switching states that "Les verbes français s'enchâssent donc en se soumettant à la morphologie de la langue matric: l'arabe marocain" (2002: 150). This patterning is the same not only for the perfect and imperfect, but for all MA verbal types including the medio-passive where Moroccan Arabic prefixes which normally form the medio-passive voice are attached to the English stem and adopt the same Moroccan Arabic morphological patterning: "*ttensit*" 'I got tense', "*ttensisti*" 'you got tense', "*ttensina*" 'we got tense'. Caubet et al. (2000) add that:

In Algerian Arabic / French, French transitive, intransitive and reflexive verbs are often assigned to the same conjugation paradigm and the medio-passive prefix it scarcely uses (2000: 159).

Therefore, the above corpus provides more than adequate support for the MLF model and Ziamari also found in her Moroccan Arabic / French code switching data that her corpus supports the basic premise of the MLF Model:

L'étude de l'insertion des verbes français a permis de vérifier ce concept clé du Matrix Language Frame: la langue matrice. En effet, les verbes français suivent les règles morphologiques de l'arabe marocain (2007: 156).

---

<sup>225</sup> cf. Caubet (1993) for further comments on French verb insertions in Moroccan Arabic matrix frames.

This analysis shows that the Matrix Language not only provides the syntactic frame but also as has been evidenced, the morphological frame upon which the embedded elements uniformly integrate. Therefore, the empirical evidence in all examples thus far in the code-switched corpus validates the concept of the MLF in that the embedded verbal elements which occur in mixed clauses will assume the syntactic and morphological qualities of the Matrix Language and the inflectional affixes are always that of the system morphemes. This then coincides with the content morpheme – system morpheme asymmetry as presented in Chapter Three.

## **6.7 Conclusion**

The main focus of this chapter is that of the MLF model and its application and placement with regards to intra-sentential code-switched data from English to Moroccan Arabic. It examines grammatical categories which incorporate lexical insertions within code-switched data. This provides evidence to show that speakers employ certain discourse strategies to further insert lexical items within a Moroccan Arabic Matrix frame. Grammatical gender and concord where speakers of both generation groups engage in the semantic application of lexical convergence is essential amongst Arabic speakers. This has been analyzed in detail with certain generalizations made, most important of which are the distinctions between phonological and semantic differences in grammatical gender, a concept barely touched upon in the literature. Data recorded shows, however, that Moroccan-born migrants are more skilled at this than British-born bilinguals and this is to be expected given levels of proficiency. It has been shown in this phenomenon, whereby semantic gender overrides the morpho-phonological caters for a wider perspective on accommodation hypotheses and concord in general. There is sometimes evidence of

concord mismatch in this levelling phenomenon where speakers of different generations are unsure of gender assignment in MA and/or whether to apply phonological or semantic agreement particularly as convergence in MA is learnt very early on. This, at times, trial and error process is most evident amongst second generation British-born Moroccans as opposed to the more fluent MA speakers.

Concord across generations then is produced either semantically or phonologically. This is a crucial and salient aspect of this thesis and has been little explored in previous research. The issue is whether the agreement relation is determined as semantic or phonological. Such gender agreement and the matching of features and morphological systems in natural bilingual discourse has been little researched in the general domain of code switching but is fast gaining interest in linguistics as a grammatical issue in its own right. The use of Moroccan Arabic diminutives conveyed how speakers use certain discourse strategies so that as well as Moroccan Arabic diminutives, English diminutives can also be structured in a 'Moroccan Arabic environment' as long as the noun adheres to the MA monosyllabic shape. In terms code switching, insofar as the morpheme does not contravene the MLF or Uniform Structure Principle, diminutive forms are constructed not only with ease, but with a certain degree of frequency amongst bilingual speakers. In addition, this chapter discusses quantifiers and numerals in Moroccan Arabic and the way in which their insertion into ML frames provides further support for the MLF model and how examples in the Moroccan Arabic / English corpus show complete integration in code-switched discourse within the MLF frame which does not impede well-formedness constraints nor does it impede intelligibility where an English substantive can be directly adjoined to an MA quantifier.

The final part of the chapter discusses English verb insertion where it has been shown that certain constructions in the perfect and imperfect tense and medio-passive

voice show how the stem of the English verb is maintained as an embedded element and Moroccan Arabic inflectional affixes, or system morphemes are attached to this stem. In essence this further validates Myers-Scotton's (1993, 2002) MLF Model. Certain generalizations have been made throughout this chapter in line with generalizations made in previous chapters which further corroborate the premise of grammatical outcomes in bilingual encounters and bilingual discourse in general. The bilingual outcome of code switched data not only overtly highlights the theoretical construct of the MLF model but the juxtaposition of the two grammars from typologically diverse languages further validates the theoretical orientation of the MLF model and its supporting model, the 4-M Model.

## CHAPTER SEVEN

### CODE SWITCHING OF CONSITUENTS - THE MLF AS A SUITABLE SYNTACTIC VEHICLE

This chapter is concerned with further morphological and syntactic processes in intra-sentential code switching and these are examined in light of the following objectives which are twofold: first, the formal description of the morphological and syntactic processes of Moroccan Arabic and English within a code-switched paradigm are analyzed and secondly an examination of certain grammatical categories in bilingual discourse is concentrated upon in order to further validate the application and viability of the MLF model. Further data will be postulated which analyzes the grammatical categories of Moroccan Arabic and English in bilingual frames within CP clauses as used by respondents in their bilingual clauses. Given the frequency of prepositional insertion in the data collected, these embedded elements are also examined together with their integration in a bilingual environment.

An analysis of certain productive adverbs and the insertion of conjunctions and conditionals makes up the central part of the chapter as the penultimate section are devoted to the phenomenon of pronoun doubling which is a systematic syntactic result of the interaction of two grammars and speakers' attempts at preserving the integrity of both grammars (Eid 1996). The data provide important insights into current views of the syntactic and morphological processes involved in code switching and bilingual discourse in general. As has been previously discussed, there are certain syntactic strategies that bilingual speakers use in order to incorporate certain noun phrases or morphemes into the lexicon and/or structure of either language and this is further analyzed and explored in this chapter, together with further enlightenment and motivation for the innovative concept of *Reactive Syntax* and its placement amongst certain generation groups. The main thrust then of this

chapter is whether embedded Moroccan Arabic nouns and other grammatical categories adhere to the main issues and principles of the MLF model (Myers-Scotton 2002) and whether this is in essence a viable vehicle for the linguistic analysis of Moroccan Arabic and English code switching. The chapter concludes with an account of non-switchable Moroccan Arabic nominal constituents which are loaded with cultural meaning which are deemed non-transferrable in terms of a lexical equivalent in English.

### 7.1 Moroccan Arabic Prepositional Insertion

Participles play an important role in Moroccan Arabic due to their syntactic importance and most notably in code-switched discourse, to their frequency of usage. These particles have the expected characteristics, namely that they are short and invariable and syntactically always function as subordinate modifiers of nouns and pronouns (Harrell 2004: 208) and in essence govern nouns or pronouns. Before analyzing the data, we look at the most common Moroccan Arabic prepositional particles which are embedded in grammatical structures. See Figure 7.1 below:

Figure 7.1: Moroccan Arabic prepositions<sup>226</sup>

MA preposition	Gloss
<i>b-, bi-</i>	‘with, by means of’, <i>bi-</i> is only used before suffixed pronouns
<i>bin, binat</i>	‘between’, the form <i>binat</i> is used before plural suffixes
<i>bla</i>	‘without’ – <i>bla</i> does not take pronoun suffixes
<i>bħal</i>	‘like, as’
<i>beħd</i>	‘after’ – usually used with <i>men</i> ‘from’ preceding

<sup>226</sup> cf. Harrell (2004: 210) for a more detailed overview.

<i>d-, dyał</i>	‘to belong to’ (see Chapter Four)
<i>f-, fi</i>	‘in, amongst’ – only used before pronoun suffixes
<i>fuq</i>	‘above, over’
<i>l-, li-, lil</i>	‘to, for’ – the forms <i>li-</i> and <i>lil-</i> are used before pronoun suffixes in independent forms
<i>men, menn, mn-</i>	‘from, of, than’ – the form <i>menn-</i> is used before pronoun suffixes beginning with a vowel e.g. <i>menni</i> ‘from me’
<i>muř</i>	‘behind, after’
<i>mɛa</i>	‘with’ – indicates accompaniment
<i>qbel</i>	‘before’
<i>qeddām</i>	‘in front of, facing’
<i>teħt</i>	‘under, beneath, below’
<i>ğir</i>	‘except, other than, nothing but’
<i>ħða</i>	‘next to, beside’
<i>ɛla, ɛel</i>	‘on’
<i>ɛnd</i>	‘at, with, at the place of’

The above prepositions illustrate the most frequently used and most commonly inserted particles in natural discourse. Prepositions as a group, typically take nouns or pronouns as objects. However, they are also found with clauses or prepositional phrases as objects.<sup>227</sup> Pronominal clitics attached to prepositions are similar in form to those attached to verbs and nouns. The one difference between the prepositional object suffixes and those used with nouns and verbs is in the first person; with prepositions the suffix /+ya/ is found when immediately following a vowel. Further, certain prepositions notably *fi* ‘in’, *bi* ‘by’ *l* ‘to’ behave as prefixes when preceding a

<sup>227</sup> Harrell (2004: 210).

noun phrase. Numerous examples of prepositional usage in code-switched discourse have been cited in the literature in code-switched discourse:

(291) smeɣt- ha fe r rādiu  
Heard 3F SG on DEF radio  
'I heard it on the radio'  
(Moroccan Arabic / *French*, Nortier, 1990: 29)

(292) bhēl daba hnaya *met kerstmis*  
Like now here with Christmas  
'Like here now at Christmas'  
(Moroccan Arabic / *Dutch*, Nortier, 1990: 135)

Ziamari (2007) in her corpus of code-switched data in Moroccan Arabic and French found that use of prepositional particles was very frequent: "Les syntagmes prépositionnels mixtes sont très fréquents dans notre corpus" (2002: 146). It was also found that insertion of a Moroccan Arabic preposition introduced a nominal mixed clause:

(293) ka-nedxel nišan l l-*poste*  
Go 1 SG straight to DEF post office  
'I go straight to the post office'  
(Moroccan Arabic / *French*, Ziamari, 2007: 146)

(294) ɣrafti ana *parfois* ka-tetra l-i qeddām l- *poste*  
Know PAST 2SG I sometimes arrive 3SG to me in front DEF post office  
'You know sometimes it arrives for me in front of the post office'  
(Moroccan Arabic / *French*, Ziamari, 2007: 146)

In line with the MLF model, prepositions should conform to the language providing the grammatical frame, in other words, the matrix language. Recall that the content morpheme – system morpheme asymmetry is crucial in the model's basic premise and also in clearly identifying which is the matrix language in mixed clauses. As previously illustrated, content morphemes, e.g. nouns, verbs and some prepositions express semantic and pragmatic aspects and assign or receive thematic roles. System

morphemes, however, operate as function words, e.g. inflections and do not assign or receive thematic roles. In bilingual discourse, system morphemes must come from the matrix language with content morphemes emanating from either matrix or embedded language. We refer to the basic principles as set out by Myers-Scotton 1993):

*Morpheme Order Principle (1993: 83)*

1 *In ML + EL constituents consisting of singly-occurring EL lexemes and any number of ML morphemes, surface morpheme order will be that of the ML.*

*The System Morpheme Principle (1993: 83):*

1 *In ML+EL constituents, all system morphemes which have grammatical relations external to their head constituent will come from the ML.*

Prepositional phrases in mixed discourse adhere to the maxims of the MLF and the principles as outlined above. Also, under the MLF, prepositions are classed as not only content morphemes but also as they belong to the content morphemes hierarchy within the asymmetrical assignment, they are also matrix language islands in mixed clauses. According to Myers-Scotton's revised edition of the MLF (2002) and as previously discussed, matrix language islands are full constituents consisting only of matrix language morphemes occurring in a bilingual CP that is otherwise framed by the matrix language. A matrix language island shows structural dependency relations; minimally it can be two content morphemes (e.g. noun and modifier) or a content morpheme and a non-derivational system morpheme (2002: 139). Therefore, in effect, such an island represents a break or a switch in the language whereby for that particular micro frame, the grammatical frame is led by the matrix language within the CP and its grammatical frame. They are in essence exposed to ML grammatical

constraints and show the internal structural dependency relations of the guest language. Prepositional phrases are easily inserted into the CP phrase in bilingual discourse. Other examples which highlight prepositional insertion which function entirely as pro-clitics in the corpus are as follows. The pro-clitic preposition attaches itself to the subsequent object clitic or noun:

(295) *ma dert ha -š fi rās -i* at all  
 NEG1 put it FEM NEG 2 in head my at all  
 'I didn't think of it at all'

(296) *mša l -l football match beš y-play šwiya*  
 Went to DEF football match so that play 3SG a bit  
 'He went to the football match to play a bit'

Of the various studies on code switching in general (Bentahila & Davies, 1983; Heath 1989), the most detailed research on Moroccan Arabic code switching has been that of Nortier, (1990) and Boumans, (1998) but it is Abbassi (1997) who posits three constraints on code switching in his corpus of Moroccan Arabic and French and these are listed as follows:

1. The relative pronoun constraint
2. The preposition constraint
3. the wh-constraint

It is the second constraint which is of interest at this juncture of the research in analyzing prepositional insertion in Moroccan Arabic and English code switching. In this constraint Abbassi (1997) prohibits code switching between prepositions expressing purpose and their complements and he gives the following examples:

(297a) *mšina l -l qehwa pour boire un pot*  
 Went 1PL to DEF cafe in order to drink a drink

'We went to a cafe to get a drink'  
(Moroccan Arabic / *French*, Abbassi, 1997)

(297b) \* mšina l -l qehwa pour nšerbu un pot  
Went 1PL to DEF cafe in order drink 1PL a drink  
'We went to a cafe to get a drink'  
(Moroccan Arabic / *French*, Abbassi, 1997)

In the above examples, Abbassi prohibits switching as outlined in (297b) and as described by Aabi (1999), Bentahila & Davies (1983) describe the prohibition in [297b] above as being due largely to the constrained nature of code switching and in this case in terms of sub-categorisation rules. These conditions and rules as have been previously discussed are completely parametric between Moroccan Arabic and the respective guest language, whichever operates as the matrix language. As described by Aabi (1999), the above examples (297b) is very 'odd-sounding' and this is due to the condition that the category *baš* 'so that' in MA c-selects a TNSP with the grammatical feature [+subjunctive] and this is violated in (297b) above. Notice that in the case where the category *baš* 'so that' is followed by a French verb, bilingual Moroccan Arabic speakers compensate for the parameterisation by attaching a Moroccan IP to the French verb lexeme. In this way, the condition violated in (298b) becomes satisfied in (297c) below:

(297c) tayhāwlu daba beš y develop-iw un nouveau système  
Try 3PL now so that develop 3PL INDEF new system  
'They are now trying to develop a new system'  
(Moroccan Arabic / *French*, Abbassi, 1997)

The non-occurrence of switching between the French category 'pour' and an MA IP complement can also be explained by the fact that its property of [+infinitive] cannot be fulfilled. Therefore, the MA syntax does not have the property [+infinitive], and switches like (297b) above, are anticipated not to be possible. Is this assumption

correct? Example (297b) in the corpus corroborates the view that this is indeed a correct assumption such that insertion of the morpheme 'y' in the CP neutralizes any potential violation and renders the clause both possible and grammatical in bilingual discourse. Interestingly, this falls under the rubric of UG where bilingual speakers are aware of potential syntactic conflicts and therefore employ discourse strategies in order to by-pass any potential conflict. This then leads to another generalization (9):

*Generalization 9*

*Switching between certain MA categories namely baš + an infinitive are not possible. These must be followed by a subjunctive morpheme in order to neutralize the violation and render grammaticality.*

Therefore, the above generalization and analysis is firmly in line with the MLF in that it is further evidence for the matrix – embedded hierarchy and its identification is realized through the insertional approach of certain morphemes and lexical items which further corroborate the model. As described, the matrix language is a device which not only sets the grammatical frame for all morphemes in bilingual discourse but is also used to describe grammatical structures containing morphemes from more than one language. In this case, the examination of prepositional phrases provides clear evidence of the grammatical placement of certain morphemes. As Boumans (1998) describes it, on the constituent level, the matrix language is the language to which the internal structure of the constituent as expressed by the distribution of all morphemes within the constituent can be attributed. The distribution of a morpheme concerns both its occurrence and its order relative to other morphemes that make up the constituent (1998: 66). Are adverbial phrases also governed in the same manner? This is examined in the following section.

### 7.1.2 Moroccan Arabic / English Adverbial Insertions

The insertion of adverbs in bilingual discourse as presented in the corpus is both frequent and varied.<sup>228</sup> The main adverbs in Moroccan Arabic are particles of spatial relations as well as those of time and space. Embedded adverbs and ADvPs constitute a notoriously heterogeneous word class both notionally and grammatically (Boumans, 1998: 279). As Schachter (1985) describes it, adverbs are “*modifiers of constituents other than nouns*” (1985: 20). MacSwan (1997) states that adverbs are generally regarded as adjuncts so they do not interact with other elements in any obvious way (1997: 172). The particles below are some of the most frequently used:

Figure 7.2: Moroccan Arabic adverbs<sup>229</sup>

MA Adverb	Gloss
<i>daba</i>	‘now’
<i>ḡād</i>	‘after’
<i>la ḡedda</i>	‘tomorrow’
<i>berra</i>	‘outside’
<i>daxel</i>	‘inside’
<i>fuq</i>	‘above’
<i>l-lūr</i>	‘behind, the rear’
<i>murr(a)</i>	‘behind, after’
<i>teħt</i>	‘under, below’
<i>Ṫamma , temma, temaya</i>	‘there’

<sup>228</sup> However, Tuc (2003) in his analysis of Vietnamese and English code switching found that the proportion of switches that are adverbs was low compared to that of other word classes.

<sup>229</sup> The particles *lūr* and *mur* are originally one and the same. The *l-* of *lūr* is originally the definite article and *mur* represents a fusion from *mur* ‘from’ with *wra* ‘behind’. The forms *daxel*, *fuq* and *teħt* take the definite article always with the form *-l* (Harrell 2004: 211).

<i>bezzaf</i>	'a lot'
<i>hna</i>	'here'

Such single word insertions form part of the category of the most switched items in code-switched utterances and this is largely due, as previously discussed to the ease of lexical insertion of single lexical items as opposed to strings of bilingual discourse. Adverbs, together with nouns make up the largest switchable category of all grammatical categories. The following are examples of Moroccan adverbs inserted in bilingual frames:

(298) *daba je vais te parler un peu de ce type*  
 Now I am going you to talk a bit of DEM type  
 'Now I am going to talk to you a bit of this type'  
 (Moroccan Arabic / French, Ziamari, 2007)

(299) *ila rah hna guli lu y-come in fisaġ*  
 COND is 3SG here tell him come in 3SG quickly  
 'If he's here, tell him to come in quickly'

In terms of English adverbs inserted in MA frames there were numerous examples which were easily inserted and posed no problems for the bilingual speaker in terms of insertional placement or grammaticality:

(300) I really *bġit nemši Oamma* but Dad said no  
 I really wanted go 1SG there but Dad said no  
 'I really wanted to go there but Dad said no'

(301) *y-safir tomorrow u me y-wsel-š hetta la ġde li-h*  
 Travels 3SG tomorrow and NEG1 arrive NEG2 until DEF tomorrow to it  
 'He travels tomorrow and he doesn't arrive until the following day'

(302) *parfois yemken ma nebgi -š*  
 Sometimes possible NEG1 like 1SG NEG2  
 'Sometimes it's possible I don't want to'  
 (Moroccan Arabic / French, Ziamari, 2007: 157)

(303) personally *ana ma gultlu -š, šāf ha ʔamma maʕh*  
 Personally I NEG1 said NEG2, saw 3SG her there with him  
 ‘Personally I didn’t say anything, he saw her there with him’

The majority of adverbs occupy the same position as that of head phrase. The problem lies in their classification within the MLF as at times they are regarded as content morphemes and at other times, they are categorised as system morphemes. Is this a violation of the MLF? For example, the adverb *bezzaf* ‘a lot’ is classed as a system morpheme, but the English adverb ‘really’ is classed as a content morpheme. Just as Boumans (1998) questions within his MSA model<sup>230</sup> whether embedded Dutch adverbs are EL constituents or EL content morphemes. Ziamari (2007) also raises the issue of whether adverbs pose problems for Myers-Scotton’s MLF as adverbs are at times content morphemes and at other times they are regarded as system morphemes. Boumans (1998) discusses this point in his Dutch and Moroccan Arabic corpus and states that if we disregard adverbially used nouns, all the single embedded adverbs constitute independent constituents on their own. They are well-formed constituents in Dutch (the EL) and they function as adverbial constituents in the MA (ML clause). They never combine with MA morphemes to form MA adverbial constituents, unlike embedded nouns for instance which can be part of a larger matrix language NP. For this reason, singly embedded Dutch adverbs may be considered EL constituents but they are also EL content morphemes at the same time (1998: 280).

In conclusion, Boumans (1998) summarizes that while not all embedded adverbs can be analysed as instances of content word insertion, we can lump together singly embedded adverbs and adverbs modified by a degree adverb as being EL constituents (1998). The difficulty then lies in the classification of adverbs as not all

---

<sup>230</sup> See Chapter Three for a full analysis of Bouman’s (1998) MSA model.

adverbs operate in the same way. There are adverbs of place and time, adverbs of frequency and degree, manner adverbs and modal and conjunctive adverbs. Such adverbs have a certain flexibility in terms of positioning within a CP frame as they can occur in front, medial or final positions in bilingual frames without effecting grammaticality or well-formedness conditions. The following manner adverbs can both precede or follow the verb:

(304) kan ɛlik bɣiti tɣul -u ignorant tdewwwez *implicitment des message*  
 Was on you want 2SG to tell him ignorant to pass implicitly some message  
 ‘You wanted to treat him as ignorant in passing implicitly some messages’  
 (Moroccan Arabic / French, Ziamari, 2007: 158)

(305) walakin *implicitement implicitement* nta xess-k tefhem  
 But implicitly implicitly you have 2SG to understand  
 ‘But implicitly, implicitly you have to understand’  
 (Moroccan Arabic / French, Ziamari, 2007: 158)

Boumans (1998) Monolingual Structure Approach recognises this flexibility and states that such adverbs have their own syntactic rules of distribution which operate independently of the matrix language. Some of the most frequently inserted adverbs are those of manner which are frequently inserted in bilingual frames:

(306) *haði* normally doesn’t matter, *bessah al yawm* it does  
 DEM normally doesn’t matter but DEF day it does  
 ‘This doesn’t normally matter but today it does’

The main difference is where the adverb occurs in a matrix clause as an island in a matrix clause or whether it occurs as part of a mixed clause and is an internal embedded language island. It can then be said that adverbs of manner are content morphemes and this is perhaps also due to their high frequency in bilingual speech as they operate in the same way as singly inserted lexical items. The following examples highlight the distinction where the adverb is considered a system morpheme as opposed to a content morpheme:

(307) *bezzef d ness y-like-iw haða al* place  
Many of people like 3PL DEM DEF place  
'Lots of people like this place'

(308) *ħett-ih* *Əamma, xessna n-carr-iw bezzef d* things  
Put it (MASC) there, we have to carry 3PL lots of things  
'Put is down, we have to carry lots of things'

Example (307) shows how the adverb *bezzef* is an embedded language island and adheres to that particular structure within the mixed frame. The adverb *bezzef* in example (308) however is a matrix language island which functions in the normal manner and observes the grammatical frame as the matrix language. In both cases, the adverb is in Moroccan Arabic, the matrix language and they are both then system morphemes activated within the same maximal projection. Another element worthy of observation involves that of adverbs and word order. If the adverb occurs sentence initially or sentence finally then it is beyond the boundary of word order constraints and acts freely. However, sentence medially, it is difficult to determine the word order.

As noted by Boumans (1998):

To my knowledge there is as yet no detailed description of word order properties of the various types of adverbs in Moroccan Arabic. Moreover, the word order of adverbs is very complicated and because of this it is not possible to draw firm conclusions as to whether embedded Dutch adverbs in MA follow Dutch or MA word order (1998: 281).

The same analysis can be said for the current corpus and the status within Moroccan Arabic and English code-switched intra-sentential CPs whereby the complex word ordering of adverbs in general renders syntactic analysis difficult to define.

So do embedded language islands, and in this case, adverbs represent a problem for the MLF? As Myers-Scotton (2002) notes:

Embedded language islands represent a break in the Matrix Language Frame. Is this a problem for the model? No, but it does indicate that characterizing the

relationship of the participating languages and their level of activation during production is more complex than simply stating that the Matrix Language is in continuous control of the morpho-syntax in the bilingual CP (2002: 140).

In this regard then, adverbs in bilingual clauses do not pose a problem for the MLF and in fact Boumans (1998) describes how the MLF and in particular the: “CP analysis thus offers an elegant way out for some intricate problems concerning the status of adverbs.” (1998: 136). Myers-Scotton (2002) describes how even though the evidence provided by embedded language islands is that matrix language procedures are at the very least inhibited when embedded language islands are produced, it is important to note two characteristics of the bulk of embedded language islands across diverse data sets that imply a relatively lower level of activation for embedded language islands than for the mixed constituents under matrix language control. Further, as discussed above, many embedded language islands are adverbial phrases of time or place; that is they are adjuncts. This means that they are outside the predicate-argument structure projected by the main clause verb (2002: 140-141). Furthermore, adverbs as a property of categories which are [+ quantification] shows that they pattern as system morphemes in classic code switching data. Myers-Scotton (2002) states that:

If a quantifier is to be produced in the Embedded Language, then the phrase must be ‘finished’ in the Embedded Language too<sup>231</sup>. One finds Embedded Language islands beginning with a degree adverb e.g. English ‘very’ and the island, ‘very big’. There are instances as \*very nan ‘very difficult’ in the corpus (2002: 70).

In his analysis of adverbs in Dutch and Moroccan Arabic, Boumans (1998) finds that: “Various adverbs have been reported to maintain their source language word order properties when they are part of a finite clause in another language” (1998: 110-111). The same can also be said for present corpus of Moroccan Arabic and

---

<sup>231</sup> This is also supported by Wei (1998) in her Chinese / English corpus.

English intra-sentential clauses such that there is no mixing or overlap in bilingual clauses as follows:

Figure 7.3: Moroccan Arabic Classification of adverbs

MA / Eng Adverbial Clause	Gloss
*He is very <i>kbir</i>	'He is very big'
He is <i>kbir bezzef</i> <i>Huwwa</i> very big	'He is very big'
*She is very <i>mezyena</i>	'She is very nice'
She is <i>mezyena bezzef</i> <i>Hiya</i> very nice	'She is very nice'

Quite simply then, switching between adverbs in MA or English is ill-formed and is not produced in the corpus nor is it grammatical from an introspective analysis. Even within my innovative *Reactive Syntax* frame where second and third generations of Moroccan Arabic and English speakers use a new speech style, the above forms have not been recorded as this would syntactically be deemed a violation of the principles and parameters of language usage, in this case Moroccan Arabic and English. This then gives rise to the following generalization based on the literature review, data analysis and grammatical evidence:

*Generalization 10*

*There cannot be switching of mixed constituents, namely MA or English adverbs within a maximal projection.*

Now that we have established the suitability of the MLF in adverbial phrases as clearly-defined system morphemes, how does the model fare in other grammatical categories? The next section evaluates the insertion of Moroccan Arabic conjunctions in intra-sentential clauses.

### 7.1.3 Moroccan Arabic Inserted Conjunctions

The use of conjunctions from one language in the context of another is a recurrent characteristic of bilingual conversations. Conjunctions from the ‘other’ language are particularly favoured as discourse markers, perhaps because they are more salient (De Rooij, 1996), Boumans (1998: 117).

Amongst the category of most frequently inserted switches together with nouns and adverbs are conjunctions. The most commonly used of these are as follows:

Figure 7.4: Moroccan Arabic conjunctions

MA Conjunction	Gloss
<i>wella</i>	‘or’
<i>walakin</i>	‘but’
<i>ɛla heqqeš</i>	‘because’
<i>be l heqq</i>	‘but’
<i>beš</i>	‘so that, in order to’
<i>ila</i>	‘if’
<i>mnin</i>	‘since’
<i>gir</i>	‘only’
<i>hit</i>	‘since, when’
<i>hetta</i>	‘when, until, so that’

Within the MLF, conjunctions are content morphemes in the embedded language frame and as discussed previously, they are defined by the feature [+ thematic role assigner / receive] and are the main elements conveying semantic and pragmatic aspects of messages (2002: 15). They are embedded language islands which are full constituents consisting only of singly occurring embedded language morphemes in a bilingual CP that is otherwise framed by the matrix language. The embedded language island shows structural dependency relations; minimally it can be two content morphemes (e.g. noun and modifier) or a content morpheme and a non-derivational system morpheme (2002: 139). Examples of Moroccan Arabic conjunctions within bilingual clauses are as follows:

(309) *Ik ben dokter wella ik ben ingénieur*  
 I am doctor or I am engineer  
 ‘I am a doctor or I am an engineer’  
 (Moroccan Arabic / Dutch, Nortier, 1990: 142)

(310) *ħett-ih guddam -i beš n-find* it in the morning  
 Put it (MASC) in front me so that find-1SG it in the morning  
 ‘Put it in front of me so I can find it in the morning’

(311) *ħit min kunt fi school, ma qrit -š mezyen*  
 Since when was in school, NEG1 study NEG2 good  
 ‘When I was at school I didn’t study well’

In the first example above (309) *wella* ‘or’ is one of the most frequently inserted of all the MA conjunctions and conjoins two Dutch clauses. It would seem, given the literature, that insertions of conjunctions are very common and bilingual speakers insert them not only with ease but also with accuracy. Lipski (1978) in code-switched analysis of English and Spanish gave numerous examples, one of which is presented below:

(312) *no créian en jésus and then he sent este hombre*  
 No believed in Jesus and then he sent DEM man

‘They didn’t believe in Jesus and then he sent this man’  
(*English / Spanish*, Lipski, 1978: 256)

Boumans (1998) in his analysis of conjunction insertion in mixed clauses states:

One of the motivations for qualifying complementizers and subordinate conjunctions as governors in GB syntax, is that these elements coincide with a subjunctive mood on the verb in the verb in the following clause, or in some languages like Dutch and German, a distinct word order. The subordinate clause is regarded as the complement of C. CS data seem to reinforce this idea to some extent. In German and Dutch, verb-final order does not follow after conjunctions from other languages such as French or English. This might be explained as a property of the French and English conjunctions since there is no subordinate clause word order in these languages. On the other hand, Dutch and German subordinate conjunctions typically do not have any structural effect on the clause they mark. This is due either to their status as ‘unintegrated’ foreign forms, or to their serving discourse functions that are not structurally marked inside their clause. In any case, the fact that ‘code-switched’ conjunctions fail to produce structural impact on the clauses they conjoin makes it less appealing to associate them with the C node in the hierarchical structure CP (1998: 141).

English conjunctions operate in the same manner in terms of their placement within the CP in code-switched clauses:

(313) *la, bga yamši* because *ma kammal -š* homework *dyalu*  
No, wants go 3SG because NEG1 finish NEG2 homework POSS  
‘No, he wants to go because he didn’t finish his homework’

(314) *šāf al warqa* but *ma zed -š beš yadfeḡ d- dossi*  
Saw DEF paper but NEG1 go forward NEG2 so that push 3SG DEF dossier  
‘He saw the paper but he didn’t go forward to push ahead with the file’

(315) *parce que* huwwa kān bi xir  
Because he was with good  
‘Because he was fine’

(Moroccan Arabic / *French*, Ziamari, 2007: 163)

According to Slaoui (1998), the placement and categorisation of such markers including conjunctions is difficult to analyse as:

They are elements of nature which function differently, which seem interesting from a structural perspective, articulation and organisation of the discourse. (1998: 62-63).<sup>232</sup>

Ziamari (2007) also cites the difficulty of classifying certain insertional particles, namely discourse markers, adverbs and conjunctions:

L'étude linguistique de l'insertion des conjonctions, locutions conjonctives et marqueurs énonciatifs peut s'aligner sur celle des adverbs. Difficiles à classer, ils ne trouvent pas toujours leur place dans tous les modèles syntaxiques (2001: 162).

Boumans (1998) states that:

Conjunctions in particular have often been cited as counter-examples to matrix language models which, like Myers-Scotton's MLF model, predict that function (system) morphemes originate from the MLF (e.g. Eliasson, 1995: 56; De Rooij, 1996: 161). There are of course valid arguments supporting the classification of conjunctions with function rather than content morphemes; their use is better explained in terms of referential meaning (1998: 109).

Within the domain of morpho-syntactic theory, it has been clearly stated within this thesis and in Myers-Scotton's (2002) MLF model that conjunctions are opaquely identifiable as content morphemes as they assign thematic roles on the syntactic discourse level. This then cancels out the above claim by Boumans. In further analysis of Moroccan Arabic conjunctions, we refer to Bentahila & Davies (1983) and their sub-categorization constraint which is formulated as follows:

All items must be used in such a way as to satisfy the (language-particular) sub-categorization restrictions imposed on them (1983: 329).

Therefore, this constraint rules that code-switching may take place provided that the language-specific sub-categorization selection requirements of all categories (lexical

---

<sup>232</sup> MacSwan (1999) also cites difficulties in classification of certain conjunctions whereby he states that "With respect to conjunctions involving 'because' boundaries are more difficult to determine for South East Puebla Nahuatl" (1999: 136).

and functional alike) are satisfied. In the case of *beš* as described above, the Moroccan Arabic complementizer /*beš*/ always selects finite clauses as complements. This requirement also holds in the following example where /*beš*/ takes an English finite clause:

(316) *xessu y-go beš y-get there on time*  
Must go 3SG so that get 3SG there on time  
'He has to go so he gets there on time'

(317) *gra beš njeħ*  
Study PAST 3SG so that pass PAST 3SG  
'He studied and so he passed'

(318) \* *gra beš pass*  
Study so that pass

In (318) above, the constraint is violated as the Moroccan Arabic complementizer is not followed by a finite clause as exemplified in (316) and (317) above and this is clearly the basic premise of the Sub-categorization constraint. Although this constraint is valid for the data in this thesis as logged by the respondents in Moroccan Arabic and English discourse, Nortier (1990) recorded nine violations of the Sub-categorization constraint in her Dutch and Moroccan Arabic data and states that:

The sub-categorization constraint as it has been formulated by Benthila and Davies (1983) cannot be adopted as it stands in order to account for Dutch-Moroccan Arabic code switching. In fact, the sub-categorization constraint is concerned with government phenomena so it is not surprising that an analysis of both the government and the sub-categorization constraints lead to the same conclusion (1990: 178).

The data and conclusion as presented by Nortier (1990) seem odd in the sense that such markers as recorded by my informants and also introspectively are always followed by a finite clause and in the example below, it seems syntactically to be ill-formed and it is questioned whether the speaker is fluent in both Moroccan Arabic and Dutch:

(319) yeqraw beš vooruitgaan  
Study 3PL so that progress  
'They study in to order to progress'  
(*Dutch / Moroccan Arabic*, Nortier, 1990: 178)

Furthermore, in analysis of intra-sentential code-switched data an alternative constraint as posited by Poplack (1980), Sridhar and Sridhar (1980) and Berk-Seligson (1986) is that of the size-of-constituent constraint which is interpreted by Sridhar and Sridhar (1980) as:

Although elements from practically every syntactic category (including purely grammatical morphemes such as determiners) occur in code mixed sentences, it has been found that certain types of elements are more likely to be mixed than others. In general, except for single words especially nouns, the higher the constituency of the element, the more likely it is to be mixed: thus, conjoined sentences, main clauses, subordinate clauses including relative clauses, major constituents such as noun phrases, verb phrases and prepositional phrases are among the most frequently mixed elements. Among single words, nouns outrank all others in frequency of mixing, followed by adjectives, adverbs and verbs. Grammatical items such as articles, quantifiers, auxiliaries, prepositions and clitics are least likely to be mixed by themselves (1980: 205).<sup>233</sup>

This constraint grades the level of constituents and their frequency of usage in switched discourse. However, it must be noted in general terms as well in the data presented that code-switching can occur at different levels with low as well as with high level constituents so then this constraint is only valid if analyzed from the perspective of level of constituents together with frequency of insertion and this is a general observation rather than a syntactic theory of detailed constraint.

However, given seemingly difficult placement of discourse markers in syntactic structures, there is no difficulty in the placement or categorisation of conjunctions or any discourse marker within the MLF and this can be said for Moroccan particles inserted in English matrix clauses and also English particles

---

<sup>233</sup> Poplack (1980) has found that the size-of-constituent constraint leads to the conclusions that the ability to code-switch intra-sententially is a measure of bilingualism.

inserted in Moroccan Arabic matrix clauses. Can the same be said of general discourse markers and do they adhere to the MLF model and its parameters? This is examined and answered in the following section.

## 7.2 MA Discourse Marker Insertions

Discourse markers are a frequently inserted category of marker in the corpus due their ease of insertion and their classification as a very heterogeneous type of morpheme.

Boumans (1998) describes discourse markers as being a group of markers which:

Form a heterogeneous group of expressions that include members from various word classes. The linguistics properties are at least partly dependent on their word class membership, and because of this it is difficult to come up with a clear definition of discourse markers (1998: 106-107).

Discourse particles play a key role in forming coherence and cohesion in natural segments of discourse.<sup>234</sup> Structurally, such markers can be found turn initially, medially and finally with no recourse for categorical placement and are in essence the ‘wild card’ of spoken discourse. Their highly multifunctional status provides an interesting backdrop when analysing two typologically different languages as I do in this thesis, namely, Moroccan Arabic or *darija* as it is known locally and English. I analyse the use of three of the most salient discourse markers in *darija* MA in a code-switched domain; namely, *iwa* ‘well’, *zaɣma* ‘in other words’ and *bassaħ* ‘but’. *Darija* discourse markers<sup>235</sup> occur in prototypical placement in natural, unhinged bilingual discourse and as described below are multifunctional whether employed to seek listener agreement, highlight a point, move on to a new topic and other such paralinguistic purposes.<sup>236</sup> The particles in question are some of the most commonly used both amongst British born and Moroccan born speakers of *darija* inserted at

---

<sup>234</sup> See Benchiba (2008).

<sup>235</sup> *Zaɣma* ‘so to speak’ is actually classified as an adverb of modality.

<sup>236</sup> Ziamari (2007) found that the majority of insertions using *zaɣma* were made by females in her corpus of Moroccan Arabic and French code-switched discourse, which is an interesting observation.

syntactically convenient points of utterance that either add contextualizing clues to an utterance or as will be addressed, act as cultural banners amongst second generation British Moroccan bilingual speakers. Discourse markers are independent particles and are a law unto themselves in the sense that at both deep and surface structure syntactic analysis, they do not have to adhere to any internal force for grammaticality and can 'opt in' or 'opt out' of structures without recourse for ungrammaticality. Such markers do not need to conform to any syntactic or functional category and are in fact polysemous with no real attested cases of misuse or anomalies in natural discourse. Discourse markers may be considered as contextual coordinates of talk with a special kind of contextualization cues (Gumperz 1982). Saeed (2003) describes how the marking of epistemic modality may not affect the truth value of the propositional content, but it does mark the speaker's assessment of the validity of the propositional content. Therefore, the inclusion or exclusion of discourse markers does not alter the truth state or value of a sentence as also further discussed.

Discourse markers act as quasi-categorical and lie on the periphery of grammatical structure, proffering cohesion and greater coherence to interlocutors within a natural speech domain. Such connectivity enables the speaker to better manage his discourse, interject, mark attitude, agreement, validate an idea or thought, and in general manage the interlocutor's conversation.<sup>237</sup> Givón (1993) in a theory-driven approach to discourse particles attests that the propositional modality associated with a clause may be likened to a shell that encases it but does not tamper with the kernel inside. The propositional frame of clauses...as well as the actual lexical items that fill the various slots in the frame, remain largely unaffected by the modality wrapped around it. Rather, the modality codes the speaker's attitude toward the proposition (1993 I: 169). Although Givón's (1993) work was based on an

---

<sup>237</sup> For a more functional analysis see Givón (1979) for a broad syntactic, theory-based syntactic analysis.

abstract theory and not any specific data sets, I concur with the notion that discourse markers are extensions of a speaker's mindset and attitude, a pausing tactic to digest, deliberate then respond and are only associated with natural unwritten conversation and the more informal the conversation and setting, the greater the usage. The heterogeneity of their class and lack of grammatical categorization speaks volumes for the true independence of these particles and can be considered a partial semantic invariant. Studies on Arabic discourse particles are starting to gain momentum Holes (1995),<sup>238</sup> Ennaji (2005) where in translation, many of the Arabic discourse markers have a zero value due to heavy use in the Arabic code. Such discourse are not heavily used nor are they inserted as often as possible but fulfil the role of contextualizing discourse. Discourse markers, or *darija particles* under analysis are the most frequently inserted in the corpus, namely *iwa*, *zaɣma* and *bassaħ* and these were elicited when transcribing the recorded dialogue.

### 7.2.1 Discourse Markers: Uses, Sequencing and Turn-Taking

*Darija* particles can be used in multifarious contexts serving a wide range of functions. Intonation, tempo and turn-taking cues provide the backdrop to where, syntactically, they can be placed in natural dialogue. The below Figure 7.5 sets out the structural positioning of the three particles in question:

---

<sup>238</sup> For an analysis on the Arabic discourse particles '*fa*' and '*wa*' see Holes (1995) who explains the consequential relations between discourse units.

Figure 7.5: Moroccan Arabic discourse markers

	Interrogative	Imperative/ Exclamatory	Initial	Medial	Final
<i>Iwa</i>	✓	✓	✓	✓	*
<i>Bassaḥ</i>	✓	*	✓	✓	✓
<i>Zaḡma</i>	*	*	✓	✓	✓

Prosody within natural conversation then is of prime importance when analysing the *darija* particles as prosodic cues then gives licence to the hearer/interlocutor to decide which marker to use if any. Changes in topic, and subject reorientation can also motivate the use of different markers within different discourse environments. In light of the above discourse markers as set out in Figure 7.5, the following gives an example of subject reorientation using *zaḡma*:

(320) (a) - *u wašta ngulu then? Rani qlil -š fin n-šufu fin aḡd*  
 And what say 3PL then? I little thing where see him 1SGwhere just  
*na-talk 3PL ma'ḡh.*  
 talk with him.  
 'And what should we say then? I rarely see him let alone speak with him'

(b) - *iwa, bint-i, inti taḡrfi, best ila ḥdārt u gult ši hejja,*  
 well, girl POSS, you know 2FSG best, if say 1SG and say some thing,  
*ana* my fault  
 I my fault.  
 'Well my daughter, you know, if I talk and say something, it'll be my fault'

(a) - *zaḡma* I'll leave it then..?  
 So I'll leave it then..?

(b) - yes, *bint -i, xall-ih.*  
 Yes, daughter POSS, leave it (MASC).  
 'Yes, may daughter, leave it'

(a) - *wašta ta fanci-w* for supper?  
 what you fancy 2PL for supper?  
 'What do you fancy for supper?'

*Zaġma* above suggested to the hearer that the content of the conversation had to move on as did the contact with the person in question, and also that there was nothing more to say on the subject and so the conversation moves on. In terms of positioning, (see Figure 7.5 above), the *darija* particles can occur in either front, medial or final positions except *iwa* which does not occur sentence or clause finally. This finding is corroborated by Ziamari (2007) who also found that such particles can be placed as head phrase, sentence medially or finally. It can be said that such particles have a positional flexibility. Some examples of different positions are as follows:

(321) *zaġma on a trop formé les yeux*  
 So to say one has a lot formed DEF eyes  
 ‘One has made up eyes a lot’  
 (Moroccan Arabic / *French*, Ziamari, 2007: 217)

(322) *bon je te parle zaġma de mon premier semestre*  
 Good I you speak so to say of my first semester  
 Ok, I will tell you so to say, about my first semester  
 (Moroccan Arabic / *French*, Ziamari, 2007: 217)

(323) *šufi maġh, zaġma xessu yafhem walla what?*  
 See you 2SG, with him so to say, has to understand 3SG or what?  
 ‘See what he says, so to speak, he has to understand or not?’

Content largely dictates where they can and are positioned as does intonation of the interlocutor, contextual cues, acceptance of content of speech, disagreement, pausing and agreement, all of which are interrelated. Also there is the function of reported speech where *zaġma* can be used to paraphrase, to highlight a point that was misunderstood or to make a point. The below example illustrates this:

(324) (a) – *gel -li bassah ma taqtuš* and I told Mum about it.  
 told 3SG to me but NEG1 believe him NEG2 and I told Mum about it.  
 ‘He told me but I didn’t believe him and I told Mum about it’

(b) – about what?

- (a) – *iwa kamli min rās -ak*, what do you think!  
 well finish from head POSS, what do you think!  
 ‘Well put two and two together, what do you think!’
- (b) – *zaɣma ħta l ɣam jāy.. fi l* summer.  
 in other words until DEF year coming, in the summer  
 ‘In other words, until next year, in the summer’

The above conversation between two British-born second generation Moroccans is revealing with regard to the usage of the *darija* particles. In the corpus, there is usually greater propensity for discourse particles amongst second generation speakers in the following scenarios:

- Subject matter is culturally/ethnically related
- In group is of similar Moroccan background and so speakers feel ‘free’ to use the *darija* particles
- Emblematic usage where the lexical items are used to show ‘Moroccan-ness’
- Stating or reinforcing the cultural stamp of the group and so is inclusionary
- In mixed groups, i.e. with non-Moroccans and so is exclusionary
- To show off and portray mastery of the MA amongst peers and first generation speakers
- In using ‘flashes’ of the *darija* particles to show that a greater fluency of the MA is known.

As other research has shown (Nortier 1990),<sup>239</sup> second generation MA/English speakers often use *darija* particles to coordinate conjunctions linking English sentences and the reason for this is as attested above. The above list of observed uses has not been recorded amongst first generation speakers and the use of the *darija* particles are used as naturally as they are in general conversation. However, using the particles to conjoin English sentences has been evidenced amongst all informants for this research.

<sup>239</sup> For further analysis see Myers-Scotton (1993b) on Swahili code switching, and Nortier (1990) on Moroccan Arabic and Dutch intra-sentential bilingual discourse.

## 7.2.2 Functionality amongst *Darija* Particles

The *darija* particles *iwa*, *zaʕma* and *bassaħ* in general can be described as semantically independent particles in monolingual discourse and even more so in bilingual discourse in the sense that there can be no need for any inherent *silsila* or chaining to any other particle or even a discourse unit. In other words, they can and do function as stand alone particles, syntactically peripheral and this very much depends on intonation, tempo and turn-taking cues. Of the three markers in question *iwa* and *bassaħ* are very flexible in that they can be employed in different linguistic scenarios, interrogatives and imperatives. *Zaʕma*. However, is less flexible in that it is not used in imperatives and is more used as a pausing tactic and it also has the slowest prosodic features in that it can drawn out for much longer than *iwa* and *bassaħ*. The below samples illustrate some examples taken from various recordings:

(325)

- (a) - *u zaʕma, kifeš?* Did he want to come *wella ba gal-u la, hadri!*  
and, well, what? Did he want to come or Dad tell him no, talk!  
'And well, what? Did he want to come or did Dad tell him not to? Talk!
- (b) - *Iwa, intiya shufti maħ-h, illi smʕtu* I've told you, *wallah!*  
well, you see with him, REL heard 1SG I've told you, really  
'Well, you see with him, what I have heard, I've told you, I swear!'

In (325) above, speaker (a) employs *zaʕma* at a slower tempo than the rest of the discourse, thus allowing for time and space for the rest of the sentence, syntactically inserting it in a slot that allows for such independent particles. However *iwa* above, can be used at the fastest speed and is very effective in an exclamatory sense, but here, with a rising intonation it asks the interlocutor to take action and so is emphatically used to incite action and invite a response. These markers can be and are used in a number of different environments which also serve to highlight their poly-functionality and pluralistic nature. Interestingly, the above conversation was

recorded of two British-born MA speakers and such discourse markers were used in all recordings.<sup>240</sup> Further examples of two in-group speakers highlight the naturalness of the particles:

(326)

(a) - *za'ama. Wallah*, that's what I heard. *Bint le hrām*, after all  
Well. I swear, that's what I hear. Daughter DEF badness, after all

*illi deret u mazal* ta-carry on.

REL did and still carry on 3SG

'Well, well. I swear, that's what I heard. Bitch, after all she did and she still carries on.'

(b) - *iwa, xas-ha ta-calm down swiywa, bassah, ðork rabbi*  
Well has 3SG (FEM) calm down 3SG a bit, but, now God

*ya-calmi -ha, futu ši yemet u aqli fi hadar*  
make 3SG calm her pass some days and remember 2SG in word –  
*ti.*

POSS.

'Well she has to calm down a bit, but soon God will make her calm down. Let a few days pass as remember my warning'

In the above example, the girl in question is a British-born Moroccan and her actions belie what should be normal practice for a Moroccan girl hence such usage of the *darija* particles. My corpus research, further examples of which are given in (327) to (329) below, shows that subject matter is of great importance and a catalyst for using such particles and provides great motivation for inserting *darija* markers into English conversation. Also, there are strong links with aspects of the 'bled' or the 'motherland' that serves as another very strong motivational factor amongst British-born Moroccans. Furthermore, the use of such discourse markers speaks volumes about speaker competence and proficiency:

(327) Hiya, *iwa*, how are you? *La* telephone, *la briya*, oh well, I expect you're  
Hiya, well how are you? No telephone, no letter, oh well, I expect you're

<sup>240</sup> I find the use of discourse markers in second language learners as well as second generation speakers provide further emphasis and evidence of fluency. The greater the knowledge of discourse markers, the greater the knowledge and fluency in a second language.

busy these days, *bessah* don't forget me *safi*?  
 busy these days, but don't forget me enough  
 'Hi, well, how are you? No telephone, no letter, oh well, I expect you're busy  
 these days, but don't forget me ok?'

(328) I miss *al Maghrib* so much, *al familia njatmaξu* around *bered atay*, *al*  
 I miss DEFMorocco so much, DEL family gather 3PL around pot tea, DEF

*aḏān*, *bessah ma ξ alineš*, next year *inshaAllah*.  
 Call to prayer but, NEG1 on us next year *inshaAllah*.  
 'I miss al Maghrib so much, the family gathering around a pot of tea, the call  
 to prayer, but never mind, next year *inshaAllah*'

(329) *š-hel min marra wa na-an ξawanak, zaξama ma h-šemt -š*  
 How many from time and help you 1SG, so to say NEG1 ashamed NEG2  
 or what?  
 or what?  
 'How many times did I help you, so to say, and you are not ashamed or  
 what?'

As Backus (1996) states:

Speakers who are dominant in one of the languages, mainly have insertional  
 CS with that language as the ML (1996: 389).

This then clearly highlights structural differences between inserted items  
 within the matrix language and the peripheral embedded elements. As has been stated,  
 the matrix language serves as the fulcrum upon which insertional elements hinge and  
 this in essence is the intrinsic importance of its asymmetry. The notion of a matrix  
 element should also be credited to Hasselmo (1972) who "*was ahead of his time*"  
 (Boumans 1998: 29) in that his model provides a uniform and systematic principle  
 for the insertion of single content words, inflected content words, function words and  
 constituents in general. His Ordered Selection hypothesis, as described in Chapter  
 Three, is built upon by Petersen (1988) who states:

The dominant language hypothesis states that in word-internal code switching,  
 grammatical morphemes of the dominant language may co-occur with lexical

morphemes of either the dominant language or the non-dominant language. However, grammatical morphemes of the non-dominant language may co-occur only with lexical morphemes of the non-dominant language (1988: 486).

This then was a pre-cursor to the MLF in the sense that the notion of asymmetry was born through the study of his Swedish and English data and has a lot in common with the matrix language – embedded language hypothesis but does not provide a detailed account of diverse language groups, types and different levels of fluency in the same way as the MLF. Given the above, how then do we address discourse markers within code-switched clauses and within the MLF?

### **7.2.3 Discourse Markers within an Asymmetrical Framework**

On a morpho-syntactic level, Further, Boumans, in his description of the placement of discourse markers within mixed CP clauses cites that:

The ML does not predict from which language discourse markers will be drawn, nor does it predict anything about the distribution of these markers. More generally discourse grammar interferes with the finite clause level in such a way that it cannot be handled satisfactorily within clause syntax (1998: 109).

However, Boumans, in citing that the MLF model is not a suitable vehicle upon which to evaluate certain discourse markers then concedes that:

It is important to note that even in monolingual contexts the syntactic distribution of discourse markers cannot be fully explained within sentence grammar since they function entirely or partly on the level of discourse organisation. Related to this is the problematic classification of various markers into word classes such as conjunctions, adverbs, interjections or simply discourse markers. The distinction between these classes is not always clear-cut and each word class tends to be rather heterogeneous in itself (1998: 106).

While it is true to say that in Myers-Scotton's original MLF model (1993b), discourse markers were categorised as system morphemes, in a later revision (1997), on further

analysis, Myers-Scotton classified them as content morphemes. This is still valid in the current MLF model together with the 4-M Model (2002) and overrides the original model as such discourse markers assign thematic roles at the level of discourse (1997a: 255). Furthermore, as Myers-Scotton (2002) states:

Another issue is the status of discourse markers. I argue that they can be considered content morphemes at the discourse level. We argue that whatever can appear in the position of Comp in a CP and whatever can occur in the position of Spec of Comp can also be a discourse-thematic element. (However, in some languages, e.g. Arabic, complementizers may be multi-morphemic, including a system morpheme in addition to the discourse marker itself) (2002: 70).<sup>241</sup>

Within the MLF model and on the basis of further research, discourse markers have been analysed as operating on the syntactic level as content morphemes. As Boumans (1998) states “Since many discourse marking particles do not display clear syntactic properties, they generally do not challenge the ML concept” (1998: 110). Syntactically, then it is difficult to clearly define and categorise discourse markers at this level. However, the 4-M Model (2002) not only accounts for a wide range of code switching examples, but can also explain most of them under the Matrix Language – Embedded Language and the content-system morpheme oppositions.<sup>242</sup> Can the 4-M then also account for the phenomenon of pronoun doubling?

---

<sup>241</sup> In further citing the distinction between the content-system morpheme hierarchy, Myers-Scotton states that “A feature [+/- quantification] also distinguishes the two types of morpheme (cf. Myers-Scotton 1993a [1997], although the [+/- thematic role assigner/receiver] feature is sufficient alone. Under the notion of quantification, one can argue that ‘any lexical item belonging to a syntactic category which involves quantification across variables is a system morpheme’ (1993a: 100)” 2002: 70)).

<sup>242</sup> Recall that the 4M Model divides morphemes into four types, content morphemes and three types of system morpheme. The distribution of these morphemes across diverse data sets motivates hypotheses about how the different types are activated in language production. (cf. Myers-Scotton 2002: 194).

### 7.3 Moroccan Arabic Pronoun Doubling in Intra-Sentential Discourse

Pronoun doubling is a phenomenon whereby two pronouns with the same pronominal features occur adjacent to each other in what appears to be the same syntactic position.<sup>243</sup> In bilingual data, the pronouns are drawn from different languages and are subject to directionality constraints (Eid, 1994). This phenomenon is described by Myers-Scotton (1996) who refers to a specific pattern which occurs in CS data sets involving Arabic: two pronouns appear before a verb in CS discourse (1996: 25). Linguists (Bentahila & Davies 1983; Eid 1992) have mainly focused on the pattern whereby the first pronoun is in Arabic, which forms the matrix language and the second subsequent pronoun is in the embedded language. Personal pronoun doubling may emphasize the subject, object or predicate in any given clause. Cowell (1964) makes a noteworthy comment when he cites that the use of the extra-posed pronoun makes a predicate stand out from its context and sound more insistent (1964: 549). Is this why pronoun doubling is so productive in all Arabic dialects and also within code-switched discourse? As Eid (1994) notes, explaining these constraints, pronoun doubling is analyzed as being the result of Spec-Head Agreement and Case assignment properties of the languages involved. The analysis is proposed to account for doubling in bilingual code switching and in monolingual data as well. The following highlights an example of a code-switched data set and the pronoun doubling in question:

- (330)...*inn humma* they become conscious of it  
That 3PL they become conscious of it  
'That they become conscious of it'  
(*Egyptian Arabic / English*, Eid, 1994: 8)

---

<sup>243</sup> This was originally termed 'Arabic repetition' by early linguists. cf. Holes (1995a) Eid 1996) and also referred to as asymmetries in pronoun distribution (Bentahila & Davies 1983; Eid 1992; Azuma 1991, 1993).

(331) *ana j'approve ton point de vue*  
I approve 1SG your point of view  
'I agree with you'  
(Moroccan Arabic / *French*, Ziamari, 2007: 146)

(332) *huwwa he's never around beš na-talk mašh*  
He he's never around so that talk 1SG with him  
'He's never around so I can talk with him'

In the above examples, the subject pronoun is followed by English / French pronouns and agrees in number and where necessary, gender as in (331) above. Therefore, Moroccan Arabic, as well as other forms of Arabic pronouns agree in their entirety and are in essence mirror forms of each other. Bentahila & Davies (1983) also analyze the use of pronoun doubling in Moroccan Arabic and English data:

(333) *yašni ana* I was really lucky  
Meaning I I was really lucky  
'I was really lucky'  
(Moroccan Arabic / *English*, Bentahila & Davies, 1983: 313)

In analysis of bilingual clauses where pronoun doubling is examined, Eid (1994) states that:

The data are in fact perfect examples of bilinguals maintaining the syntactic integrity of both languages and [such] structures do in fact satisfy the requirements of both Arabic and English grammars simultaneously. (1994: 16).

Following Chomsky (1986) where C is head of CP this therefore gives rise to the following:

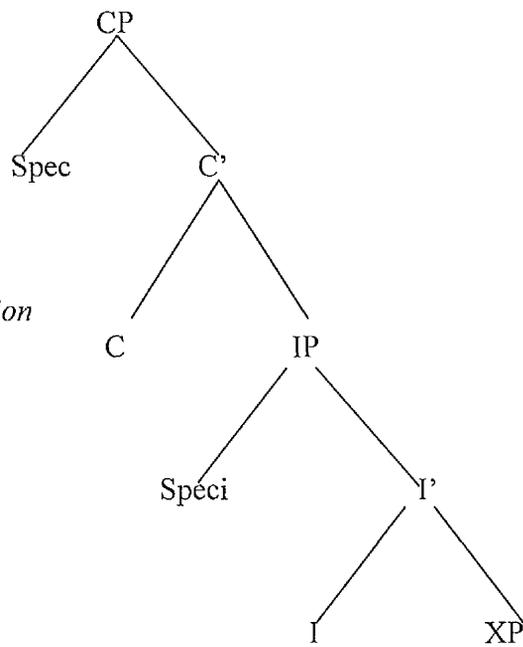


Figure 7.6: Chomsky's (1986) representation

Therefore, in light of the above model, the derivation of pronoun doubling in bilingual intra-sentential clauses shows how the English pronoun occurs in the position of Spec of IP in the manner such that there is full agreement of features within Spec-Head agreement:

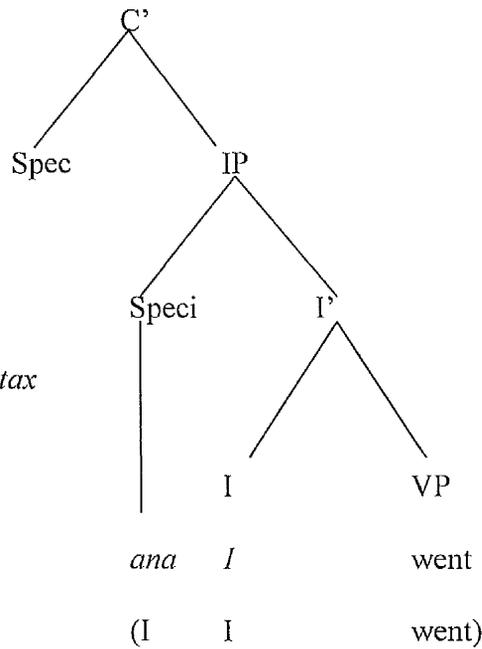


Figure 7.7: Moroccan Arabic Syntax

As Eid (1994) describes it, the above example highlights how the Moroccan Arabic pronoun *ana* occurs in Spec of IP and the English pronoun occurs in Infl and is

the result of lexicalization of Agreement features. Boumans (2000) rejects Eid's concept of the doubling of subject pronouns and instead affirms that we cannot generalize with regards to the co-indexation of the doubling of topic and subject pronouns. In typologically different languages such as Moroccan Arabic and English, if for example the MA provides the topic pronoun (*ana* 'I', *nta* 'you', *ħna* 'we') then English will provide clitic pronouns (I, you, he, she etc) as outlined below:

- (334) *ana* I was really happy to go there  
 I I was really happy to go there  
 'I was really happy to go there'

Interpreted as:

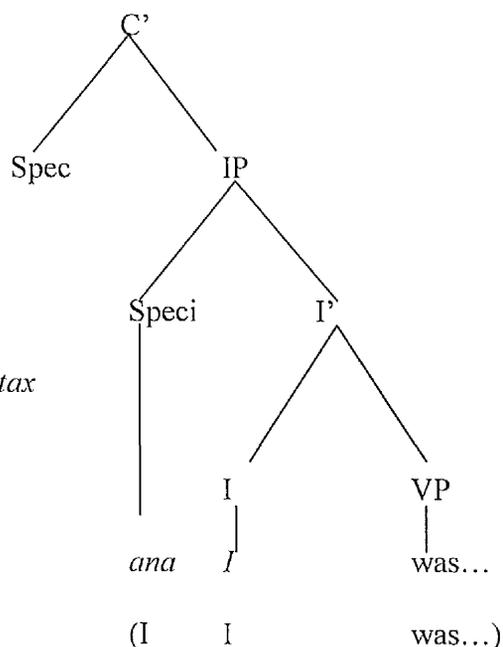


Figure 7.8: Moroccan Arabic Syntax

Therefore, contrary to Boumans, such syntactic structures in fact favour pronoun doubling in subordinate clauses and this perhaps why they are so productive in both Arabic monolingual and bilingual discourse. Eid states that the nature of pronoun doubling in Arabic dialects is very systematic and that it is the result of interaction between certain principles of the grammatical systems of the two languages involved whereby code-switchers attempt as much as they can to preserve

the ‘integrity’ of both grammars (1994: 28). In code-switched corpora, how does the phenomenon of pronoun doubling relate to the MLF and 4-M Models?

### 7.3.1 Moroccan Arabic Pronoun Doubling within the MLF

Given the above data and explanations, it will be shown that pronoun doubling occurs within the parameters of the MLF model and clearly within the hierarchy of the asymmetrical framework. It has already been shown that structurally, pronoun doubling is predicted with an Arabic topic pronoun if Moroccan Arabic is the matrix language and the following verb is in the embedded language island or Arabic is the embedded language and the verb is in mixed constituents. The ordering of the pronouns by reference to language where Moroccan Arabic is always first opaquely refers to the asymmetry and the roles of the ML – EL distinction within the MLF. In addition, the ML always provides the grammatical frame in bilingual clauses and the inserted EL islands must be well-formed according to the embedded language. Wherever there is insufficient congruence between the two (or more) respective language, the speaker is required to draw upon certain compromise strategies.

Pronoun doubling occurs when there is a lack of congruence between the agreement systems of two languages, in this case, Moroccan Arabic and English and it is then predicted to occur with an Arabic topic pronoun in first position if Moroccan Arabic is the matrix language and the verb is in the embedded language island or Arabic is the embedded language and verb is in mixed constituents. Myers-Scotton (1996) describes how “An Arabic full form pronoun can be followed by a pronoun from the EL.” (1996: 26). This is further illustrated as follows:

(335) *huwwa* he’s always going on and on!  
He he’s always going on and on!  
‘He’s always going on and on’

(336) *ħna* we can go *besseħ xessnana* na-come back early  
We we can go but must 1PL come 1PL back early  
'We can go but we must come back early'

The syntactic patterning is of interest as the first pronoun which is in Arabic, the matrix language and is followed by the second pronoun which is in the embedded language is rarely produced in the opposite form, namely with the first pronoun in the embedded language, in this case English followed by the subsequent pronoun in Moroccan Arabic. We re-visit (336) above and draw a hypothetical comparison:

(337) \*we *ħna* can go *besseħ xessna* na-come back early  
We we can go but must 1PL come 1PL back early  
'We can go but we must come back early'

The above is clearly ill-formed as the matrix language is Moroccan Arabic and not English and the ordering of the pronouns must be that Moroccan Arabic is always first as this sets the grammatical frame of the whole clause and therefore it is an Arabic full-form pronoun which should be followed by a pronoun from the EL, English in the case above. As explained by Myers-Scotton (1996):

If Arabic is not the ML, then pronoun doubling is not possible since even if Arabic supplied the topic pronoun, there is a null in argument position of the IP E1 island since it is framed by Arabic. Further if Arabic is the EL and supplies the topic pronoun, doubling is still not possible. Such examples are not unattested (1996: 37).

Furthermore, Myers-Scotton adds:

If Arabic is the ML and Arabic supplies the topic pronoun, then pronoun doubling must occur. The topic pronoun is in Spec of CP, while the second pronoun satisfies the requirements of the EL for the realization of its subject/ Following the Arabic topic pronoun, any of the following may occur in the EL island; a full form pronoun in argument position from [English] or a null in argument position but a subject clitic from [another language]. If Arabic is the ML and Arabic does not supply the topic pronoun, then doubling is unlikely. The only way for doubling to occur would be for the EL to supply both the

topic pronoun and the argument position pronoun or clitic pronoun of the verb in IP (1996: 37).

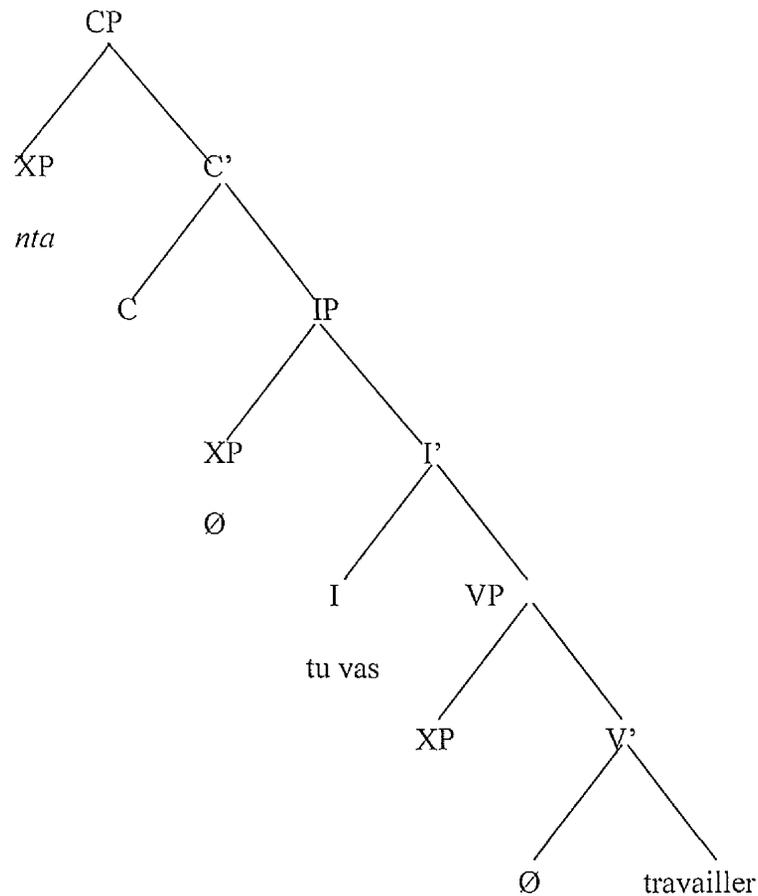
Therefore, as exemplified below, Arabic is the matrix language and contributes the first pronoun followed by a second subsequent pronoun from the EL, in this case French and therefore as Moroccan Arabic provides the topic pronoun pronoun doubling is obligatory:

(338) *nta tu vas travailler*  
You you go 2PL work INF  
'You, you are going to work'  
(Moroccan Arabic / *French*, Bentahila & Davies, 1983: 313)

(339) *nti* you're always saying that  
You FEM you're always saying that  
'You're always saying that'

Therefore, within the scope of the MLF, namely the ML - EL distinction, certain morphological realization patterns unfold whereby when pronoun doubling occurs there are two syntactic positions available. The first is where the pronoun is outside the argument structure of the IP in which the second subsequent pronoun and verb are both located and the second pronoun is thus contained within the IP. This is shown as detailed below in the following tree structure:

Figure 7.9: 'nta tu vas travailler'



In analyzing the above structure, with Moroccan Arabic as the ML framing the CP, the first pronoun *nta* fills the Spec of the overall CP. The second part *tu vas travailler* occurs within the IP frame and is an EL island. As a consequence, in order for well-formedness requirements of the EL in French (as well as other languages that are positioned within the IP and are EL islands), there is an absolute requirement for a pronoun to occur in the argument position.

In terms of the MLF and 4-Model, how are pronouns within the pronoun doubling paradigm assigned in the content morpheme – system morpheme hierarchy?

Myers-Scotton (1996) states that:

Even though they occur outside the argument positions projected by a particular predicate, topic pronouns can be viewed as receiving a thematic role; therefore they are considered to be content morphemes. They receive the thematic role of topic, a role projected by the discourse. Again, in GB terms, topic pronouns occur in Spec of CP. Thus, in Arabic and the other relevant languages [here], topic pronouns are content morphemes (1996: 31).<sup>244</sup>

In terms of morphemehood, this distinction is salient as the status of pronouns as content morphemes or system morphemes needs to be clearly defined within the MLF because while “Content morphemes from the EL which are congruent with an ML counterpart can appear in a mixed constituent framed by the ML, system morphemes cannot” (1996: 31).<sup>245</sup> In general terms pronoun doubling is possible where the ML frames the CP and EL islands are well-formed within the grammatical framework as set by the ML. Therefore, within the MLF framework we can make the following generalization:

*Generalization 11*

*Within the realms of pronoun doubling, if Moroccan Arabic supplies the ‘leader’ pronoun in Spec of CP, the ‘led’ pronoun must be in the opposing language and pronoun doubling must occur.*

Boumans (1998) in his syntactic analysis of certain discourse markers and in evaluating the phenomenon of Arabic pronoun doubling within the Myers-Scotton MLF frame states:

The CP analysis is a meaningful contribution to the development of a model because it associates morphemes that function in the domain of discourse marking with a matrix structure (CP) derived from their source language. This

---

<sup>244</sup> Jake (1994) in analysis of pronoun doubling states that “The syntactic category pronoun in any one language can consist of both content (i.e. thematic) and system (i.e. non-thematic) morphemes” (1994: 278). Therefore, not all languages pattern the same way regarding the content morpheme – system morpheme distinction.

<sup>245</sup> In Moroccan Arabic, pronominal inflections are system morphemes (non-thematic) and topic pronouns (*ana* ‘I’, *nta* ‘you’) are content morphemes (thematic).

matrix language determines the distribution of these discourse marking/organising morphemes on both the syntactic level and the level of discourse structure. Now this gives the CP analysis a certain appeal since the overall impression is that discourse markers tend to function in accordance with their source language rather than with the language of the finite verb (IP) (1998: 135-136).

Furthermore, and notwithstanding the fact that the MLF was specifically designed in order to suitably explain structural configurations in code-switched discourse and in light of the above data and explanations, it can be stated that not only does the MLF as a model of bilingual discourse suitably account for intra-sentential switching but that the 4-M Model, which is a refined and supportive version of the MLF, is able to account for a wider range of data and in particular new data that becomes evident and is produced by researchers. Pronoun doubling is such data. This phenomenon of pronoun doubling as outlined above shows how Moroccan Arabic, which is often the ML in code-switched data sets patterns in a manner whereby the pronouns supplied are usually Moroccan Arabic in the first position which act as the 'leader' pronoun in Spec of CP and the second subsequent pronoun in 'led' position will be in a language other than Moroccan Arabic. Given the suitability of the MLF to analyse Moroccan Arabic pronoun doubling, how does it fare with regard to conditional clauses in bilingual discourse? The final section of this chapter addresses the use of Moroccan Arabic and English conditionals in intra-sentential paradigms.

#### **7.4 Moroccan Arabic Embedded Conditional Clauses**

In Moroccan Arabic there are two basic condition types dependent upon whether the conditional 'if' clause is a possible or impossible condition and both condition types are introduced by set clauses.<sup>246</sup> To introduce the impossible condition, the particle

---

<sup>246</sup> Nortier (1990) distinguishes between the two types of conditions by referring to them as counterfactual and non-counterfactual where counterfactual conditions are introduced by *kun* and non-counterfactual conditions are introduced by *ila*.

*kun* ‘if’ is used.<sup>247</sup> An illustration in monolingual Moroccan Arabic shows how the verb of if-clauses introduced by *kun* is regularly in the perfect tense and if the result clause, or subsequent clause is also in the perfect tense, then it is also introduced by the *kun* particle (cf. Harrell 2004):<sup>248</sup>

(340) *kun sket, kun flet*  
 If quiet, if run away  
 ‘If he had kept quiet, he would have gotten away’  
 (Moroccan Arabic, Harrell 2004: 168)

(341) *kun ma šāf-ha -š, kun xrej ašql -u*  
 Had NEG1 see her NEG2, if go out 3SG mind POSS  
 ‘If he hadn’t seen her, he would have lost his mind’

(342) *kun kan ɛndi flūs, kun mšit haða šhēl*  
 If were have have 1SG money, if went DEM long time  
 ‘If I had the money, I would’ve gone ages ago’

The second type of condition is the non-counterfactual or possible condition and this is introduced by the particle *ila* ‘if’ and this is followed by the subsequent or result clause in the perfect or imperfect tense.<sup>249</sup> As described in the impossible or counter-factual condition above, the possible condition also does not distinguish between the gradients of conditional clauses (zero, first, second etc) as in ‘If she came’ or ‘If she comes’ as these are both rendered *ila zet* in Moroccan Arabic. Before turning our attention to the process within code-switched clauses, a few illustrations below highlight the basic formation of the possible condition:

(343) *ila šāt ši, ka-yeɣtiweh l -l masakīn*  
 If left over some give 3PL to DEL poor PL  
 ‘If anything is left over, they give it to the poor’  
 (Moroccan Arabic, Harrell, 2004: 170)

<sup>247</sup> Compare this with the English conditional clauses from zero to third, however, Moroccan Arabic does not distinguish between past unfulfilled ‘If I had, had I been and present conditions ‘If I were’.

<sup>248</sup> There is also the hortatory use of *kun* which is formed without a result clause (cf. Harrell 2004) for more examples.

<sup>249</sup> In certain Moroccan Arabic dialects, the conditional particle *ila* is reduced to *la*. This is not to be confused with the negative particle *la* and in any event is comprehensible through context.

- (344) *ila seqs -ek, gul -u*  
 If ask 3SG 2SG, tell him  
 ‘If he asks you, tell him’

Conditional sentences in Moroccan Arabic are easily formed with two clauses; the first sets the conditional clause and the second is the result clause in the perfect or imperfect tense. In the corpus, there are examples of the conditional in code-switched discourse and it appears that the favoured direction when employing the conditional is Moroccan Arabic conditions in English which were far more frequent than English conditions in Moroccan Arabic. Conditional clauses in Moroccan Arabic and English intra-sentential discourse, given the frequency of MA inserted conditions in English, occupy the first position in the English main clause. Consider the following examples:

- (345) *ila qrit mezyen, you will get it*  
 If study 2SG good, you will get it  
 ‘If you study well, you will get it’

- (346) *kun me mše -š, we could’ve carried on with the dinner, besseħ never mind*  
 If NEG1 go NEG2, we could’ve carried on with the dinner, but never mind  
 ‘If he didn’t go, we could’ve carried on with dinner, but never mind’

- (347) *u ila dewwez-ti l- uɣtla terzeħ hna, dan heb je het zefelde*  
 And if pass 2SG DEF holiday 2 return here then have you the same  
*problem*  
 problem  
 ‘And if you pass your holiday in [Morocco] and you come back here, you have the same problem’

(*Dutch / Moroccan Arabic, Boumans, 1998: 323*)

In the above example (347) Boumans (1998) also notes that in MA conditional code-switched clauses: “The MA conditional clause occupies the first constituent position in the Dutch main clause, thus triggering the inversion of the finite verb and the Subject in the main clause” (1998: 323). This is also the case for the present data, and

the same observation is made in Nortier (1990) (although she does not mention it but simply presents the data):

(348) *ila šuft ana f- rāsi, ka nqul aḡleš*  
If see 1SG I in head I say 1SG why  
'When I see myself, I say: why?'  
(Moroccan Arabic, Nortier, 1990: 40)

(349) *kun ken aḡnd -i patroondiploma, rah neqder nħell l maħall dyalli*  
If be 3SG with POSS, master certificate look can open DEF shop POSS  
'If I would have had a master-certificate, then I could have opened my own shop'  
(Dutch / Moroccan Arabic, Nortier, 1990: 40)

(350) *si j'avais la maison, maḡmmri na-kul temma*  
If have 1SG DEF house, never eat 1SG there  
'If I had the house, I would never eat there'  
(Moroccan Arabic / French, Bentahila & Davies, 1982: 309)

It is evident then that code-switching of subordinate clauses is frequent amongst language pairs and as in the conditional examples above, the second clause is an implication from the first one which introduced by 'if' states the premise (cf. Chan 2003). This is also in line with Bentahila and Davies (1983) and their findings where switching at clause boundaries was found to be common and productive in their Moroccan Arabic and French data. Furthermore, Tuc (2003) in his analysis of switching between main and dependent clauses observed that none of the switched sentences violate the syntactic structure of either language, and this may be due to the similar word order of Vietnamese and English as there is an equivalence of word classes in both languages (cf. Tuc 2003). In this regard, Poplack and Sankoff's (1981) Equivalence Constraint which as described in Chapter Two highlights the notion of categorical equivalence, is valid for the above examples in conditional clauses.

However, as previously discussed, the Equivalence Constraint does not hold true for all languages, all categories and in all cases. Countless counter-examples, already listed have proven this to be the case. However, in terms of the MLF and its

domain within classic code-switching, the formation of conditional clauses is a straightforward process when switching between subordinate clauses as long as the necessary well-formedness conditions are met and maintained throughout. As Tuc (2003) notes: “Myers-Scotton’s model is specifically useful” (2003: 134). This then is a universally observed assumption which in general terms is in line with the principles and parameters of UG, whereby any monolingual or bilingual clause must be grammatical and well-formed. Recall then that the MLF model implies that the matrix language – embedded language opposition and the content morpheme – system morpheme distinction are universal features underlying language production when classic code switching data as in the conditionals as described above are involved (cf. Myers-Scotton 2002: 85). None of the data presented with regard to MA conditionals provides any evidence to the contrary. Therefore, conditional clauses illustrate typical code-switched constituents which support the MLF model and in doing so also support the Morpheme Order Principle and the System Morpheme Principle as well as the 4-M Model which is an extension of the MLF as it can account for a wider range of data and shows how language production actually operates on a micro level. The final section of this chapter is devoted to Moroccan and Islamic terminology or ‘non-switchable’ nominal constituents.

#### **7.4.1 Non-Switchable Nominal Constituents**

Bilingual speakers also switch when there are items of no equivalence in the respective language of which they speak. More specifically, there are certain ‘non-switchable’ lexical items, i.e. cultural nominals which have no lexical equivalent in English. Both from the data and introspectively, it has been observed that these word classes are not normally switched unless they are being directly translated to a non-Arabic speaking person. In terms of syntactic structure, they are considered singly

inserted lexical items embedded in a matrix frame.<sup>250</sup> This has also been analyzed by

Boumans (198) who states:

Once could claim that many of the Dutch insertions in MA also refer to concepts related to Dutch society and the speaker's experiences in the Netherlands. So Dutch *dorp* 'village' refers to a particular village in the Netherlands and various terms that refer to the Dutch educational system feature in conversations in which the respondents discuss their studies, which also take place in the Netherlands. However, these terms generally have translation equivalents in MA and the respondents often know the MA words for these concepts, although it can be argued that the MA terms never have quite the same associations as the Dutch ones. Yet the MA nouns in Dutch are on average much more culturally specific and less translatable than the Dutch nouns (1998: 302).

It is the last sentence above which is of relevance with regard to the current analysis as certain Moroccan Arabic words simply do not have English equivalents, therefore giving rise to speakers using compromise strategies to insert single lexical items in an English matrix frame, adding the usual agreement features where necessary. Furthermore, as embedded nominal constituents, or copulas, they must adhere to the respective syntactic rules of the matrix language. Nortier (1990) refers to such nouns as belonging to specific 'semantic fields' (1990: 196) and found that single nouns are more frequently used because of lexical need. She also found that there are not only more single nouns without a Moroccan Arabic equivalent but that they are also divided into semantic fields. Some examples of Moroccan Arabic terminology are listed below:

---

<sup>250</sup> It must be noted that I refer to such singly inserted lexical items as examples of intra-sentential switches and not borrowings.

Figure 7.10: Moroccan Arabic cultural constituents

MA 'Maghrebi' nouns	Gloss
<i>salam aḡlaikom</i>	'hello'
<i>yarḥmak Allah</i>	'bless you'
<i>bi ṣaḥa</i>	'with health'
<i>t-hāra</i>	'male circumcision'
<i>baraka</i>	'a blessing from God'
<i>ḡurba</i>	'state of loneliness'
<i>marḍi</i>	'blessed'
<i>niyya</i>	'one's true intention'
<i>ṣābekiyya</i>	'Moroccan sweet'
<i>negeffa</i>	'woman who assists the bride'

With respect to these notions, there are also semantic or conceptual gaps in the English lexicon meaning that English cannot substitute the Moroccan Arabic and no equivalent exists that can provide the exact meaning. The following illustrates some of the most common Moroccan Arabic verbs which are maintained and not switched in bilingual discourse:

Figure 7.11: Moroccan Arabic cultural verbs

MA 'Maghrebi' verbs	Gloss
<i>slex</i>	'to slaughter sheep, skin it and then prepare for cooking'
<i>riyyeš</i>	'to pluck a bird when preparing to cook it'
<i>kuwwer</i>	'make into small balls'
<i>sejjed</i>	'prostrate in Islamic prayer'
<i>kehhel</i>	'to put kohl in one's eyes'
<i>jerri</i>	'to make dough runny'
<i>hejji</i>	'to tell a proverb at night'
<i>hetref</i>	'to have a bad premonitory dream'

As expected, we find references to Islamic practices and terminology as in *Ramadan*, *laylat ul qadr* 'night of power', *ḥāfiẓ Qur'an* 'one who has memorized the Qur'an' etc. Some examples are as follows:

(351) No, it's my Dad who usually *slex-s*            *the houli* and Mum cooks it  
 No, it's my Dad who usually slaughter 3SG the sheep and Mum cooks it  
 'No, it's my Dad who usually slaughters the sheep and Mum cooks it'

In (351) above, the third generation speaker has attached the English third person singular marker /s/ to a Moroccan Arabic verb which is another example of *Reactive Syntax* discussed earlier in the thesis. This is quite extraordinary as it highlights the way in which third generation Moroccan Arabic speakers integrate their discourse with fully-fledged English suffixes to Moroccan Arabic verbs. This would never be uttered by a first generation speaker and is illustrative of a new and innovative speech style. Other examples of cultural constituent insertions are as follows:

(352) He had his *thāra*            *u š-hel*            *kān* y-scream  
 He had his circumcision and how much was scream 3SG  
 'He was circumcised and he screamed a lot'

(353) *laylat l -qadr us op de zevenen twintgste dag dyal Ramadan*  
 Night DEF power is on the seven twenty day of Ramadan  
 ‘Laylat al Qadr is on the twenty seventh day of Ramadan’  
 (*Dutch / Moroccan Arabic, Boumans, 1998: 302*)

(354) *kānet ta-hejji -ni all night about this šerfa*  
 Was 3SG tell story me all night about this old lady  
 ‘She was telling me this story all night about this old lady’

In the above examples, *thāra*, *laylat ul-qadr* and *hejji* are all semantic concepts which do not exist in the English lexicon as ‘ready’ lexical items but in fact require lengthy translations. It is with this in mind that they are referred to as ‘non-switchable’ as bilingual speakers would not use an English code-switched item in their place. These then are embedded language islands if they are singly inserted within an English matrix frame as in (354) above where the MA verb even has an English inflection. This is explained, within the MLF model, as an embedded language verb which received matrix language inflection, in this case English with no apparent problems and these can include languages which are typologically different.<sup>251</sup> We re-visit the Embedded Language Islandhood discussed in Chapter Three.

#### 7.4.2 Embedded Elements of Cultural Constituents

Can such culturally embedded constituents be accounted for within the MLF? Recall that Myers-Scotton (2002) cites the following reason for the insertion of embedded elements as detailed below:

Singly occurring Embedded Language forms require less checking and undergo less processing at all levels of abstract grammatical structure than do Embedded Language phrases (phrases that will include grammatical as well as lexical elements). Singly occurring nouns simply must be sufficiently congruent with a

<sup>251</sup> Myers-Scotton states that previously embedded language verbs with matrix language inflections were rare, however “New corpora, especially those from speakers who are very aware of their language use (e.g. university graduate students) report that they increasingly are using English verbs inflected with [Spanish] in their informal speech” (2002: 138-139). These new corpora then are an example of *Reactive Syntax*, which I disagree, is not only exemplified by well-educated graduate students, but is an increasingly common speech style amongst third generation speakers.

Matrix Language morpheme counterpart at the level of lexical-conceptual structure to occur. (If there is not a counterpart, then under the Uniform Structure Principle--even this congruence is not necessary as long as the Embedded Language nouns follow the relevant language's procedures regarding potential Matrix Language predicate-argument structure and morphological realization patterns.) In the case of verbs, congruence may be more difficult to establish with a Matrix Language counterpart, since predicate-argument structure is at issue too (2002: 144).

Given the lack of equivalence in English, speakers need to maintain a semantic association with the lexical item and the cultural concept and therefore maintain the lemma. At the point of spell-out, dependent upon subject matter, speaker intention, fluency and proficiency, code-switchers are in essence at liberty to lexically insert the cultural noun or noun phrase as well as cultural verbs as long as grammaticality is maintained. Further, within the rubric of the MLF, such Embedded Language islands occur because of structural or lack of equivalence mismatches, and in some data sets they are as frequent as pragmatically motivated islands, if not more so (cf. Myers-Scotton 2002). Does this also pertain to English cultural constituents? Quite simply 'yes' as every culture has culturally-specific items which are specific to that particular culture and are untranslatable, even with the best of intentions. English examples which simply do not exist in Moroccan Arabic are 'sushi' and 'surf the web'.

Therefore as Myers-Scotton states:

Embedded Language islands are full constituents consisting only of Embedded Language morphemes occurring in a bilingual CP that is otherwise framed by the Matrix Language. An Embedded Language island shows structural dependency relations (2002: 141).

Contrary to Boumans (1998), the majority of the Moroccan Arabic constituents inserted in English matrix frames are not solely culturally related but it has been observed that English / Moroccan Arabic bilingual speakers are far more able and proficient in the art of merging Moroccan Arabic verbs with English inflections than speakers for any other data set that has been reviewed thus far. This is

informative as it implies that there is a speech style which has emerged in the last decade where speakers are more confident and proficient in merging the syntaxes of two typologically different languages. Furthermore, there appears to be a correlation between the use and incorporation of embedded language islands and proficiency as stated by Myers-Scotton (2002):

The relative presence of Embedded Language islands seems to tell us different contradictory conclusions about the proficiency of the speakers depending on the community (i) when the overall prevailing pattern includes many bilingual CPs (with many mixed constituents), singly occurring forms (typically nouns) prevail. If speakers employ relatively many Embedded Language islands, they seem to be among the more proficient speakers. That is, it seems that higher language proficiency in the Embedded Language is necessary to feel at home producing islands (2002: 148).

We can thus summarize that the non-switchable cultural constituents are maintained due to requirement and lack of lexical equivalence in the L1. This is socio-pragmatically driven and this motivation triggers the embedded language island within a matrix frame. Given new corpora and data amongst third generation speakers, it seems that although grammaticality must be maintained, not only are lexical items being uttered at spell-out, but the cultural concept, or largely formulaic expressions are being relayed in bilingual discourse.

## **7.5 Conclusion**

In tandem with Chapter Three, the present chapter evaluates certain grammatical categories and analyzes their suitability within the theoretical framework of the MLF model with regard to certain syntactic and morphological processes in Moroccan Arabic and English intra-sentential discourse. It has been shown as Ziamari (2007) states: "That the MLF's specificity lies in the value attributed to the ML which is seen as extremely dynamic. It changes both diachronically and synchronically" (2008: 279).

It has been evidenced that the frequently inserted prepositional phrases conform to the language providing the grammatical frame, in other words, the matrix language and are always uniformly grammatical in full prepositional phrases. The same is also true of adverbial phrases where a generalization was made whereby there cannot be switching of mixed constituents, namely MA or English adverbs within a maximal projection. This was found to conform to the maxims and basic premise of the matrix-embedded language hierarchy.

This was then followed by an examination of conjunctions in mixed discourse where in bilingual syntactic structures, there is no difficulty in the categorisation of conjunctions or any discourse marker within the MLF. This is true of Moroccan particles, which are frequently inserted in English matrix clauses and also English particles inserted in Moroccan Arabic matrix clauses. Such speakers often use *darija* particles to coordinate conjunctions linking English sentences. The subsequent section in the chapter is devoted to the phenomenon of pronoun doubling whereby in Arabic plus other languages, two pronouns with the same pronominal features occur adjacent to each other in what appears to be the same syntactic position. The final part of this section is devoted to cultural constituents which speakers insert, usually as singly occurring lexical items within a matrix frame. Syntactically, they are well-formed embedded units inserted at grammatically suitable slots in the overall frame due to semantic or conceptual gaps in the English lexicon which cannot substitute the concept or semantic field of the Moroccan Arabic item.

The overriding premise of the data presented in this chapter, and in code-switched intra-sentential discourse in general is that when two (or more) languages are juxtaposed, in all cases, one language will always assume the matrix position and provide the grammatical framework. This then is the theoretical orientation behind the

Matrix Language Principle and its supporting models, the 4-M Model and the Abstract Level Model where there is always both an analyzable and resolvable frame structuring the morpho-syntax of any CP and as a consequence, the two languages never participate equally as the source of this matrix language (cf.. Myers-Scotton 2002). This is further evidenced in Chapter Eight which provides an analysis of all eleven generalizations or hypotheses presented in this thesis and answers the research questions posed in Chapter Five.

## CHAPTER EIGHT

### EVALUATION OF THE MATRIX LANGUAGE MODEL AND GENERALIZATIONS

As discussed, code switching as a discourse strategy is by no means a novel or innovative concept and has attracted a lot of attention over the last few decades on both sociolinguistic and psycholinguistic levels (Belazi, Rubin, Toribion 1994, Bentahila & Davies 1983, Di Sciullo, Muysken & Singh 1986, Myers-Scotton 1993b, Poplack 1980, Sankoff & Poplack 1981, Pfaff 1979, Santorini & Mahootian 1995). There are those who claim that there are no universal, purely syntactically driven constraints on code switching thus setting it apart from monolingual usage (Bokamba 1989, Clyne 1987 and MacSwan 1999, 2005). As previously presented, the current study focuses on the morphological and syntactic analyses of such bilingual communication. The data strongly corroborates the claim that code switching as a discursive tool is rule-governed and adheres to certain principles and parameters not only in language production but also within the realms of UG.

However, as noted throughout this thesis, there is still much controversy and debate surrounding certain theoretical approaches to code switching and their suitability within a whole array of typologically similar and dissimilar data proving or disproving certain theories. Nevertheless, it is certain that the insertional approach has become the favoured approach due to its overriding compatibility with most data sets as opposed to linear or alternational methods and models (cf. Poplack 1980). Since the 1990s, Myers-Scotton with her influential and authoritative oeuvre on English and Swahili code switching set the stage for one of the most influential insertional models presented within the field of linguistic disciplines. Of all the models discussed involving the asymmetrical matrix language – embedded language hierarchy and the

content morpheme – system morpheme dynamic, it is the MLF which has been shown in this thesis to be the most comprehensive in dealing with two typologically dissimilar languages; here Moroccan Arabic and English. The main facts of syntax within code-switched data have been described using the principles of interpretation and general classification within the MLF model, 4-M Model and Abstract Level Model, and it is this basic asymmetrical model which has identified uniform regularities in the data presented and in bilingual discourse in general.

This chapter then within this uniform descriptive framework commences with the over-arching generalization and summary of my approach to code switching and bilingual discourse in general, particularly in intra-sentential clauses. It then re-addresses the eleven generalizations / hypotheses as postulated in this thesis and further analyzes their suitability in light of the data presented and in comparison with historical data analyses. Throughout this thesis an explanation of the code switching facts has been pursued through the MLF language production model in terms of conflicts in the lexical items of words which are in essence code switching-specific mechanisms. We commence now with a review of the most salient factors in code switching theory and summarize their validation, falsification and any supporting evidence. Problematic data for the MLF and code switching in general is addressed and discussed in full. Furthermore, an in-depth analysis of problematic data is made and the matrix language hypothesis is presented and evaluated in light of this data. We commence, then with the highly salient generalization which forms the main thrust of the thesis and lend value to its conclusion.

### **8.1 Evaluation of the Main Generalization**

The main evaluation throughout this thesis is the data collated which have led to the eleven generalizations. This evaluation is central to the thesis as the basic construct of

the asymmetrical model with the flexibility and feasibility of its application to all data sets is that both theoretically, socio-linguistically and pragmatically, the grammatical parameters as set by the matrix language provide the frame for the morpho-syntax of mixed clauses. A recapitulation of the first fundamental generalization and my overriding conclusion is as follows below:

*Generalization 1:*

*The Leader (of two or more languages) is that which contributes word order in the CP frame where the Led (embedded) variety adheres to the morpho-syntactic frame provided by the Leader.*

The above generalization, which is my contribution to intra-sentential code switching analysis, forms the premise of an asymmetrical model. The data presented in this thesis corroborates this hypothesis within an insertional framework where the matrix language sets the grammatical frame and the embedded variety adheres to the principles and parameters within the morpho-syntax. There are no other revisions, amendments or further sub-models postulated in order to account for any potential problematic data as the essence of this approach is the fact that the fewer constraints posited, the more data can be incorporated. In order to recapitulate, the following examples clearly set out Generalization 1 in the the bilingual intra-sentential as shown below:

(355) *ila ma bga -š,* then *šufi* somewhere else  
 COND NEG1 want NEG2, then see somewhere else  
 ‘If he doesn’t want to then see somewhere else’

(356) Here we go again with that *marð*  
 Here we go again with DEM illness  
 ‘Here we go again with that rubbish’

(357) *book-it maξ -h u fi -l* end *ma mšinā -š*  
 Book 1 SG with him and in DEF end NEG1 went 3PL NEG2

'I booked with him and in the end we didn't go'

For the most part, the data can more than adequately be described in terms of insertion patterns. Most examples of insertions are those of content nouns in Moroccan Arabic or English frames. In terms of non-nominal clauses, integrated English verbs which are '*Moroccanised*' or bear Moroccan affixation are quite frequent in the corpus and were found in all generation groups. Adjectives, prepositions and general nominal constituents occur freely in the data and are a symbol of natural bilingual proficiency and in most cases fluency.<sup>252</sup> It is then in terms of the generalization as made above and certainly in terms of the theoretical concept of the basic MLF that the data as presented in this thesis are adequately described and presented, with the asymmetrical model serving as a clear and precise vehicle. Therefore, it can be said that a general pattern has indeed emerged where all data analysed in this thesis veers towards an asymmetrical model as it has been clearly shown that other more specific and stringent constraints are over-restrictive for most data sets where the languages involved are typologically dissimilar. In essence, none of the constraints, bar the model of asymmetry, can satisfactorily account for all data sets in their entirety.

### **8.1.1 Asymmetry and Basic Premise of Generalization 1**

Of course not all code switching data slot into place within one model in its entirety as certain lexical categories, insertions and verb forms etc fall outside the scope of any one model. This is further evidence for Generalization 1 above. There will always be counter-examples to any one model which attempts to provide a sole constraint theory

---

<sup>252</sup> As noted previously, bilingual fluency is not a pre-requisite for any code switchers as after analysis of the data and also introspectively, it has been evidenced that some speakers who are receptively bilingual as opposed to productively bilingual, do not make any errors of judgement in their grammatical choices, nor do they go against any UG principles.

for all data in all languages. This has proven to be the Achilles heel of many approaches in the last few decades and is a further endorsement of both Generalization 1 and also the Myers Scotton's MLF model as, in its basic asymmetrical form, it provides a coherent and theoretical framework for the analysis of bilingual discourse and in essence has applicability to a wider range of data than any other model postulated thus far. It must be noted however that establishing a generic theory to account for all data is perhaps impossible as discussed by Alvarez-Caccamo (1998), who states:

In order to argue convincingly for or against the existence of 'code-switching constraints' and 'code-switching grammars' based on the two monolingual ones (...) research should first convincingly prove that (a) speakers who code-switch possess two (or more) identifiable linguistic systems or languages, each with its identifiable grammatical rules and lexicon; and (b) 'code-switched' speech results from the predictable interaction between lexical elements and grammatical rules from these languages. None of these assumptions, I believe it proven yet (1998: 36).

Furthermore, Boumans (1998) adds that code switching should be approached from different perspectives in order to fully understand both the regularities and irregularities in the data:

Explanations at this level can be approached from two different angles. On one side, CS patterns can be related to sociolinguistic variables in order to account for patterns which are characteristic for entire bilingual speech communities or even types of contact situations. From another angle, the matter can be viewed from the perspective of the individual speaker and explanations can be sought in theories on speech production and on the organisation of the (bilingual) mental lexicon. These two angles may be termed the sociolinguistic and the psycholinguistic explanatory framework. Explanations that relate regularities in CS patterns to language-specific factors would fall under the psycholinguistic type (1998: 365).

Given the aptly phrased quote by Romaine (1995): "Where more than one language exists in a community, they are rarely equal in status," (1995: xiv) it lends currency to the fact there is always one language which provides the overriding

grammatical frame enabling the pre-dominant language to embed itself within that structure and this has been the conceptual tone throughout this thesis. Nishimura (1986) notes that researchers who work on typologically similar language pairs such as Spanish and English tend to assume symmetrical and/or linear models of switching which depart from alternation and only ever fit the data in which they work (Poplack, 1980, 1981). In essence their models are not universally valid, whereas those working on typologically different language such as Turkish and Dutch develop asymmetrical models of insertion which emphasize the contrast in terms of the degree of participation of the languages involved (e.g. Backus, 1996; Myers-Scotton, 1992, 1993a, 2002).<sup>253</sup> This is a valid assumption and given the data analysis as presented most notably in Chapter Two and subsequent Chapters thereafter is certainly an apt conclusion. It appears that certain constraints as presented historically have been born out of the data sets as detailed by the researcher in question and on the particular languages analysed and not from a detached and abstract perspective. Generalization 1 then is an abstract and workable construct, and due to its abstract nature, it is universally valid as there will always be one of the two participating languages which provides the morpho-syntactic frame. This is also the attraction and validation of Myers-Scotton's approach as this is a model of language production and there is always a discernible morpho-syntactic frame in bilingual discourse. Hasselmo (1970, 1972) in his analysis of English and Swedish code switching also referred to the fact that languages do not participate equally and therefore there was a need to find a suitable asymmetrical model. Boumans (1998) writes also that "Asymmetry is a

---

<sup>253</sup> This is an apt observation as previous studies have shown that in code-switched data examined by certain linguists., those that worked on typologically similar languages, presented linear models, and those that worked on typologically dissimilar languages presented insertional-type models. See Chapters two and three for a full analysis of the most salient contributions in code switching analysis.

striking feature in many CS varieties” (1998: 366) and this is a fact supported by most data in bilingual discourse.<sup>254</sup>

This is corroborated by the data presented in this thesis and the summary of hypotheses and/or generalizations to be listed and analysed below. As discussed previously, more recently different approaches (MacSwan, 1999) have addressed code switching in terms of a structural application of Chomsky’s (1995) minimalist framework. Given the corpus as presented in Chapters Three to Seven, in this chapter, a brief recapitulation will be presented of the corpus description as a whole and its placement with regard to the hypotheses and generalizations postulated. It is within this paradigm that the MLF is evaluated and its performance and results further demonstrated and validated. Syntactic analysis is integral in attempting to both determine and decipher the manner in which language alternation operates.<sup>255</sup>

However, Gardner-Chloros and Edwards (2004) state that:

Although syntax plays an important role in CS, it cannot be summed up a priori that the constructs of syntacticians are the best means for characterising the processes of performance data such as CS. The possibility of throwing light on this question depends partly on whether or not it is right to assume that all bilinguals alternate in some meaningful way between two clearly distinguishable sets of rules – and this is a question which manifestly cannot be decided by grammar analysis alone (2004: 126).

In their sociolinguistic analysis of French outside France,<sup>256</sup> Gaudet & Jones (2008)

conclude that:

---

<sup>254</sup> Boumans (1998) in his evaluation of asymmetry in code switching, adds: “How do we account for the asymmetry, on the one hand, and the fact that Arabic has the same role in both CS varieties? Unless we want to believe that there is something specific about Arabic that makes it function as the matrix language, the explanation must be sought in the sociolinguistic situation that accounts for the asymmetry in CS. The comparison of Moroccan Arabic/Dutch and Algerian Arabic/French to code switching in closely parallel sociolinguistic circumstances with other language pairs illustrates this” (1998: 367).

<sup>255</sup> Syntactic analysis together with sociolinguistic variables should both be borne in mind when positing constraints on code switching, as Gardner-Chloros & Edwards (2004) suggest: “The role of sociolinguistic factors is neglected, although studies have shown that CS between the same two languages in different contexts can produce significantly different grammatical results” (2004: 104).

<sup>256</sup> As previously detailed, due to space and time constraints, this thesis does not centre on intricate sociolinguistic approaches to code switching and instead focuses on an in-depth structural analysis.

Sociolinguistic arguments appears to be more clear-cut than structural ones, as agency (namely the role of the speaker as agent), is a leading force in the process of [borrowing] and change. Contact usually affects everyday spoken varieties, which are mostly used and transmitted via social networks, and their existence outside formal teaching and writing means that they are able to escape normative pressures. These varieties display considerable amounts of variation and it is always necessary to establish the identity of each speaker in a given corpus in terms of age, education mode of acquisition, frequency of use and so forth and also that of his/her network(s)...Sociolinguistic considerations involve representations and ideology as much as they do actual socio-demographic (quantifiable) facts. Speakers have feelings about their language(s), for example loyalty towards their mother tongue, or the desire to resist the use of a dominant language, and these feelings are central to their language attitudes (2008: 245).<sup>257</sup>

While it is true that a pluralist approach and other important factors should be borne in mind when analysing code-switched data, it is correct to predict that in any code-switched string, there will always be one of two (or more) language(s) which provides the grammatical frame where speakers as previously evidenced, adhere to the principles of universal grammar where phrase structure rules and grammar are adhered to as well as criteria of uniformity and well-formedness conditions which are *always* maintained. This is the basic premise of my first hypothesis out of a total of eleven generalizations posited throughout this thesis and my over-riding statement in analysing classic code-switched data.

### 8.1.2 Generalizations Evaluated

Before examining and summarising my own results from the corpus for this thesis, it is important to note the historical development and progress in code switching analysis in the last few decades and also to compare my research results with those of other researchers. The following Figure 8.1 then lists the main constraints on code switching over the last thirty years and certainly before the MLF came to the fore.<sup>258</sup>

---

<sup>257</sup> cf. Gaudet, F and Jones M (2008) 'Variation, contact and convergence in French outside France.' *Journal of Language contact* – *THEMA* 2.

<sup>258</sup> cf. Nortier (1990) for a further analysis.

Figure 8.1: Contributions to Constraints on Code Switching

	Researcher	Summary of Constraint	Language Pair	Validated by own data ✓ / x
1	Gumperz	A conjunction must be in the L of the 2 <sup>nd</sup> of the conjoined Sentences	Spanish – Eng	x
2	Kachru	A conjunction must be in the same L as both conjoined sentences	Hindi - Eng	x
3	Abbassi	A relative pronoun must be in the same L as the rest of the clause	Arabic - French	x
4	Abbassi	In an interrogative switching after the wh- element is forbidden	Arabic – French	x
5	Timm; Gumperz, Pfaff, McClure and Wentz	Clitic pronoun objects are in the same L as the V to which they are cliticized	Spanish – Eng	✓
6	Timm; Lipski	Switching between subject or object pronoun and V is impossible	Spanish - Eng	x
7	Pfaff	Switching at PP boundary is rare; it never involves locative PPs	Spanish – Eng	x
8	Timm; Abbassi; Lipski	Switching cannot occur between fin V and infl compl or Aux and main V	Spanish – Eng and Arabic - French	x
9	Abbassi	Switching between Fr Prep and Ar NP is not possible	Arabic – French	x
10	Pfaff; McClure and Wentz	Switching between Det and N is seldom found, or even ungrammatical Switching between Fr Det and Ar NP is not possible	Arabic – Arabic	x
11	Abbassi	Switching between Fr Det and Ar NP, the reverse is OK	Arabic - French	✓

As previously discussed in Chapter Two onwards, most of the claims as outlined above can no longer be regarded as valid and this is largely due to progress made in research methods and approaches. An obviously counter-invalid claim is that of Kachru (1977) who claims in (2) above that a conjunction must be in the same language as both conjoined sentences:

	Reported in:	Descriptive boundaries (+ = code switch)	In disagreement with:
1a	Gumperz (1976)	Because + CP	Poplack (1981) Sankoff and Poplack (1981) Mahootian (1993)
1b	Gumperz (1976)	Conj + CP	Poplack (1977) McClure (1981)
2	Belazi, Rubin & Toribio (1994)	That + IP	Bentahila & Davies (1983) Mahootian (1993)
3a	Belazi, Rubin & Toribio (1994)	Have + VP	Di Sciullo, Muysken & Singh (1986)
3b	Belazi, Rubin & Toribio (1994)	Modal + VP	Di Sciullo, Muysken & Singh (1986)
3c	Timm (1975)	To + VP	Lipski (1978) Poplack (1981) McClure (1981)
3d	Timm (1975)	Aux + V	Lipski (1978) Poplack (1981) McClure (1981) Mahootian (1993)
3e	Timm (1975)	Neg + V	undisputed
4a	Belazi, Rubin & Toribio (1994)	Q + NP	Bentahila & Davies (1992) Mahootian (1993)
4b	Belazi, Rubin & Toribio (1994)	Demonstrative + NP	Nishimura (1985) Bentahila & Davies (1992) Mahootian (1993)
4c	Belazi, Rubin & Toribio (1994)	Article + NP	Brown (1986) Bentahila & Davies (1992) Mahootian (1993)
4d	Wentz (1977)	Complex D + NP	Poplack (1981)
5a	Gumperz (1976) Lipski (1978) Belazi, Rubin & Toribio (1994)	N + Adj (Adj from Adj-N language, N from Adj-N language)	Bokamba (1989) Mahootian & Santorini (1996)
5b	Belazi, Rubin & Toribio (1994)	Adj + N (Adj from N-Adj language, N from Adj-N language)	Poplack (1981)
6a	Timm (1977) Gumperz (1976) Lipski (1978)	Subject pronoun + V	Poplack (1981) Mahootian (1993) Bentahila & Davies (1993)
6b	Timm (1975) Gumperz (1976) Lipski (1978)	V + object pronoun	Poplack (1981) Mahootian (1993)
6c	Timm (1975)	Clitic + clitic	undisputed
6d	Gumperz (1976)	Gapping constructions	Poplack (1981)

		with Aux second V switched (marginal)	
7	Poplack (1981) Sankoff & Poplack (1981)	A switch involving a bound morpheme	Nishimura (1985) Mahootian (1993) Myers-Scotton (1993b)

(358) He said he was going *walakin* he didn't in the end  
 He said he was going but he didn't in the end  
 'He said he was going but he didn't in the end'

(359) *ila me ken -š Əamma* and *me walle -š, xəlli -h*  
 COND NEG1 was NEG2 there and NEG1 return NEG2, leave him  
 'If he isn't there and doesn't' return, leave him'

(360) No, in general we go to Morocco twice a year, *hit*, we love going home  
 No, in general we go to Morocco twice a year, because, we love going home  
 'No in general we go to Morocco twice a year, because, we love going home'

The examples above give obvious counter-examples to Kachru (1978) and also Gumperz (1976) who also states that a conjunction must be in the language of the second of the conjoined sentences. Clearly, this is not the case.<sup>259</sup> These results are also compared with more recent analysis on code switching using at times, typologically different language pairs. Figure 8.2 below gives an outline of the major contributions.

*Figure 8.2: Contributions to Constraints on Code Switching*<sup>260</sup>

Figure 8.2 above as listed in MacSwan (1999: 66) shows that developments have been made over the last few decades but as previously discussed, the majority of these assumptions and hypotheses are no longer valid. However, two which are still valid and have already been addressed in this thesis are that of the Clitic + construct and the Neg +V, both of which are maintained to this day. It is interesting to note that of all the disputed constraints posited thus far, it is these two which have stood the test of

<sup>259</sup> It is not intended to list all the counter-examples to Figure 8.2 as these claims should now be considered obsolete as research has since then progressed as my hypotheses and data have shown. We are more concerned with contemporary models and examples.

<sup>260</sup> See MacSwan (1999) for a further analysis.

time and remain valid given empirical evidence and data. Further examples of negation are as follows:

(361) *me gelli -š fi -l awwal hetta* I asked him directly  
NEG1 tell NEG2 in DEF beginning until I asked him directly  
'He didn't tell me in the beginning until I asked him directly'

(362) *me like -ha -š bezzeŋ*  
NEG1 like 3SG it FEM NEG2 a lot  
'He didn't like it a lot'

The *ma-...š* structure is the most common configuration in Moroccan Arabic negation of verbal forms. The above examples, as detailed in Chapter Four, show that splitting the two negative particles within an intra-sentential CP and suffixing the *-š* to an English verbal form is not possible. Furthermore, the interesting Moroccan Arabic and English data analysed highlight how there is an aspect of manoeuvrability in terms of adding the suffixed *-š* negative particle to an English verbal stem. It is evident then that MA / Eng code switching allows for the evolution of speech styles where the *ma-* is the head of the NEGP and *-š* as its Specifier attaches itself to an English verb. Other generalizations and conclusions are summarised below in the following figure which lists all generalizations / hypotheses resulting from the data corpus as detailed throughout this thesis and further analysis and examination thereafter. This is an important list, and contribution to knowledge as it not only provides new insights into the examination of code switching in general terms but is an original contribution to linguistics and to the study of bilingual discourse:

Figure 8.3: Generalization / Hypothesis Overview

Generalization as listed in thesis	Name	Generalization	Validated by data ✓ / x
1	<i>Leader vs Led</i>	<i>The Leader (of two or more languages) is that which contributes word order in the CP frame where the Led (embedded) variety adheres to the morpho-syntactic frame provided by the Leader.</i>	✓
2	<i>Possessive marker 'dya'</i>	<i>"Dya" as a possessive marker can only take a direct Moroccan Arabic suffix in Moroccan Arabic.</i>	✓
3	<i>Suffixation and 'End'</i>	<i>"End" as a preposition can only take a direct Moroccan Arabic suffix in Moroccan Arabic</i>	✓
4	<i>Generation factors</i>	<i>Amongst singly occurring Moroccan Arabic lexical insertions where the Leader is English and provides the morpho-syntactic frame; these will mainly be from second and third generation Moroccan Arabic speakers</i>	✓
5	<i>MA negation</i>	<i>In MA negation, the MA suffixed negative particle -s does not also have to be in MA.</i>	✓
6	<i>Indefinite Zero</i>	<i>Indefinite zero articles will be</i>	

		<i>mainly used by first generation speakers or those with a less proficient command of an L2 variety.</i>	✓
7	<i>Grammatical Gender - phonology</i>	<i>Phonology takes precedence over semantic association in determining noun gender in English only if both the English noun and the MA noun map phonologically.</i>	✓
8	<i>Grammatical Gender - semantic</i>	<i>MA Semantic considerations may override the morpho-phonological make-up of an English noun provided there is a direct semantic concord in between the MA and English words.</i>	✓
9	<i>Discourse marker 'beš'</i>	<i>Switching between certain MA categories namely beš + an infinitive is not possible. These must be followed by a subjunctive morpheme in order to neutralize the violation and render grammaticality</i>	✓
10	<i>Adverbial Phrases</i>	<i>There cannot be switching of mixed constituents, namely MA or English adverbs within a maximal projection.</i>	✓
11	<i>Pronoun Doubling</i>	<i>Within the realms of pronoun doubling, if Moroccan Arabic supplies the 'leader' pronoun in Spec of</i>	✓

		<i>CP, the 'led'  pronoun must be in  the opposing  language and  pronoun doubling  must occur.</i>	
--	--	---	--

The above figure gives an overview of the generalizations made in bilingual intra-sentential Moroccan Arabic - English code switching and is a summary of the findings from Chapter Two to Chapter Seven. After extensive examination of the literature, it has become evident that none of the specific code switching constraints could encompass the data findings and summary as listed above in Figure 8.3 which gives the hypotheses of the data examined. Generalization 1 as outlined above provides the motivation to explain not only Moroccan Arabic and English data as a combined data set, but would naturally extend to other corpora also. Generalization 1 however is universally valid in its abstract nature. MacSwan (1999) in his analysis of Spanish and Nahuatl also concludes that:

Since it has been shown that code switching-specific constraints cannot account for the data under analysis, and since the data under analysis may be explained without reference to such constraints, they may be assumed not to exist by general principles of scientific parsimony (1999: 312).

Furthermore, Myers-Scotton (2008) adds in her latest paper on code switching:

What does the MLF model accomplish? In a word – and the word is “asymmetry” – the MLF model offers a synthetic characterization of the constraints that apply to bilingual distributions in CS (2008: 26).

The above statement is a fundamental categorisation of a generic approach to code switching. Therefore, the main theoretical premises of both Generalization 1 and the basic construct of the MLF model in its basic asymmetrical form are maintained, in

that they both set out a general construct for classic code switching, an evaluation of which follows in the next section.

## 8.2 Classic versus Composite Code Switching

An important aspect of code switching in terms of the MLF model which needs to be highlighted is that of 'Classic' versus 'Composite' code switching mechanisms. The distinction between 'Classic' and 'Composite' code switching involves an asymmetrical relationship between two languages used in the same clause by bilingual speakers. Classic code switching as described by Myers-Scotton is initially defined as follows:

The term 'classic code switching' refers to speech for which the speakers are proficient enough in the participating languages that they can produce well-formed monolingual utterances in the variety which becomes the source of what is called the Matrix Language, the abstract morpho-syntactic frame of bilingual utterances (2002: 8)

Myers-Scotton explains further the notion of classic code switching:

It is characterized as bilingual speech within a bilingual CP, with the morpho-syntactic frame derived from only one of the participating languages. Classic code switching occurs when speakers have full access to the morpho-syntactic frame of one of the participating languages (the source of the Matrix Language). Speakers also have enough proficiency in the other language (the Embedded Language) either to (i) insert Embedded Language content morphemes into mixed constituents framed by the Matrix Language or (ii) produce well-formed Embedded Language islands, or both (i) and (ii) (i.e. produce both mixed constituents and Matrix Language islands) (2002: 105)

'Composite' code switching however is a new concept developed by Myers-Scotton, again born out of further research into different data sets. This is where the constraints on insertions posited by the MLF are no longer strictly adhered to and elements then 'split and recombine' and results in system or functional morphemes from both

contributing languages made available. Myers-Scotton (2002) details Composite code switching as follows:

The term composite code switching is new. It can be characterised as a phenomenon with morphemes from two languages within a bilingual CP, and with the abstract morpho-syntactic frame derived from more than one source language. Composite code switching occurs when speakers – because of psycholinguistic or socio-political factors – do not have full access to the morpho-syntactic frame of the participating language that is the desired source of the Matrix Language. Or, possibly the notion of a target Matrix Language is not clear to the speakers themselves. The result is that a composite Matrix Language frames the resulting bilingual CP. Thus, in effect, composite code switching necessarily entails convergence (2002: 105).

Therefore, language proficiency and competence are factors in distinguishing between classic and composite code switching, where the former is used by speakers who have full access to the morpho-syntactic frame, namely, they are fluent in at least the Matrix Language, and the latter is a by-product of speakers who do not have full access to the morpho-syntactic frame of the Matrix Language. It is also noted that given the prominence of bilingualism in the world today, composite code switching may be more prevalent than classic code switching. However, a full explanation of composite code switching has yet to be fully developed by Myers-Scotton and as it is not as yet fully defined. This then leads us on to examine the Matrix Language Turnover which occurs when late system morphemes, as outlined in the 4-Model, come into a language when its own morpho-syntactic frame undergoes a 'reconfiguration.'

### **8.3 Matrix Language Turnover**

To describe changing of the Matrix Language, Myers-Scotton has developed the Matrix Language Turnover Hypothesis as a by-product of the MLF. This is in essence,

a change of dominance of the Matrix Language and subsequently, this is termed as the ML turnover.<sup>261</sup>

This is described by Myers-Scotton as follows:

An ML Turnover means that the main language which had structured constituents becomes the minor or Embedded Language (EL); in turn, the language which has been the minor language regarding structure becomes the ML. The result is that grammatical structuring the CPs showing CS is now the task of the new ML (1998: 299).

Therefore, ML language turnover is evident in certain sociolinguistic contexts and occurs where speakers change the matrix language within a single conversation due to certain factors. Ziamari (2007: 115) gives the following example of ML Turnover:

(363) f s-*sociologie* ka-yqerrew-na  
In sociology teach 3PL us  
'In sociology they teach us [that]'

(364) *les relations* humaines sont mebniiyyīn sur les intérêts égocentriques des  
DEF relations human are built on DEF interests egocentric of  
*individus*  
individuals  
'Human relations are built on the selfish interests of individuals'

(365) daba nti ma tzewwiji b rajel  
Now you FEM NEG marry with man  
'Now you get married to a man'

(366) *c'est sûr il va y avoir des problèmes* mġa hmāt -k  
It is certain it going and have DEF problems with mother-in-law POSS  
'It's certain that it's going to cause problems with your mother-in-law'  
(*French / Moroccan Arabic*, Ziamari, 2007: 115)

In the above example, the matrix language changes from French to Moroccan Arabic and back to French throughout the conversation between young bilingual Moroccans.

Generally, in code-switched sentences where Moroccan Arabic is one of the

---

<sup>261</sup>This was firstly established in Myers-Scotton's paper '*A way to dusty death in the Matrix Language Turnover Hypothesis*' (Myers-Scotton, 1998b).

participating languages, Moroccan Arabic is usually the matrix language, depending on fluency and competency, but in the above example and in light of composite code switching, the dynamic is certainly changing amongst different generations of speakers, where fast code switching is becoming the norm. In such a situation where certain surface-level morphemes from two languages are used, a certain type of code switching is formulated, namely that of composite code switching. In this regard, the CP is bilingual because there are surface level morphemes from both participating languages and also grammatical structure from both languages. In essence, this shows that there is a composite matrix structure as the grammatical frame of the CP. In effect, in this hypothesis, there are two models which compete for the position of matrix language and, a single unified matrix language does not always prevail.

Both classic and composite code switching adhere to the important Uniform Structure Principle re-outlined below (2002):

*Uniform Structure Principle*

*A given constituent type in any language has a uniform abstract structure and the requirements of well-formedness for this constituent type must be observed whenever the constituent type occurs. (Myers-Scotton 2002: 121).*

As Myers-Scotton (2002) states of the above Principle:

This Principle formalizes the observation that languages preserve grammatical structure. In reference to bilingual speech, the preference is still to preserve uniformity. This leads to a preference for the structures of a Matrix Language as opposed to an Embedded Language. This preference is especially obvious in classic code switching, but is also clear even in split languages where the structures of one contributing language dominate. This principle implies explanations for several outcomes relevant to language contact. Specifically, in reference to split languages, the principle helps explain a number of outcomes. First, the principle predicts resistance to breaking the uniformity of a morpho-syntactic frame from one language. That is, there is resistance to replacing late system morphemes in language X with those of another language when language X is in contact with that other language, even under socio-political conditions favouring the other language's cultural dominance.

That is the Uniform Structure Principle helps explain why there are few split languages (2002: 100).<sup>262</sup>

Therefore, from 2002 onwards Myers-Scotton refers to code switching either in its classic form or its composite counterpart. Is this then problematic for the base notion of the MLF which, in its original form, calls for the identification of a sole matrix language? Certainly, in bilingual discourse, the onus is on language stability and maintenance and with regard to the above, could the flipping from one Matrix Language to another within a single conversation be the undoing of the MLF? Not according to explanations provided by Myers-Scotton (2002), who suggests that such composite code switching is an example of language contact phenomena in general and is a result of language attrition and language shift. However, given the above, Thomason (2003) finds that the constraints of the MLF may now not always be maintained:

A combination of both the MLF model and the Abstract model may be useful in explaining structures in composite code switching. However, as the composite nature of the frame becomes greater, the constraints of the MLF do not always hold. For example, morpheme order may not come entirely from what would have been the Matrix Language under classic code switching. Still, one language always seems to dominate in structuring the morpho-syntactic frame involved and so the labels Matrix Language and Embedded Language remain useful (2003: 89)

Thomason goes on to give her own Complete Matrix Language Hypothesis, which she maintains is *only* a hypothesis:

The former Embedded Language gradually becomes the Matrix Language in a bilingual corpus. That is, the morpho-syntactic frame changes from that derived from one participating language to that derived from the other (2003: 89).

---

<sup>262</sup> cf. *The Mixed Language Debate: Theoretical and Empirical Advances*. Matras, Y. & Bakker, P. (2003).

The whole nature of Matrix Language Turnover as posited by Myers-Scotton (2002) is yet another revision borne of data sets she has researched.<sup>263</sup> Certain researchers in fact claim that too many revisions have been made to her original work and this has proven problematic.<sup>264</sup> The following section addresses problematic data for the MLF and its supporting models.

#### 8.4 Problematic Data for the MLF Model

It is clear and well-understood that Myers-Scotton's work since her original presentation of the MLF (1993a) has undergone some, at times, extensive changes, modifications and additions. Given the increase in data sets including, counter-examples to some of her analysis, therefore it is of no surprise that certain aspects of the original model would be revised. The MLF, as Myers-Scotton details (2008),<sup>265</sup> now does not apply to all data and also now carries a disclaimer which was not applicable in previous revised attempts:

---

<sup>263</sup> This thesis, however, concerns itself only with code switching in the classic sense and does not differentiate between the revised category of switching but it must be noted that new and innovative forms of switching in this thesis fall under the rubric of *Reactive Syntax*. As outlined throughout this thesis occurs where speakers tag Moroccan Arabic affixes to English stems, use a 'youth' speech style and employ linguistic characteristics typical of second and third generation styles as previously discussed.

<sup>264</sup> Myers-Scotton (2002) in explicating the distinction between classic and composite code switching adds basic assumptions about code switching and convergence: "Assumption 1 – Either code switching or convergence, or both, are the contact phenomena that are entailed by other forms of contact phenomena. Assumption 2 - here is not a categorical link between the presence of classic code switching in a community and language shift. That is, structural contact developments resulting in language shift are not necessarily outcomes of classic code switching. What is more certain is that classic code switching frequently leads to increased lexical borrowing and often leads to convergence. Assumption 3 – Neither is shift a necessary outcome of convergence. However, convergence is more likely than classic code switching to lead to further structural contact developments resulting in language shift. Assumption 4 – Composite code switching need not lead to shift. However, it more frequently leads to shift than either classic code switching or convergence. The rationale is that composite code switching includes two ways in which the putative Matrix Language is compromised. First, if code switching includes many Embedded Language islands, this means more switches in the activation level of the participating languages. Second, convergence means incursions from a putative Embedded Language into any or all of the three levels of abstract grammatical structure of the morpho-syntactic frame" (2002: 298). Therefore the rationale leads to the conclusion that if there is indeed language shift, a certain hierarchy entails:

Classic code switching < Convergence < Composite code switching.

<sup>265</sup> cf. 'Language contact: why outsider system morphemes resist transfer.' *Journal of Language Contact – THEMA 2* (2008: 26).

The MLF model offers a synthetic characterization of the constraints that apply to bilingual distributions in CS. A few disclaimers clarify the limits on the model. First, the model applies only to classic CS; this is CS in which the source of the morpho-syntactic frame of the bilingual clause clearly and consistently is only one of the participating languages. Thus, the model was not designed to explicate all language contact phenomena. Still, portions of the model do apply to what I call composite CS, as well as to other contact phenomena (2008: 26).

It must be re-stated that this thesis adopts the MLF in its basic form, namely the asymmetrical dynamic which exists between two (or more) participating languages in a code-switched intra-sentential clause. Code-switching here is analysed in its 'classic' sense, with any revisionist, innovative and new forms of code switching referred to as my innovative concept of *Reactive Syntax* which I developed after examination and research of the data. In essence, the MLF in its basic form provides a universal template for the world's languages and corpora in general. Although critics have cited too many revisions and amendments as further complicating the model(s) and at times making it difficult to follow, overall this detracts from the research Myers-Scotton has achieved since 1993.<sup>266</sup> Boumans (1998) states that: "*The model seems to be supported by various data sets.*" (1998: 37). In describing the MLF, Boumans (cf. 1998) recognises the complexities going on to state that:

Differentiating system and content morphemes in a principled way that is valid for all languages is an extremely complicated task (1998: 39).

Further, Boumans criticizes the 'apparent' flaws in the (1993) MLF model with regard to the status of the matrix language and embedded language constituents:

Another flaw of the MLF model concerns the definition of ML and EL constituents, also called islands. The model does not provide for the possibility of single morpheme constituents since "all islands must be composed of at

---

<sup>266</sup> It must be noted that many of the revisions of Myers-Scotton's models have been borne out of necessity due to new data sets and criticisms thus making necessary the need to add revised models to an already well-established and functional asymmetrical language production model(s).

least two lexemes/morphemes in a hierarchical relationship” (1993b: 78). (1998: 41).

Boumans adds in a footnote:

If there are no single morpheme EL or ML constituents in the MLF model, single morpheme ML + EL constituents seem even more unlikely. ML + EL constituents “typically consist of morphemes from both the ML and EL” (1993b: 77)” (1998: 41).

As discussed previously, I sympathise with Boumans in that the model from (2002) onwards has become rather complex. However, the analysis by Boumans of single morphemes and the number of morphemes is contentious as Myers-Scotton in later revised analysis of her own work, dispensed with the number-of-morphemes approach. This was also the case with regard to identifying the matrix language which initially Myers-Scotton based on the number of morphemes or the ‘language of the majority of morphemes.’ Later however, she retracted this view and in (2002) said: “That claim [the number of morphemes approach] was abandoned” (2002: 61) and based her revised analysis firmly on an abstract, theoretical concept:

The Matrix Language is an abstract construct because, although it is empirically verifiable, it is only indirectly verifiable. (For example, the Morpheme Order and System Morpheme Principles directly support the claim that both morpheme order and one type of system morpheme come from the same language in intra CP code switching; it is this finding that indirectly supports the claim of a single, unified frame in such code switching). Like many other theoretical constructs, the Matrix Language refers to an abstract architecture. The Matrix Language is an abstract frame. This means it does not include actual morphemes nor is it isomorphic with any fully fleshed-out linguistic variety. Instead, it includes *specifications about slots* and how they are to be filled, based on directions from lemmas in the mental lexicon (2002: 68).

Further revisions by Myers-Scotton resulted in making it ever more difficult for the reader to keep up and follow her approach. In a recent paper (Myers-Scotton 2008), she dismisses the base relevance of earlier models in favour of highlighting new

additions and concepts. The following is an example of how both the MLF and 4-M model now can no longer fully justify all data:

Even though the MLF model's recognition of the asymmetries in CS leads to a synthetic analysis of morpheme distribution in CS, the model does not explain why its System Morpheme Principle applies to only one type of system morpheme. Also, even though the 4-M model offers a useful classification of morpheme type, again, it does not explain why any of the morpheme types show the distributions they do, not just in CS, but across contact phenomena in general. The 4-M model does imply a hypothesis about language production that offers an explanation. The Different Access Hypothesis<sup>267</sup> is one clue as to the extent to which different morpheme types are available for transfer to, or reconfiguration in, another language. (2008: 30)

Gardner-Chloros & Edwards (2004) in their analysis of code-switching as bilingual discourse and the MLF model add that:

The arguments involved, and the elaborations of the model itself, are of increasing complexity and appear to serve mainly to maintain the viability of the model. They have no direct bearing on the Base Language (2004: 119).

The model clearly has 'a lot to do with the Base Language.' However, I sympathise with Gardner-Chloros and Edwards in their criticism of the increasing complexity of the models involved.

However, Gardner-Chloros & Edwards (2004) concede that identifying a base language is of use:

For example some notion of base language can be of some practical use as a means of sifting the data and correlating the patterns found with sociolinguistic parameters: Rindler-Schjerve (1998), using Myers-Scotton's quantitative criterion, refers to a change of ML among the younger generation, which is symptomatic of language shift. At a grammatical level, however, instances of CS which contradict the MLF are found in her data (1998: 243). This is not surprising in that it is a big leap from using the notion of quantitative preponderance of morphemes from one variety to asserting that at

---

<sup>267</sup> The Differential Hypothesis as posited by Myers-Scotton is as follows: "The different types of morpheme under the 4-M are differentially accessed in the abstract levels of the production process. Specifically, content morphemes and early system morphemes are accessed at the level of the mental lexicon, but late system morphemes do not become salient until the level of the formulator" (2002: 2005a, 2008: 30).

an abstract frame provided by that variety provides a grammatical template for bilingual language production. (2004: 120).

They further question the viability of the model:

Although the grammatical details of Myers-Scotton's system have repeatedly been amended, the definition of the Base Language continues to be based on non-grammatical criteria – a fact which raises questions for the viability of the grammatical claims embodied in the model (2004: 119).

In concluding their analysis, Gardner-Chloros & Edwards (2004) summarise:

One of the greatest difficulties with existing models is in accounting for the role of CS in language change. If CS consists in a combination of two discrete systems, based on specific grammatical principles, then there is no clear place for the variation which precedes and underlies the refocusing of norms. Myers-Scotton's suggestion that the ML in a community may change over time (the "ML turnover hypothesis"), or even in extreme cases within a conversation, fails to account for the gradualness and irregularities of this process. In fact, there is ample evidence that the assumption that two distinct systems interact in CS, while at the same time retaining their separate identities, is an oversimplification, applying probably only to a minority of instances of CS (2004: 126).

Myers-Scotton (2002) in defence of the many revisions and amendments to her initial model adds:

However, what has changed is that I now benefit from the research and commentary by others in the code switching literature, as well as my own data collection and deliberations since the early 1990s. That is this work has given me corpora and analyses to consider, making it possible for me to reckon with the nature of these 'outlaw' constructions in a more informed way (2002: 109).<sup>268</sup>

---

<sup>268</sup> Myers-Scotton (2002) further explicates that: "The MLF model was drawn up to explain code switching patterns of speakers whose linguistic proficiencies in the language involved were fairly stable. Of course, the condition of a bilingual's abilities to speak the different languages in his or her repertoire is never really stable! For a variety of internal motivations and conditions external to the speaker, a bilingual's control over linguistic varieties varies. However, the main data set exemplifying the MLF model in Myers-Scotton (1993a) of Swahili/English bilingualism comes from a relatively stable bilingual condition....In its original formulation, the MLF model cannot account for all the structures in code switching of speakers in those communities where the relative status of the languages – in terms of both speaker proficiency and socio-political prestige – is more fluid than not. A composite source for the frame structuring bilingual CPs must be posited and the Morpheme Order and System Morpheme Principles do not apply categorically" (2002: 111-112).

Other problem areas in the MLF concern the association between content morphemes and system morphemes. As previously discussed, in the MLF model, content/system morphemes are distinguished according to whether they assign/receive a thematic role or not. To recapitulate, system morphemes are defined as having one of the following features:

#### Defining features of system morphemes

[+Quantificational] – System morphemes are quantificational, such as quantifiers, determiners and possessive adjectives.

[-Thematic Role Assigner] – System morphemes do not assign thematic roles.

[-Thematic Role Receiver] – System morphemes do not receive thematic roles.

(Myers-Scotton, 1993: 99-101)

However, both categorising grammatical forms into two concise content morpheme or system morpheme categories is at times difficult to distinguish, giving impetus to criticisms of the MLF model. Longobardi (1994) analyses that pronouns as a functional, system morpheme, category, whereas Myers-Scotton considers two types of pronouns as [-Thematic Role Receiver] (1993: 126) and certain system morphemes, namely ‘dummy’ pronouns (such as *it* and *there* in English) and clitics. Chan (1998) in his research on Cantonese and English code switching criticizes the model due to the ultimately non-distinct nature of content and system morphemes:

On the other hand, content morphemes are non-quantificational; they assign or receive thematic roles. More concretely, content morphemes cover major word classes such as nouns, verbs, pronouns, most adjectives and prepositions. In Myers-Scotton (1995), some conjunctions (e.g. *because*) are considered content morphemes as well because they assign thematic roles to clauses at the discourse level (e.g. “*because*” assigns the role ‘cause’ to the clause it introduces). In a nutshell, most of the system morphemes are what are assumed to be functional categories-determiners, quantifiers, modal verbs, inflectional affixes and complementizers. Some counter-examples have been attested to the above two principles. In one case, what are classified as EL content morphemes, for example, *pronouns*, do not always appear in mixed constituents. To deal with this particular problem, *the Matrix Language Blocking hypothesis* has been proposed” (1998: 5).

The ML Blocking Hypothesis as conceptualised by Myers-Scotton (1993) is as follows:

*The ML Blocking hypothesis*

*In ML + EL constituents, a blocking filter blocks any EL content morpheme which is not congruent with the ML with respect to three levels of abstraction regarding subcategorization (Myers-Scotton 1993: 120).*

Chan (2003) adds that it is not sufficient to 'label' certain constituents as content or system morphemes to fit in with certain data. The Blocking Hypothesis may be problematic as it further restricts the role of the embedded language as it only allows certain embedded language content morphemes in mixed discourse:

For an EL morpheme (say, a pronoun) to appear in a mixed constituent, it is not sufficient that the EL pronoun is a content morpheme. According to the ML Blocking hypothesis, the ML counterpart of this EL morpheme (say, the ML pronoun) has to be "congruent". In other words, the ML pronoun has to be a content morpheme as well. In other words, the ML Blocking hypothesis blocks EL single pronouns (even when these EL pronouns are content morphemes) if the pronouns in ML are system morphemes (Chan 1998: 5).

The content morpheme – system morpheme distinction also proves problematic also for Gardner-Chloros & Edwards (2004):

In recent formulations, the ML is said to provide the majority of system morphemes. The division between system and content morphemes is, however problematic. Firstly, as Muysken (2000) points out, there are at least four different criteria relevant to this kind of classification in different languages; also, the distinction does not operate in the same way across languages. Indeed in a later paper, Myers-Scotton's collaborator Jake writes that "there is variation across languages in the assignment of particular lexical "concepts" to content or system morpheme status" (1998: 354). There are also many examples of CS in which it is function words on their own that are the switched elements. This makes it difficult to see how the language of the function words could in itself determine the ML (Gardner-Chloros and Edwards 2004: 118).

The following example shows, according to Gardner-Chloros and Edwards that certain single function words are themselves switched and in essence proves

problematic in trying to establish how the identification of system morphemes can determine the matrix language overall:

- (367) *Et lui qui n 'est la que trois mois odder deux mois odder quatre*  
And him REL NEG is DEF except three months older two months older four  
*mois*  
months  
'And with him being there only three months or two months or four months'  
*French/Alsatian, Gardner-Chloros, (1991: 169).*

Furthermore, Boumans (1998) describes the flaw in attempting to strait-jacket descriptions in terms of content and system morphemes:

One of the major tasks of the MSA and of matrix language approaches generally is the classification of EL material that is inserted in mixed constituents. The examples [just] discussed show that the idea of content morpheme insertion as advanced in Myers-Scotton's MLF model is too restricted to describe the attested insertion patterns. Furthermore, embedded compound words cannot be explained by content morpheme insertion alone, unless the ML and EL share the same structures of compound words. 'Content word' instead of 'content morpheme' is a broader term that covers the attested derivation and inflected forms as well as EL compounds. However, 'content morpheme' is rather a discriminate expression introducing its own problems of demarcation. This is an even more serious drawback of the term 'EL island' or even 'EL internal island' used in the MLF model since these terms can refer to any combination of two or more EL elements which are formed by a larger ML constituent (1998: 68-69).

#### 8.4.1 Further Problematic Areas

Other problematic areas for the MLF are addressed by Myers-Scotton (2002), who states that: "Admittedly, not every single example in the literature is amenable to a satisfactory analysis under the MLF model" (2002: 108).

As has previously been examined in the literature review in Chapter Two, other researchers are also not without their own criticisms either, although over time and with further research, their work on the whole has been revised. Nortier (1990) in her

analysis of Moroccan Arabic and Dutch gives a now out-moded and obsolete working definition on aspects which define code switching:

- The first words of the sentence determine the matrix language
- If there is one switched constituent in a mixed sentence, the language of the majority of the constituents is the matrix language
- If the syntax of the sentence is Dutch, the matrix language is Dutch; if the syntax is Moroccan Arabic, the matrix language is Moroccan Arabic
- If there is no difference between Dutch and Moroccan Arabic syntax, the language of the majority of the constituents – among which preferably, but not necessarily, the main verb – is the matrix language
- If it still is doubtful which of the two languages is the matrix language, I assume that there is no matrix language at all. (Nortier, 1990: 160)

It is clear, given the data analysis and research progression over the last two decades that the above statements are not valid. This example has been given to highlight the fact that criticisms of the MLF (1993) onwards focus mainly on the revisions and amendments made to the original model and not the basic asymmetrical construct itself. The MLF has not been shown to be invalid in the same way that Nortier's (1990) approach has been invalid. Nortier (1990) summarises her account of certain constraints analysed, determining that there are three possible ways to account for violations in her data. These are as outlined below:

1. None of the constraints is valid (except perhaps the free morpheme constraint and the size-of-constituent constraint to a certain extent) as they are all violated in Dutch - Moroccan Arabic.

2. The constraints are valid, but the Dutch-Moroccan Arabic language pair is different from all other language pairs investigated so far.
3. The constraints should be reformulated on the basis of data from other language pairs (1990: 182).

Certainly, (1.) and (2.) above are no longer relevant to today's code-switched corpora as empirical evidence points to a more asymmetrical model. The third assumption (3.) is the most valid and relevant as this is the base hypothesis which should be adopted by linguists when attempting to identify a universally valid constraint. Nortier (1990) goes on to state:

If linguists ever want to formulate generally valid constraints on code switching, it is absolutely necessary to lay the proper groundwork before moving on to the formulation of constraints. They must clearly state whether or not a matrix language is assumed and if so, how it is selected (1990: 182).

In defence of criticisms of the MLF, Myers-Scotton (2002) in her approach to considering problematic data states that:

However, most of what appear to be problems can be accommodated under the provision of the model, or expansions of how the provisions can be interpreted (2002: 108).<sup>269</sup>

The above provides a further catalyst for my own conclusions and in essence the overriding result that is Generalization 1, as this does not overly explicate data and in general terms, is a well-sourced, abstract construct for code switching data. This in turn is the basic format for an all-encompassing and accommodating approach to code switching irrespective of language type, insertions, amalgamations and speech styles.

---

<sup>269</sup> Myers-Scotton (2002) comments on the changes to her theory since the original model was put forward in 1993 as follows: "What has changed is that I now benefit from research and commentary by others in the code switching literature, as well as my own data collection and deliberations since the early 1990s. That is, this work has given me corpora and analyses to consider, making it possible for me to reckon with the nature of these 'outlaw' constructions in a more informed way" (2002: 109).

That is that there will be asymmetry between the matrix language and its embedded variety each fitting into the morpho-syntactic frame as set by the matrix language within any code-switched clause. It is this abstract nature and flexibility together with the base asymmetry which renders Generalization 1 successful, as well as the original MLF model and applicable to most, if not all data in classic code switching varieties. This has been addressed throughout this thesis and in essence answers the research questions and statements as set out in Chapter Five.

### **8.5 Research Questions Revisited**

The research questions as posed in Chapter Five were based on the following criteria:

- Age – all three generation groups must be represented
- Both genders must be as proportionately represented as possible
- Informant must be bilingual or at least have a proficient command of Moroccan Arabic as well as English
- Informants must be from different regions of Morocco
- Only open-ended questions may be asked so as to maximize language responses

Certain essential research points with regard to the data were borne in mind whilst (a) recording the informants, (b) during the transcribing stage and (c) focusing on the possible intra-sentential switch sites. The issue of whether these have been satisfactorily answered and addressed is outlined as follows:

Figure 8.4: Research questions summary

	Research Questions		Validated by data ✓ / x
1	Moroccan Arabic to English code switching?		✓
2	English to Moroccan Arabic code switching?		✓
3	Single Moroccan Arabic noun insertions (islands)?		✓
4	Single English noun insertions (islands)?		✓
5	Consistency of Moroccan Arabic as the matrix language?		✓
6	Consistency of English as the matrix language?		✓
7	Consistency of Moroccan Arabic as the embedded language?		✓
8	Consistency of English as the embedded language?		✓
9	Moroccan Arabic affixation to English stems?		✓
10	English affixation to Moroccan Arabic stems?		✓
11	Grammaticality?		✓
12	Validation of the MLF model?		✓

The validity of the MLF and Generalization 1 largely depend upon the principles and hypotheses which they support; namely, the flexibility and abstract nature of the matrix language – embedded language distinction. Therefore, the basic tenor of this thesis is such as has been stated by Myers-Scotton (2002), namely that:

The theoretical notion [is] that the same principles and processes underlie all language contact phenomena...that these principles and processes are not basically different from what structures language in a general sense. That is what happens to languages when their speakers know and use two or more languages is not really unusual if one considers the larger linguistic picture. The fact that we are dealing with bilingual speech does not mean certain other provisions need to be made, but they are not all that remarkable and certainly not curious and never bizarre...we can see how the division, as well as the interface, between the basic roles of the lexicon and structurally assigned grammatical specifications is apparent in linguistic phenomena (2004: 295).

## 8.6 Conclusion

The main concern of this chapter is the theoretical generalizations / hypotheses posited throughout the thesis. These form the bedrock of the entire research as they highlight certain principles and processes which underlie not only Moroccan Arabic

and English code switching, but in the case of Generalization 1 should in theory apply to all language contact phenomena. This is the general linguistic theory behind the MLF model which renders it feasible and the most suitable current model for describing language contact phenomena today. Further generalizations and hypotheses are summarised and compared with historical accounts of code switching over the last thirty years and the approach is corroborated by the Moroccan Arabic and English corpus. It is hoped that this will be a useful stimulus for further linguistic research.

The MLF and its supporting models, the 4-M Model and the Abstract Level model, both add precision to the earlier MLF edition (1993). Within the parameters of classic code switching, the MLF accounts for a wide range of data, whereas the revised models are able to account for more composite and innovative types of code switching. This chapter has also analysed certain problematic data for the MLF and code switching in general which are addressed and discussed in full. It is shown that the problematic data mainly stem from the many revisions made by Myers-Scotton (2002) to the original templatic model. It was shown how categorising the content morpheme – system morpheme distinction has proven to be problematic for certain researchers. Categorising grammatical forms into a single content morpheme versus system morpheme dichotomy is at times difficult to do, giving impetus to criticisms of the MLF model. The chapter concludes below with a summary of the research questions as set out in Chapter Five and the corroboration of the hypotheses by the data examined throughout the thesis.

## CONCLUSION

This thesis is explorative and original in its nature as it is one of the first concerning intra-sentential code switching of Moroccan Arabic and English in the UK. Furthermore, it presents the original code-switched data and validated this phenomenon under the rubric of the Matrix Language Frame model, as it has transpired that the other models and approaches to code switching were not a valid vehicle for my data. On the basis of discussion of previous Chapters, the Matrix Language Frame Model as an asymmetrical model (Myers-Scotton 1993b, 2002) has been evaluated and I have proposed this model to account for Moroccan Arabic and English intra-sentential code switching as evaluated in detail in Chapter Three. This model, which involves a matrix language and an embedded language has proven to be the most suitable for code switching in its basic form and can also be utilized for typologically similar or dissimilar data in all language categories (cf. Boumans 1998, Ziamari 2007).

On the basis of an analysis and assessment of the most prominent theoretical models and data as detailed in previous chapters, the Matrix Language Frame model linguistic model has been proposed and presented to account for most code switching data in natural bilingual intra-sentential discourse. The principal approaches to such linguistic exchanges as detailed are linear models, Government and Binding theory, other asymmetric models, constraints-based models, non-constraints-based models and the minimalist approach. This research has attempted to locate the most viable and all-encompassing solution to language contact and the grammatical constraints governing intra-sentential code switching using the corpus as presented. The principal constraints-based models analysed (Belazi, Rubin & Toribio 1994, Bentahila & Davies 1983, Di Sciullo, Muysken & Singh 1986, Poplack 1980, Sankoff & Poplack 1981, Pfaff 1979, Santorini & Mahootian 1995) as well as those approaches and those

which reject the idea of universal, purely syntactically driven constraints on code switching (Bokamba 1989, Clyne 1987, MacSwan 1999, 2000) have all been found to be too stringent.

I have also attempted to show that MacSwan's (1999) notion of his monolingual structural constraints is valuable with its insight that no code switching-specific constraints which have been posited in the literature can account for the range of facts as evidenced in code-switched discourse. In essence, nothing constrains code switching apart from the requirements of mixed grammars. This will be a useful premise for further research.

In addition, I presented a set of generalizations which emerge from an analysis of the data and the main code switching theories and models. The first Generalization as initially detailed in Chapter One and at various points thereafter has been proposed to account for most data on a morphological and syntactic level. This is recapitulated below:

*Generalization 1:*

*The Leader (of two or more languages) is that which contributes word order in the CP frame where the Led (embedded) variety adheres to the morpho-syntactic frame provided by the Leader.*

After analysis of the data, I postulate a further ten generalizations which account for Moroccan Arabic and English code switched complex data. These are summarized in Chapter Eight.

In evaluating code switching, it has been shown how lexical insertions give us an important insight into code switching practices and regularities amongst bilingual speakers. Asymmetry as discussed is a striking feature of code switching in numerous typologically similar and dissimilar languages asymmetry is valid for all data sets.

The generalization that there will always be a matrix language which provides the grammatical frame with an embedded variety adhering to the grammatical parameters as fixed by the matrix variety has been validated in data presented in this thesis as well as in data by other researchers. Furthermore, the notion of asymmetry and the syntactic dynamic in the content morpheme versus system morpheme distinction has been validated by the corpus presented. In this regard, the theoretical concept of the MLF as posited by Myers-Scotton (1993b, 2002) as well as the 4-M Model and Abstract Level Model remain not only valid but also a suitable vehicle for analysis of both inter-sentential and intra-sentential data sets.

Upon examination of the data, a new and emerging speech style amongst second but most notably third generation Moroccan Arabic speakers was noted. This I termed *Reactive Syntax*, as the syntax of these generational groups is clearly distinct from that of first generation speakers. Moroccan Arabic has only been in intensive contact with English over the last thirty years or so and we predict that certain patterns and trends will emerge, namely novel speech styles, more mixing of clauses such that this becomes conventional usage, with conversion rules applying in affixing certain Moroccan morphemes to English stems, and conversely, affixing English morphemes to Moroccan stems, as shown and discussed in chapters Six to Eight. These all come under the new concept of *Reactive Syntax*. This innovative notion will need to be corroborated with further data and research which spans all generational groups in order to attain comprehensive insights into different speech styles and syntactic variations.

The findings of this research have some implications for the analysis of bilingual discourse, since it has been found that it is not possible to formulate absolute constraints that are applicable in code-switched discourse (cf. Tuc 2003). This is also shown in the work by Romaine (1995) who states:

More collaborative work is needed between psycholinguists and sociolinguists to develop models of processing and production which can handle code switching. Such models should then be used to inform and test grammatical theories (1995: 180).

This lends support to the suitability of the MLF model in its basic form as an asymmetrical model which covers all language groups irrespective of typology, there always being a matrix and an embedded variety at any given time. A further point of importance is that this thesis does not provide a sociolinguistic analysis of Moroccan Arabic and English code switching, due to time and space constraints. This will be a useful area of further research as it will add a new dimension to the analysis postulated thus far. It may contribute to the development of a comprehensive theoretical speech model which can account for language contact in general.

It is anticipated that the findings and data of this thesis will provide a solid basis for further research not only of intra-sentential code switching but also of Arabic linguistics in general.



## BIBLIOGRAPHY

- Aabi, M. (1999). *The Syntax of Moroccan Arabic/French and Moroccan Arabic/Standard Code Switching*. PhD Thesis. Sheffield.
- Abbassi, A. (1977). *A Sociolinguistic Analysis of Multilingualism in Morocco*. PhD Thesis. University of Texas.
- Abboud, P. F. (1970). Spoken Arabic. *Current Trends in Linguistics VI: Linguistics in So. West Asia and No. Africa*. T. A. Sebeok. The Hague, Mouton. VI: 439-466.
- Abdel Massih, E. (1973). *A Course in Moroccan Arabic*. Michigan.
- Abu-Absi, S. (1986). 'The Modernization of Arabic: Problems and Prospects.' *Anthropological Linguistics* 28. Indiana University. 337-348.
- Agar, Michael, H. (1980). *The Professional Stranger: An Informal Introduction to Ethnography*. New York: Academic Press.
- Aissati, A. El. (1993). 'Berber in Morocco and Algeria: Revival or Decay?' *AILA Review*, 10: 88-109.
- Aissati, A. El. & K. De Bot. (1994). 'Moroccan Arabic in the Netherlands: Acquisition and Loss.' *Indian Journal of Applied Linguistics*, 20: 1&2, 177-192.
- Al Khatib, M. A. & M. N. El Ali. (2005). 'Language and Cultural Maintenance Among the Gypsies of Jordan.' *Journal of Multilingual and Multicultural Development*. Vol. 26 (3): 187-215.
- Alfonzetti, G. (1992). *Il discorso bilingue. Italiano e Dialetto a Catania*. Milano. Franco Angeli.
- (1998). 'The conversational dimension in code-switching between Italian and dialect in Sicily.' In, P. Auer (ed.) *Code-switching in Conversation*. London: Routledge. 180-214.
- Andersen, E. S. (1982). 'Determining the Linguistic Attributes of Language Attrition.' In, R. Lambert & B. Freed (eds.) *The Loss of Language Skills*. Rowley-Massachusetts: Newbury House Publishers. 83-118.
- (1990). *Speaking with Style: The Sociolinguistic Stem of Children*. London and New York: Routledge.
- Anderson, S. R. (1985a). 'Inflectional Morphology', in T. Shopen (ed.) *Language and Typology and Syntactic Description. Vol. 3: Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press. 150-201.
- Annamalai, E. (1989). 'The Language Factor in Code Mixing.' *International Journal of the Sociology of Language*. Mouton de Gruyter.

- Appel, R. & P. Muysken. (1987). *Language Contact and Bilingualism*. London: Edward Arnold.
- Argyle, M. (1969). *Social Interaction*. London. Methuen.
- Arndt, W.N. 1970. 'Nonrandom Assignment of Loanwords: German Noun Gender,' *Word*. 26, 244-253.
- Assad, M. (1978). *Le Parler Arabe de Tanger*. PhD Thesis. University of Gothenburg, Sweden.
- Atkinson, J. M. & J. Heritage. (1984). *Structures of Social Action*. Cambridge: Cambridge University Press.
- Auer, J, C. P. (1984). *Bilingual Conversation*. Amsterdam: John Benjamins.
- (1998). *Code Switching in Conversation: Language, Interaction and Identity*. London: Routledge.
- Azuma, S. (1991a). *Processing and Intrasentential Code-Switching*. PhD Thesis. University of Texas, Austin.
- (1991b). 'Two-Level Process Hypothesis in Speech Production: Evidence from Intrasentential Code-Switching.' Paper presented at the 27<sup>th</sup> Chicago Linguistic Society Meeting, May 1991.
- (1993). 'The Frame Content Hypothesis in Speech Production: Evidence from Intra-Sentential Code Switching.' *Linguistics*, 31. 1071-93.
- Bach, K. & Harnich, R. M. (1979). *Linguistic Communication and Speech Acts*. Cambridge, Massachussets: MIT Press.
- Backus, A. (1990). 'Turkish-Dutch Codeswitching and the Frame Process Model.' Paper presented at the International Workshop on Ethnic Minority Languages in Europe, Tilburg, Netherlands, December 1990.
- (1996). *Two in One. Bilingual Speech of Turkish Immigrants in the Netherlands*. (Studies in Multilingualism 1). Tilburg: Tilburg University Press.
- (2000). 'Code-Switching in Conversation: Language, Interaction and Identity.' *Journal of Pragmatics*. Elsevier. Vol. 32(6): 831-838.
- Baetens-Beardsmore, H. (1988). 'Bilingualism, Basic Principles.' Clevedon, Avon: *Multilingual Matters*.
- (1992). 'Code-Switching and Language Dominance.' In, R.J Harris (ed.) *Cognitive Processing in Bilinguals*, Amsterdam: Elsevier. 443-58.
- Baker, A. (1998). *Voices of Resistance: Oral Histories of Moroccan Women*. Albany, NY: State University of New York Press.
- Barion, C. (1981). *Islam, Société et Communauté: Anthropologies du Maghreb*.

Paris: Editions du Centre National de la Recherche Scientifique.

- Bassiouney, R. (2006). *Functions of Code Switching in Egypt: Evidence from Monologues*. Leiden; Boston: Brill.
- Bates, E. & B. MacWhinney. (1981). 'Second Language Acquisition from a Functionalist Perspective: Pragmatic, Semantic and perceptual Strategies.' In, H. Winitz (ed.) *Native Language and Foreign Language Acquisition. Annals of the New York Academy of Sciences*. Vol. 379. New York: New York Academy of Sciences Press. 190-214.
- Bauman, R. (1986). *Story, Performance and Event*. Cambridge: Cambridge University Press.
- Bautista, M. L. (1975). 'A Model of Bilingual Competence Based on an Analysis of Tagalog-English Code-switching.' *Philippine Journal of Linguistics* 6(1). 51-89.
- Bean, S. (1978). *Symbolic and Pragmatic Semantics*. University of Chicago Press.
- Belazi, H. M, E.J. Rubin & A. J. Toribio. (1994). 'Code Switching and X-Bar Theory: The Functional Head Constraint.' *Linguistic Inquiry* 25, S. 221-37.
- Benchiba, N. (2008). 'Discourse 'Darija' Particles in Moroccan Arabic/English Code Switching.' In, proceedings of *Association Internationale de Dialectologie Arabe*. Paper presented at (AIDA 7), Vienna, 5-9 September 2006.
- Benchiba, N. (2007). 'Language & Gender in Moroccan Arabic and English Codeswitching. A Levelling Phenomenon.' In proceedings, *International Perspectives on Gender*. Paper presented at Fourth International Conference Language and Gender Association. University of Valencia, 8-10 November 2006.
- Benhallam, A. (1990). 'Native Speakers' Intuitions About Moroccan Arabic Stress.' In, J. Pleines (ed.) *La Linguistique au Maghreb*. Rabat: Editions Okad. 91-110.
- Benkaddour, A. (1982). *A Non-Linear Analysis of Some Aspects of the Phonology and Non-Concatenative Morphology of Arabic*. PhD Thesis. School of Oriental and African Studies, London.
- (1990). 'Semi-vowels and high vowels in Moroccan Arabic: A Nonlinear Approach.' In, In, J. Pleines (ed.) *La Linguistique au Maghreb*. Rabat: Éditions Okad. 127-143.
- Bentahila, A. (1983). 'Motivations for Code-Switching Among Arabic-French Bilinguals in Morocco.' *Lingua*, 59: 301-330.
- (1983). *Language Attitudes Among Arabic-French Bilinguals in Morocco*. Clevedon: Multilingual Matters.

- Bentahila, A. & E. Davies. (1983). 'The Syntax of Arabic-French Code Switching.' *Lingua*, 59. 301-30.
- (1988). 'Convergence and Divergence: Two Cases of Language Shift in Morocco.' Paper presented at the International Conference on Maintenance and Loss of Ethnic Minority Languages, Noordwijkerhout, August 1998.
- (1991). 'Constraints on Code-switching : A Look Beyond Grammar.' *Papers for the Symposium on Code-switching in Bilingual Studies*, Vol. 2. 369-403.
- (1995). Patterns of Code-Switching and Patterns of Language Contact.' In, *Lingua*. Vol. 96(2). Elsevier Science. 75-93.
- (2002). 'Language Mixing in Rai Music: Localisation or Globalisation?' *Language and Communication*. Vol. 22(2): Strasbourg: European Science Foundation. 187-207.
- Berk-Seligson, S. (1986). 'Linguistic Constraints on Intrasentential Code-Switching. A Study of Spanish-Hebrew Bilingualism.' *Language in Society*. Cambridge. Cambridge University Press. 15:313-48.
- Bhatia, T. K. & W. C. Richie. (1996). *The Handbook of Bilingualism*. Blackwell.
- Bhatt, R. M. (1997). 'Code-Switching, Constraints and Optimal Grammars.' *Lingua*. Vol. 102(4). Elsevier Science. 223-251.
- Bickmore, L.S. (1985). *Hausa-English code-switching*. MA Thesis. University of California: Los Angeles.
- Bilmes, J. (1988a). 'The Concept of Preference in Conversation Analysis.' *Language in Society*, 17(2). Cambridge: Cambridge University Press. 161-81.
- Black, M. (1962). *Language & Philosophy: a Symposium*. New York: NYU Press.
- Blaikie, N. (2006). *Designing Social Research*. Blackwell.
- Bloch, M. (1998). 'Language, Anthropology & Cognitive Science.' *How We Think They Think*. Boulder: Westview Press. 39-53.
- Blom, J.P. & J. J. Gumperz. (1972). 'Social Meaning in Structure: Sode-switching in Norway.' In J.J. Gumperz and D. Hymes (eds.) *Directions in Sociolinguistics*. New York: Holt, Rinehart, and Winston. 407-34.
- Bock, J.K. (1987). 'Co-ordinating Words and Syntax in Speech Plans.' In, A.W. Ellis (ed.). *Progress in the Psychology of Language*, iii. London: Erlbaum. 337-90.
- Boeschoten, H. (1987). 'Are There Syntactic Constraints on Code Mixing?' In, K. Denning (ed.) *Variation in Language*. Stanford, California: Stanford University Press. 35-51.
- (1988). 'Code-mixing, Language Variation and Linguistic Theory: Evidence from Bantu Languages.' *Lingua*. Elsevier Science. 76:21-62.

- (1991). 'Asymmetrical Code-switching in Immigrant Communities.' *Papers for the Workshop on Constraints, Conditions and Models*, Strasbourg: European Science Foundation. 85-100.
- Bokamba, E. (1989). 'Are there Syntactic Constraints on Code-mixing?' *World Englishes*. 8(3).
- Bolinger, D. (1968). *Aspects of Language*. New York: Holt, Rinehart, and Wilson.
- Bos, P. & L. Verhoeven. (1994). 'Moroccan-Arabic Dutch Bilingual Development,' *Indian Journal of Applied Linguistics*. 20: 1&2. Bahri Publications. 120-150.
- Boudali, B. (1984). *Syntax: Some Aspects of Sentential Complements in Moroccan Arabic*. Technical Report, Mohamed V University. Faculté des Lettres.
- Boukos, A. (1995). *Société, Langues et Cultures au Maroc: Enjeux Symboliques*. Faculté des Lettres et des Sciences Humaines. Rabat.
- Boumans, L. (1998). *The Syntax of Code Switching: Analysing Moroccan Arabic/Dutch conversations*; Tilburg: Tilburg UP.
- (1996). 'Embedding Verbs and Collocations in Moroccan Arabic/Dutch Codeswitching.' In, Eid, M & Parkinson, D. *Perspectives on Arabic Linguistics IX*. (Current Issues in Linguistic Theory 134). Amsterdam/Philadelphia: Benajmins. 45-67.
- Boumans, L & D. Caubet. (2000). 'A Comparative Study of Code Switching in Algeria and in the Netherlands.' In, J. Owens, *Arabic as a Minority Language*. Berlin: Mouton de Gruyter.
- Bourdieu, P. & L. J. D. Wacquant. (1992). *An Invitation to Reflexive Sociology*. University of Chicago Press.
- Bourquia, R. (1996). *Femmes et Fécondité*. Casablanca: Afrique Orient.
- Bouwhius, C. B. (2002). 'Determination of Ethnicity in the Netherlands: Two Methods Compared.' *European Journal of Epidemiology*. Vol. 18(5). Netherlands: Springer. 385-388.
- Boyer, P. (1990). *Tradition as Truth and Communication: A Cognitive Description of Traditional Discourse*. Cambridge: Cambridge University Press.
- Bresnan, J. (1978). 'A Realistic Transformational Grammar.' In, M. Hale, J. Bresnan, and G. Miller (eds.) *Linguistic Theory and Psychological Reality*. Cambridge, Massachussets: MIT Press. 1-59.
- (1982). *The Mental Representation of Grammatical Relations*. Cambridge, Massachussets: MIT Press.

- Brown, B. (1986). 'Cajun/English Code-Switching: A Test of Formal Models.' In D. Sankoff (ed.) *Diversity and Diachrony*. Amsterdam: Benjamins. 399-406.
- Brown, P. & S. Levinson. (1987). *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Brunot, L. (1950). *Introduction a l'arabe Marocain*. Paris: G. P. Maisonneuve. Brunswick, N.J. (USA). London: Rutgers University Press.
- Brustad, K. E. (2000). *The Syntax of Spoken Arabic*. Georgetown University Press.
- Bull, P. (1991). *The Non-Linear Phonological Structure of Moroccan Colloquial Arabic*. PhD Thesis. University College, London.
- Caccamo, A. (1998). From 'Switching Code' to 'Code-Switching': Towards a Reconceptualisation of Communicative Codes.' In, P. Auer (ed.) *Code-Switching in Conversation*. London: Routledge. 29-50.
- Cannell, C. F. & R. L. Kahn (1968). 'Interviewing.' In, G. Lindzey & A. Aronson (eds.) *The Handbook of Social Psychology, Vol 2. Research Methods*. New York: Addison Wesley.
- Caubet, D. (1993). *L'arabe Marocain*. Paris/Louvin: Peeters.
- Chan, B. (2003). *Aspects of the Syntax, the Pragmatics and the Production of Code Switching*. New York: Peter Lang.
- Cherti, M. (2006) *Personal Communication*.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- (1965). *Aspects of the Theory of Syntax*. Cambridge, Massachussets: MIT Press.
- (1968). *Language and Mind*. New York: Harcourt Brace Jovanovich.
- (1973). *Cognitive Sociology*. Harmondsworth: Penguin.
- (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- (1986). *Barriers*. Cambridge, Massachussets: MIT Press.
- (1986). *Knowledge of Language: Its Nature, Origin and Use*. New York: Praeger.
- (1995) *The Minimalist Program*. Cambridge: MA. MIT Press.
- Clark, J. (1982). 'Measurement Considerations in Language Attrition Research.' In, R. Lambert & B. F. Freed (eds.) Rowley, MA: Newbury House. 138-152.
- Clarke, M.A., M.D. Losoff, M.D McCracken & J. Still. (1981) 'Gender Perception in Arabic and English.' *Language Learning*, 31.159-67.

- Clyne, M. G. (1967). *Transference and Triggering*. The Hague; Nijhoff.
- (1987). 'Constraints on Code-switching: How Universal Are They?' *Linguistics*. The Hague: Paris: Mouton Publishers. 25: 739-64.
- Cole, P. & J. L. Morgan. (eds.) (1975). *Syntax and Semantics. Vol 3: Speech Acts*. New York: Academic Press.
- Colin, G. S. (1999). *Arabe Marocain /Inédits de Georges S. Colin; (ed.) Dominique Caubet & Zakia Iraqui-Sinaceur; Paris: Édisud.*
- Collyer, M. (2004). *The Development Impact of Temporary International Labour Migration on Southern Mediterranean Sending Countries: Contrasting Examples of Morocco and Egypt*. Development Research Centre on Migration, Globalisation and Poverty, University of Sussex.
- Comrie, B (1989). 'Genetic Classification; Contact and Variation.' In, T.J. Walsh (ed.) *Synchronic and Diachronic Approaches to Linguistic Variation and Change*. Washington, D.C: Georgetown University Press. 81-93.
- Corder, S. P. (1981). *Error Analysis and Interlanguage*. London: Oxford University Press.
- Coulthard, M. (1977). *An Introduction to Discourse Analysis*. London: Longman.
- Couper, K, E. Selting & M. Selting, M. (eds.) (1996). *Prosody in Conversation: Interactional Studies*. Cambridge: Cambridge University Press.
- Dabene, L. & J. Billiez. (1986). 'Code-Switching in the Speech of Adolescents Born of Immigrant Parents.' *Studies on Second Language Acquisition*. Oxford: Oxford University Press. 8: 309-325.
- Dahbi, M. & M. Ezroua & L. Haddad. (1996). *Cultural Studies, Interdisciplinarity, and the University*. PhD Thesis. Al-Rabat: Faculty of Letters and Human Sciences.
- Davis, S. S. (1982). *Patience and Power: Women's Lives in a Moroccan Village*. Rochester, VT (USA): Schenkman Books.
- (1989). *Adolescence in a Moroccan Town: Making Social Sense*. New York.
- De Bot, K. (1992) A Bilingual Processing Model: Levelt's 'Speaking' Model Adapted. *Applied Linguistics*, 13: 1-24.
- De Haas, H.(2003). *Migration and Development in Southern Morocco. The Disparate Socio-Economic Impacts of Out-Migration on the Todgha Oasis Valley*. PhD Thesis. Radboud University, Nijmegen
- (2005). 'Morocco: From Emigration Country to Africa's Migration Passage to Europe. Country Profile Morocco.' *Migration Information Source*, October 2005.

- Dell, G.S. (1990). *'Frame Constraints and Phonological Speech Errors.'* Paper presented at CUNY Conference on Human Sentence-Processing, New York.
- Denison, N. (1977). 'Language Death or Language Suicide?' *International Journal of the Sociology of Language*. Vol. 12. SanDiego, California: UCSD Institute for Cultural Pluralism. 13-22.
- Di Sciullo, A. M., P. Muysken, & R. Singh. (1986). 'Government and Code-Mixing.' *Journal of linguistics*. Cambridge: Cambridge University Press. 22: 1-24.
- Di Sciullo, A. M. & Williams, E. (1987). *On the Definition of Word*. MIT Press.
- Dorian, N. (1977). 'The Problem of the Semi-Speaker in Language Death' (ed.) (1989) *Investigating Obsolescence: Studies in Language Contraction and Death*. Cambridge: Cambridge University Press.
- (1982). 'Language Loss and Maintenance in Language Contact Situations,' in R. Lambert and B. Freed (eds.). *The Loss of Language Skills*. Rowley, MA: Newbury House. 44-59.
- Doron, E. (1983). 'On a Formal Model of Code-Switching.' *Texas Linguistic Forum*, 22: 35-59.
- Douglas, J. D. (1976). *Investigative Social Research*. Beverly Hills, California: Sage.
- Duran, L. (1994). 'Toward a Better Understanding of Code-switching and Interlanguage in Bilinguality: Implications for Bilingual Instruction.' *The Journal of Educational Issues of Language Minority Students*. 14. 69-88.
- Duranti, A. (ed.) (1985). 'Sociocultural Dimensions of Discourse.' In, T.A. V. Dijk (ed.) *Handbook of Discourse Analysis. Vol. 1. Disciplines of Discourse*. New York: Academic Press. 193-230.
- (1988a). 'The Ethnography of Speaking: Toward Linguistics of the Praxis.' In, F. Newmeyer (ed.), *Linguistics: The Cambridge Survey. Vol. 4. Language: The Socio-Cultural Context*,. Cambridge: Cambridge University Press. 210-28.
- (1991). 'Four Properties of Speech-in-Interaction and the Notion of Translocutionary Act', in J. Verschueren (ed.) *Pragmatics at Issue: Selected Papers of the 1987 International Pragmatics Conference*. Antwerp, August 1991. Amsterdam: Benjamins. 17-22.
- (1997). *Linguistic Anthropology*. Cambridge University Press.
- (2001). *Key Terms in Language and Culture*. Oxford: Blackwell.
- Dwyer, D. H. (1978). *Images and Self-Images: Male and Female in Morocco*. New York; Guildford: Columbia University Press.
- Dwyer, K. (1982). *Moroccan Dialogues: Anthropology in Question*; Baltimore; London: Johns Hopkins University Press.

- Dzameshie, A. (1989). 'Code-Switching in Ewe and English.' Paper presented at African Linguistics Conference, Urbana, III.
- Eckert, P. & S. McConnell-Ginet. (2003). *Language and Gender*. Cambridge: Cambridge University Press.
- Eid, M. (1992). 'Directionality in Arabic-English Code-Switching.' In, A. Rouchdy (ed.) *The Arabic Language in America: A Sociolinguistic Study of a Growing Bilingual Community*. Detroit: Wayne State University Press. 50-71.
- (1996). 'Things Are Not What They Seem: Pronoun Doubling in Bilingual and Monolingual Grammars.' *Perspectives on Arabic Linguistics VIII*, (ed.) by M. Eid. (Current Issues in Linguistic Theory 134). Amsterdam / Philadelphia: Benjamins. 7-29.
- El Aissati, A. (1997). *Language Loss Among Native Speakers of Moroccan Arabic in the Netherlands*. Tolburg: Tilburg UP.
- Eliasson, S. (1991). 'Models and Constraints in Code-Switching Theory.' *Papers for the Workshop on Constraints, Conditions and Models*. Strasbourg: European Science Foundation. 17-50.
- El Idrissi, A. (1990). *Verbes Opérateurs et Nominalisation en Arabe Marocain (Parler de Fès)*. PhD Thesis. Université de Jussieu (Paris VII).
- El Mansouri. (1996) Basfao & Taarji 1994 (IT 1982, 1990); *National Statistical Services* (B and F 1998; N, G, ES, IT 1998 and 2005); López García 1999 (ES 1968-1990); IOM and Fondation Hassan II 2003 (B and F 2002).
- Ennaji, M. (1982). *A Comparative Analysis of the Complex Sentence in English, Moroccan Arabic and Berber*. PhD Thesis. University of Essex.
- (1985). *Contrastive Syntax: English, Moroccan Arabic and Berber Complex Sentences*. Würzburg: Königshausen + Neumann, 1985.
- (1994). *Applications of Modern Linguistics*. Casablanca: Afrique-Orient.
- (2005). *Multilingualism, Cultural Identity and Education in Morocco*. New York: Springer.
- Ervin-Tripp, S. (1972). 'On Sociolinguistic Rules: Alternation and Co-occurrence.' In, J. J. Gumperz and D. Hymes (eds.) *Directions in Sociolinguistics: The Ethnography of Communication*. New York: Holt. 213-50.
- (1973). 'The Structure of Communicative Choice,' in A. S. Dil (ed.) *Language, Acquisition and Communicative Choice: Essays by Susan Ervin-Tripp*. Stanford: Stanford University Press. 302-73.
- (1984). 'Polish-English Code-Switching: A Clue to Constituent Structure and Processing Mechanisms.' In, J. Drogo et al. (eds.) *Papers from the Regional Meeting of the Chicago Linguistic Society*. Chicago: CLS. 52-64.

- Extra, G. & L. Verhoeven. (1993). 'A Bilingual Perspective on Turkish and Moroccan Children and Adults in the Netherlands.' In, G. Extra & L. Verhoeven (eds.). *Immigrant Languages in Europe*. Clevedon: Multilingual Matters. 67-100.
- Fasold, R. W. (1977). *Studies in Language Variation : Semantics, Syntax, Phonology, Pragmatics, Social Situations, Ethnographic Approaches. Colloquium on New Ways of Analyzing Variation* (3rd) Georgetown University. Washington: Georgetown University School of Languages and Linguistics.
- Fehri, A. (1996). *Al-Lisaniyat al-muqaranah wa-al-lughat fi al-Maghrib = Linguistique comparée au Maroc / coordination scientifique*. Rabat: Faculté des Lettres et des Sciences Humaines.
- Ferguson, C. A. (1959). 'Diglossia.' *Word*. New York: International Linguistic Association.
- Ferguson, G. (2003). 'Classroom Codeswitching in Post-Colonial Contexts: Functions, attitudes and Policies.' In, *AILA Review*. Vol. 16(1): 38-51.
- Fielding, N. G. & J. L. Fielding. (1986). *Linking Data*. London: Sage.
- Fillmore, C. J. (1977b). 'Topics in Lexical Semantics,' in R. Cole (ed.) *Current Issues in Linguistic Theory*. Bloomington: University of Indiana Press. 76-138.
- Fishman, J A. (1968). *Readings in the Sociology of Language*. The Hague: Mouton.
- (1972). *The Sociology of Language: An Interdisciplinary Social Science Approach to Language in Society*. Rowley, Massachussets: Newbury House Publishers.
- (1980). 'Bilingualism and Biculturalism as Individual and Societal Phenomena.' *Journal of Multilingual and Multicultural Development*, 1: 3-17.
- (1990). 'What is Reversing Language Shift (RLS) and How Can it Succeed?' *Journal of Multilingual and Multicultural Development*. Clevedon: Multilingual Matters. 11: 5-36.
- Fodor, J. (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.
- Foley, W. A. (2000). *Anthropological Linguistics*. Oxford: Blackwell.
- Fowely, W. & R. Van Valin. (1984). *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Fowler, C. (1985). 'Current Perspectives on Language and Speech Production: A Critical Overview.' In, R. Daniloff (ed.) *Speech Science*, 193-278. San Diego: College Hill Press.
- Gaudet, F. & M. C. Jones. (2008). 'Variation, Contact and Convergence in French Spoken Outside France.' *Journal of Language Contact*. THEMA 2.

- Gal, S. (1984). 'Phonological Style in Bilingualism: The Interaction of Structure and Use.' In, D. Schiffrin, Tannen, D & Hamilton H. (eds.) *The Handbook of Discourse Analysis*. Blackwell. 290-302.
- Gallagher, C. (1963). *The United States and North Africa*. Cambridge, Massachussets: Harvard University Press.
- Gardner-Chloros, P. (1985). *Choix et Alternance des Langues a Strasbourg*. PhD Thesis. University Louis Pasteur, Strasbourg.
- (1987). 'Code-Switching in Relation to Language Contact and Convergence.' In, G. Lüdi (ed) *Devenir Bilingue-parler Bilingue*. Tübingen: Niemeyer.
- Gardner-Chloros, P. & M. Edwards. (2002). 'Touching Base: The Relevance of Grammatical Models to Code Switching.' *Actas/Proceedings II II Simposio Internacional*. 1433-1452.
- Garrett, M. F. (1975). Garrett, M. F. (1975). 'The Analysis of Sentence Production.' In, G. H. Bower (Ed.). *The Psychology of Learning and Motivation*. New York: Academic Press. 133-177.
- (1988). 'Process in Sentence Production,' in F. Newmeyer (ed.) *The Cambridge Linguistics Survey*, iii. Cambridge: Cambridge University Press. 69-96.
- (1990). 'Sentence Processing.' In, D. Osherson and H. Lasnik (eds.) *An Invitation to Cognitive Science*, i/1: Cambridge, Massachussets: MIT Press. 133-75.
- (1991). *Language Selection and Switching in Strasbourg*. Oxford: Clarendon Press.
- Geertz, C. (1973). *The Interpretation of Cultures*. New York: Basic Books.
- George, A. (1990). 'Whose Language is it Anyway? Some Notes on Idiolects.' *The Philosophical Quarterly*. 40, 275-295.
- Gibbons, J. (1987). *Code-Mixing and Code Choice: A Hong Kong Case Study*. Clevedon: England.
- Gingras, R. (1974). 'Problems in the Description of Spanish-English Intrasentential Code-Switching.' In, G. D. Bills (ed.) *Southwest Areal Linguistics*, 167-74.
- Givon, T. (1979). *On Understanding Grammar*. New York: Academic Press.
- (1989). *Mind, Code, Context: Essays in Pragmatics*. Hillsdale, NJ: Erlbaum.
- Goffman, E. (1981). *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Gold, R. (1969). 'Roles in Sociological Field Observation.' In, G. J. McCall and J. L. Simmons (eds.) *Issues in Participant Observation*. Reading, Massachussets: Addison-Wesley. 30-39.

- Gonzo, S. & Saltarelli, M. (1983). 'Pidginization and Linguistic Change in Emigrant Languages.' In, R. Anderson (ed.) *Managing Linguistic Diversity*. Multilingual Matters. 181-97.
- Goodenough, W. H. (1956). 'Componential Analysis and the Study of Meaning.' *Language*. Linguistic Society of America. 32: 195-216.
- (1964). 'Cultural Anthropology and Linguistics.' In, D. Hymes (ed.) *Language in Cultures and Society: a Reader in Linguistics and Anthropology*, New York: Harper & Row. 36-39.
- Goodwin, C. (1993). 'Recording Human Interaction in Natural Settings.' *Pragmatics*. Oxford: Oxford University Press. 3(2): 181-209.
- Goodwin, C., and J. Heritage. (1990). 'Conversation Analysis.' *Annual Reviews of Anthropology*, 19: 283-307.
- Grabe, W. (1979). 'Morocco: Diglossia, Triglossia? The Language Situation in Morocco.' *Working Papers in Applied Linguistics*. Columbia University Press. 6: 11-21.
- Gravel, L. A. (1979). *A Sociolinguistic Investigation of Multilingualism in Morocco*. PhD Thesis. Columbia University, New York.
- Green, D. W. (1986). 'Control, Activation, and Resource: A Framework and a Model for the Control of Speech in Bilinguals.' *Brain and Language*. Elsevier. 27: 210-23.
- Greenberg, J. (1963). *Universals of Languages*. Cambridge, Massachussets: MIT Press.
- Grice, H. (1975). 'Logic and conversation.' In, P.Cole (Ed.), *Syntax and Semantics*, Vol. 3. New York: Academic Press. 41-58.
- (1989). *Studies in the Way of Words*. Cambridge, Massachussets: Harvard University Press.
- Grosjean, F. (1982). *Life with Two Languages*. Cambridge, Massachussets: Harvard University Press.
- (1988). 'Exploring the Recognition of Guest Words in Bilingual Speech.' *Language and Cognitive Progress*. Elsevier. 3: 233-74.
- (1995). 'A Psycholinguistic Approach to Code Switching: the Recognition of Guest Words by Bilinguals.' In, L. Milroy & P. Muysken, P. (eds.) *One Speaker, Two Languages*. Cambridge. 259-275.
- Grosjean, F. & C. Soares. (1986). 'Processing Mixed Language: Some Preliminary Findings.' In, J. Vaid (ed.), *Language Processing in Bilinguals: Psycholinguistic and Neurolinguistic Perspectives*. Hillsdale, NJ: Erlbaum. 145-79.

- Gumperz, J. J. (1964). 'Linguistic and Social Interaction in Two Communities.' *American Anthropologist*. Blackwell. 66(6): 137-53.
- (1968). Types of Linguistic Communities.' In, J.A. Fishman (ed.) *Readings in the Sociology of Language*. The Hague: Mouton. 460-72.
- (1971). *Language in Social Groups: Essays*. Stanford, California: Stanford University Press.
- (1976). 'The Sociolinguistic Significance of Conversational Code-Switching'. *Working Papers of the Language Research Laboratory 46*, University of California: Berkeley. 1-46.
- (1982a). *Discourse Strategies*. Cambridge: Cambridge University Press.
- (1982b). *Language and Social Identity*. Cambridge: Cambridge University Press.
- (1996). 'The Linguistic and Cultural Relativity of Conversational Inference.' In, J. J. Gumperz and S. C. Levinson (eds.) *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press. 374-407.
- (1982). 'Conversational Code-Switching,' in *Discourse Strategies*. Cambridge: Cambridge University Press. 55-99.
- Gumperz, J. J. & D. Hymes. (1972). *Directions in Sociolinguistics*. New York: Holt, Rinehart, and Winston.
- Gumperz, J. J., & E. Hernández-Chavez. (1975). Bilingualism, Bidialectalism, and Classroom Interaction. In C. Cazden, V. P. John, & D. Hymes (eds.) *Functions of Language in the Classroom*. New York: Teachers College Press. 84-107.
- Gumperz, J. J. & S. Levinson (eds.), (1996). *Rethinking Language Relativity*. Cambridge: Cambridge University Press.
- Gyori, G. (2005). 'The Adaptive Nature of 'Meaning' as Understanding.' In, *Acta Linguistica Hungarica*. Vol. 52. Akademiai Kiado. (2-3): 199-220.
- Haas, M. R. (1978). *Language, Culture, and History: Essays*. Stanford, California: Stanford University Press.
- Halfpenny, P. (1996). 'The Relationship Between Qualitative and Quantitative Research: Paradigm Loyalty Versus Methodological Eclecticism.' In, J.E.T. Richardson (ed.) *Handbook of Qualitative Research Methods for Psychology and the Social Sciences*. Leicester: British Psychological Society. 159-72.
- Halmari, H. (1997). *Government and Codeswitching: Explaining American Finnish*. Amsterdam/Philadelphia: John Benjamins.
- Hamers, J. & M. Blanc. (1989). *Bilinguality and Bilingualism*. Cambridge: Cambridge University Press.

- Hankamer, J. (1989). Morphological Parsing and the Lexicon.' In, W. D. Marlsen-Wilson (ed.) *Lexical Representation and Process*. Cambridge, Massachussets: MIT Press. 392-408.
- Harrell, R. S. (2004) *A Short Reference Grammar of Moroccan Arabic*. Georgetown University Press.
- Hasselmo, N. (1972). 'Code-switching as Ordered Selection.' In, Evelyn Sherabon Firchow, Kaaren Grimstad, Nils Hasselmo, and Wayne A. O'Neil (eds.), *Studies for Einar Haugen Presented by Friends and Colleagues*, The Hague and Paris: Mouton. 261-280.
- Haugen, E. (1950). 'The Analysis of Linguistic Borrowing.' *Language*. Washington DC: The Society. 26: 210-31.
- (1956). *Bilingualism in the Americas. A Bibliography and Research Guide*. Alabama: University of Alabama Press.
- (1973). 'Bilingualism, Language Contact and Immigrant Languages in the United States: A Research Report 1956-1970.' In, T.A. Sebeok (ed.) *Current Trends in Linguistics*. The Hague and Paris. 505-591.
- (1973). 'The Analysis of Linguistic Borrowing,' in E. Haughen (ed.) *The Ecology of Language*. Stanford, California: Stanford University Press. 80-109.
- (1992). 'Borrowing', in W. Bright (ed.) *Oxford Encyclopaedia of Linguistics*, i. New York: Oxford University Press. 197-79.
- Hawkins, R. & Y. H. Chan. (1997). 'The Partial Availability of Universal Grammar in Second Language Acquisition: The 'Failed Functional Feature Hypothesis.' *Second Language Research*, 13(3). 187-226.
- Heath, J. (1987). *Ablaut and Ambiguity: Phonology of a Moroccan Arabic Dialect*. Albany: State University of New York Press.
- (1989). *From Code-Switching to Borrowing: Foreign and Diglossic Mixing in Moroccan Arabic*. London: Kegan Paul Intenational.
- (1989). 'A Case of Intensive Lexical Diffusion: Arnhem Land, Australia.' *Language*. Washington DC: The Society. 57: 335-67
- (2002). *Jewish and Muslim Dialects of Moroccan Arabic*. London: New York: Routledge: Curzon.
- Heller, M. (1982). 'Negotiation of Language Choice in Montreal.' In, J. J. Gumperz (ed.) *Language and Social Identity*. Cambridge: Cambridge University Press.
- (1988). *Code-Switching: Anthropological and Sociolinguistic Perspectives*. Berlin: Mouton de Gruyter.

- (1995). 'Language Choice, Social Institutions, and Symbolic Domination.' *Language in Society*. Cambridge: Cambridge University Press. 24(3): 373-405.
- Henton, C. (1995). 'Cross-Language Variation in the Vowels of Female and Male Speakers.' *Proc. 13 ICPHS*. Stockholm. Vol. 4: 420-3.
- Hill, J. H. & K. C. Hill. (1986). *Speaking Mexican*. Tucson, University of Arizona Press.
- Hitchcock, G. & D. Hughes. (1989). *Research and the Teacher: A Qualitative Introduction to School-based Research*. London: Routledge.
- Hoffman, B. G. (1967). *The Structure of Traditional Moroccan Rural Society*. The Hague: Mouton.
- Hockett, Charles, F. (1958). *A Course in Modern Linguistics*. New York: Macmillan.
- Holes, C. (1993). 'Perspectives on Arabic linguistics V.' *Papers from the Fifth Annual Symposium on Arabic Linguistics*. Amsterdam: John Benjamins Publishing Company.
- (1995). *Modern Arabic: Structures, Functions and Varieties*. Longman, London.
- (2000). *Dialect, Culture, and Society in Eastern Arabia*. Leiden: Brill.
- Hollan, D. (1992). 'Cross-Cultural Differences in the Self.' *Journal of Anthropological Research*. Albuquerque, NM: University of New Mexico Press. 48(4): 283-300.
- Hornstein, N., Nunes, H. & Grohmann, K. (2005). *Understanding Minimalism*. Cambridge: Cambridge University Press.
- Hussein, R. F. (1999). 'Code-Alteration Amongst Arab College Students.' *World Englishes*. London: Routledge. Vol. 18(2): 281-289.
- Hymes, D. (1972). 'Models of Interaction of Language and Social Life.' *Directions in Sociolinguistics: The Ethnography of Communication*. New York: Holt, Rinehart and Winston
- (1974a). *Foundations in Sociolinguistics: An Ehtnographic Approach*. Philadelphia: University of Pennsylvania Press.
- Ingham, B. (1994). *Najdi Arabic Central Arabian*. Amsterdam / Phildaelphia: John Benjamins.
- Jackendoff, R. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge, Massachussets: MIT Press.
- (1977). 'The Social Implications of Intrasentential Code-Switching.' In, R. Romo and R. Paredes (eds.) *New Directions in Chicano Scholarship* (special edition of *The New Scholar*). University of California at San Diego. 227-56.

- Jakendoff, R. (1977). *X-Bar Syntax: A Study of Phrase Structure*. Cambridge, Massachusetts: MIT Press.
- (1987). 'The Status of Thematic Relations in Linguistic Theory.' *Linguistic Inquiry*. Walter de Gruyter. Vol. 18: 369-411.
- (1990). 'Socioeconomic Status as a Factor in the Selection of Encoding Strategies in Mixed Discourse.' In, R. Jacobson (ed.) *Code-Switching as a Worldwide Phenomenon*. New York: Peter Lang. 111-40.
- Jake, J. L. (1994). 'Intrasentential Codeswitching and Pronouns: On the Categorical Status of Functional Elements.' *Linguistics*, 32. 271-98.
- Jake, J. L. & C. Myers-Scotton. (1992). 'Testing the Fit: Syntactic Theory and Intrasentential Codeswitching,' Paper presented at Annual Meeting of the Linguistic Society of America, January 1992.
- Jakobson, R. & M. Halle. (1956). *Fundamentals of Language*. The Hague: Mouton.
- Jaspert, K. & S. Kroon. (1989). 'Social Determinants of Language Loss,' *ITL Review of Applied Linguistics*. New York: Harper & Row. 83/84: 75-98.
- Jorgensen, J. N. (1998). 'Children's Acquisition of Code-Switching for Power Wielding.' In: Auer, P. (Ed.). *Code-Switching in Conversation: Language, Interaction and Identity*. Routledge, London. 237-258.
- Joshi, A. (1981). 'Some Problems in Processing Sentences with Intrasentential Code Switching.' Paper presented at University of Texas Parsing Workshop, Austin.
- (1983). 'How Much Context-Sensitivity is Necessary for Characterizing Structural Descriptions - Tree Adjoining Grammars.' In, D. Dowty, L. Karttunen, A. Zwicky (eds.) *Natural Language Processing - Theoretical, Computational, and Psychological Perspectives*. Cambridge University Press.
- (1985). 'Processing of Sentences with Intrasentential Code Switching', in D. R. Dowty, L. Karttunen & A. Zwicky (eds.) *Natural Language Parsing*. Cambridge: Cambridge University Press. 190-205.
- Kachru, B. B. (1978). 'Code-Mixing as a Communicative Strategy in India.' In, J. E. Alatis (ed.) *International Dimensions of Bilingual Education*. Washington, DC: Georgetown University Press. 107-24.
- (1982). 'The Bilingual's Linguistic Repertoire.' In, A. Valdman, B. Hartford & C. Foster (eds.) *Issues in International Bilingual Education*. Plenum: New York. 25-52.
- Kamwangamalu, N. M. (1989a). 'Theory and Method of Code-Mixing: A Cross Linguistic Study.' PhD Thesis. University of Illinois, Urbana.

- (1989b). 'Some Morphosyntactic Aspects of French/English-Bantu Code-Mixing', in B. Music et al. (eds.) *Parasession on Language Contact: Regional Meeting of the Chicago Linguistic Society*. Chicago: CLS. 157-70.
- (1990). '*Bilingual Code-Mixing: Syntax and Constraints*.' Revised version of paper given at African Linguistics Conference, University of Georgia, Athens, April 1990.
- Kamwangamalu, N. M. & C. L. Lee. (1991). 'Mixers and Mixing: English Across Cultures.' *World Englishes*. London: Routledge. 10: 247-61.
- Kapchan, D. A. (1996). *Gender on the Market: Moroccan Women and the Revoicing of Tradition*. Philadelphia: University of Philadelphia Press.
- Katan, Y. (1990). *Oujda, Une Ville Frontière du Maroc (1907-1956): Musulmans, Juifs et Chrétiens en Milieu Colonial*. Paris: L'Harmattan.
- Kaye, A. (1994). 'Formal vs. Informal in Arabic: Diglossia, Triglossia, Tetraglossia, etc. Polyglossia-Multiglossia Views as a Continuum.' *Zeitschrift für Arabische Linguistik*. BRILL. 20: 47-66.
- Keegan, J. M. (1986). '*The Phonology and Morphology of Moroccan Arabic*.' PhD Thesis. City University of New York.
- Kempen, G. & Hoenkamp, E. (1987). 'An Incremental Procedural Grammar for Sentence Formulation.' *Cognitive Science*. Taylor & Francis. 11: 201-58.
- Kendon, A. (1977). *Studies in the Behavior of Social Interaction*. Lisse: The Peter de Ridder Press.
- (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters*. Cambridge: Cambridge University Press.
- Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Cambridge: Blackwell.
- Kifleyesus, A. (2002). 'Muslims and Meals: The Social and Symbolic Function in Changing Socio-Economic Environments.' *Africa*. Oxford: Oxford University Press. Vol. 72(2): 245.
- Kiraz, G. (1994). *Computational Analyses of Arabic Morphology*. PhD Thesis. University of Cambridge.
- Klavans, J. L. (1983). 'The Syntax of Code-Switching: Spanish and English.' *Proceedings of the Linguistic Colloquium on Romance Languages, xiv*. Amsterdam: Benjamins.
- (1985). 'The Syntax of Code-Switching: Spanish and English.' In L.D. King & C.A. Maley (eds.). *Selected Papers from the XIIIth Linguistic Symposium on Romance Languages*. Amsterdam: Benjamins. 213-21.
- Laala Hafdane, H. (2003). *Les Femmes Marocaines: Une Société en Movement*. Paris: L'Harmattan.

- Labov, W. (1966). *The Social Stratification of English in New York City*. Arlington: Center for Applied Linguistics.
- (1971). 'Methodology.' In, W. Dingwall (ed.), *A Survey of Linguistic Science*. College Park, MD: University of Maryland. 412-497.
- (1972). *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.
- (1977). *Language in the Inner City: Studies in the Black English Vernacular*. Oxford: Blackwell.
- (1984). 'Field Methods of the Project on Linguistic Change and Variation.' In, J. Baugh & J. Sherzer (eds.). *Language in Use: Readings in Sociolinguistics*, 28-53. Englewood Cliffs, NJ: Prentice Hall.
- (1994). *Principles of Linguistic Change*. Oxford: Blackwell.
- Lagnaoui, A. (1999). *Aux Frontières de la Tradition Marocaine*. Casablanca, Maroc: Editions Toubkal.
- Lahlou, M. (1989). 'Arabic-French Codeswitching in Morocco.' Paper presented at African Linguistics Conference, University of Illinois, Urbana.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*. University of Chicago Press.
- Lakoff, R. T. (1975). *Language and Women's Place*. New York: Harper & Row.
- Lambert, R. (1989). 'Language Attrition,' *ITL Review of Applied Linguistics*. Institute Toegepaster Lingvistiek. 83/84: 1-18.
- Lambert, W. E. (1975) 'Culture and Language as Factors in Learning and Education.' In, A. Wolfgang (ed.) *Education of Immigrant Students*. Toronto: Ontario Institute for Studies in Education.
- Lane, E. W. (1836). *The Manners and Customs of the Modern Egyptians*. London: K Baedeker.
- Levelt, W. J. M. (1989). *Speaking from Intention to Articulation*. Cambridge, Massachussets: MIT Press.
- Levin, A. (1998). *Arabic Linguistic Thought and Dialectology*. Jerusalem: Max Schloessinger Memorial Foundation. Hebrew University.
- Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Levi-Strauss, C. (1963). *Structural Anthropology*. New York: Basic Books.
- (1978). *Myth and Meaning*. New York: Schocken Books.
- Lipski, J. (1977). 'Code-Switching and the Problem of Bilingual Competence.' In, M. Paradis (ed.). *Aspects of Bilingualism*. Columbia, SC: Hornbeam Press. 250-63.

- (1982). 'Bilingual Language-Switching: Models and Methods.' *Bilingual Review*. 9. 191-212.
- Liu, H., Bates, E. & P. Li. (1992). 'Sentence Interpretation in Bilingual Speakers of English and Chinese'. *Applied Psycholinguistics*. Cambridge: Cambridge University Press. Vol. 13: 451-484.
- Loebel, H. (1990). 'Framing Sentences.' *Cognition*. London: Blackwell. 35:1-39.
- Ludi, G. (1986). 'Forms and Functions of Bilingual Speech in Pluricultural Migrant Communities in Switzerland,' in J.A. Fishman et al. (eds.) *The Fergusonian Impact*, ii. Berlin: Mouton de Gruyter. 217-36.
- Mackey, W. (1968). 'The Description of Bilingualism.' In J. A. Fishman (ed.) *Readings in the Sociology of Language*. The Hague: Mouton.
- MacSwan, J. (1999). *A Minimalist Approach to Intrasentential Codeswitching*. New York: Garland Press.
- (2000). 'The Architecture of the Bilingual Language Faculty: Evidence from Codeswitching.' *Bilingualism: Language and Cognition*, 3(1), 37-54
- (2005). 'Comments on Jake, Myers-Scotton and Gross's Response: There is no Matrix language.' *Bilingualism: Language and Cognition*, 8(2). 277-284.
- MacWhinney, N. & Bates, E. (eds.) (1989). *The Cross-Linguistic Study of Sentence Processing*. New York: Cambridge University Press.
- Maher, J. (1991). 'A Cross-Linguistic Study of Language Contact and Language Attrition,' In, H. Seliger and R. Vago (eds.). *Change and Language*. Multilingual Matters. 67-84.
- Mahootian, S. (1993). *A Null Theory of Code Switching*. PhD Thesis. North Western University.
- Marini, A. (1993). *Dalil al-Mar'ah al-Maghribiyah*. al-Rabat: Dar Nashr al-Ma`rifah.
- Maykut, P. & R. Morehoiuse. (1994). *Beginning Qualitative Research - A Philosophic and Practical Guide*. London: Falmer Press.
- McClure, E. (1981). 'Formal and Functional Aspects of the Code-Switched Discourse of Bilingual Children.' In, R. Duran (ed.) *Latino Language and Communicative Behavior*. Norwood, NJ: Abex. 69-92.
- (1998). 'The Relationship Between Form and Function in Written National Language - English Codeswitching: Evidence from Mexico, Spain and Bulgaria.' In, Jacobson, Rodolfo (Ed.) *Codeswitching Worldwide*. Berlin, New York: Mouton de Gruyter. 125-150.
- McDonald, J. (1987). 'Sentence Interpretation in Bilingual Speakers of English and Dutch.' *Applied Psycholinguistics*. Cambridge: Cambridge University Press. Vol. 8: 379-413.

- Menyhart, K. (2002). 'The Temporal Organisation of Speech in Monolingual and Bilingual Children.' In, *Acta Linguistica Hungarica*. Institut. Vol. 49(304): 347-362.
- Mernissi, F. (1987). *Beyond the Veil: Male-Female Dynamics in Modern Muslim Society*. Bloomington: Indiana University Press.
- (1988). *Doing Daily Battle: Interviews with Moroccan Women*. London: Women's Press.
- Milroy, J. (1980). *Language and Social Networks*. Baltimore: University Park Press.
- Mohamed, F. (1983). *Arabic-English Code-switching in the Speech of a six-year-old Child*. Prince & Pintzuk.
- Moravcsik, E. (1978). 'Language Contact.' In, J.H. Greenberg (ed.) *Universals of Human Language*, i. Stanford, California Stanford University Press. 93-122.
- Mougeon, R., and E. Beniak. (1987). 'The Extralinguistic Correlates of Core Lexical Borrowing.' In, K. Denning et al. (eds.) *Variation in Language: New Ways of Analyzing Variation*, xv. Stanford, California: Stanford University Press. 337-47.
- (1991). *Linguistic Consequences of Language Contact and Restriction* (Oxford Studies in Language Contact). Oxford: Clarendon Press.
- Moyer, M. (1995). 'On Defining Matrix Language in Sentential Code-switching: a Syntactic Approach.' In, *Summer School Code-Switching and Language Contact*. 192-204.
- Mühlhäusler, P. (1985). 'Patterns of Contact, Mixture, Creation and Nativization: Their Contribution to a General Theory of Language.' In, C. J. Bailey, and R. Harris (eds.) *Developmental Mechanisms of Language*. Oxford: Pergamon. 51-87.
- Muysken, P. (1990). 'Concepts, Methodology, and Data in Language Contact Research: Ten Remarks from the Perspective of Grammatical Theory.' In, *Papers for the Workshop Concepts, Methodology and Data*. Strasbourg: European Science Foundation. 15-30.
- (2000). *Bilingual Speech: A Typology of Code Mixing*. Cambridge: Cambridge University Press.
- Muysken, P. & L. Milroy. (1995). 1995. *One Speaker, Two languages: Cross-Disciplinary Perspectives on Code-Switching*. New York, NY: Cambridge University Press.
- Myers-Scotton, C. (1982a). 'The Possibility of Code-Switching: Motivation for Maintaining Multilingualism.' *Anthropological Linguistics*. Bloomington, IN: University of Indiana. Vol. 14: 432-44.

- (1983). 'The Negation of Identities in Conversation: A Theory of Markedness and Code Choice,' *International Journal of the Sociology of Language*. Berlin: Mouton de Gruyter. Vol. 44: 115-36.
- (1986). 'Codeswitching and Diglossia.' In, J. A. Fishman et al. (eds.) *The Fergusonian Impact*. ii Berlin: Mouton de Gruyter. 403-15.
- (1988a). 'Codeswitching as Indexical of Social Relationships.' In, M. Heller (ed.) *Codeswitching: Anthropological and Sociolinguistic Perspectives*. Berlin: Mouton de Gruyter. 151-86.
- (1988b). 'Code-Switching and Types of Multilingual Communities.' In, P. Lowenberg (ed.). *Language Spread and Language Policy*. Washington, DC: Georgetown University Press. 61-82.
- (1988c). 'Differentiating Borrowing from Code-Switching.' In, K. Ferrara et al. (eds.) *Linguistic Change and Contact: New Ways of Analyzing Variation*, xvi. Austin: University of Texas Press. 318-25.
- (1989). 'Code-Switching with English: Types of Switching, Types of Communities.' *World Englishes*. London: Routledge. Vol. 8: 333-46.
- (1991a). 'Intersections between Social Motivations and Structural Processing in Code-Switching.' In, *Papers for the Workshop on Constraints, Conditions and Models*. Strasbourg: European Science Foundation. 57-82.
- (1991b). 'Whither Code-Switching? Prospects for Cross-Field Collaboration: Production based-models of Code-Switching.' In, *Papers for the Symposium on Code-Switching in Bilingual Studies: Theory, Significance and Perspectives*. Strasbourg: European Science Foundation. 207-232
- (1992b). 'Comparing Codeswitching and Borrowing.' *Journal of Multilingual and Multicultural Development*. Multilingual Matters. 13: 19-39.
- (1992a). 'Constructing the Frame in Intrasentential Codeswitching.' *Multilingua*. Berlin: Mouton de Gruyter. Vol. 11: 101-27.
- (1992c). 'Codeswitching as a Mechanism of Deep Borrowing, Language Shift and Language Death.' In, M. Brenzinger (ed.) *Language Death in East Africa*. Berlin: Mouton de Gruyter. 31-58.
- (1993). 'English loans in Shona: Consequences for linguistic systems.' *International Journal of the Sociology of Language*. Vol. 100. Mouton. 125-48.
- (1993a) [1997]. *Duelling languages. Grammatical Structure in Codeswitching*. Clarendon Press: Oxford.
- (1993b [1997]). *Social Motivations for Codeswitching: Evidence from Africa*. Oxford: Oxford University Press.

- (1998) *Codes and Consequences: Choosing Linguistic Varieties*. Oxford: Oxford University Press.
- (2002). *Contact Linguistics: Bilingual Encounters and Grammatical Outcomes*. Oxford University Press.
- Myers-Scotton, C. & J. Okeju. (1973). 'Neighbors and Lexical Borrowings.' *Language*, 49:871-89.
- Myers-Scotton & W. Ury. (1977). 'Bilingual Strategies. The Social Function of Code Switching.' *Linguistics*. 19, 20-35.
- Myers-Scotton, C. & S. Azuma. (1990). 'A Frame-Based Process Model of Codeswitching.' In, M. Ziolkowski, K. Deaton, and M. Noske (eds.) *Proceedings from the 26<sup>th</sup> Regional Conference*. Chicago: CLS. 307-21.
- Myers-Scotton & J. Jake. (1995). 'Matching Lemmas in a Bilingual Language Competence and Production Model.' *Linguistics* 33(3). 981-1024.
- (2000a). 'Four Types of Morpheme: Evidence from Aphasia, Codeswitching and Second Language Acquisition.' *Linguistics* 38: 6. 1053-100.
- (2000b). 'Testing a Model of Morpheme Classification with Language Contact Data.' *International Journal of Bilingualism*, 4/1. (Special Issue).
- (2001). 'Explaining Aspects of Codeswitching and Their Implications.' In, Janet Nicol (ed.), *One Mind, Two Languages: Bilingual Language Processing*. Oxford: Blackwell. 84-116.
- (2008). 'Universal Structure in Code-switching and Bilingual Language Processing and Production.' In, B. Bullock, & A. Jacqueline Toribio. (eds.) *Handbook of Code-switching*. Cambridge: Cambridge University Press.
- Myers-Scotton, C, J. Jake & M. Okasha. (1996). 'Arabic and Constraints on Codeswitching.' In, Eid, M. & D. B. Parkinson (eds.), *Perspectives on Arabic Linguistics*, 9.
- Nartey, J. (1982). 'Code-Switching, Interference or Fadism? Language Use among Educated Ghanaians.' *Anthropological Linguistics*. Indiana University Press. Vol. 24/2: 183-92.
- Navracsics, J. (2002). 'Bilingual Semantic Representation and Lexical Access.' In *Acta Linguistica Hungarica*. Springer. Vol. 49(2): 225-247.
- Newmeyer, Fredrick J. 1986. *The Politics of Linguistics*. Chicago: University of Chicago Press.
- Nishimura, M. (1986). 'Intrasentential Code-Switching: the Case of Language Assignment.' In, J. Vaid (ed.). *Language Processing in Bilinguals: Psycholinguistic and Neuropsychological Perspectives*. Hillsdale, NJ: Erlbaum. 123-144.

- (1997). *Japanese/English Code-switching: Syntax and Pragmatics*. New York: Peter Lang.
- Nortier, J. (1990). *Dutch-Moroccan Arabic Code-switching Among Moroccans in the Netherlands*. Dordrecht: Foris Publications.
- Ouhalla, J. (1994). *Introducing Transformational Grammar: From Rules to Principles and Parameters*. London: Edward Arnold.
- Owens, J. (2005). 'Pre-diaspora Arabic: Dialects, statistics and historical reconstruction.' *Diachronica*. 22:2, 271-308.
- Pandharipande, R. (1990). 'Formal and Functional Constraints on Code-Mixing.' In, R. Jacobson (ed.), *Codeswitching as a Worldwide Phenomenon*. New York: Peter Lang. 15-32.
- Pandit, I. (1990). 'Grammaticality in Code-switching.' In, R. Jacobson (ed.), *Codeswitching as a Worldwide Phenomenon*. New York: Peter Lang. 33-69.
- Paradis, M. (ed.). (1978). *Aspects of Bilingualism*. Columbia, SC: Hornbeam Press.
- (1981). 'Neurolinguistic Organization of a Bilingualism.' *LACUS Forum*, 7: 486-494.
- (1987). *The Assessment of Bilingual Aphasia*. Hillside, NJ: Lawrence Erlbaum.
- Petersen, J. (1988). 'Word-Internal Code-Switching Constraints on a Child's Grammar.' *Linguistics*. Vol. 26:479-93.
- Pfaff, C. (1979). 'Constraints on Language Mixing: Intrasentential Code-Switching and Borrowing in Spanish/English.' *Language*, 55: 291-318.
- Pinker, S. (1994). *The Language Instinct: How the Mind Creates Language*. New York: William Morrow & Co.
- (2003). 'Language as an Adaptation to the Cognitive Niche.' In, M. Christiansen & S. Kirby (Eds.), *Language Evolution: States of the Art*. New York: Oxford University Press.
- Pleines, J. (ed.) (1990). *Maghreb Linguistics*. Rabat: Editions Okad.
- Poplack, S. (1980). 'Sometimes I'll Start a Sentence in Spanish y Termino en Español': Toward a Typology of Code-Switching.' *Linguistics*. Vol. 18: 581-618.
- (1981). 'Syntactic Structure and Social Function of Code-Switching.' In, R. Duran (ed.), *Latino Language and Communicative Behavior*. Norwood, NJ: Ablex. 169-84.
- (1987). 'Contrasting Patterns of Code-Switching in two Communities.' In, E. Wande, J. Anward, B. Nordberg, L. Steensland & M Thelander (Eds.), *Aspects of Bilingualism: proceedings from the Fourth Nordic Symposium on Bilingualism*. Uppsala: Borgstrom, Motala. 51-76.

- (1988a). 'Contrasting Patterns of Code-Switching in Two Communities.' In, M. Heller (ed.), *Codeswitching: Anthropological and Sociolinguistic Perspective*. Berlin: Mouton de Gruyter. 215-44.
- (1988b). 'Language Status and Language Accommodation Along a Linguistic Border.' In, P. Lowenberg (ed.) *Language Spread and Language Policy*. Washington, DC: Georgetown University Press. 90-118.
- (1990). 'Variation Theory and Language Contact: Conceptions, Methods and Data,' in *Papers for the Workshop on Concepts, Methodology and Data*. Strasbourg: European Science Foundation. 33-66.
- (2004). 'Code-switching Soziolinguistik.' In, U. Ammon, N. Dittmar, K.J, Mattheier & P. Trudgill (eds.) *An International Handbook of the Science of Language*. 2nd edition. Berlin: Walter de Gruyter.
- Poplack, S. & D. Sankoff. (1981). 'A Formal Grammar for Code-switching.' *Papers in Linguistics* 14 (1): 3-46.
- Poplack, S., S. Wheeler & A. Westwood. (1987). 'Distinguishing Language Contact Phenomena: Evidence from Finnish-English Bilingualism.' In, P. Lilius & M. Saari (ed.). *The Nordic Languages and Modern Linguistics*, 6. Helsinki: University of Helsinki Press. 33-563.
- Poplack, S. & C. Miller. (1988). 'The Social Correlates and Linguistic Processes of Lexical Borrowing and Assimilation.' *Linguistics*. Vol. 26: 47-104.
- Poplack, S. & D. Sankoff. (1988). 'Codeswitching.' In, A. Ulrich et al. (eds.) *Sociolinguistics/Soziolinguistik*. Berlin: Mouton de Gruyter. 1174-80.
- Poplack, S. & M. Meecham. (1995). 'Patterns of Language Mixture: Nominal Structure in Wolof-French and Fongbe-French Bilingual Discourse.' In, L. Milroy and P. Muysken (eds.) *One Speaker, Two Languages*. Cambridge: Cambridge University Press
- Poplack, S. & M. Meecham. (1998). 'Introduction: How Languages Fit Together in Code-mixing.' In, S. Poplack & M. Meecham (Guest Eds), *International Journal of Bilingualism - Instant Loans, Easy Conditions: The Productivity of Bilingual Borrowing*, 2: 127-139.
- Prince, E. & S. Pintzuk. (1984) *Bilingual Code-switching and the Open/Closed Class Distinction*. University of Pennsylvania.
- Radford, A. 1988. *Transformational Grammar*. Cambridge: Cambridge University Press.
- Rabinow, P. (1977). *Reflections on Fieldwork in Morocco*. Berkeley; London: University of California Press.
- Ramat, A. (1991). 'Code-Switching in Dialectal Communities: Effects on Language Shift.' *Papers for the Workshop on Impact and Consequences: Broader Considerations*. Strasbourg: European Science Foundation. 189-224.

- Reyes, R. (1974). *'Studies in Chicago Spanish.'* PhD Thesis. Cambridge, Massachussets: Harvard University Press.
- (1976). 'Language Mixing in Chicano Bilingual Speech.' In, J.D. Bowen & J. Ornstein (eds.) *Studies in Southwest Spanish.* Rowley, Massachussets: Newbury House. 182-188.
- Romaine, S. (1982). *Sociolinguistic Variation in Speech Communities.* London: Edward Arnold.
- (1995). *Bilingualism.* Oxford: Blackwell.
- Rosenhouse, J. (1984). *The Bedouin Arabic Dialects: General Problems and a Close Analysis of North Israel Bedouin Dialects.* Wiesbaden: Otto Harrassowitz.
- Rouchdy, A. (2002). 'Language Conflict and Identity: Arabic in the American Diaspora.' In, *Language Contact and Language Conflict in Arabic: Variations on a Sociolinguistic Theme.* New York: Routledge Curzon. 133-148.
- Ruiter, J. J. de. (1989). *Young Moroccans in the Netherlands: An Integral Approach to their Language Situation and Acquisition of Dutch.* PhD Thesis, University of Utrecht.
- Sadiqi, F. (1997). *Grammaire du Berbère.* L'Harmattan.
- (2003). *Women, Gender and Language in Morocco.* Leiden: Brill.
- Saeed, J. I. (2003). *Semantics.* Malden, MA; Oxford: Blackwell Publishing.
- Saffran, J. R., R. N. Aslin & E. L. Newport. (1996). *Science* 274. 1926-1928.
- Saib, J. (1988). 'Code-Mixing and Multilingual Competence in Morocco.' Paper presented at the Second Dutch-Moroccan Symposium, Amsterdam: Leiden, April 1988.
- Sankoff, D. & Poplack, S. (1981). 'A Formal Grammar for Code-Switching.' *Papers in Linguistics*, 14: 3-43.
- Sankoff, G. (1980). *The Social Life of Language.* Philadelphia: University of Pennsylvania Press.
- Sankoff, G. (1986). *Diversity and Diachrony* (ed.) Amsterdam : John Benjamin
- Santorini, B. & S. Mahootian. (1995). 'Code-switching and the Syntactic Status of Adnominal Adjectives.' *Lingua*, 96. 1-27.
- Sapir, E. (1949). 'The Status of Linguistics as a Science and Culture, Genuine & Spurious.' In, D. G. Mandelbaum (ed.) *Culture, Language & Personality.* London: University of California Press. 78-119.
- Sayed, A. A. (1981). *The Phonology of Moroccan Arabic: a Generative Phonological Approach.* PhD Thesis. Austin: University of Texas.

- Schegloff, E.A. (1991). 'Reflections on Talk and Social Structure.' In, D. Boden & D. H. Zimmerman (eds.) *Talk and Social Structure*. Berkeley and Los Angeles: University of California Press.
- Schieffelin, B. B. (1994). 'Code-Switching and Language Socialization: Some Probable Relationships.' In, J. Duchan, L.E. Hewitt & R. M. Sonnenmeier (eds.) *Pragmatics: From Theory to Therapy*. New York: Prentice Hall. 20-42.
- Schieffelin, B. B. & E. Ochs. (1986). *Language Socialization Across Cultures*. Cambridge: Cambridge University Press.
- Schwanenflugel, P.J. & M. Rey. (1986). 'Interlingual Semantic Facilitation: Evidence for a Common Representational System in the Bilingual Lexicon.' *Journal of Memory and Language*. Academic Press. Vol. 25: 605-18.
- Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Seliger, H. (1985). 'Primary Language Attrition in the Context of Other Language Loss and Mixing.' Manuscript. Department of Linguistics. City University of New York.
- Silva-Corvalán, C. (1994). *Language Contact and Change: Spanish in Los Angeles*. Oxford: Clarendon Press.
- Shaffer, D. (1977). 'The Place of Code-Switching in Linguistic Contact.' In, M. Paradis (ed.) *Aspects of Bilingualism*. Columbia, SC: Hornbeam Press. 265-74.
- Sheynin, H.Y. (2003). 'The Syntax of Spoken Arabic: A Comprehensive Study of Moroccan, Egyptian, Syrian and Kuwaiti Dialects.' In, *Studies in Language*. Amsterdam: John Benjamins. Vol. 27(3): 709-712.
- Slaoui, H. (1986). *Problèmes de Transitions de Langues Posés par le Discours Bilingue Arabe-Marocain/Français*. Thèse de Doctorat de 3ème cycle, Université de Paris VII.
- Soares, C. & F. Grosjean. (1984). 'Bilinguals in a Monolingual and a Bilingual Speech Mode: The Effect on Lexical Access.' *Memory and Cognition*. Psychonomic Society. Vol. 12: 390-6.
- Solivetti, Luigi, M. (1994). 'Family, Marriage and Divorce in a Hausa Community: A Sociological Model.' In, *Africa*. Edinburgh University Press. Vol. 64(2): 252.
- Sperber, D. & D. Wilson, (1995). *Relevance: Communication & Cognition*. 2<sup>nd</sup> Edition. Oxford: Blackwell.
- Sridhar, K. (1996). 'Societal Multilingualism.' In, McKeay, L., H. Horberger (eds.). *Sociolinguistics and Language Teaching*. Cambridge: Cambridge University Press. 47-70.
- Sridhar, S. N. & K. Sridhar. (1980). 'The Syntax and Psycholinguistics of Bilingual Code-Mixing.' *Canadian Journal of Psychology*. Sage. Vol. 34: 401-16.

- Stacey, M. (1969). *Methods of Social Research*. Oxford: Pergamon.
- Stemberger, J. P. (1985). 'An Interactive Activation Model of Language Production', in A. W. Ellis (ed.) *Progress in the Psychology of Language*. Norwood, NJ: Erlbaum. 143-86.
- Stevens, R. (1983). A Theoretical Model of the Language Learning/Language Teaching Process. *Working Papers On Bilingualism*, 11, 129-152.
- Taman, H. (1993). 'The Utilization of Syntactic, Semantic, and Pragmatic Cues of Assignment of Subject Role in Arabic.' *Applied Psycholinguistics*, 14: 299-317.
- Tannen, D. (1990). *You Just Don't Understand*. London: Virago.
- (1993a). *Gender and Conversational Interaction*. New York: Oxford University Press.
- (1994). *Gender and Discourse*. New York; Oxford: Oxford University Press
- Thomason, S. G. & Kaufman, T. (1975). 'Toward an Adequate Definition of Creolization.' Paper presented at the 1975 International Conference on Pidgins and Creoles: Honolulu.
- Timm, L. (1975). 'Spanish-English Code-Switching: El Porque y How-not-to.' *Romance Philology*. Vol. 28:473-82.
- Toribio, A. J. (2001). 'Accessing Bilingual Code-Switching Competence.' In, *International Journal of Bilingualism*. Vol. 5(4): 403-436.
- Tosi, A. (1984). *Immigration and Bilingual Education*. Oxford: Pergamon.
- Treffers-Daller, J. (1991a). 'Towards a Uniform Approach to Code-Switching and Borrowing.' In, *Papers for the Workshop on Constraints, Conditions and Models*. Strasbourg: European Science Foundation. 259-79.
- (1991b). 'French-Dutch Language Mixture in Brussels.' PhD Thesis. University of Amsterdam.
- (1994). 'Mixing Two Languages: French-Dutch Contact in a Comparative Perspective.' *Topics in Sociolinguistics*, 9. Berlin/New York: Mouton de Gruyter.
- Trudgill, P. (1983). *On Dialect: Social and Geographical Perspectives*. Oxford: Blackwell.
- (1994). *Dialects*. London: Routledge.
- Trudgill, P & J. K. Chambers. (1980). *Dialectology*. Cambridge: Cambridge University Press
- Tuc, H (2003). *Vietnamese-English Bilingualism*. Routledge: Curzon.

- Tucker, G.R., W.E. Lambert & A. A. Rigault. (1977). *The French Speaker's Skill with Grammatical Gender: An Example of Rule-Governed Behaviour*. The Hague: Mouton.
- Turner, B. A. (1994). 'Patterns of Crisis Behaviour: A Qualitative Inquiry.' In, A. Bryman and R.G. Burgess (eds.) *Analyzing Qualitative Data*. London: Routledge.
- Vaid, J. (1980). 'The Form and Function of Code-Mixing in Indian Films: The Case of Hindi and English.' *Indian Linguistics*, 41: 37-44.
- Van Coetsem, F. (1988). *Loan Phonology and Two Transfer Types in Language Contact*. Dordrecht: Foris.
- Veltman, C. (1983). *Language Shift in the United States*. New York: Mouton.
- Versteegh, K. (1984). 'Pidginization and Creolization: The Case of Arabic.' *Current Issues in Linguistic Theory*, 33. Amsterdam & Philadelphia: John Benjamins.
- (2001) *The Arabic Language*. Edinburgh University Press.
- Wacksler, R. (1999). 'Cross-Linguistic Evidence for Morphological Representation in the Mental Lexicon,' In *Brain and Language*. Elsevier. Vol. 68(1): 68-74.
- Webb, E. J., D. T. Campbell, R. D. Schwartz & L. Sechrest. (1966). *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago Illinois: Rand McNally.
- Weinreich, U. (1967). *Languages in Contact, Findings and Problems*. New York: Linguistic Circle of New York. First Edition 1953.
- Weinreich, U, W. Labov & M. Herzong. (1968). 'Empirical Foundations for a Theory of Language Change.' In, W. P. Lehmann & Y. Malkeil (Eds.). *Directions for Historical Linguistics: A Symposium*. Austin: University of Texas Press. 95-188.
- Wentz, J. & E. McClure. (1976). 'Ellipsis in Bilingual Discourse.' *Papers from the Regional Meeting of the Chicago Linguistic Society*. Chicago: CLS. 656-665.
- Whorf, B. L. (1956). 'The Relation of Habitual Thought & Behaviour to Language.' In, J. B. Carroll (ed.) *Language, Thought & Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge, Massachusets: MIT Press.
- (1956c). 'Linguistics as an Exact Science.' In, J.B. Carroll (ed.) *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*. Cambridge, Massachusets: MIT Press. 220-32.
- Wolfson, N. (1976). 'Speech Events and Natural Speech: Some Implications for Sociolinguistic Methodology.' *Language in Society*. Cambridge: Cambridge University Press. Vol. 5: 189-209.
- Woolford, E. (1983). 'Bilingual Code-Switching and Syntactic Theory.' *Linguistic Inquiry*, 14: 520-36.

- Wei, L. (1994). *Three Generations Two Language One Family: Language Choice and Language Shift in a Chinese Community in Britain*. Clevedon: Multilingual Matters.
- (2000) 'Dimensions of Bilingualism.' In, Wei, L. (ed.) *The Bilingualism Reader*. London: Routledge.
- Youssi, A. (1992) *Grammaire et Lexique de L'Arabe Marocain Moderne*. Casablanca: Wallada.
- Zeggaf, A. (1994). *L'interculturel au Maroc: Arts, Langues, Littératures et Traditions Populaires*. Casablanca: Afrique Orient.
- Zentella, A.C. (1981) 'Tá bien, You could Answer Me en Cualquieridioma: Puerto Rican Codeswitching in Bilingual Classrooms.' In, R.P. Durán (ed.), *Latino Language and Communicative Behavior*. Norwood, NJ: Ablex. 109-132.
- (1997). *Growing up Bilingual*. Maiden, MA: Blackwell.
- Ziamari, K. (2007). *Moroccan Arabic-French Codeswitching*. PhD Thesis. University of Meknes.
- Zurif, E. B. (1990). 'Language and the Brain,' in D. Osherson and H. Lasnick (eds.) *An Invitation to Cognitive Science, i*. 177-98. Cambridge, Massachussets: MIT Press.

Electronic Resources:

[http://en.wikipedia.org/wiki/Semitic\\_languages](http://en.wikipedia.org/wiki/Semitic_languages)

[http://www.intersolinc.com/newsletters/newsletter\\_32.htm](http://www.intersolinc.com/newsletters/newsletter_32.htm)

<http://ec.europa.eu/eurostat>

<http://www.ingentaconnect.com/content/oeed/16080157;jsessionid=23d9ajp1tfi9f.alexandra>

<http://www.migrationinformation.org/Profiles/display.cf.m?ID=339>

[http://www.eds-destatis.de/en/downloads/sif/nj\\_03\\_20.pdf](http://www.eds-destatis.de/en/downloads/sif/nj_03_20.pdf)

APPENDIX 1

SECTION A

PHASE 1 AND 2 –

LANGUAGE ATTITUDE SURVEY

Name: Interview #

1. Personal Details:

Place of birth  Morocco  UK  Other  
Parents profession in Morocco  Father: .....  Mother: .....  
Parents profession in the UK  Father: .....  Mother: .....

2. Level of Education

School  College/University  Qur'anic school  
Parents level of education  Father: .....  Mother: .....

3. Language Situation

Languages used at home (UK):

With parents  
 Darija  English  Berber varieties  French/Spanish

With siblings  
 Darija  English  Berber varieties  French/Spanish

Age when language was used? .....  
Do you consider yourself bilingual?  Yes  No

4. Which languages do you use outside home?

English  Darija  Other

5. Which languages do you prefer to use in general?

Darija  English  
Why .....

Which are you more confident speaking?  
 Darija  English  
Why .....

Which language do you prefer to:  
Watch a film:  Darija  English  
Chat with friends:  Darija  English

**6. Code switching – using both English and Darija varieties at the same time.**

Do you usually use the two languages at the same time in a conversation?

- Yes  No  Sometimes  Rarely

Do you usually switch from English to Darija or Darija to English?

- English to Darija
- Darija to English

**7. Attitudes to both languages**

Why do you switch from one language to another?

- Subject of conversation
- Incompetence in one of the languages
- Audience
- Family
- Other reason .....

**8. Language dominance**

Which language do normally switch from:

- Darija to English
- English to Darija

Why .....

**9. What do you think when people switch between the two languages**

- Well-skilled  Incompetent in either language(s)  Indifferent

Why .....

**10. Are you for or against using two languages at the same time**

- For, because .....
- Against, because .....

**11. Classify the following languages in order of importance**

- Darija 1/2/3  English 1/2/3  Modern Standard Arabic 1/2/3

**12. Future**

Which language do you class as most important to your daily life?

- English  Darija  Other

Why .....

Would you prefer to be taught in Arabic or English?

- English  Darija  Other

Why .....

Which language do you see as important for your future (children)?

- English  Darija  Other

## APPENDIX 2

### SECTION B

#### INTERVIEW - MOROCCAN-BORN MIGRANTS

Name:

Age:

Date of interview:

Interview No:

Interview length:

1. Date and place of birth.
2. Family make-up.
3. Schooling in Morocco. Level?
4. Date of migration to the UK.
5. Why did you choose the UK? What was the process like to come across?
6. What do you think of those that come across illegally? Do you feel sorry for them?
7. What were your first impressions when you first arrived? Who met you? Did you have somewhere to stay?
8. Did/do you have family here in the UK?
9. How did you cope with English? Did you have any previous knowledge of the language? Did anyone help you?
10. At that time, were there any organisations, mosques, *halal* shops around? If not, what did you do?
11. What was your first job here?
12. Did you get married in Morocco or in the UK? How did your partner cope?
13. Where do you live at the moment? Do you get on with the neighbours?
14. Have you had any difficulties since leaving Morocco?
15. What language do you speak at home? With your children?
16. How did you learn English? Through TV, friends, work?
17. Describe a typical day in your routine.
18. What do you do at weekends or when you are not working?
19. Do you feel there is a real sense of community amongst the Moroccans?
20. What do you think are the main problems Moroccans face here in the UK?
21. Do you agree with mixed marriages at all?
22. Where are you from in Morocco? Do you miss it? How often do you go back?
23. Do you ever consider going back to live there?
24. How does your life in the UK compare to your life in Morocco? Which is better for you?
25. Do you regret leaving Morocco?
26. Do you consider yourself more British than Moroccan?
27. How do you keep your Moroccan identity alive? Do you pass this onto your children?
28. Do you worry about their future in the UK? Their Arabic language?
29. Any future plans?

Notes:

---

---

### APPENDIX 3

#### SECTION C

#### INTERVIEW – BRITISH-BORN MOROCCANS

**Name:**

**Age:**

**Date of interview:**

**Interview No:**

**Interview length:**

1. Date and place of birth.
2. Family make-up.
3. Schooling in the UK. Level?
4. When did your parents migrate to the UK?
5. Do you think their life would have been different had they stayed?
6. Do you wish you were born in Morocco and lived there?
7. Do you consider yourself more British or Moroccan?
8. Do you have lots of Moroccan friends here or a social network?
9. Do you attend any Moroccan social functions, parties? Do you feel more Moroccan at these events?
10. Do you attend the local mosque or any other community events? Do they help the Moroccan community?
11. Do you feel there is a real sense of community amongst the Moroccans?
12. Are you married? Yes - Did you get married in Morocco or in the UK? How did your partner cope?
13. Are you married? No – would you want to marry someone from Morocco and bring them across?
14. Do you agree with mixed marriages at all?
15. What language do you speak at home? With parents? With Moroccan friends?
16. Describe a typical day.
17. What's your favourite film?
18. What do you do at weekends or when you are not working?
19. What do you think are the main problems Moroccans face here in the UK?
20. Do you think that second generation Moroccans are not doing as well as they should?
21. What are the main reasons for this?
22. Where are you from in Morocco? Do you miss it? How often do you go back?
23. Would you ever consider going back to live there?
24. What makes you Moroccan do you think? Your language, ethnic background?
25. Do you worry about the future in the UK? Loss of your Arabic language?
26. Any future plans?

Notes:

---

---

## APPENDIX 4

### DATA SET PHASE 1 – SELECTED STRUCTURED INTERVIEW

The below is a selection of transcribed and translated parts of data recorded by British Moroccans during Phase One of data collection.

Key: A: Interviewer  
B: Interviewee

INTERVIEW # [6] Ali

- 1 A: *fin fi Casa?*  
 2 B: *fi Derb Sultan..*  
 3 A: *wa ʔndak šhēl ʔā ila baqiya ʔamma?*  
 4 A: *iyih, baqi al ab dyali wa ixwān dyalli...*  
 5 A: *al ʔā la baqiya ʔamma..?*  
 5 B: *iyih*  
 6 A: *xwettet walu?*  
 7 B: *ih, zuj xwettet u kleʔa dyal xūt*  
 8 A: Are you the oldest, youngest?  
 9 B: No, I'm in the middle.  
 10: A: third, fourth?  
 11 B: I would say third...  
 12 A: *u fin grit fil magrib?*  
 13 B: *grit fi school local school..and college as well..u ʔndi nivea de baccalaureat u hbest feš xrajt li hnaya..*  
 14 A: *u taʔraf elglenziya, ul aʔrbiya, ul français..*  
 15 B: *eh.....*  
 16 A: *fuqes jiti...kunt xaddem fi Casa??*  
 17 B: *la, ma aʔmarni ma xdəmt..kunt kenaqra safi..u xdit aʔa:m break..hiya feš xrajt...kannu teydiru dek offer dyal inter-rail, wahed el ticket illi ketserra fiha lil Europe..u xdittha, hedik xrejt..jit li hnaya aʔjebni al hēl hnaya..u bqit hnaya...iyih..*  
 18 A: *Jiti birāsak?*  
 19 B: *Jit a rassi...alone..*  
 20 A: *Iwa? Taʔqel aʔla the day you came? Do you remember the day you came?*  
 21 B: *naʔqel aʔlih mazyen!*  
 22 A: What happened?  
 23 B: What happened..well the journey took like four days because I have to stop in Spain and in France as well..yeah, about four or five days..and kan surprise li wahed šahbi dyall li anna kunt nagulu raba ma ma njiš hetta aʔyit lu min Earl's Court station u gult lu, I'm here, yeah..u gelli ya bqa ʔamma I'm coming u fi dek saʔa kən xaddem fi Mayfair and within, what, ten or fifteen minutes he was there..safi min ʔamma, min baʔd usbuʔin, ihna ten hadru aʔliha bil aʔrbiya walla bi l'anglais..hedə huwwa al beht hada! Hedik two weeks later  
 24 A:

- 10 A: *yallah, zid..*  
 11 B: *fin kunna?* Yeah, that's how it starts, you know...After the hotel, I got married, I met someone, you know,  
 12 A: *Hna fi London?*  
 13 B: Yes...because I started going to school to be honest with you just for the visa, you know, not for English language. In order to extend your visa, you need to like, um, register with a school and that and they take care of your visa ..kind of thing but you..have to use a private...yeah..and that was for six month [sic]...yeah..and I got married like I said..
- 14 A: *zuwajti hna walla fil mağrib?*  
 15 B: *La, la hna..English girl..*  
 16 A: *Hm..ok*  
 17 B: For five years...got divorced and..had a break for a few years..  
 18 A: *šnu gellu walidik ʕla..?*  
 19 B: *walidiya waš aʕla zawwaj..*was a bit shock..the were like..I was young at that time..but main reason I did it to be honest with you..was just to get...because it was so expensive...to start paying for a silly...for the visa... that the only ways..kind of...what she, she..it was her idea..you know, she said what's the best way we can..because we started living together in Camden town..she moved in with me..and she...we got married...and that's it really..but when I spoke to the family..that explained everything..you know, and they said well..
- 20 A: *ʕndak maʕha šu wlēd?*  
 21 B: No, no, no, unfortunately, no  
 22 A: *u maʕ a zawja a ʔāniya?*  
 23 B: *ʕndi zuj*  
 24 A: *ki smiyethum?*  
 25 B: *Adam wa Iman*  
 26 A: *Tbark Allah. Yʕrfu al aʕrbiya?*  
 27 B: *yaʕrfu šwiya, yallah daxalnehum lil Islamic school now..yeah..*  
 28 A: *Fin?*  
 29 B: In Stanmore, where I live yeah...and but Iman she's only two so...but she will eventually go to nursery soon..
- 30 A: *ket walli lil mağrib? Ketwalliw l Casa?*  
 31 B: *eh, kullu šayf..*  
 32 A: *fi šayf..iwa?*  
 33 B: *Iwa,* we just stay and travel..and see the family like...you know, busy visiting families, and taht's it and have a break..that's nice..
- 34 A: *u nta xeddem hna?*  
 35 B: *eh, xeddem hna..kunt kenħal un restau..u wine bars fi City of London..eh, kunt ken-run his bars..restaurants, wine bars..and I just got fed up with it..it's too much and..I don't know...I was doing it for 18 years and I got fed up with it..and long hours, and now you've got family now so it's not like..when you're single..or when you've got kids..it's not fair..now we're starting this chauffeuring company..and waht's good about it because you can balance between work and the family kind of, you know..so and I'm enjoying it you know..*
- 36 A: *Gulli al routine dyallak?*  
 37 B: *Eh?*

- 38 A: *Gulli al routine dyallak?*
- 39 B: *Al routine..between now walla..? Routine, na fiq fi şabah u ken di a drari lil madrasa..ken naxuðhum lil madrasa..afterwards we have breakfast..with the family and afterwards pick him up again..do a bit of shopping and get ready for work..so I do like, late shifts, you know, because of the kids, take to school..u fi, bñel weekend..I take the family out, you know,..day out..shopping and everything..*
- 40 A: *Martak katexdem?*
- 41 B: *Non, ma texdemš..*
- 42 A: *Ma tebgis?*
- 43 B: *La, mağa al family it's difficult..yeah, I would love her to work..but you know, if you have kids, it starts getting, you know, like those...looking after the kids, how much is it going to cost..might as well she stay at home and look after them...and I believe, women's job, it depends on, what the best thing that woman can do..as like to her home, job..better than, you know..aħsən hejja..illi kenət al xedma dyalha fi dār dyalha dek aħsən šay lil mar'a..fhemti..*
- 44 A: *wa lil atfāl..*
- 45 B: *eh,,*
- 46 A: *Əndak ši şhābak fi London?*
- 47 B: *ah, Əndi..mağerba şha:bi..u geğ jina fi waqt wahed..kullu, kullu Əndu family dyallu..al waqt ma kaysmeħš beš kantlaqaw..walakin, kanħāwlu marra marra..ntaslu u kantlaqaw..wa waqt, I bağd al awqāt fi jumuəa..bağd salat kantlaqaw u namšiw naxurju natlaqaw..*
- 48 A: *taşalli??*
- 49 B: *kunt kån şalli fi Ladbroke Grove deba šwiya bağida ağliya kanmši... East London..*
- 50 A: *East London..*
- 51 B: *East London, ah..west London, nine, kan şalli Əmma u bağd al aħyən, fi Baker Street...*
- 52 A: *Kaynīn bezzef dyal mağerba fi Goldborne Road..*
- 53 B: *kaynīn al mağerba, iyih, iwa hedek, bñel gulti, al area dyalhum..waxa..ma kån jiš hna bezzef bi siraha*
- 54 A: *ağleš?*
- 55 B: *ma kan teši hejja...*
- 56 A: *ma tabğis??*
- 57 B: *Non, non, ma teši hejja...tet-offer dek al blasa..heya blasa dyal mağerba walakin me kan teši hejja...waš na gullik, ma..even if the shops, walla heða and the restaurants, meši ši hejja interesting..walla yxellik, maybe the wrong people there, I don't know..it should be a really nice area..and I don't know..what's there, you know..it's nice to meet friends there sometimes, you know, because of the mosque..it's just the gossip, you know, you don't get no help no more..*
- 58 A: *waš al mağerba ybğiw al gossip?*
- 59 B: *ybğiw al gossip, fin mšiti, eš kayen..u ma kan teši hejja xra lil musağd..ma kaysağduš binethum..şufi, ana ağla ma smağt..geğ dek al maħellet illi ağtaħhum lil mağerba..maš lil Council, nađuw ..mšew bağuhum peanut money..wa ağd deba bğew ybiğuhum wa yašriħum back...fhemti...*
- 60 A: *heði awwal marra nasməğ heði..*
- 61 B: *ih, wallah.. ağtaħhum..shop-a*

## APPENDIX 5

### DATA SET PHASE 2 - SELECTED OPEN CONVERSATIONS

The below is a selection of transcribed and translated parts of data recorded by British Moroccans during Phase 2 of data collection.

General Discussion # 1:

- 1 *na warikum wašta šrit al yaum..šrit waħəd al bargain, innit Mum, innit Mum..?*
- 2 *wallah a bentu ma aġqalt, yeah..ih, ih, sorry a bentu..wallah ranni....*
- 3 that's really cute...
- 4 that's beautiful..
- 5 *meši i cute, lovely..*
- 6 *min šritihum??*
- 7 expensive..
- 8 *šrit set..look*
- 9 that's really nice
- 10 fine bone china
- 11 that's really beautiful
- 12 *xfiġin i riša*
- 13 *heđuk nice for black coffee*
- 14 exactly, nice for coffee..beautiful..
- 15 *taġraf..very, very expensive heđuk, China....*
- 16 *me bġitakš tadi i xemsa walla ndihum for myself..*
- 17 *nadi xemsa, kifeš?*
- 18 you must be joking..
- 18 *zaġma ana nadi waħəd wa nti tadi xemsa...*
- 19 *wa humma ruħhum i xemsa..*
- 20 beautiful innit aba, really aġjbūni..
- 21 *aġndi setta dyal hađu, setta dyal saucers..wa hetta tisiyēt dyal gateaux, innit Mum...*
- 22 yeah..it's really nice, *bi šaħtak..*
- 23 *ysalmak, aba ana našri i quality..*
- 24 *heđi aġmrat li al dar i bi xurda..šhel ramina..*
- 25 *ħwijjet bħel hađu nice, bnedum yqđər ytfaraj fihum walla yšrəb fihum..*
- 26 Don't you feel embarrassed?
- 27 *warilu al jipa li šriti..*
- 28 *hm..xallit-ha Əamma...*
- 29 *šrit waħəd al jipa li..*
- 30 *li ruħak..*
- 31 that's not for you...
- 32 no, *šrit-ha li Faiza..*
- 33 *bešaħ taji geddi, I tried it on..innit ama, taji fenna..*
- 34 *ma mašitiš ġnd fulena al yaum?*
- 35 *kenat fi jamaġ*
- 36 *ntuma ma mašituš?*
- 37 *waš all day wa ħna maġ l l'auto MOT, maġ l town, maġ this, aġyit...*
- 38 *w huma ma ywasluš ħta zūj, w anti jebti al MOT maġ Ənāš walla l waħda..*
- 39 *al waħda..*
- 40 *al waħda u rbaġ kunna xrajna aġla kamertu..*

- 41 *al muhim, he nti qđiti swalhək.*
- 42 *ah, relieved..mđit li hađak al garage hđe BMW, you know, fil corner hđe Bowyers..nice garage...gelli can I see your log book, your service history...gult-lu I've got full BMW service history u bqe yđuf y-look u gelli at my car u gelli..I've never seen a BMW bhel heđi in excellent condition, gelli Endak kul ři excellent, gult-lu thank you very much, u gelli I will ring you and let you know..heđik l'auto řuft-ha aba it's nice..bassađ jerya řwiya, 84,000, u four years old..really nice*
- 43 *u zađma u heđi, yđriha ađlik?*
- 44 *kan gedi yrđd ađliya u ygulli řh yađđiwni fi my car.. bařađ ma derř tilifun yet..na xaliha Endi anyway, ndir-ha for sale..ila ma derř tilifun Monday, nasməđ..gedya na biđha anyway..hađak a rajul ma gelliř expensive walla, ma gelliř nothing..řhel fiha mileage, weř fiha MOT..nothing..walla service history, ma gelliř gelya bezzef..*
- 45 *mul garage?*
- 46 *Eh? No, hađak a rajul illi kan fi l motorway, al motorway man..*
- 47 *gellak ma jeryeř bəzzef u zina..u cute, meři whopper kbira u meři řđira..*
- 48 *yes, really nice..*
- 49 *heđik illi řufnāha, heđik silver one..really big*
- 50 *řət gbila bint Rashid řret wađəd al jeep really nice..weř dyalha, dyalha...*
- 51 *aba humma lebes ađlihum..*
- 52 *getlak gedya taji now?*
- 53 *yeah*
- 54 *that's all I need..*
- 55 *u ken taji bekri beř ywalliw bekri, hiya me tajiř hetta half past eight..*
- 

