



**DISCOURSE MARKERS AND CODE-SWITCHING:
ACADEMIC MEDICAL LECTURES IN SAUDI ARABIA
USING ENGLISH AS THE MEDIUM OF INSTRUCTION**

Manal A. Al Makoshi

For the degree of Doctor of Philosophy

Department of English Language and Applied Linguistics
School of English, Drama and American & Canadian Studies
College of Arts and Law

June 2014

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

This thesis is a corpus-based study of two spoken academic corpora in an English as the (foreign) medium of instruction (EMI) context. The first corpus is compiled of transcripts of academic lectures by non-native speakers (NNS) from an EMI medical college in Saudi Arabia. To compare the data, a second corpus is compiled of similar transcripts by native speakers (NS) taken from the British Academic Spoken English (BASE) corpus.

The first part of the research qualitatively and quantitatively investigates the use of English discourse markers (DMs) on two levels: Structural (e.g. *okay, so, because*) and Interactional (e.g. *okay?, I mean, any questions?*). Structural DMs are found to function frequently as Topic Initiators, Topic Developers, Summarizers, and Closers, and occur more frequently in NS lectures' discourse. Interactional DMs, which function as Confirmation Checks, Rephrasers and Elicitors, are found to occur more frequently in the NNS lectures. This thesis demonstrates that the uses of DMs by the NS and NNS lecturers are affected by discourse context, pedagogic goals, personal lecturing styles, interaction with students and the need to create a conducive learning environment.

The second part explores the use of Arabic discourse markers (ADMs) in the NNS lecture discourse on similar Structural and Interactional levels. Interactional ADMs (e.g. *ya3ni* {means}, *mafهوم?* {understood}) have a higher overall frequency than Structural ADMs (*fa* {so}, *laanu* {because}).

The third part of this thesis explores the pedagogical functions of English-Arabic code-switching (CS) in the NNS lectures. When the purpose of CS is to make meaning clearer and convey knowledge more efficiently, it is not a language barrier but an effective communicative strategy. The data shows that CS is used mainly in seven roles in the NNS lecture discourse: (1) solidarity, (2) reiteration, (3) elaboration, (4) topic, (5) elicitation, (6) checking comprehension and (7) classroom management.

ACKNOWLEDGMENTS

First and foremost, I would like to thank the King Saud bin AbdulAziz University for Health Sciences and the UK Saudi Arabian Cultural Bureau for the academic scholarship and the support for me to complete my PhD.

I must thank my supervisors, Paul Thompson and Crayton Walker, for their endless patience, extensive knowledge, continued support and good humor. I am especially grateful for their helpful comments, instruction and constant encouragement throughout the research and writing process. I am indebted to them both for everything they have taught me during my time at the University of Birmingham. I could not have asked or wished for better supervisors than Paul and Crayton and know that my success in this program and beyond is a product of their guidance.

I also wish to acknowledge the support of Professor Abdulrahman Abouammoh, a senior consultant to the Centre for Higher Education Research and Studies (CHERS) in the Saudi Ministry of Higher Education. I am grateful for his assistance in accessing information on higher education in Saudi Arabia and providing me with a copy of his co-edited book.

I am grateful to all those in King Saud University and King Saud bin AbdulAziz University for Health Sciences for their willingness to open their lecture halls and allow me an observer's peek into the academic discourse of medicine. Without the help of the students, lecturers and administration, this thesis would not have been possible.

A special thanks to Katie Brady Barker for her help proofreading my many drafts, to Salmah Yakoob and Yueh Lee for their advice and support and to Sheng Li for his brilliance helping me resolve my many computer related crashes and lost files. And to the English department's PG-Tips student community for providing a forum to develop, share and encourage research knowledge.

I would like to thank the many friends and the supportive community I was blessed to have; those who made the PhD process much more interesting and manageable than it otherwise would have been.

Lastly, I give my thanks and appreciation to my viva committee members, Professor Hilary Nesi and Dr. Nicholas Groom, for their comments and suggestions.

DEDICATION

This thesis is dedicated to my parents, siblings and their beautiful children. I am truly grateful and fortunate to have parents who not only encouraged me to complete my PhD, but were there throughout with advice, support and late night Facetime calls. I could not have come as far as I have without my parents, Olfat Fodah and Abdullah Al Makoshi, and I shall be forever grateful to them both.

I am absolutely lucky and grateful to have such extraordinary siblings. Thanks must be given first to my older sister, Ghadah, for her unconditional support, endless proofreading and humor (sarcastic and otherwise). Knowing you are just a phone call away has made this PhD process bearable, as well as being able to talk to my nephew Laith (a two-time heart transplant survivor), whose charm and innocence always brings a smile to my face.

My sister Amel, a fellow PhDer, knows this process is not easy and being in the same boat, was able to give great advice and encouragement. My nephew, Khalid AlHumaid, and niece, Jude AlHumaid, have my love for just being their cute, funny selves.

I am also thankful for the information and direction given by my brother, Ziyad, and sister, Latifah, both of whom are graduates of KSU medical college. Their humor and bluntness of advice was much appreciated and enjoyed.

Last but certainly not least, I thank my baby sister, Sarah, for the love, random calls, visits and entertainment that only she can provide. And most recently for sharing her experiences as she enters medical school.

I also dedicate this thesis in memory of my aunt, Lyla, my grandmother, Sakina and others I lost during my PhD program. You are all dearly missed and always in my thoughts and prayers.

TABLE OF CONTENTS

List of Abbreviations	x
Romanized Transcription Characters	xi
List of Figures	xii
List of Tables	xiv
List of Mind-Maps	xv
CHAPTER 1 INTRODUCTION	2
1.1 Background	2
1.2 Why Academic Medical Lectures?	3
1.3 Why Discourse Markers?	5
1.4 Why Code-Switching?	6
1.5 Medical Education	6
1.5.1 Medical Education in Britain	6
1.5.2 Medical Education in Saudi Arabia	7
1.6 The Study	10
1.6.1 Aims and Objectives	11
1.6.2 Significance of the Study	11
1.7 Summary of Thesis Structure	12
CHAPTER 2 DISCOURSE MARKERS	16
2.1 Introduction	16
2.2 Academic Discourse	16
2.2.1 The Lecture as a Genre	16
2.2.2 Lecturing Styles	18
2.2.3 Features in Lectures: Skills and Strategies	21
2.3 Corpus Linguistics	22
2.3.1 Corpora for Language Research	23
2.4 Discourse Markers	26
2.4.1 Discourse Markers: Terminology Reviewed	26
2.4.2 Characteristics & Definitions of Discourse Markers	28
2.4.3 Approach to DM Analysis in this Study	38
2.4.4 Functional Paradigms of Discourse Markers	39
2.5 Discourse Markers in Pedagogic Settings	50
2.5.1 Introduction	50
2.5.2 Comprehension of Lectures: The Case of Discourse Markers	51
2.5.3 DMs in Academic Discourse	54
2.6 Summary	59
CHAPTER 3 CODE-SWITCHING	61
3.1 Introduction	61
3.2 Purpose of the Study: The Case of Arabic Code-Switching and Discourse Markers	61
3.3 Background	61
3.4 Definitions of Code-Switching	63

3.5 Characteristics and Terminology	65
3.6 Theory of Code-Switching	67
3.7 Stance Towards Code-Switching	68
3.8 Roles of Code-Switching	70
3.8.1 Code-Switching in the Middle East	72
3.8.2 Code-Switching in Academic Discourse	74
3.8.3 Code-Switching with Discourse Markers	77
CHAPTER 4 METHODOLOGY	81
4.1 Introduction	81
4.2 The Study	81
4.2.1 Context of the Study	86
4.2.2 English as a Foreign Medium of Instruction	87
4.3 Corpus Approach	87
4.4 English, Medicine and Lectures	89
4.5 Academic Medical Discourse	90
4.5.1 Exploratory Study	90
4.5.2 Results of Exploratory Study	93
4.5.3 Transcription Style	94
4.5.4 Arabic Transcription Style	96
4.6 Corpus Selection	96
4.6.1 Saudi Academic Spoken Medical English (SASME) Corpus	96
4.6.2 British Academic Spoken English (BASE) Corpus	99
4.7 Corpus Description	102
4.7.1 SASME	102
4.7.2 BASME	103
4.8 Framework for Analysis	103
4.8.1 DM Framework Development	103
4.8.2 English Discourse Markers Classification	104
4.8.3 Interactional Discourse Markers	108
4.8.4 Inter-rater Agreement	112
4.8.5 Arabic Discourse Markers Classification	112
4.9 Conclusion	116
CHAPTER 5 ANALYSIS OF RESULTS I: STRUCTURAL DMS	119
5.1 Introduction	119
5.2 Results of Discourse Markers in SASME and BASME: An Overview	120
5.2.1 SASME Lectures	122
5.2.1.1 Non-native Arabic Lecturers	123
5.2.1.2 Non-Arab Lecturers	124
5.2.2 BASME lecturers	125
5.3 Structural Discourse Markers	128
5.4 Structural Discourse Markers in SASME	129
5.4.1 Topic Initiator (so, now, okay)	129

5.4.1.1 Topic Initiator ‘so’	129
5.4.1.2 Topic Initiator ‘now’	134
5.4.2 Topic Developer (so, because, 'cause, also).....	139
5.4.2.1 Topic Developer ‘so’.....	139
5.4.2.2 Topic Developer ‘because / ‘cause’	142
5.4.2.3 Topic Developer ‘also’	145
5.4.3 Summarizer (so, therefore)	148
5.4.3.1 Summarizer ‘so’.....	148
5.4.3.2 Summarizer ‘therefore’	151
5.4.4 Topic Closer (okay)	153
5.5 Structural Discourse Markers in BASME	156
5.5.1 Topic Initiator (so, now, okay).....	157
5.5.1.1 Topic Initiator ‘so’	157
5.5.1.2 Topic Initiator ‘now’	158
5.5.1.3 Topic Initiator ‘okay’	160
5.5.2 Topic Developer (so, because, ‘cause, also).....	162
5.5.2.1 Topic Developer ‘so’.....	163
5.5.2.2 Topic Developer ‘because / ‘cause’	164
5.5.2.3 Topic Developer ‘also’	167
5.5.3 Summarizer (so, therefore)	168
5.5.3.1 Summarizer ‘so’	169
5.5.3.2 Summarizer ‘therefore’	170
5.5.4 Topic Closer (okay)	172
5.6 Overview of Structural DMs in SASME and BASME	176
5.7 Results I: Structural Discourse Markers Chapter Summary.....	177
CHAPTER 6 ANALYSIS OF RESULTS II: INTERACTIONAL & ARABIC DMS _____	180
6.1 Introduction	180
6.2 Interactional Discourse Markers.....	181
6.3 Interactional DMs in SASME.....	181
6.3.1 Confirmation Checks (okay? right? clear?).....	181
6.3.1.2 Confirmation Checks ‘right?’	186
6.3.1.3 Confirmation Check ‘clear?’	189
6.3.2 Elicitors (So-prefaced question; Any questions?)	191
6.3.2.1 Elicitors ‘So-preface question’	191
6.3.2.2 Elicitors ‘any questions?’.....	195
6.3.3 Rephrasers (I mean; in other words).....	197
6.3.3.1 Rephraser ‘I mean’.....	198
6.3.3.2 Rephraser ‘In other words’	199
6.4 Interactional Discourse Markers in BASME.....	202
6.4.1 Confirmation Checks (okay? right? clear?).....	202
6.4.1.1 Confirmation Checks ‘okay?’	203

6.4.1.2 Confirmation Checks ‘right?’	205
6.4.1.3 Confirmation Checks ‘clear?’	205
6.4.2 Elicitors (so-prefaced question?; any questions?)	206
6.4.2.1 Elicitors ‘So-preface question’	207
6.4.2.2 Elicitors ‘Any questions?’	208
6.4.3 Rephrasers (I mean; in other words).....	210
6.4.3.1 Rephrasers ‘I mean’	211
6.4.3.2 Rephrasers ‘In other words’	212
6.5 Overview of Interactional DMs in SASME and BASME	214
6.6 Arabic Discourse Markers	215
6.6.1 Introductions	215
6.6.2 Code-switching and ADMs: Structural Discourse Markers	216
6.6.3 Code-switching and ADMs: Interactional Arabic Discourse Markers	223
6.6.4 Code-switching: Its Role in Academic Medical Discourse	228
6.6.5 Code-switching and ADMs: Summary	235
6.7 Results II: Interactional & Arabic DMs Chapter Summary	237
CHAPTER 7 DISCUSSION	241
7.1 Introduction	241
7.2 Comments on Structural Discourse Markers	242
7.2.1 Structure in Medical Lecture Discourse	242
7.2.2 The Lecturers’ ‘okay’ vs. ‘now’	244
7.2.3 The High Frequency Structural DM ‘so’	246
7.2.3.1 Use of ‘so’ as Topic Initiators and Topic Developers’	247
7.2.3.2 Use of ‘so’ vs. ‘therefore’	248
7.2.4 Topic Developer ‘also’	250
7.2.5 Formal ‘because’ vs. Non-Formal ‘cause’	251
7.3 Comments on Interactional Discourse Markers	252
7.3.1 Interaction in Medical Lectures	253
7.3.2 Lecturers’ Use of Confirmation Checks	257
7.3.3 The Use of Questions to Elicit Responses	258
7.3.4 Rephrasers: Native vs. Non-Native and the Use of ADMs	259
7.3.5 IDMs in Native and Non-Native Medical Lectures.....	260
7.4 Comments on Arabic Discourse Markers and Lecturers’ Code-Switching	263
7.4.1 ADMs in English Medium Lectures	263
7.4.2 Code-Switching in English Medium Lectures.....	266
CHAPTER 8 CONCLUSION	271
8.1 Introduction	271
8.2 Thesis Summary	272
8.3 Contribution in Relation to Research Findings	274
8.3.1 English DMs in Academic Medical Lectures	274
8.3.2 Arabic DMs in Academic Medical Lectures.....	276
8.3.3 Implications for Lectures	277

8.4 Implications of Code-Switching in Academic Lectures	278
8.5 Real World Implications of Research Findings	279
8.5.1 Pedagogic Implications	279
8.5.2 Post-Academia Implications.....	281
8.6 Limitations of the Study	282
8.7 Further Research.....	283
8.8 Concluding Remarks	284
LIST OF REFERENCES	286
APPENDICES	305
APPENDIX A: Informed Consent Letter	305
APPENDIX B: Exploratory Study	306
APPENDIX C: Methodology	312
APPENDIX D: Results.....	313

LIST OF ABBREVIATIONS

ADM	Arabic Discourse Marker
BASE	British Academic Spoken English (corpus)
BASEX	British Academic Spoken English Exploratory (sub-corpus)
BASME	British Academic Spoken Medical English (sub-corpus)
CS	Code-Switching
CL	Corpus Linguistics
DM	Discourse Marker
EL	Embedded language
ELF	English as a lingua franca
ESP	English for Specific Purposes
EMI	English Medium Instruction
S1	First segment (sentence or utterance)
KSA	Kingdom of Saudi Arabia
MWU	Multi-Word Unit
NS Arabic	Native Arabic speaker(s)
NS English	Native English speaker(s)
L1	Native or first language
NS	Native speaker(s)
NNS	Non-native speaker(s)
‰	Occurrences per thousand (normalization)
SAMEX	Saudi Academic Medical English Exploratory (corpus)
SASME	Saudi Academic Spoken Medical English (corpus)
L2	Second or other language
S2	Second segment (sentence or utterance)
Token	The individual units of a corpus (i.e. ‘words’)

ROMANIZED TRANSCRIPTION CHARACTERS

Arabic Letter	IPA	Transcription character	Example
ا	[a:]	a	<u>a</u> lso
ب	[b]	b	<u>b</u> oy
ت	[t]	t	<u>t</u> able
ث	[θ]	th	<u>th</u> at
ج	[dʒ]	j	<u>j</u> em
ح	[h]	ha	(emphatic h)
خ	[x]	kh	Scottish <u>Loch</u>
د	[d]	d	<u>d</u> ark
ذ	[ð]	dh	<u>th</u> en
ر	[r]	r	<u>R</u> amadan
ز	[z]	z	<u>z</u> ero
س	[s]	s	<u>s</u> ummer
ش	[ʃ]	sh	<u>sh</u> elf
ص	[sʰ]	s	ps <u>a</u> lm
ض	[dʰ]	dʰ	emphatic d in <u>d</u> ock
ط	[tʰ]	ty	<u>t</u> uck
ظ	[zʰ]	th	emphatic th in <u>th</u> is
ع	[ʕ]	ʕ	(no equivalent) guttural ‘a’
غ	[ɣ]	gh	(like French <i>r</i> in <u>Pa</u> ris)
ف	[f]	f	<u>f</u> ort
ق	[q]	q	<u>Q</u> uraan
ك	[k]	k	k as in <u>ca</u> rpet
ل	[l]	l	<u>l</u> eft
م	[m]	m	<u>m</u> elon
ن	[n]	n	<u>n</u> an
و	[w] / [u:]	wa / ou	<u>w</u> and / <u>w</u> ound
ي / ي	[j]	y / i	<u>y</u> ank / <u>coo</u> kie
ه / ه	[h]	ha	<u>h</u> alf

LIST OF FIGURES

Figure 4.1 Screenshot of Concordance Lines ‘right’	99
Figure 5.1 Overall Comparative Frequency of DM Functions ‰	120
Figure 5.2 Overall Comparative Frequency of Structural Discourse Markers ‰	121
Figure 5.3 Overall Comparative Frequency of Interactional DMs ‰	121
Figure 5.4 Overall DM Category Distribution ‰ across SASME	122
Figure 5.5 Non-Arab Lecturers’ DM Distribution ‰	124
Figure 5.6 Arab Lecturers’ DM Distribution ‰	125
Figure 5.7 BASME Overall DM Distribution across Lectures ‰	126
Figure 5.8 BASME Medical Lecturers’ DM Category Distribution ‰	127
Figure 5.9 BASME Biological Sciences Lecturers’ DM Category Distribution ‰	127
Figure 5.10. SASME Frequency ‰ of Topic Initiator ‘so’	130
Figure 5.11. SASME Frequency ‰ of Topic Initiator ‘now’	136
Figure 5.12 Distribution of Topic Initiator DMs ‰ across SASME	139
Figure 5.13 SASME Frequency ‰ of Topic Developer ‘so’	142
Figure 5.14 SASME Frequency ‰ of Topic Developer ‘because’	145
Figure 5.15 SASME Frequency ‰ of Topic Developer ‘also’	147
Figure 5.16 Summarizer Category Use ‰ across SASME Lecturers’	148
Figure 5.17 SASME Frequency ‰ of Summarizer ‘so’	151
Figure 5.18 SASME Frequency ‰ of Summarizer ‘therefore’	152
Figure 5.19 SASME Frequency ‰ of Closer ‘okay’	155
Figure 5.20. BASME Lecturers’ Structural DM Use ‰	156
Figure 5.21 Topic Initiator Use ‰ across BASME Lectures	157
Figure 5.22 Frequency ‰ of BASME Topic Initiator ‘now’	160
Figure 5.23 Comparative Frequency ‰ of Topic Initiator Category	162
Figure 5.24 Topic Developer Category Use ‰ across BASME	162
Figure 5.25 Comparative Frequency ‰ of Topic Developer Category	168
Figure 5.26 Summarizer Category Use ‰ across BASME Lecturers	169

Figure 5.27 Frequency ‰ of BASME Summarizer ‘therefore’	171
Figure 5.28 Comparative Frequency ‰ of Summarizer Category	172
Figure 5.29 Closer Category Use ‰ across BASME Lecturers	173
Figure 5.30. Comparative Frequency ‰ of Closer Category	175
Figure 6.1 Use of Confirmation Checks ‰ across SASME	182
Figure 6.2 SASME Frequency ‰ of Confirmation Check ‘okay’	185
Figure 6.3 SASME Frequency ‰ of Confirmation Check ‘right’	188
Figure 6.4 SASME Frequency ‰ of Confirmation Check ‘clear?’	190
Figure 6.5 Frequency ‰ of SASME Elicitation ‘So-Preface Question’	194
Figure 6.6 Frequency ‰ of SASME Category Elicitors	197
Figure 6.7 Frequency ‰ of SASME Category Rephrasers	201
Figure 6.8 BASME Lecturers’ Interactive DM Use ‰	202
Figure 6.9 Confirmation Check Category Use ‰ across BASME Lecturers	203
Figure 6.10. Use of Elicitor Function ‰ across BASME	206
Figure 6.11 Comparative Frequency ‰ of Elicitor Function	210
Figure 6.12 Rephraser Function Use ‰ across BASME	210
Figure 6.13 Comparative Frequency ‰ of Rephraser Category	213
Figure 7.1 Structural Discourse Marker Functions ‰	243
Figure 7.2 Topic Initiators ‘now’ and ‘okay’ ‰	244
Figure 7.3 Overall Frequency ‰ of Structural DM ‘so’	247
Figure 7.4 Dispersion Plot of Summarizer ‘therefore’ in SASME corpus	249
Figure 7.5 Dispersion Plots of DM ‘and’ in SASME and BASME corpus	251
Figure 7.6 SASME: No. Interactions vs. Lecture Duration vs. Total DM ‰	254
Figure 7.7 BASME: No. Interactions vs. Lecture Duration vs. Total DM ‰	255
Figure 7.8 Frequency ‰ of Confirmation Checks in the Corpora	257

LIST OF TABLES

Table 1.1 King Saud University Preparatory Year Program	8
Table 2.1 Dudley-Evans and Johns' (1981: 134) Academic Lecturing Styles	19
Table 2.2 Variety of 'DM' Terms	27
Table 2.3 Schiffrin's (1987) Initial Word Classes of Discourse Markers	35
Table 2.4 Bellés-Fortuño's (2006) Micro-Marker Category of DMs	44
Table 2.5 Schiffrin's (1987) DMs Multifunctionality on Planes of Discourse	46
Table 2.6 Discourse Marker Functional Categories (Fung and Carter, 2007: 418)	49
Table 2.7a Textual Functions based on Brinton (1998: 12)	49
Table 2.7b Interpersonal Functions based on Brinton (1998: 12)	50
Table 2.8a Metadiscoursal Phases Developed by Young (1994: 166-167)	54
Table 2.8b Non-Metadiscoursal Phases Developed by Young (1994: 166-167)	55
Table 3.1 Code-Switching Roles in CS Literature	71
Table 3.2 Primary Roles of Code-Switching (Abalhasan and Alshalawi, 2000: 184)	72
Table 4.1 Transcription Summary	94
Table 4.2 Tagging Method for SASME Lectures	95
Table 4.3 Breakdown of SASME Corpus Lectures	98
Table 4.4 Breakdown of BASME Corpus Lectures selected for Comparison	101
Table 4.5 Discourse Marker Classification Framework for Study	111
Table 4.6 Breakdown of SASME Corpus Arabic Tokens	113
Table 4.7 Arabic Discourse Marker Classification Framework	115
Table 4.8 Classification of Code-Switching Roles in SASME Lectures	116
Table 5.1 SASME Structural DMs Summary	176
Table 5.2 BASME Structural DMs Summary	176
Table 5.3 Structural DMs Overview	176
Table 6.1 SASME Interactional DMs Summary	214
Table 6.2 BASME Interactional DMs Summary	214
Table 6.3 Interactional DMs Overview	214

Table 7.1 Speaker-Hearer Exchanges in the NS and NNS corpora	256
Table 7.2 Multi-Word Units used by NS Lecturers of BASME	260
Table 7.3 Multi-Word Units used by NNS Lecturers of SASME	261
Table 7.4 Multi-Word Units used by only by NNS Lecturers of SASME	261

LIST OF MIND-MAPS

Mind-Map 4.1 Lectures included in Exploratory Corpora BASEX & SAMEX	92
Mind-Map 4.2 Functions of Structural Discourse Markers	105
Mind-Map 4.3 Functions of Interactional Discourse Markers	109
Mind-Map 5.1 Structural Discourse Markers	128
Mind-Map 6.1. Interactional Discourse Marker Functions	181
Mind-Map 6.2 Arabic Structural Discourse Markers by Function	217
Mind-Map 6.3 Arabic Interactional Discourse Markers by Function	224

CHAPTER 1
INTRODUCTION

“Words do not have meanings, meanings have words”

(Williams, 2003: 91)

Chapter 1 Introduction

1.1 Background

“To have English is to have access to the wealth of the world that is otherwise obscured behind linguistic barriers” (Seargeant, 2009: 8). This desire to access otherwise unattainable knowledge is often the driving force for individuals to obtain the linguistic tools needed “especially in such areas as science and technology” (Crystal, 2003: 110). The English language’s reach as *the* international language has created ripples that travel to nearly every corner of the globe and ensures its status as *lingua franca* - the common language of communication - in many facets of global society.

The demands placed on learners to acquire English skills goes hand in hand with its dominance globally as a medium of international communication in most political, economic, technological and scientific exchanges. This is seen by some as an inevitable consequence of economic globalization (Graddol and Meinhof, 1999: 1). According to Crystal (2003: 110), English is the chief foreign language being taught in schools in over a hundred countries including China, Germany, Egypt and Brazil; in many cases, displacing another language in the process and firmly strengthening its hold as the peoples’ common language. The globalization of English has become a *fait accompli* in the current century (Bellés-Fortuño, 2006) and now with at least four times as many non-native speakers (NNS) of English as there are native speakers (NS) (House, 2010: 363), the impact is gradually becoming clear.

In the field of education, “English has become the normal medium of instruction in higher education for many countries - including several where the language has no official status” (Crystal, 1999: 6). This adoption of English, often as the official medium, in NNS countries forces us to reflect on its effects on the academic discourse in lectures, which are “the central instructional activity of higher education institutions” (Bellés-Fortuño, 2006: 4). The communicative demands placed on both the speaker and learners during instruction at university are already cognitively demanding and high stakes (e.g., pass exams) without the added challenge

of a foreign medium of instruction. This is why research of academic English is important and one of the focuses of this study.

Another noticeable feature of English contact in foreign countries is the creation of a phenomenon referred to as code-switching (CS) in which “interactants rely on two or more languages to communicate with one another” (Crystal, 2003: 164). This phenomenon has made its way into the academic classroom drawing over time its supporters and opponents along the way. One of the main criticisms of using the native language (L1) in instructional discourse is that it prevents the creation of pure foreign language environments and is said to undermine the learning of the target language (see Chaudron, 1984; Chambers, 1991; Lightbown, 2001). Furthermore, critics often viewed CS as evidence of a gap in linguistic knowledge when, in fact, studies have begun to consider it a pragmatic strategy and a sign of a speaker’s abilities in more than one language (Myers-Scotton, 1997: 224). This strategy and creative use of a speaker’s linguistic resources to enable and support communication in academic discourse is also the focus of this study.

1.2 Why Academic Medical Lectures?

My interest in academic lectures is based on my involvement as an English language teacher in the King Saud bin Abdulaziz University for Health Sciences (KSAU-HS), where I first began working as a teaching assistant (TA) and following my masters degree, as an instructor. My experiences as a TA and as an instructor to both undergraduate and graduate students in their foundation English courses, now a part of the preparatory year program, demonstrated the importance of lectures for students.

Academic lectures are still the main method of transmitting information to students and play a “central educational role in many universities” (Swales, 2001: 34). As such, it is important that it is researched and investigated for both the benefit of teachers (faculty) and for the student body. The fact that many universities globally have adopted English as the (foreign) medium of

instruction (hereafter, EMI) at their medical schools only further supports the need for increased research into academic discourse.

For students, “understanding lectures is critical to academic success” (Olsen and Huckin, 1990: 33) and providing them with the tools to achieve the greatest success in their studies, and consequently their futures, is paramount. Some of the tools that students require are the development of listening comprehension skills and an increased awareness of lecture discourse. Considering English is used as the EMI in most university colleges in areas such as the sciences, medicine, engineering, allied health and other technical subjects (Al-Seghayer, 2012), improving medical students’ language skills should be foremost in the minds of educators.

Students in Saudi Arabia often report that one of the greatest difficulties they face is with English as the medium of instruction. Almoallim et al. (2010) reported that 62% of medical students in Saudi Arabia claimed they struggled to comprehend and write notes in lectures where they were dually challenged by new information and the EMI in which it was given. This struggle can also be explained by students’ deficiency of “basic learning skills such as listening and note-taking” and difficulty “keeping up with the required reading” (Almoallim et al., 2010: 1276). Furthermore, around 50% reported that they did not understand what they were supposed to have learnt by the end of some lessons and felt that the language barrier contributed to their lack of understanding (Almoallim et al., 2010).

Students in medical programs must absorb a great deal of information in a set amount of time from both written and spoken texts. In the case of written texts, students have the option of rereading the textbook, notes or even referencing other sources. Spoken texts, on the other hand, do not provide that same luxury. Students are under considerable stress to comprehend the spoken discourse of academic lectures in real time, while simultaneously following the often dense subject matter and writing notes. In order for students to derive the maximum intake from lectures, they must be able to perceive not only the overall structure, but also when direction is changed and new aspects of information are introduced (Smit, 2009). Recognition of linguistic discourse cues can play a role in the successful utilization of academic lectures by students.

1.3 Why Discourse Markers?

Crystal (1988: 48) describes discourse markers (DMs) as the “oil which helps us perform the complex task of spontaneous speech production and interaction smoothly and efficiently”. However, its importance in academic discourse is often overlooked and discounted by learners and speakers alike in academia. Studies such as Flowerdew and Tauroza (1995), Nattinger and DeCarrico (1988), Fung (2003) and Bellés-Fortuño (2006) among others, have supported the necessity of DM awareness for the successful coherence and retention of information.

More studies must be done on academic discourse, specifically those dealing with a foreign medium of instructions at the tertiary level. More focus should be placed on exploring the features of this discourse and the palpable, but often neglected, presence of DMs in academic lectures. As both a learner and a teacher, I view DMs as important linguistic cues that function like a road map in discourse, providing signals of what to expect next and which direction to follow. Without linguistic cues, connections between sentences or in the case of spoken discourse, utterances, can be ambiguous, which is especially problematic in academic discourse. So the question is how exactly do DMs function in academic discourse and do they function in a comparable manner across lectures in a similar field? Furthermore, do NS and NNS lecturers use DMs in the same manner? And, are the functions of DMs comparable with their English usage in the lectures?

In this research, I will attempt to answer such questions. To my knowledge, no other studies have explored the functions and uses of DMs in EMI medical academic lectures in Saudi Arabia, with NNS lecturers and a homogenous student body of NNS. More importantly, no research to my knowledge have taken into account the use of Arabic DMs, or code-switching itself, in the spoken discourse of EMI lectures and this too, is a gap in the field.

Andersson and Trudgill (1990) point out that DMs “may pass unnoticed for a long period, but once they are discovered (or, rather, brought to our awareness) we hear them all the time” (ibid: 94). The importance of DM awareness, for both lecturers and students, will be discussed and supported in this thesis. Furthermore, the results of this study are intended to demonstrate the value of DMs in spoken medical academic discourse.

1.4 Why Code-Switching?

Current literature of the pedagogic uses and functions of CS are mainly concerned with its role in English as a Second or Foreign language classrooms (ESL/EFL), where the subject matter is also the medium of instruction. My interest lies in the use of CS when the subject matter (e.g. medicine) differs from the (foreign) EMI.

International universities often have official regulations that strongly prohibit the use of the L1 in English as the Second or Foreign medium classrooms. Academic lecturers are also often led to believe that the inclusion of their native tongue in the classroom is indicative of their weakness in the language and their inability to communicate coherently in the L2. This research investigates its role in academic medical lectures and attempts to explore whether CS is a benefit or impediment in academic discourse.

1.5 Medical Education

In order to better understand the background of the academic medical lectures used in this study, a brief overview of medical programs in the UK and Saudi Arabia will be presented in the two following subsections.

1.5.1 Medical Education in Britain

There are 33 undergraduate medical schools in the UK and each vary with the admission process. For this reason, information regarding medical education in Britain will be given in respect to one school, the University of Leicester¹, from which the NS academic medical lectures were collected².

¹ Recordings of lectures were from the Leicester-Warwick Medical School, established in 2000 and separated to two individual institutions in 2007 (<http://www2.warwick.ac.uk/newsandevents/news/medicals10years/>)

² BASE corpus developed by Hilary Nesi and Paul Thompson (<http://www.coventry.ac.uk/research/research-directory/art-design/british-academic-spoken-english-corpus-base/contents/>)

According to the Leicester Medical Program (LMP), admission for international students requires that they complete the equivalent of four “AS levels including Biology (or Human Biology) & Chemistry and to continue three of these subjects, including Chemistry, to A2 level” (UoL, 2014). Alternatively, students may apply with the International Baccalaureate³ which requires “36 points overall, with three Higher Level subjects (including Chemistry and Biology), and three Standard Level subjects” (UoL, 2014).

International students, who have not completed their education in countries which LMP accepts being taught in the medium of English (no countries listed), are required to complete an English language test. Current admission requirements for the IELTS are 7.5 overall, with no component below 7.0 (UoL, 2014).

Regarding the program itself, the first phase (five semesters) of the five year Leicester Medical Program includes approximately 20 hours per week of contact teaching. Further stating that it would be “unusual for more than 10 hours to be taken up by lectures” (UoL, 2014). Students are also expected to undertake at least 18 hours of self-directed learning and progressively taught to develop their self study skills during the first phase of the program (ibid).

1.5.2 Medical Education in Saudi Arabia

In recent years (1997-2008), Saudi Arabia has increased the number of undergraduate medical schools from four to sixteen in an attempt to bridge the gap in the shortage of Saudi physicians, which in a recent consensus constituted less than 20% of physicians working in the country (Hamdy et al., 2010). While admission requirements are similar for medical schools in Saudi Arabia, the program details may differ and therefore, information regarding medical education in KSA will be given in respect to one program, the King Saud University medical college.

Admission to medical colleges was originally based solely on a student’s high school grades (GPA) where students with the top grades were selected and enrolled. Note that the Leicester Medical Program admission’s page specifically states that the KSA’s General Secondary

³ First authorized in 1999, there are now nine schools in Saudi Arabia which offer the International Baccalaureate (<https://www.ibo.org/country/SA/>)

Education Certificate is insufficient to make a competitive application to their program (UoL, 2014). In recognition of the weakness of this single variable selection process, a written science exam, English test and an interview were added; and in 2002, the National Center for Assessment in Higher Education was established. All Saudi high school graduates are now required to sit for a College Aptitude Test (CAT) before applying for admission to the universities. The total intake for all the medical colleges in Saudi Arabia is around 2500⁴ students per year. The average annual intake per college ranges from 80 to 300.

The average number of student intakes in the KSU medical college is approximately 300 students (200 males and a 100 females). During the first three years students are divided into four groups of 40 to 50 students and during the last two years of the medical program, they are divided up into 10 groups of 20 to 30 students. Prior to admission to the medical college, students are required to complete a preparatory year which they must pass with highly competitive grades (minimum cumulative GPA 4.5/5.0) to even be considered for acceptance into the program.

The preparatory year is a two semester foundation program similar to A level admission requirements by the Leicester Medical Program. Students in the preparatory year⁵ have the following subjects in English as the medium of instruction:

Semester 1	Credit hours	Semester 2	Credit hours
English	8	English (Medical Terminology)	8
Introduction to Mathematics	2	Organic Chemistry for Health Sciences	2
Communication Skills	2	General Physics	3
IT skills	3	Biostatistics	2
		General Biology	3
<i>Cumulative hours</i>	15	<i>Cumulative hours</i>	18

Table 1.1 King Saud University Preparatory Year Program

⁴ According to a commissioned report by the Higher Education Funding Council for England and the Department of Health, the total intake of pre-clinical medical students in England in 2011/2012 was 6,377 students. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213236/medical-and-dental-school-intakes.pdf

⁵ <http://ksu.edu.sa/en>

English language classes are only given explicitly in the first term of the preparatory year and are not given at any other point in the medical program. Students have eight credit hours of English language; however, depending upon the level they are placed in (e.g., beginner, intermediate, advanced) they may attend between eight to twenty hours of lectures. Recently, students have been given the option to forgo the final exam for their English course⁶, if they demonstrate their proficiency level with an IELTS minimum score of 5.0. Consider this in contrast to the Leicester Medical Program's requirement of a minimum of 7.5.

With regard to the medical school, the program traditionally consisted of three years basic sciences followed by three years clinical disciplines until recently⁷. This curriculum had problems of overcrowding in subjects (see Al-Gindan et al., 2000; Elfaki, 2004), which reflected negatively on the academic and educational environment (Al-Hazimi et al., 2004). This raised concerns and has since created a spirit of reform in most of the medical colleges toward a more student-centered curriculum (Khalid, 2008).

However, unlike the 10 hours of lectures at Leicester Medical School, the King Saud Medical School averages 30 to 40 lecture hours per week. Moreover, self-directed learning may currently be an ambitious skill in the current program. A Saudi lecturer in my NNS corpus states the value of the lectures by warning students to not to depend on the powerpoint slides and instead explains “as you know, course guidelines, course contents is the source of the questions... ninety per cent of the questions are from whatever I say in the lecture” (*DBT, SASME*).

Students often rely on teachers as the main source and providers of information and tend to reduce what they have learned into manageable, and sometimes isolated, chunks of information to be memorized (Gallagher, 1985). The learning task, as viewed by students and often reinforced by the education system, is to learn with the aim of reproducing the subject matter in the final exam (Al-Hazimi et al., 2004). Additionally, while the medical colleges recruit the very top high school graduates, the attrition rate is still high. This high attrition rate is multi-factorial but the

⁶ For the first semester English course only.

⁷ The medical college is a six year course: the first year is preparatory (pre-med), the next two years are devoted to basic medical sciences and the last three are for clinical rotations (Al-Ayed and Sheik, 2008)

development and use of the English language as the medium of instruction is a main factor (Al-Hazimi et al., 2004).

The faculty in most of these medical schools are reported to be of a high academic standard, with many of the Saudi faculty trained in Western countries (Al-Hazimi et al., 2004). Unfortunately, with the increase in the number of colleges, there is now a shortage of qualified staff which has led to a higher student to teacher ratio in some institutions.

Medical colleges in Saudi Arabia, higher education in general, continue to develop and improve the educational programs and methodology every year. There are a large group of educators in the country who are dedicated to improving and contributing to the growth of the higher education system. My intention with this research is to help contribute to this growth.

1.6 The Study

The main purpose of this corpus-based study is to analyze spoken academic discourse, by means of a contrastive analysis between NS English and NNS English lectures. In both contexts, English is used as the official medium of instruction (EMI) at the tertiary level in the medical schools. Furthermore, a supplementary analysis of the native Arabic discourse used in the EMI medical lectures will be studied.

I focus my analysis on the role and function of selected DMs (see 4.8) in academic medical lectures from Britain and Saudi Arabia. Research presented in Chapter 2 (see 2.5) investigates the impact of DMs regarding second language comprehension (Olsen and Huckin, 1990; Flowerdew and Tauroza, 1995; Jung, 2003; Yang, 2007) and provides an awareness of the effect DMs have within academic discourse (Young, 1994; Swales and Malczewski, 2001; Thompson, 2003; Yu, 2008). This study, similar to Bellés-Fortuño (2006), approaches the research with the intention of providing insights into how DMs are used and their roles by NS and NNS of English in a first and foreign medium context at two medical colleges.

1.6.1 Aims and Objectives

The aims and objectives of this study are to analyze the linguistic features of academic lectures that have been reported to aid in the retention and recall of information (Nattinger and DeCarrico, 1988), namely, discourse markers (DMs). The objectives are to investigate the uses and functions of DMs in a contrastive study of British and Saudi-based medical academic lectures. I will also investigate the uses and functions of Arabic DMs and the roles of code-switching that occur in the Saudi-based lectures.

The specific aims of the study are:

- (a) To investigate the roles of DMs in academic medical lecture discourse using data from a non-native English academic corpus and a native English academic sub-corpus.
- (b) To compare and contrast qualitative and quantitative differences in the use of spoken DMs in academic medical settings between Saudi-based NNS English speakers and British NS speakers.
- (c) To identify the functions and roles of speaker code-switching in the non-native English medium medical lectures.

1.6.2 Significance of the Study

This research is significant because it is one of the first studies to investigate DMs in the NNS speech of medical academic lecturers. It contributes to a small but growing amount of research into English as a medium of instruction in foreign settings (e.g. China, Turkey, Egypt, etc.) and is the first, to my knowledge, that investigates NNS speakers in a Saudi medical college. It is innovative because of the use of corpus-based methodologies to explore the uses and functions of English and Arabic DMs in academic medical speech and additionally, the inclusion of code-switching in EMI discourse.

The research will benefit English for Specific Purposes (ESP) programs for native Arabic speakers of English, help raise awareness among NNS and help facilitate bilingual

communication in the academic medical classroom. In a broader sense, this research will add to the knowledge of these under-researched fields and, it is hoped, will make a contribution to the development of NNS English in foreign medium programs and bilingual education.

1.7 Summary of Thesis Structure

This introductory chapter provides an overview of the research undertaken in this thesis, laying out the background for the study, reasons behind the data and context choices, aims and objectives and the significance of the research on spoken academic medical lectures. The following sections are general overviews of each chapter in this thesis:

Chapter 2 Academic Discourse

This chapter presents the literature of two distinct research areas that pertain to this study: academic/pedagogic classrooms/lectures and what we refer to as discourse markers (DMs). The review of the first research area outlines the predominant theories related to the pedagogy and the impact of research in the native and foreign medium academic classroom. The chapter will then relate this to research on DMs and will focus on the theories, definitions and taxonomies related to spoken discourse.

Chapter 3 Code-Switching

This third chapter presents literature related to CS as a social phenomenon and its pedagogic implications. It describes current theoretical views of CS from Gumperz (1982) and Myers-Scotton (1983, 1993) to Poplack (1980), Auer (1984) and Ferguson (2003). It provides current definitions and theories of CS in various languages, including English and Arabic. Examples are given to illustrate the function of CS and the reported roles it plays in social and academic discourse.

Chapter 4 Methodology

This chapter describes the research context in which this study is carried out in both Saudi Arabia (KSA) and in the United Kingdom (UK). It introduces two sub-corpora of spoken academic discourse: non-native speaker EMI medical lectures (KSA) and a comparative NS medical lectures (UK). The Methodology chapter presents and discusses the corpus selection, collection methods, transcription process, instruments and data analysis procedures. It also presents and discusses the categorization of the DMs analyzed in the study in both English and Arabic* (*solely with NNS corpus data).

Chapter 5 Analysis of Results I: Structural DMS

This chapter and the following (chapter 6), describe the results of the data analysis of the study. The Results I chapter presents and briefly introduces the findings of the data sets overall, before moving onto a detailed presentation of the DMs identified as Structural DMs: Topic Initiators, Topic Developers, Summarizers and Closers. Each function under the classification of Structural DMs is described and illustrated with excerpts from the data sets. This chapter includes only the English Structural DMs results.

Chapter 6 Analysis of Results II: Interactional DMS

This chapter continues the presentation of the results of the data analysis. The first section presents and briefly introduces the DMs, before providing a detailed presentation of the second classification which are identified as Interactional DMs: Confirmations Checks, Elicitors and Rephrasers. Each function under this classification is described and illustrated with excerpts from the data sets. The second section of the chapter introduces the Arabic discourse markers (ADMs) in the Saudi-based academic medical lectures. Each function under the classifications is described and is illustrated with excerpts from the data set. A further section is included which describes the roles that pedagogic code-switching may play in academic medical lectures.

Chapter 7 Discussion

This chapter provides a discussion of the data results presented in the previous two chapters with more speculation of the results found between the NS and the NNS lecturers. The reasons for the differences and similarities will be expanded upon and studies which relate to these results will be discussed.

Chapter 8 Conclusion

This final chapter provides a summation of the study and describes the most important findings with a look towards its pedagogic value. It also discusses the limitations and challenges faced with a spoken data corpus and the future implications this and other studies may have in the academic classroom.

CHAPTER 2

DISCOURSE MARKERS

“The lecture class is changing (Waggoner, 1984), so that traditional methods of learning coexist with newer interactive methods; both lecturers and students feel the influence of a greater egalitarianism than in former times.”

(Fortanet-Gómez and Bellés-Fortuño, 2005: 162)

Chapter 2 Discourse Markers

2.1 Introduction

This chapter provides the basis for the theoretical and analytical framework of this study. The literature presented addresses four main issues: the study of (spoken) academic lectures, the study of Corpus Linguistics, the study of discourse markers (DMs) and the study of DMs used in pedagogic settings.

The first part of this chapter provides background of the lecture genre and provides the foundation for my analysis of two sub-corpora of spoken academic medical lectures. The second part of this chapter provides background into the development and a discussion of the advantages of using a corpus approach to data analysis. The third part of this chapter begins with the various terminology and classifications of DMs and then develops a working definition of DMs through a review of the main research in the field. The section then continues with a review of currently cited taxonomies for DMs and their development from past to present. Additionally, this third section provides the basis for my analytical framework, which identifies the uses and functions of DMs in monologic medical lectures. The fourth and final part of the chapter describes two categories of pedagogic studies of DMs based on the different aims and purposes of each study.

2.2 Academic Discourse

2.2.1 The Lecture as a Genre

The term *genre* is generally defined as a text or discourse type that can be recognized by its features and particular functions as belonging to a similar group (e.g., prayers, poems or legal reports). Genre can be regarded, according to Swales's (1990) and Mauranen's (1993) views, as formed by both the members of the discourse community as well as defined by the genre itself (Suviniitty, 2012). Genre is realized in discourse communities as a mutual communicative goal

(Swales, 1990). And while there may be cultural differences regarding how students and lecturers perceive lectures there are also common features found in most lectures (Suviniitty, 2012) for example, the use of DMs, which are amongst the top 10 word forms found in an analyses of corpora of spoken interaction (Allwood, 1996; Fung and Carter, 2007). Additionally, Jucker and Smith (1998: 176) found that DMs (e.g., *yeah, okay, I mean, well, etc.*) occurred roughly every four to five seconds in a three and a half hour recorded spoken interaction between participants⁸. This suggests that interactive academic discourse may include a comparable frequency of DMs to spoken conversation, although occurrences may be lower in lecture monologues.

Bhatia (1993) elaborates on Swales' definition stating that genre is a socially recognized, highly structured communicative discourse event or activity that aims to fulfill the requirements of a community for a particular communicative need. Since the early 1980s, interest in genre-centered approaches to the analysis of written and spoken discourse has been greatly motivated by the need to provide NNS students with satisfactory models and descriptions of academic and scientific texts in order to enhance their ability to understand and produce them (Holmes, 1997).

Swales (1990) describes genre analysis as a means of studying how language is used within a particular setting to derive a rationale as to why certain genres display certain features. Henry and Roseberry (2001) describe the general aim of genre analysis as a means of identifying the moves and strategies within a genre, the accepted order of moves within it and the key linguistic features.

As an established genre in tertiary education the lecture is considered to be a major part of university instruction (Chang, 2012; Flowerdew, 1994). Students' ability to internalize and comprehend the content of lectures is key to their success at university. For this reason, most research concerned with spoken academic discourse focuses on the lecture, especially on linguistic aspects (Chaudron and Richards, 1986; Flowerdew, 1994; Morell, 2004; Khuwaileh, 1999; Bellés-Fortuño, 2006; Fung and Carter, 2007). Academic discourse differs from everyday social interactions in that it is more carefully organized, structured and includes a higher density of specialized lexis of the specific academic discipline.

⁸ The 15 participants recruited from Long Beach University took part in conversations with various topics for 5 minutes each; 9 participants were strangers and 6 were friends (Jucker and Smith, 1998: 175).

Flowerdew (1994: 14) proposes that lecture discourse research is important as it:

“[...] can indicate to teachers and course designers what linguistic and discorsal features learners need to be familiar with in order to understand a lecture and what, therefore, should be incorporated into ESL courses. In addition a knowledge of the linguistic/ discourse structure of lectures will be of value to content lecturers in potentially enabling them to structure their own lectures in an optimally effective way”

Flowerdew (1994) describes the valuable and enabling resource of lectures and yet, as Swales (1990) points out, little research has been based on an analysis of the lecture as a genre⁹. Most studies have limited their research to *written* academic language (Biber, 1988, 1990, 1995; Biber et al., 1999; Conrad, 1996; Grabe, 1987; Johns, 1997) because *spoken* academic language has never been as easily accessible. The main focus of the studies carried out in the 70s and 80s, was to improve training in English as a Second or Foreign language (ESL/EFL) and focused on how information is organized in a lecture through micro- and macro- structures (Cook, 1975; Murphy and Candlin, 1979; Coulthard and Montgomery, 1981; Chaudron and Richards 1986; Nattinger and DeCarrico, 1988). Since the 1990s, a shift occurred and a slow but steady increase in spoken academic studies have been conducted. The studies focuses on a range of topics from lexical, topical, discourse patterns of lectures, lecture introductions, to DMs (Dudley-Evans, 1994; Nattinger and DeCarrico, 1992; Young, 1994; Olsen and Huckin, 1990; Flowerdew, 1994; Thompson, 2003).

2.2.2 Lecturing Styles

Spoken text has its own particular lexis and grammatical characteristics (Flowerdew, 1994) and the academic lecture places greater demands on NNS as it requires real time processing of lecture content, which is often informationally dense, especially in the scientific fields (Olsen and Huckin, 1990; Flowerdew, 1994; Nattinger and DeCarrico, 1988; Chang, 2012). Academic lectures are high in informational content and pre-planned monologic lectures provide little to no opportunity for negotiations of meaning and thus suggest a greater pedagogic need for discourse

⁹ ‘Genre’ is identified in terms of its communicative purpose within the context of a particular discourse community and its characteristic content and style (cited in Thompson, 1996)

structuring devices that make connections between propositions clear (Nesi and Basturkmen, 2006).

The academic style which a lecturer adopts is an individual choice that can often be divided into a number of categories based on common features. One of the first classifications of academic lectures resulted from Morrison's (1974, reported in Jordan 1989: 153) research which analyzed science lectures. The lectures were divided into two kinds: (1) formal, which referred to "close spoken prose"; and (2) informal, which is defined as "high informational content" though not necessarily in "high formal register". However, Morrison's (1974) categories were viewed as highly simplistic and a more complete classification was proposed by Goffman (1981) who divided academic lectures into three modes, namely, *memorization*, reciting a text from memory; *aloud reading*, reading aloud from a prepared text; and *fresh talk*, which are utterances spontaneously composed on the spot. Dudley-Evans and Johns (1981: 134), in turn, also identified three academic lecturing styles which gradually move from more formal and controlled discourse to more informal:

Style	Characteristics	Intonation
<i>i. Reading style</i>	Characterized by short tone groups and narrowness of intonational range, in which falling tone predominates; although level tone may also occur.	Lecturer reads from text or sounds as though they are reading
<i>ii. Conversational style</i>	Characterized by longer tone groups and key-sequences from high to low.	Informal with marked increase in tempo and vowel reduction and reduced intensity.
<i>iii. Rhetorical style</i>	Characterized by its wide intonational range. The lecturer often exploits high key and a 'boosted high key'.	There are frequent asides and digressions marked by key and tempo shift- sometimes also by voice quality-shift

Table 2.1 Dudley-Evans and Johns' (1981: 134) Academic Lecturing Styles

Dudley-Evans (1994) points out that the key to understanding a lecture lies in appreciating a lecturer's individual style. While there is no general agreement on the preferred style of lecturing, it has been suggested that most lectures, especially at English-speaking universities, are becoming less formal and more interactive (Benson, 1994; Swales, 2004). Swales (2004) defines this more interactive style as "open style" lecturing and characterizes it as lecturers not reading from their

notes, but rather working from an outline or lecture slides as guides. Goffman (1981) refers to this open style as "fresh talk" and Dudley-Evans and Johns (1981) as "conversational style".

This trend towards a more interactive style of teaching (Benson, 1994) may be problematic for NNS English students who are used to a more traditional style of lecturing. Bellés-Fortuño suggests that international students may have cultural difficulty with the role and status of lecturers in regards to deference and may be alarmed with the egalitarian concept of lecturer as guide or facilitator (2006; Benson, 1994). Additionally, NNS students' unfamiliarity with varying styles of lectures may result in a failure to recognize key points despite adequate English proficiency (Olsen and Huckin, 1990) and in differentiating between the main and subsidiary points when note-taking (Clerehan, 1995).

As previously stated, little evidence exists about the type of lectures that currently predominate in university contexts and admittedly, lectures and lecturers can be rather idiosyncratic (Schourup, 1999; Bellés-Fortuño, 2006). There seems to be tacit agreement that a more interactive lecturing style is becoming the norm across universities (see Lynch, 1994; Morell, 2004; Fortanet-Gómez and Bellés-Fortuño, 2005; Dafouz et al., 2007).

With respect to Saudi Arabia, recent publications indicate that universities have acknowledged the challenge of improving the quality of teaching and learning especially with academic lectures which have been the traditional center of university teaching (Alnassar and Dow, 2013). In the health sciences, there has been a shift from 'large group teaching', which has been criticized as "one-way communication, focused on imparting information, sometimes dull, boring and repetitive", to small group teaching (Alnassar and Dow, 2013: 53) and the adoption of a more student-centered approach (Al-Shehri et al., 2013). These smaller classes provide teachers with the opportunity to implement more interactive styles of teaching where they can check students' understanding of concepts and gradually progress in the "building of bodies of knowledge" (Alnassar and Dow, 2013: 53).

2.2.3 Features in Lectures: Skills and Strategies

Bellés-Fortuño (2006) reports that the features of academic speech which are often hard to identify by NNS who have had little to no exposure to English lecture discourse are generally phonological. These features include unit boundaries, intonation, false starts, irregular pausing, hesitations, stress and intonation patterns (Bellés-Fortuño, 2006). Phonological boundaries in written texts are often marked by means of punctuation (Flowerdew, 1994), whereas in spoken texts these boundaries are guided by intonation patterns. To illustrate, the following excerpt (*VitA*) from my study demonstrates the marked phonological boundaries of three highlighted DMs ‘*okay*’, ‘*alright*’ and ‘*so*’. The first DM ‘*okay?*’ functions as a Confirmation Check normally used by the speaker to confirm whether listeners have understood the preceding information and is characterized by the rising intonation of a question. The second marker ‘*alright*’ is marked with falling pitch - associated with a closing statement and end of a proposition and functions like a closer, ending the previous topic. The third DM ‘*so*’ functions as a Topic Initiator - starting a new topic and is normally characterized by a flat intonation.

“Okay, we’ll take it step by step, okay? Alright. So what happens here – so, this is our retina and it contains rod and cone cells, okay? So the rod and cone cells are basically all over...”

[*Vit A* - NNS, SASME Corpus]

The characteristics of academic lectures, according to Biber (1995), have a high informational focus, comparatively low interactivity and involve advanced planning (i.e. to prepare the lecture). He also suggests that academic lectures may share several situational characteristics with face-to-face conversations as they are both produced in spoken mode, with the addresser and addressee sharing time and space and are both interactive. The following Excerpt (*CTB*) illustrates the characteristics of academic lectures as interpreted using Biber (1995).

“{Why doesn’t it grow in the dish? Because Allah has made it this way not to grow externally, but only to grow in the body where it can multiply}¹⁰ *So, partially treated meningitis [...] Someone with chronic meningitis mitatawal* {chronic}, *please think first [...] the wrong dose, by the wrong route suh?* {right?}. *You understand? Fahmeen?* {you understand?} *Man’tum fahmeen?* {You don’t understand?} *Tayib* {Okay}”

[*CTB - NNS, SASME Corpus*]

The *CTB* Excerpt, taken from the NNS academic medical corpus, illustrates high informational content, involves advance planning for lecture presentation and projects an informal style associated with Dudley-Evans and Johns’s (1981) *conversational* style (see 2.2.2). Moreover, while the *CTB* Excerpt does include a number of eliciting questions in both English (L2) and Arabic (L1), there are no incidences of student interactions during or immediately following this, thus we could label these questions as pseudo-elicitors and the lecture, at this point, as having comparatively low interactivity.

Students in academic lectures are tasked with actively constructing a coherent interpretation of what is usually a complex, informationally dense and cognitively demanding spoken message (Thompson, 1994). This active and continuous struggle by listeners to build a representative mental model “is the result of our combining the new information in what we have just heard with our previous knowledge and experience” (Anderson and Lynch, 1988: 11). NNS students frequently face considerable difficulty in following academic lectures delivered in a foreign medium and the language barrier often results in missing crucial details in the lecture which may compromise their ability to successfully build a mental model (Thompson, 1994).

2.3 Corpus Linguistics

The computer revolution of the 1980s gave rise to what is now known as the modern form of Corpus Linguistics (CL). Corpus Linguistics is the study of language through a method of linguistic analysis using a large, principled collection of naturally occurring examples of written and/or transcribed spoken discourse that is stored electronically (Bennett, 2010). This electronic

¹⁰ {English translation} of lecturer’s Arabic discourse

collection of text is known as a corpus or corpora (plural). Sinclair (1994) defines a corpus as “a collection of pieces of language, selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language” (cited in Pearson 1998: 42). More recently, McEnery et al. (2006) have altered the definition of a modern corpus as being “a collection of sampled texts, written or spoken, in machine-readable form which may be annotated with various forms of linguistic information” (ibid: 4).

With the computer revolution, the emergence of modern CL allows for the analysis of language with increased speed, greater accuracy, accountability, replicability and larger volumes of better annotated data. Researchers using CL are able to explore a number of linguistic questions and offer unique insight into language, language patterns and rules “which has made it one of the most widely used linguistic methodologies” (CL, 2012).

Corpus Linguistics as a methodology allows for a more scientific and empirical analysis of language, rather than relying on language intuition alone. CL provides the resources for exploring theoretical issues through a corpus and allows the verification of the “falsifiability, completeness, simplicity, strength and objectivity of linguistic hypotheses” (Leech, 1992: 112-113)

Corpus-based research is described by Aston and Burnard (1998: 123) as research which naturally grounds its "theorizing in empirical observation rather than in appeals to linguistic intuition or expert knowledge".

2.3.1 Corpora for Language Research

The aim of a corpus-based analysis is not usually to generate theories of what may be in the language, but rather provide insight into patterns and what this shows us about language behavior (Krieger, 2003; Mauranen, 2003). Aijmer (2002) describes the advantage of using corpora for linguistic research as one which provides the opportunity to study the distribution and function of language in extensive text extracts representing different registers (ibid: 3). Furthermore, it allows researchers to analyze the functions of linguistic expressions in their social and situational context (ibid: 277).

A corpus-based analysis can be used to investigate virtually all language patterns - lexical, structural, lexico-grammatical, phonological, morphological, stylistic, etc. - with specialized focus. For example, Biber et al. (1994) use corpora for comparing registers within a language; specifically, they classify text types according to six communicative “dimensions” (e.g. “narrative vs. non-narrative”, “explicit vs. situation-dependent reference”, “abstract vs. non-abstract information”, etc.).

With the advancement of computer technology, the exploitation of large-scale corpora became a reality. From the 1980s onwards, the number of corpora and corpus-based studies have grown substantially with more examples of written texts than spoken, as the latter involves transcription of audio recordings and are substantially more time consuming than the former.

The following are a brief list of significant and specialized corpora which includes or are solely, spoken text:

- 1) *Cambridge and Nottingham Corpus of Discourse in English (CANCODE)*: A specialized corpus of informal spoken English was established at the University of Nottingham and completed in 1996. It contains about five million tokens and consists of five broad contexts types, namely - transactional, professional, pedagogical, social and intimate.
- 2) *International Corpus of English (ICE)*: The first systematic parallel corpora of one million tokens from at least 18 nations or regions globally, where English is spoken as a first, second or other language (Greenbaum, 1996). It contains about 20 million tokens in total.
- 3) *Michigan Corpus of Academic Spoken English (MICASE)*: A specialized spoken corpus of transcribed academic speech events recorded at the University of Michigan between 1997 and 2001. It currently has a 152 transcripts totaling 1.8 million tokens and divides the speech events, among other variables, by academic division: Humanities & Arts, Social Sciences & Education, Biological & Health Sciences and Physical Sciences & Engineering.

- 4) *British Academic Spoken English Corpus (BASE)*¹¹: A specialized spoken corpus of transcribed academic discourse - 160 lectures and 39 seminars recorded between 2000-2005 across two universities, Reading and Warwick, totaling 1.64 million tokens. Lectures and seminars are subdivided by academic division: Arts & Humanities, Social Studies & Sciences, Physical Sciences and Life & Medical Sciences. The BASE corpus gives researchers the opportunity to investigate spoken academic English at both a lexical and structural level and provides a British counterpart to the American MICASE.
- 5) *Vienna-Oxford International Corpus of English (VOICE)*: A specialized spoken corpus of naturally-occurring, non-scripted face-to-face interactions involving 753 identified individuals from 49 different first language backgrounds using English as a lingua franca¹² (ELF). VOICE was compiled at the University of Vienna and recordings were carried out between 2001-2007; it contains over a million tokens. The transcriptions are complete speech events from different domains (e.g. educational, leisure, professional) and of different speech event types (VOICE, 2013).
- 6) *Louvain International Database of Spoken English Interlanguage (LINDSEI)*: A specialized corpus containing spoken informal interviews produced by higher intermediate to advanced learners of English from several mother tongue backgrounds (e.g., Chinese, Dutch, French, German, Greek, Italian, Japanese, Polish, Spanish, etc.). It contains over a million tokens and is a collaborative project between several universities internationally, that is coordinated by the University of Louvain.

The use of computerized CL as a basis for language analysis and for various other applications has dramatically altered the study of language (Fung, 2003). The development of corpora such as those listed above, have provided immeasurable documentation of language use in pedagogic and multilingual environments.

¹¹ The recordings and transcriptions used in this study come from the British Academic Spoken English (BASE) corpus. The corpus was developed at the Universities of Warwick and Reading under the directorship of Hilary Nesi and Paul Thompson. Corpus development was assisted by funding from BALEAP, EURALEX, the British Academy and the Arts and Humanities Research Council.

¹² *Lingua franca* is the use of a language, in this case English, as a common means of communication among speakers from different first-language backgrounds

CL based methodologies allow research, such as the present study, to observe the frequency of DM occurrences in spoken academic lectures. Its advantages in analysis are, first and foremost, its relatively immediate display of the presence or absence of DMs in corpora. Second, it reflects the frequency of occurrences per lexical unit and per lecture. Third, word lists allow the comparison of individual and overall frequency of occurrences in a data set, such as my NNS English corpus, against the occurrences in another data set such as a NS English corpus like BASE presented above. Lastly, concordance lists and searches allow further analysis of DM meanings and uses with respect to context.

2.4 Discourse Markers

2.4.1 Discourse Markers: Terminology Reviewed

One of the earliest suggestions that DMs are worthy of study on their own merits came from Levinson (1983). He observes:

“... there are many words and phrases in English, and no doubt most languages, that indicate the relationship between an utterance and the prior discourse. Examples are utterance-initial usages of *but, therefore, in conclusion, to the contrary, still, however, anyway, well, besides, actually, all in all, so, after all*, and so on. It is generally conceded that such words have at least a component of meaning that resists truth-conditional treatment... what they seem to do is indicate, often in very complex ways, just how the utterance that contains them is a response to, or a continuation of, some portion of the prior discourse.”

(Levinson, 1983: 87-88)

Levinson (1983) did not go so far as to name this group nor did he discuss the topic further than the above quote. However, it can be deduced that he regards these “words and phrases” as operating on the level of discourse structure to signal the relationship between utterances and organize the discourse.

The most frequently used, and often preferred, term to describe these lexical expressions is ‘discourse markers’ (Schiffrin, 1985, Brinton, 1990, Blackmore, 2002; Fuller, 2003, Müller, 2005,

Fung and Carter, 2007) though Levinson himself does not use this term. A wide range of possible terms and definitions exist and have been described by researchers as being "as pervasive in language as they are difficult to define for the linguist" (Taboada, 2006: 568). Some of these terms can be seen in Table 2.2.

Term	Researchers
Pragmatic Devices	van Dijk, 1979; Stubbs, 1983
Discourse Particles	Goldberg, 1980; Schourup, 1985
Discourse Connectives	Blackmore, 1987, 1992
Pragmatic Markers	Fraser, 1988, 1990; Schiffrin, 1987; Brinton, 1996
Utterance Particles	Luke, 1990
Discourse Operators	Redeker, 1990, 1991
Cue Phrases / Cue Words	Knott and Dale, 1994 / Horne et al., 2001
Continuatives	Romero-Trillo, 1997
Backchannel Cues	Verschuren, 1999

Table 2.2 Variety of 'DM' Terms

The term 'discourse markers' is generally used as an umbrella term to cover a wide range of linguistic expressions by researchers such as Schiffrin (1987), while others such as Fraser (1990, 1999) treat discourse markers as a sub-set of the cover term 'pragmatic markers'. There is no consensus in the literature on what the term means precisely, which members make up the class of DMs, nor, how these expressions should be classified. Schourup (1999) describes DMs as inherently problematic and difficult to define or classify and suggests the lack of consensus in classification is rooted in the multiplicity of definitions. This diversity is reflective of the wide range of linguistic approaches employed and the myriad of functions these elements are said to fulfill (Jucker and Ziv, 1998: 1).

For this study, I adopt the term Discourse Markers (DMs) because I believe it best suits the aims of this study, and is the most common label used by leading scholars in the field. In the next section, I will turn to the definition of DM.

2.4.2 Characteristics & Definitions of Discourse Markers

As with the myriad of terms to describe these linguistic features, there is also a plethora of definitions. Additionally, there is considerable variation in the classification of linguistic expressions under the umbrella of DM, with reports ranging from a dozen to five hundred (Brinton, 1998: 11).

In order to establish defining features and a classification for DMs in this study, we must first review the characteristics which have been identified in previous research. A summary based on Brinton's (1990: 46-47) and Jucker and Ziv's (1998: 3) list of features is presented with the common characteristics used to describe DMs (cf. Svartvik, 1980; Östman, 1982; Schiffrin, 1987; Fraser, 1988). The list is as follows:

- (1) They are mainly features of spoken discourse
- (2) They appear with high frequency
- (3) They are short
- (4) They are phonologically reduced
- (5) They occur either outside the syntactic structure or are loosely attached; optional
- (6) They are generally utterance-initial
- (7) They are difficult to place within a traditional word class
- (8) They have a 'core meaning'
- (9) They are multifunctional

I will now briefly discuss each of these characteristics and review whether the features are used as defining criteria to distinguish discourse marker use from non-discourse marker use in spoken discourse.

(1) Feature of Spoken Discourse

Orality (vs. written medium) has been rejected as a defining criterion of DMs by Jucker and Ziv (1998) and by Schourup (1999: 234) who states that most studies (with the exception of Fraser) are based on speech data and as a result the lexical expressions classified as DMs occur primarily in that medium. Fraser (1990: 389) asserts that certain DMs occur more frequently in

written discourse (e.g., *notwithstanding*) while others are found more frequently in conversation (e.g., *okay*). In fact, McCarthy (1993: 180) argues that our judgment of “spokenness” in a written text is largely based on the degree of spoken discourse markers in a text. Müller (2005: 7) suggests that this may contribute to the lack of interest in written DMs (e.g., *moreover*, *consequently*) which are more restricted in terms of where they can be used in context and therefore may seem less interesting.

While ‘orality’ may not be a defining feature of DMs, it remains true that they have been found to occur more frequently in spoken discourse (Müller, 2005: 7). Within academic lectures, Biber (2006) identified certain DMs as features of the genre such as the use of ‘*okay*’ and ‘*so*’ which are spread across the lecture. Based on reported studies, I anticipate that these features will be found across the spoken discourse of the research corpora.

(2) Frequency

As a feature of spoken DMs, some studies report that the frequency of DMs is higher in speech than in written discourse (Jucker and Smith, 1998; Louwse and Mitchell, 2003). Studies have also found that the frequency of occurrences varies by native and non-native speakers. Fuller (2003) reported that the frequency of DMs among non-native English speakers was lower overall, in contrast to native speakers and showed less variation in their use across speech contexts.

(3) Length

DMs as a rule tend to be short expressions; and those included by researchers generally comprise of single word units and/or two- or three-word units (cf. Schiffrin, 1987; Othman, 2010). However, there are other studies that have included various longer and even clausal expressions under the category of DMs (cf. Chaudron and Richards, 1986; Fraser, 1996; Aijmer, 2002; Fung and Carter, 2007). Nattinger and DeCarrico (1988) find in their study that the length of a lexical expression is not a reliable criterion

(4) *Prosody*

Aijmer (2002) views DMs as often characterized by separate tone units from the rest of the utterance. These tone unit boundaries may be signaled by punctuation (in transcriptions), as will be done in this study and may thus set them apart from the following utterance with punctuation marks (Lutzky, 2006). However, identifying these markers in relation to tone units is regarded as problematic as identification through sentences, propositions or actions, because the DMs “are not limited by the boundaries of the unit” (Schiffrin, 1987: 34). To illustrate her point Schiffrin (1987: 34) provides the following examples:

I mean and *but* as a preface to a tone unit

- 1) **I mean** I may be wrong, **but** I’m- **I mean** that’s what I’m- that’s my opinion.

y’know and *I mean* within a tone unit

- 2) So **I mean y’know** I just don’t know.

y’know occurring at end of tone unit

- 3) We have some **y’know**.

Schiffrin (1987) further states that DMs may also fill some tone units which do not coincide with syntactic and information boundaries (e.g., *Well I guess they will, but um... I don’t know*). She also describes DMs as having “a range of prosodic contours, e.g. tonic stress and followed by a pause, phonological reduction” (Schiffrin, 1987: 328). Phonological reduction, which can be exemplified by the reduction of ‘*because*’ → ‘*cause*’ and ‘*you know*’ → ‘*y’know*’, is also cited by Brinton (1990), Jucker and Ziv (1998) and tentatively by Sankoff et al.. However, the majority of researchers do not regard this as a defining feature and as such do not dwell on phonological features in their definitions (Müller, 2005: 5).

(5) *Independence and Optionality*

One of the key identifiers for DMs is their syntactic independence and grammatical optionality (Schiffrin, 1987; Brinton, 1996; Fraser; 1999; Aijmer, 2002). This relates to their non-truth conditionality, i.e., DMs do not affect the truth conditions of the proposition expressed by an utterance. Schiffrin (1987) describes DMs as “nonobligatory utterance-initial

items that function in relation to ongoing talk and text” (1987: 31). She claims that certain DMs “select a meaning relation from whatever potential meanings are provided through the content of talk and then display that relation” (Schiffrin, 1987: 318). She clarifies that while the absence of a DM does not affect the grammatical structure of an utterance, it does allow different potential meaning relationships to be interpreted. I believe this ambiguity in meaning would be even more profound in academic discourse, especially the hard sciences where lecturers are under time constraints to transfer large amounts of dense data to students. For example, consider the following Excerpt¹³ inclusive and then without of the highlighted DMs:

- 1) a. **So** diabetes is difficult to define **and therefore**, diabetic nephropathy is difficult to define **but** we have to work on some definitions.
- b. Diabetes is difficult to define, diabetic nephropathy is difficult to define. We have to work on some definitions.

The Excerpts (1a-b) illustrate the potential meaning relationship constraints that DMs provide in utterances, especially when we compare utterance (1a), which is inclusive of the DMs ‘*so*’, ‘*and therefore*’ and ‘*but*’. Each DM provides signposts for the listener on how to interpret the utterance. In this case, the initial DM ‘*so*’ signals to the listener the resulting consequence of a prior utterance, which is that diabetic nephropathy is difficult to define as a result of diabetes itself being difficult to define. The final DM, ‘*but*’, functions to signal a contrasting upcoming statement. The DMs used in 1a allows the listener to interpret the message that regardless of these difficulties, “we” (1a), both speaker and hearers, must continue on (with the topic) in spite of the issues of the working definitions. The lack of DMs in the second Excerpt (1b), illustrates the ambiguity in interpreting the utterance.

Fraser (1999) describes the function of DMs as conveying a message through segments (‘S’) of discourse “by S2 vis-a-vis the interpretation of S1” (Fraser, 1998: 302). However, Fraser (1999) believes that it is clear that “a DM does not ‘display’ a relationship, as Schiffrin (1987) would have it, any more than a verb displays a relationship between a subject and object”. Instead, Fraser (1999) contends that DMs “impose on S2 a certain range of interpretations,

¹³ Excerpt taken from lecture transcript *lslect032* from the Life & Health Sciences group in the BASE corpus

given the interpretation(s) of S1 and the meaning of the DM” (ibid: 942). For example, Fraser (1999) illustrates with the following sentences:

- 7) a. I want to go to the movies tonight. *After all*, it's my birthday.
 b. John will try to come on time. *All the same*, he is going to be reprimanded.
 c. A: Harry is quite tall.
 B: *In contrast*, George is quite short.

(Fraser, 1999: 944)

Fraser (1999) states that expressions which function as DMs may have several aspects of semantic meaning when relating two discourse segments as in 7a-c. Generally, DMs can be removed without affecting their host (i.e. utterance). Consider the effect of the highlighted DMs in Fraser’s examples removed from the sentences. The absence of DMs provides no linguistic cues on how to interpret the two segments, specifically the relationship of *why* they want to go to the movies tonight (7a), that John will be reprimanded tonight regardless of whether or not he comes on time or that George is “quite short”, though only in contrast to Harry. Therefore, if a DM is omitted, the relationship between segments is still available to the hearer, though it is no longer explicitly cued (Schourup, 1999: 231).

It is important to note that this optionality that Schiffrin (1987), Fraser (1999) and others refer to “only concerns grammatical well-formedness of the relevant sentence and not its pragmatic impact” (Müller, 2005: 6). Furthermore, in exceptional cases, non-compliance with this feature rarely leads a linguistic expression to be excluded as a DM (Müller, 2005: 4).

(6) Position

DMs are often claimed to occur mainly in utterance-initial positions (cf. Schiffrin, 1987, Fraser, 1990; Biber et al., 1999), however, studies have shown that certain DMs can also occur as frequently in utterance-medial or utterance-final positions or even form a discourse unit of their own (Erman, 1986; Brinton, 1990; Andersen, 2001; Jabeen et al., 2011).

Fraser (1999) states that though most DMs occur in the initial position, there are those which can occur in the medial (e.g. *however*, *although*) and those in the final position (e.g. *nevertheless*). He also proposes that all DMs share one common feature:

“they impose a relationship between some aspect of the discourse segment they are a part of, call it S2, and some aspect of a prior discourse segment, call it S1. In other words, they function like a two-place relationship, one argument lying in the segment they introduce, the other lying in the prior discourse”

(Fraser, 1999: 938)

He represents this relationship in the canonical form <S1. DM+S2>. Fraser (1999) then addresses the issue of the status of DMs which, as stated, need not necessarily be initial nor be adjacent. The four cases Fraser (1999: 938-941) identifies of the various grammatical patterns of the discourse segments that DMs relate are illustrated below:

<**S1. DM+S2**> The DM relates independent sentences S1 and S2

- 3) a. He drove the truck through the parking lot and into the street. Then he almost cut me off. After that, he ran a red light. *However*, these weren't his worst offenses.
- b. The bank has been closed all day. *Thus*, we couldn't make a withdrawal.

<**S1. DM+S2**> or <**S1, DM+S2**> Two independent clauses joined by a coordinate conjunction

- 4) a. Jack played tennis. *And* Mary read a book.
- b. Jack played tennis, *and* Mary read a book.

<**S1. DM+S2**> (5a) and two additional forms (5b and c) not canonized by Fraser (1999)

- 5) a. There was considerable flooding. *As a result* (of that), farmers went bankrupt.
- b. *As a result of* considerable flooding, farmers went bankrupt.
- c. Farmers went bankrupt *as a result of* considerable flooding.

<S1, DM+S2> and <DM+S2, S1> involves DMs which do not permit the canonical form, only these patterns.

- 6) a. Mary is angry with you *because* you ran over her cat with your car.
- b. Harry will not go, *unless* he is paid an appearance fee.
- c. *While* she is pregnant, Martha will not take a plane.

[Examples taken from Fraser (1999: 938-941)]

Jabeen et al.'s (2011) research which compared the frequency of eight English DMs (*I mean, you know, I think, kind of, sort of, well, you see* and *so*) in British and Pakistani speech supports that many DMs are flexible with utterance position. The results of their study indicate that while NS use more DMs than NNS, the Pakistani speakers used the DMs in all positions (e.g., initial, medial and final) unlike NS who, on average, used them more often at the beginning of an utterance.

While I agree that certain DMs may occur more commonly in turn - or utterance - initial positions to signal upcoming information (cf. Fraser, 1999; Othman, 2010), I would argue that they can theoretically hold any position (initial, medial or final) depending on the context of the discourse segments and the speaker's intention. Moreover, consideration must be given to the native and non-native status of speakers as with Jabeen et al.'s study. For example, the following Excerpts (2a-c), taken from my non-native Saudi-based lecturers illustrate DM positions in utterances:

- 2) a. **Because** there is a cause, the exam is a cause, fear from dentist is a cause. [DAD]
- b. Yes, true. **Okay**, we are coming to generalisation, when both the spheres...[DE]
- c. **So** dementia is a problem in the development of the memory and other cognitive function. With presentation of the consciousness, **okay**. [DBD]

(7) Word Class

The difficulty of placing DMs within a single word class was observed by Svartvik (1980) and is supported by Fraser (1990, 1999), Schiffrin (1987), Aijmer (2002), among others.

Schiffrin (1987) suggests that all DMs, based on their linguistic expressions, are comprised of various members of word classes such as those in Table 2.3.

Conjunctions	Interjections	Adverbs	Lexicalized Phrases
<i>and</i>	<i>oh</i>	<i>now</i>	<i>you know</i>
<i>but</i>	<i>well</i>	<i>then</i>	<i>i mean</i>
<i>or</i>			
<i>because</i>			
<i>so</i>			

Table 2.3 Schiffrin's (1987) Initial Word Classes of Discourse Markers

Schiffrin (1987: 328) suggests that additional word classes, such as perception verbs (*see, look, listen*), deictics (*here, there*), interjections (*gosh, boy*), meta-talk (*what I mean is*) and quantifier phrases (*anyway, anyhow*), can also be added as DM categories.

In their function as DMs, these various word classes index prior and/or upcoming utterances and typically resist truth-conditional characterization. By indexicality, we mean that DMs simply 'indicate' a relationship between or across utterances, leaving the hearer with a significant amount of interpretational work (Aijmer et al., 2005). By focusing on a functional approach to analyze DMs (cf. Schiffrin, 1987), as we do in this study, the heterogeneity of DM word classes becomes irrelevant. A functional perspective allows the description of DMs to be analyzed with the same tools; for example, conjunction '*so*' and adverb '*now*' are both listed under the sub-function Topic-Shifters in Bellés-Fortuño's (2006) study of academic lecture discourse (see 2.4.4).

(8) Core Meaning

In terms of core meaning, Schiffrin (1987) describes, for example, the conjunction '*but*' as having a pragmatic effect close to its meaning which marks speaker contrast. Its core meaning, like other DMs, does not fluctuate from use to use, but rather, what changes are the slots in which they appear in discourse (Schiffrin, 1987: 318). By 'slots', Schiffrin (1987) refers to the planes of discourse that DMs may function on (see Table 2.3).

Fraser (1999), who takes on a solely grammatical-pragmatic perspective, views DMs as linguistic expressions with a core meaning which can be enriched through context (ibid: 936). Furthermore, each DM not only functions to add textual coherence to the text but also signals the relationship that the speaker intends between utterances - introduced and foregoing (ibid: 936).

Fraser's (1990) DM framework differentiates between the content and meaning; namely that, content is "a more or less explicit representation of some state of the world that the speaker intends to bring to the hearer's attention by means of the literal interpretation of the sentence" (1990: 385). The meaning, as previously stated, is one of agreement with Schiffrin (1987) that DMs have "core meanings"; however, Fraser (2006) maintains that DMs can typically have both a conceptual meaning (Schiffrin, 1987) as well as a procedural meaning that specifies how the segment it introduces is to be interpreted relative to the prior segment. This is contrary to Fraser's (1999) initial stance and counter to Relevance Theory (cf. Blakemore, 1992, 1995), which analyzes DMs as having a solely procedural meaning (Del Saz-Rubio, 2007).

Blakemore (1987, 1992), who adopts a Relevance Theory framework from Pragmatics, refers to DMs as discourse connectives. Blakemore (1992) claims that DMs only have a *procedural* meaning and not a *conceptual* one, as Schiffrin (1987) has claimed, and are limited to specific contexts (Fraser, 1999: 936). Blakemore (1992) also argues that the types of meanings a DM encodes are mutually exclusive: procedural meanings stipulate instructions of how to manipulate conceptual interpretations and conceptual meanings specify substantive concepts (Fraser, 2006). Relevance theorists assert that as DMs encode procedures for the relation between S2 and S1, they cannot also encode a conceptual meaning. Therefore, in Blakemore's (1992) view, DM expressions such as '*in addition*', '*as a consequence*', '*as a result*' and '*contrary to expectations*', which encode conceptual meaning, cannot as such also encode a procedural meaning and thus are not DMs.

Aijmer (2002) agrees with Fraser's (1999) claim that DMs do not have propositional meanings and states that if a DM were to express a meaning it must be procedural (Aijmer,

2002: 16). She suggests that the core meaning (cf. Schiffrin, 1987; Fraser, 1999) can be assigned when the expression operates as a DM.

Defining a core meaning, also referred to as ‘basic meaning’ (cf. Fox Tree and Schrock, 2002) is regarded in this study as irrelevant to the research aims. Instead, as other researchers have done, I have chosen to take a “middle solution” (Jucker, 1993: 437) in which the different (sub)functions of a DM are listed while only pointing out the common core if determined.

(9) Multifunctionality

Research has shown that generally DMs under scrutiny fulfill more than one function or at least have sub-functions (Brinton, 1990: 8). These functions may signal a local relationship between S1 and S2; and/or, signal a global relationship across a wider span of discourse (cf. Schiffrin, 1987; Fraser, 1999). Erman (2001: 1338) states that the multifunctionality of DMs is not a controversial issue in the literature (cf. Schiffrin, 1987; Fraser, 1990; Redeker, 1990; Brinton, 1990) and this is a characteristic I accept and assume to find in my study.

Aijmer (2002: 3) states that DMs can be associated with a large number of pragmatic values in contrast to ordinary words and yet, “[native] speakers are not troubled by this multifunctionality”. She further clarifies that while native speakers do not seem to be troubled by this and use DMs frequently in their everyday conversations, non-native speakers are said to have difficulty with the multifunctionality of DMs (Aijmer, 2002: 3).

While this list includes the most common features attributed to be characteristic of DMs, it must be noted that these features alone are not sufficient for the identification of all DMs. Consequently, Jucker and Ziv (1998) and Jucker (2002) suggest viewing DMs on a scale in which these features may be used to distinguish between prototypical (i.e. exhibiting most or all features) and more peripheral members (i.e. exhibiting fewer or limited extent of features) of the DM class (Jucker 2002: 211).

2.4.3 Approach to DM Analysis in this Study

To sum up, the criteria which identify and define a lexical expression as a DM are taken, for the purposes of the present study, to be as follows:

1. A key defining feature of DMs is that they are multi-functional. DMs are used to organize and structure discourse at both a local and global level through multiple functions which include initiating, developing or closing ongoing discourse. The same DM, such as ‘*so*’, can be used to perform different functions in different contexts.

Initiating a new (sub)topic

- a) **So**, this is just the region, central nervous system, peripheral nervous system

Developing on a (sub)topic

- b) **So** similarly this light has information, right? The light ...

Summarizing a (sub)topic

So, when these inputs are received by the vestibular nuclei they are

- c) processed, then an...

2. Syntactic independence is also a key defining feature in the identification of DMs (Schiffrin, 1987; Fraser, 1999; Brinton, 1996; Aijmer, 2002). A lexical expression such as ‘*okay*’ must have a non-truth conditional sense (i.e., little to no semantic meaning) which does not contribute to the truth or falsity of the utterance in which they occur. In this way, DMs occur outside the syntactic structure or are loosely attached to an utterance and their presence functions in relation to ongoing talk and text (cf. Schiffrin, 1987) to provide linguistic cues on how the utterances are to be interpreted. By removing a lexical expression the “grammatical well-formedness” of the relevant text is not affected; however, it does remove a powerful signal on how ongoing text is to be interpreted.
3. The position of DMs is flexible - they can occur as initial, medial or final-utterance (though certain utterances may be found to occur in limited positions).
4. DMs can be drawn from a number of word classes, and their position and presence (rather than word class) can be used to identify the functional relationship of a DM in ongoing text in

context. Lexical expressions are defined as DMs when based on their surrounding context they were not used in their usual meaning (i.e., indexicality) and function to organize discourse at a local level (i.e., utterance structuring) and at a global level (i.e., overall structuring) (cf. Schiffrin, 1987).

It must be noted that no one criterion alone can be used to determine the status of a lexical expression as a DM; instead, a combination of criteria was taken into consideration.

Within the study reported in this thesis, although prosody is not identified as a defining feature, it has been used to support the identification of DMs by observing the tone units of these lexical expressions when possible. Punctuation has been employed in the transcriptions to mark for example rising intonation (*okay?*) and falling intonation (*okay.*) boundaries. Furthermore, phonological reduction has been observed with the DM ‘*because*’ (‘*cause*’); however, as Schiffrin (1987) appeared to find no functional variations between ‘*because*’ and its reduction ‘*’cause*’, this feature (phonological reduction) was regarded as a peripheral criterion.

2.4.4 Functional Paradigms of Discourse Markers

One of the first published studies to provide a model of DMs based on an analysis of spoken academic lecture discourse was by Murphy and Candlin (1979). Their study investigated how the rhetorical organization of discourse is signaled in engineering lectures and is intended to have specific application to programs for students for whom English is a Second (or foreign) Language (ESL/EFL). Murphy and Candlin's (1979) discourse analysis is developed based on the primary school model by Sinclair and Coulthard (1975) which describes the discursive and coherence features of teacher-student interaction in classroom discourse. Sinclair and Coulthard's (1975) study found that teachers regularly marked boundaries of units of talks with extra-clausal items such as ‘*right*’, ‘*now*’ and ‘*okay*’ which were accompanied by prosodic features, e.g., pitch level changes (a characteristic of DMs discussed in 2.4.2).

Murphy and Candlin (1979) proposed a classification based on academic lectures that they believed showed similarities to the classroom interactive acts from Sinclair and Coulthard's

(1975) model. Murphy and Candlin (1979) note that in both settings the situation is fairly formal and “one participant has the floor whenever he wants it” (Coulthard, 1975 cited in Murphy and Candlin 1979: 13). The following are examples they provide to illustrate the similar features:

marker:	(signaling devices that segment discourse)	<i>Well. Obviously...</i> <i>Right. Everybody...</i> <i>Now. Let me...</i>
starter:	(establish links among discourse)	<i>Well now. Let's get on with the engineering.</i>
metastatement:	(emphasize important information)	<i>let me sound reveille, I want to mention two types of generator.</i>
aside:	(attempt to deviate from the ongoing discourse)	running out of blackboard here
conclusion:	(final remarks)	So there you've got three forces which are in equilibrium.

(Murphy and Candlin (1979) examples taken from Flowerdew, 1994: 16)

Murphy and Candlin's (1979) classification focuses on how the rhetorical organizations of lectures are signaled and reflect the interactive nature of academic lectures. They note that the lecturers often speak as though involved in a dialogue “providing dummy responses and feedback by themselves” (Lebauer, 1983: 5), a feature that was noted in this study initial analysis and was regarded as pseudo-dialogue. The presence of DMs, such as those found, are indicative of a lecturing style more common with Dudley-Evans and Johns' (1981) *conversational* rather than *reading* style and are used in this study, in addition to other criteria, to determine the degree of interactivity in the lectures.

Another study often cited by researchers investigating DMs in academic lectures is that of Chaudron and Richards' (1986), which is based on an analysis testing ESL students' level of comprehension with the presence and absence of DMs in four scripted lectures on American history. The prepared scripts included: a baseline scripted version without DMs, a version of the text with macro-markers added, a version of the text with micro-markers added and a version of the text with both macro- and micro-markers added (cited in Flowerdew and Tauroza, 1995).

Chaudron and Richards' (1986) suggest that the didactic focus of academic lectures, their structuring and the organization of information within the lectures, all play an essential role in its comprehensibility to listeners. They analyze the influence of two types of DMs, micro-markers and macro-markers, on lecture comprehension. Chaudron and Richards (1986) define micro-markers (e.g., *ok, now, well, so*) as intersentential relations framing segments and function as pause fillers. The macro-markers (e.g., *Now where are we, You probably know that*) were defined as explicit expressions which signaled major transitions. Nattinger and DeCarrico's (1988) note that Chaudron and Richards' (1986) results indicate that the length of a phrase seemed to be a factor in classifying macro-markers (e.g. *on the other hand, you can see*) and show in their study that length is, in fact, not a reliable criterion for DM classification.

Chaudron and Richards (1986) suggest that the micro-markers they have identified play no significant role in learners' retention due to the function of these markers that they maintain add neither salience nor meaning to lecture content. They also suggest that the scattered nature of these micro-markers throughout the lecture "probably" makes the entire lecture seem "less well organized" (Chaudron and Richards, 1986: 123). However, one issue with their findings is that the researchers' text was scripted to be 'performed' in a reading style and the scripts included artificially inserted DMs by the researchers. This brings into question the researchers' conclusions of how accurate a scripted spoken text can be in terms of its impact on listener comprehension. Their presumption that micro-markers may affect learners is based on a notion implied by the work of Heller et al. (1969) that hesitation markers and vagueness detract from learners' retention of instructional content (cited in Chaudron and Richards, 1986: 123).

Chaudron and Richards' (1986) characterization of DMs as micro- and macro-markers are based on their view that 'micro-markers' are simply pause fillers. I view this as a weak and narrow definition. Research studies support that DMs they have listed as 'micro-markers' such as '*so*' (Rendle-Short, 2003) and '*well*' (Fuller, 2003) are multifunctional and can play a role in framing or boundary marking in discourse. Bellés-Fortuño (2006: 123) criticizes their taxonomy as based mainly on semantic categories and failing to express the pragmatic or discursive discourse relations (i.e. fails to take into consideration the speaker's shifts from segments, parts or information topics). In Schiffrin's (1987) view, DMs must be viewed as having both semantic and pragmatic meaning and serve an integrative purpose in discourse.

Morell (2000) takes Murphy and Candlin's (1979) and Chaudron and Richards' (1986) taxonomies one step further by examining these two sets of DMs in two lecture styles: a participatory (interactive) and an explanatory (non-interactive) style. With consideration to previous studies, Morell (2000: 328) investigates the role of DMs (Chaudron and Richards, 1986; Flowerdew and Tauroza, 1995; Murphy and Candlin, 1979), personal pronouns (Rounds, 1987a, 1987b; Young, 1994), questions (Athanasiadou, 1991; Brock, 1986; Long, 1981) and negotiation of meaning (Long, 1983; Pica, 1994; Pica et al., 1987) in 15 English language lectures in Spain given by both NS and NNS English lecturers.

The lectures in Morell's (2000) study can be described in her words to be “on two ends of a continuum”, meaning they can be viewed as conventional explanatory (non-interactive) or participatory (interactive) lectures. Goffman (1981: 165) defines conventional non-interactive lectures as:

“[I]nstitutionalized extended holdings of the floor in which one speaker imparts his views on a subject, these thoughts comprising what can be called a ‘text’. The style is typically serious and slightly impersonal, the controlling intent being to generate calmly considered understanding, not mere entertainment, emotional impact, or immediate action.”

Participatory lectures, on the other hand, which involve more interaction between speakers and hearers are described by Northcott (2001: 19–20) as:

“used to signify a classroom event for a large (more than 20) group of students primarily controlled and led by a lecturer and including subject input from the lecturer but also including varying degrees and types of oral participation by students.”

Morell's (2000) distinction between the two types of lectures is based on the degree of formality and number of student interactions during the lecture. She divides the functions used to classify micro-markers based on Chaudron and Richards' (1986) taxonomy (e.g., segmentation, temporal, etc.). However, her macro-marker classification is based on Murphy and Candlin's (1979) taxonomy (e.g., starter, elicitation, accept, attitudinal, metastatement, conclusion). Morell (2000) proposes that while DMs are often viewed as solely textual markers used to guide listeners

in their comprehension of spoken discourse, they also act as interpersonal features used to negotiate meaning. Pica (1994: 494) describes the process of negotiating meaning between participants as repeating discourse, adjusting syntax, changing or modifying words in form or meaning to arrive at mutual comprehension. We would expect this type of negotiation to occur in interactive lectures and therefore would expect a higher frequency of DM occurrences (in contrast to non-interactive lectures).

Morell's (2000) results indicate that interactive lectures, which are generally less formal and resemble spoken discourse more, show a greater and varied occurrence of elicitation markers in the interactive micro-marker category (e.g. *anything else?*, *what?*, *why not?*, etc.) than its non-interactive counterpart which had a greater degree of formality (e.g. *ready?*). A similar result is found between non-interactive and interactive macro-markers, in which acceptance markers, which are absent in the former, are present in the latter's interactive lectures (e.g. *that's right*, *that is true*).

Morell (2000) suggests that the increased use of interactive linguistic features in lecture discourse can shorten the distance between the lecturer and students and *personalize* the lectures (ibid: 335). Specifically, the use of elicitation questions not only signals to students that participation is welcome but that it is also requested. The use of linguistic features in lectures plays a role in the interactive and non-interactive nature of the lecture and appears to be a prerequisite for promoting student participation (Morell, 2000: 335).

Bellés-Fortuño (2006: 157) proposes a classification of DMs, with her analysis of North American and Spanish academic lectures, based on Halliday's (1994) three functional components of meanings (ideational, textual and interpersonal) and the relations they can convey along the discourse utterances. Her decision for a Hallidayan approach was stated as motivated with the intention of providing a "valid taxonomy [...] largely valid and universal to any language and social culture" (ibid: 157).

Bellés-Fortuño's (2006: 158-159) DM model consists of three categories, namely, micro-markers (ideational) (Schiffrin, 1987; Fraser, 1990), macro-markers (textual) and operators (interpersonal). Briefly, her first category *micro-markers*, a term she borrows from Chaudron and

Richards (1986), expresses the logico-semantic relations (cf. Schiffrin, 1987; Fraser, 1990) in the discourse (see Table 2.4). Her second category, *macro-markers* (also borrowed from Chaudron and Richards) conveys structural relations through five functions (see Table 2.4). The final category, *operators*, “rhetorically signal the speaker’s intention and goal (illocutionary force) as long as they play a dominant inferential role in the discourse [and] frequently [monitor] proximity between speaker-hearer and speaker-speech” (Bellés-Fortuño, 2006: 160).

Micro-markers
↓

Internal (ideational) relations

Additional	Temporal	Causal	Contrastive	Consecutive
<i>and</i>	<i>then</i>	<i>because (of)</i>	<i>however</i>	<i>so (that)</i>
<i>or</i>	<i>after</i>	<i>since</i>	<i>but</i>	<i>then</i>

Macro-markers
↓

Overall discourse structural relations

Starter	Rephraser	Organizer	Topic- Shifter	Conclusion
<i>First (of all)</i>	<i>I mean</i>	<i>let us try</i>	<i>so</i>	<i>finally</i>
<i>Let's begin</i>	<i>in other words</i>	<i>I want to discuss</i>	<i>now</i>	<i>to end/finish up/ with</i>

Operators

↙

Relation speaker-speech

Attitudinal	Pause Filler	
<i>as you know</i>	<i>and</i>	
<i>I/we think</i>	<i>well</i>	

↘

Relation speaker-hearer

Elicitation	Acceptance	Confirmation Check
<i>any questions?</i>	<i>okay</i>	<i>okay?</i>
<i>why is that?</i>	<i>right / alright</i>	<i>right? alright?</i>

Table 2.4 Bellés-Fortuño's (2006) Classification of DMs

Bellés-Fortuño's (2006) classification provides us with a well-founded report of three levels of DMs which were found to occur in North American social science lectures. It establishes firstly, that many of the DMs which occur in the present study are consistent with the results found by

Bellés-Fortuño (2006) in her North American corpus. Secondly, while I ultimately chose to limit my classification to two main categories (discussed below), Bellés-Fortuño's (2006) taxonomy provided valuable insight on an approach to academic DM analysis in tertiary education.

The previous studies discussed so far in this section have reported that the presence of macro-markers (i.e. textual markers) improves retention and recall in post-lecture tests (Chaudron and Richards, 1986; Jung, 2003), that it is generally beneficial for “activating content schemata” (Nattinger and DeCarrico, 1988) and that it has been found to help NNS students successfully follow the lecture discourse (Khuwaileh, 1999). Moreover, that interactive lectures have greater and more varied occurrences of DMs appears to be a prerequisite for promoting student participation (Morell, 2000: 335).

The taxonomies presented (Murphy and Candlin, 1979; Chaudron and Richards, 1986; Morell, 2000; Bellés-Fortuño, 2006) which analyze DMs on micro- and macro levels were informative; however, based on the initial analysis of the exploratory study (see 4.5.1 and 4.5.2), I found it sufficient to regard both micro- and macro- markers as (sub-categories of) Textual markers (cf. Aijmer, 2002). On this Textual level, DMs operate at both a *local* (ideational) and *global* (overall structuring) coherence level (Aijmer, 2002). Therefore, a revision of more sociological and coherence based functional analyses of DMs such as Aijmer (2002), Fung and Carter (2007) and Brinton (1998) was conducted and is presented as follows.

To begin, we must first review Schiffrin (1987), one of the lead researchers who adopts a coherence-based account of DMs. She maintains that DMs contribute to the coherence of spoken text by establishing coherence relations between units of discourse (Schiffrin, 1987: 9). The multi-planed discourse model she proposed reflects the different levels English DMs may (simultaneously) function at, namely, the Model of Discourse Coherence (Schiffrin, 1987: 24-28):

1. *Participation Framework* is the different ways speakers and hearers can relate to one another due to their mutual presence and shared responsibility for discourse and its production (Schiffrin, 1987: 27). For example, in this framework, speakers evaluate ideas and stances in performing actions and claiming turns. The DM ‘*I mean*’ has its primary function in the participation framework (ibid: 316).
2. *Information State* is the central roles speakers and hearers play in the cognitive capacities related to discourse interaction. This differs from the participation framework which

relates to social interactional aspects of discourse. The information state is described as interactive processes that are constantly in a flux because discourse involves an exchange of information (Schiffrin, 1987: 28). The DMs ‘*oh*’ and ‘*you know*’ have their primary use in the information state (ibid: 316).

3. *Ideational Structure* is where units within this structure are semantic (propositions). Three different relations in which propositions contribute to the overall configuration are: cohesion, topic and functional relations.
4. *Action Structure* is indicative of speech acts which are situated, not only in terms of speakers’ identities and social settings, but in terms of what speech acts precedes, what action is intended to follow and what action does in fact actually follow (Schiffrin, 1987).
5. *Exchange Structure* is described as the outcome of the decision procedures of speakers by which they alternate sequential roles and define those alternations in relation to each other (Schiffrin, 1987).

Schiffrin's (1987) model specifies the different planes of discourse and is used to illustrate how the planes are interconnected. The model focuses on the construction of local coherence through adjacent units in discourse but “can be expanded to take into account more global dimensions of coherence” (Schiffrin, 1987: 24). She lists the DMs she analyzed based on the planes of discourse they function on, noting that DMs can be used on more than one plane (Schiffrin, 1987: 316). The *asterisk in Table 2.4 highlights where the DM is used more frequently on one plane than on another (e.g. **well* is used more in the Participation Framework than any other plane).

Informational State	Participation Framework	Ideational Structure	Action Structure	Exchange Structure
<i>well</i>	<i>*well</i>	<i>well</i>	<i>well</i>	<i>well</i>
		<i>*but</i>	<i>but</i>	<i>but</i>
<i>so</i>	<i>so</i>	<i>*so</i>	<i>so</i>	<i>so</i>
		<i>*and</i>	<i>and</i>	<i>and</i>
<i>i mean</i>	<i>*i mean</i>	<i>i mean</i>		
	<i>now</i>	<i>*now</i>		
<i>because</i>		<i>*because</i>	<i>because</i>	
<i>*y’know</i>	<i>y’know</i>	<i>y’know</i>		<i>y’know</i>

Table 2.5 Schiffrin’s (1987) DMs Multifunctionality on Planes of Discourse

In the case of '*because*', the DM is said to appear on three planes of discourse: Information State, Ideational Structure and Action Structure. The DM has three different semantic realizations in discourse (examples taken from Schiffrin, 1987: 202):

On the Information State plane, '*because*' has a meaning of *warrant*, as in:

(1) John is home *because* the lights are burning. (Schiffrin, 1987: 202)

And on the Ideational Structure, '*because*' has a meaning of *cause*, as in:

(2) John is home *because* he is sick. (ibid: 202)

And on the Action Plane, '*because*' has a meaning of *motive* (of speech act), as in:

(3) Is John home? *Because* the lights are burning. (ibid: 202)

Schiffrin (1987: 57) asserts that DMs may operate on a single plane or across different planes of discourse and integrate structural, semantic, pragmatic and social factors. I agree with Schiffrin's (1987) concept of *local* and *global* coherence and agree that DMs are multifunctional on different planes of discourse which helps integrate different simultaneous processes underlying discourse and helps create coherence. Schiffrin (1987: 49) describes the analysis of DMs as part of the more general exploration of discourse coherence - how speakers and hearers jointly integrate forms, meanings and actions to make overall sense out of what is said. What this suggests is that DM identification and classification must be based on its function with its surrounding context. This is adopted in respect to analyzing the present study's DMs within their contexts and not in isolation as they "index an utterance to the local contexts in which utterances are produced and in which they are to be interpreted" (Schiffrin, 1987: 326).

Schiffrin's (1987) sociological and coherence based approach with DMs is considered a departure point for a vast number of papers on this subject. Her work has been criticized as containing numerous shortcomings, including why it was necessary to propose "five planes of talk" (Schiffrin, 1987: 24). Fraser (1990: 384) viewed Schiffrin's definition of her discourse planes as one of the main flaws because they are:

"[defined] in broad and general terms. She does not provide a systematic motivation for them and one is left with a concern that they both overlap and leave areas of interpretation unaccounted for."

In this vein, Redeker (1991) attempts to revise Schiffrin's Model to one which distinguishes three components of discourse¹⁴: Ideational Structure, Rhetorical Structure¹⁵ and Sequential Structure¹⁶. Although Redeker (1991) acknowledges that Schiffrin's approach successfully distinguishes between the ideational (representational) functions of language and the socio-pragmatic functions of exchange and structure, she criticizes the information state and participation framework planes as being somewhat unclear.

Aijmer (2002: 13), like Redeker (1991), admires Schiffrin's (1987) integrated approach which explains the behavior on five different levels, however, she finds it sufficient to restrict her DM analysis to two levels. Aijmer's (2002) account of DMs¹⁷ are indexed to attitudes, participants and to the text (cf. Schiffrin, 1987) and she therefore proposes that they function on a *Textual* and *Interpersonal* level.

Aijmer's (2002) overall approach to DMs is, as previously stated, corpus-based and uses a bottom-up framework. She bases her study on data from the London-Lund Corpus of spoken English and focuses her analysis on topic changer 'now', interjections 'oh' and 'ah', interpersonal particle 'just', the adjuster 'sort of', expectation marker 'actually' and tags such as 'and that sort of thing'. The DMs were chosen for their frequency in the London-Lund Corpus.

Fung and Carter (2007) embrace a multi-categorial framework which adopts a functionally-based account of DMs and is grounded on both Schiffrin's (1987) multi-dimensional model of coherence and Aijmer's (2002) interpersonal perspective. Fung and Carter's (2007: 415) theoretical framework categorizes DMs under four functional headings (See Table 2.6) and notes that any instance may perform one or more of these functions as shown by the previous work of Andersen (2001), Lenk (1998) and Jucker and Ziv (1998).

¹⁴ Redeker (1991: 1162) views Schiffrin's *information* structure and *participation* framework as "clearly not on par with the other three planes [*ideational* structure, *action* structure and *exchange* structure]".

¹⁵ Roughly equivalent to Schiffrin's (1987) *Action* Structure (Fraser, 1999: 935)

¹⁶ She considers two discourse units as being *ideationally* related "If their utterance in the given context entails the speaker's commitment to the existence of the relation in the world the discourse describes" and *rhetorically* related "if the strongest relation is not between the propositions expressed in the two units but between the illocutionary intentions they convey", and *sequentially* related in either a paratactic (transition between issues or topics as a preplanned list or locally occasioned) or hypotactic (leading in or out of commentary: correction, paraphrase, aside, digression or interruption segment). (Redeker, 1991: 1168)

¹⁷ Aijmer refers to DMs as *discourse particles*

Macro-level	Function	Discourse Markers
Interpersonal	Marks shared knowledge	<i>see, you see, you know, listen, etc.</i>
	Indicators of agreement, confirmation, and acknowledgement	<i>okay, oh, right, yeah, I see, etc</i>
Referential	Marks the relationship between utterances which precede and follow a DM	conjunctions: cause (<i>because/cos</i>), consequence (<i>so</i>), contrast (<i>but, and</i>), coordination (<i>and</i>), disjunction (<i>or</i>), digression (<i>anyway</i>), etc.
Structural Cognitive	Signals links and transitions between topics	<i>now, okay, right, well, let's start, let's discuss, let me conclude the discussion, first, then, finally, etc.</i>
	Marks continuation of the current topic	<i>yeah, and, cos, so</i>
Cognitive	Denotes thinking process	<i>well, I think, I see, I mean, in other words, sort of, you know, and</i>

Table 2.6 Discourse Marker Functional Categories (Fung and Carter, 2007: 418)

The work of both Aijmer (2002) and Fung and Carter (2007) acknowledges that DMs, in addition to aiding coherence, can function (sometimes simultaneously) on a personal level to indicate speaker interest or lack thereof. And while I find Fung and Carter's (2007) four functional levels of DMs to be clear and useful, the DMs in this study did not function solely as Structural Cognitive or Cognitive. Thus the categories that were proposed for this study's analysis were *Textual* and *Interpersonal* (cf. Aijmer, 2002; Müller, 2005). The following is a summary of Brinton's (1998: 12) classification of Textual and Interpersonal functions which lists the sub-functions of each in spoken discourse (see Table 2.7a-b).

Textual functions	<ol style="list-style-type: none"> 1) getting the hearer's attention 2) initiating and ending discourse 3) sustaining discourse, marking boundaries (topic shifts and episode boundaries) 4) constraining the relevance of adjoining clauses 5) repairing discourse
--------------------------	--

Table 2.7a Textual functions based on Brinton (1998: 12)

Interpersonal functions	<p>1) subjective functions such as expressing response, reactions, attitudes, understanding, tentativeness or continued attention</p> <p>2) interactive functions such as expressing intimacy, cooperation, shared knowledge, deference or face-saving (politeness).</p>
--------------------------------	--

Table 2.7b Interpersonal functions based on Brinton (1998: 12)

The paradigms discussed in this section provide a general review of relevant studies which focus on academic discourse and/or spoken DMs and illustrate the functional coherence approaches taken for DM analysis. Brinton's (1990, 1998) paradigm of discourse functions, Aijmer's (2002) and Fung and Carter's (2007) models of analyses, and Bellés-Fortuño's (2006) analysis of DMs in English/Spanish academic lectures contributed as informative points for my classification following the initial analysis.

My framework, which is discussed in more detail in the methodology chapter (see 4.8), is developed using a bottom-up approach (cf. Aijmer, 2002; Müller, 2005). Applying a bottom-up analysis provides priority to the DMs identified in my study to develop an organic paradigm as opposed to a top-down approach which gives priority to the adopted model and may result in disregarding DMs that cannot fit said model (Erman, 1987: 114). In this study, the attention was focused on first drawing out the data and then classifying the DMs sub-functions on categories labeled Structural (i.e., Textual) to Interactional (i.e., Interpersonal) Functional categories.

2.5 Discourse Markers in Pedagogic Settings

2.5.1 Introduction

In educational settings, DMs are found to have a positive role in classroom context as effective conversational endeavors (Othman, 2010); however, studies on DMs in teacher-talk are yet under-researched (Fung and Carter, 2007). In spoken discourse, the frequency and number of DMs that

speakers use are considerably greater compared with other word forms (Fung and Carter, 2007) and their presence creates a “naturalistic conversational effect” (Fox Tree and Schrock, 1999: 280). As one important element that constitutes and organizes discourse, DMs also work as effective interactional features (Schiffrin 1987; Fraser, 1999; Aijmer, 2002; Fung and Carter, 2007). The following sections review relevant studies that investigate linguistic features in spoken discourse, either focusing on, or inclusive of, research on DMs in pedagogic contexts.

2.5.2 Comprehension of Lectures: The Case of Discourse Markers

In Nattinger and DeCarrico’s (1988) study, the researchers investigated lexical phrases occurring in a variety of natural academic lectures including history, linguistics, biology, anthropology and literature among others (ibid: 93). They defined lexical phrases as representative of a higher level of information and describe them as macro-organizers. Nattinger and DeCarrico (1988) use the term macro-organizers to better illustrate the function of lexical phrases that help students mentally organize information as they listen and helps stress the importance of students’ awareness of lecture organization. Nattinger and DeCarrico (1988) divided the macro-organizers into eight categories on the basis of their function: topic markers, topic shifters, summarizers, exemplifiers, relators, evaluators, qualifiers and aside markers. They further divide these functional categories into local and global organizers (Nattinger and Decarrico, 1988: 94):

- i. *Global* Macro-organizers: indicate the overall organization of the lecture.
- ii. *Local* Macro-organizers: highlight the sequencing or importance of information within the framework at specific points set by the global organizers.

This further classification helps in distinguishing main topics from the explanations, examples, relations etc., which in turn serves as development and support for the topics (Nattinger and Decarrico, 1988). Nattinger and DeCarrico (1988) criticize Chaudron and Richards’ (1986) study for not classifying macro-markers, which they point out as being significant to lecture comprehension, as well as not considering their functions at a local and global level as they do. They report that one-word DMs such as, ‘*okay*’ and ‘*now*’ are found to function as global macro-organizers signaling a topic shift or summary (Nattinger and DeCarrico, 1988). They suggest

these lexical cues should be explicitly taught to enhance a student's ability to comprehend lectures and to help them predict what information is likely to follow next and to organize and interpret the flow of information more easily (Nattinger and DeCarrico, 1988: 91).

Olsen and Huckin (1990) conducted a similar study into the comprehension of L2 speakers of English in engineering lectures through a video recording of a lecture which built up an argument around a problem and showed how it was eventually solved. Fourteen postgraduate students watched the video and then provided an oral summary of the most important points in the lecture. Their analysis and conclusion was that while students understood "all the words of a lecture (including lexical connectives and other DMs) [they] fail[ed] to understand the lecturer's main points or logical argument" (1990: 33).

Olsen and Huckin's (1990) research suggests that students should not only be taught to recognize lexical connectives and DMs but more importantly, be taught appropriate listening strategies to distinguish between lecturers giving information (information-driven) and those that develop an argument through a problem-solving framework (point-driven). Their tentative conclusion, drawn from their study, is that while the teaching of DMs is advocated by researchers (cf. Chaudron and Richards, 1986; Nattinger and DeCarrico, 1988) and is certainly worthwhile, in and of itself is not "sufficient to make students aware of the discourse-level pragmatics of academic lectures" (Olsen and Huckin, 1990: 42).

Flowerdew and Tauroza's (1995) study investigated the impact of the presence and absence of DMs such as, '*so*', '*right*', '*well*' and '*ok*', on L2 lecture comprehension. Their results indicate that a spoken lecture with DMs present is comprehended better by listeners than the same lecture with the DMs edited out. This means that the DMs function to help listeners avoid the confusion that would arise if they tried to connect two disjointed utterances and provides cues on how to anticipate and process the relationship of utterances to a DM.

Yang (2007: 25) also examined the effects of discourse signaling cues (hereafter DMs) on the comprehension of 80 Korean EFL learners who took part in the study. Yang (2007) divided his participants into two groups with one half listening to the lecture with DMs and the other half without. The findings of the study supports previous research and indicates that the inclusion of

DMs allowed the group to perform significantly better than the non-signaled group with both micro- and macro-discourse comprehension.

A recent study in Iran (Tehrani and Dastjerdi, 2012), similar to Olsen and Huckin's (1990) study, investigated the comprehension of lecture discourse both inclusive and absent of DMs by EFL learners in a language institute in Iran. The researchers divided the 20 EFL students into two groups and gave lectures related to science and technology using DMs with one group and none with the second group. Each student was then required to write out what they understood from the lecture and these compositions were compared to analyze which group produced a more cohesive text. The results showed that the inclusive DM group used more cohesive devices in their writing and produced more coherent texts than their counterparts. This finding also supports previous research (Chaudron and Richards, 1986; Nattinger and DeCarrico, 1988; Olsen and Huckin, 1990; Flowerdew and Tauroza, 1995) that DMs are highly significant for learners' comprehension of lecture discourse.

Thompson (1994: 58) asserts that the nature of a monologic lecture places further demands on listeners' ability to coherently interpret the semantic relations lying beneath the surface of the discourse. Thompson (1994) submits that speakers can use linguistic elements such as clause relations, lexico-grammatical cohesion and intonation to help listeners by "explicitly signaling the underlying network of concepts in the monologue" (ibid: 58). She proposes that clause structures are important to show primarily the underlying semantic structure and interpretation of the text. The absence of explicit surface markers such as conjunctions in monologues, she argues, would pose difficulties for learners. Similar to Nattinger and DeCarrico (1988), Thompson (2003) argues that when both text-structuring metadiscourse and intonation are used by academic speakers, it supports the listeners' ability to form a coherent mental map and allows them to view how segments of discourse are interrelated.

In a foreign English context, Yu (2008) investigated the interpersonal meaning of DMs in a Chinese EFL classroom through a Hallidayan approach of systemic functional linguistics. In her study, Yu (2008) explored the uses and functions of DMs in the progress of teaching through six moves: opening, information checking, information clarification, responding, comment and

repetition. According to Yu (2008), the appropriate use of DMs can significantly improve the effective flow of information from teacher to students.

Thus it can be summarized that DMs are an important aspect of academic discourse and their presence serves as signals that highlight the cataphoric and anaphoric relationship between discourse utterances. As listeners process a stretch of spoken text, their interpretation depends upon a particular frame of reference or mental model (Segal et al., 1999); the role of DMs in this process provides signals of whether a text is interpreted as continuous or discontinuous with the current stretch (Flowerdew and Tauroza, 1995). Their presence in lectures has been found to help listeners extract more meaningful and comprehensible input (cf. Nattinger and DeCarrico, 1988; Flowerdew and Tauroza, 1995; Thompson, 2003; Yu, 2008).

2.5.3 DMs in Academic Discourse

Young (1997), in her study of university lectures, examined the form and content of a corpus of seven, two-hour lectures. Three of the lectures were delivered by NNS of English, while the remaining four were delivered by native English speakers in a North American university. Young (1997) distinguishes the macro-structure of the lectures as six phases and further divides the six into three *metadiscoursal* phases which “comment on the discourse itself” (ibid: 166); and three others which “mark university lectures” (ibid: 167).

METADISCOURSAL PHASES	
1. Discourse Structuring phase:	lecturer indicates the different parts or directions they will take in lecture
2. Conclusion phase:	lecturers summarizes main points they have made throughout the discourse
3. Evaluation phase:	lecturers reinforce each of the other phases through an evaluation of information which will be or has already been transmitted.

Table 2.8a. Metadiscoursal Phases Developed by Young (1994: 166-167)

According to Young (1994), these three metadiscoursal phases occur across disciplines and levels indicating that the relationship between speaker and hearers in this situation follows a particularly consistent macro-structure. The other three phases identified by Young (1994), which are referred to as *non-metadiscoursal*, are more specifically related to the actual lecture discourse:

NON-METADISCOUSAL PHASES	
4. Interaction phase:	the interpersonal strategies that the lecturer implements to establish and maintain contact with the students to ensure comprehensibility and reduce distance
5. Theory or Content phase:	where lecturers present the theories, models and definitions of the subject
6. Examples phase:	where lecturers illustrate theoretical concepts through concrete examples so that students are able to follow the ideas successfully

Table 2.8b. Non-Metadiscoursal Phases Developed by Young (1994: 166-167)

Young (1994: 173) claims - when referring to lectures - that “phasal analysis seems to offer a more realistic portrayal of the nature of this particular genre”. Young’s (1994) phasal model resembles the structure of Swales’s (1990) ‘moves’ in research article introductions. One major drawback of this type of analysis is its narrow view of the purpose of lectures beyond the conveying of theoretical information (Suviniitty, 2012). While Young’s (1994) phasal model provides an awareness of the potential structuring in lecture discourse, my study does not assume such clear structuring will be present in the corpora. An awareness of the phases though, helps serve as context for the evaluation and analysis of DM functions (e.g., introduction, conclusion, evaluation, interaction, etc.) within the academic discourse.

Swales and Malczewski (2001) studied the metadiscourse of spoken academic discourse from MICASE (Simpson et al., 2002) and focused on what they refer to as “a cluster of features that constellate around discourse management across a wide range of university speech events” (ibid: 146). They define speech events as activity types (Levinson, 1979) in which language is used to carry out a purpose such as in monologic (e.g. lectures, conference presentations, etc) or dialogic talk (e.g. seminars, interviews, etc). Swales and Malczewski’s (2001) results show how university lecturers rely on a variety of linguistic resources to mark topic and shift footing such as, group

vocatives (*folks, gang, friends, guys*), directive or vocative verbs (*say, listen, look*) and exhortative or jussive imperative (*look, let's*) and DMs (*okay, so and now*).

Fuller (2003) examined the use of the DMs 'you know', 'like', 'oh', 'well', 'yeah' and 'I mean' in two different speech events, namely, interviews and casual conversations, in order to determine the function of DMs in marking and negotiating speaker roles. The results for the DMs 'oh' and 'well' showed statistically significant differences in frequency between events, with both DMs occurring more frequently in casual conversations. Despite Fuller's (2003) initial classification of 'well' as a presentation marker, the data results indicate that 'oh' and 'well' both function primarily as reception markers and are used in casual conversations to create coherence between speaker-hearer turns. Jucker and Smith (1998: 174) distinguish between the markers by stating that "[r]eception markers [are] used to signal a reaction to information provided by another speaker" while "[p]resentation markers accompany and modify the speaker's own information." This supports Fuller's (2003) findings that the DMs 'you know', 'like', 'yeah' and 'I mean' are used at similar frequencies across speech events, indicating that the functions of these presentation markers are more universal.

Csomay's (2005) research showed that not only speech events, but also academic disciplines, could vary significantly in linguistic variation. Her findings showed that academic disciplines such as the humanities and social sciences are characterized by episodes associated with spoken discourse (e.g. contractions, that-deletion) and speaker stance (e.g. DMs: *you know, I mean*; mental verbs: *think, know*). Csomay (2005) suggests that if disciplinary variation exists within classroom discourse, it is likely to be evidenced quite strongly in linguistic features associated with spoken text and speaker stance. In an investigation of linguistic variation, Csomay (2005) compared university teaching and teaching level (e.g. graduate, undergraduate) and found that graduate courses exhibited the highest use of *contextualized, directive orientation* and *personalized framing* discourse units (Barbieri, 2013). This may be because graduate level students, as opposed to undergraduates, are more involved in dialogue with lecturers.

Since interaction varies across disciplines, level of instruction and class size (Csomay, 2002; Ferris and Tagg, 1996; Flowerdew, 1994) there is a need for students "to be made aware of the particular features that distinguish the practices of the discourse community they wish to become

members of” (Flowerdew, 1994: 157). The differences between knowledge production in the soft sciences (e.g. humanities, social sciences) and hard sciences (e.g. biology, medicine) have been shown to result in the hard sciences tending to be less dialogic in nature than the soft sciences (Chang, 2012).

On the other hand, Nesi and Basturkmen’s (2006) study, which investigated the cohesive role of lexical bundles in a range of academic disciplines in the US and UK, reported that their results showed that there are similarities with “most of the commonest bundles” (ibid: 289) across disciplines and country. Nesi and Basturkmen’s (2006) corpus consists of a 160 lectures from BASE and MICASE, selected from each of the four broad disciplinary groupings (Arts & Humanities, Social Sciences, Life & Sciences and Physical Sciences). They also reported that their findings are almost identical to that of Biber et al.’s (2004) study and that, similar to classroom teaching, academic lectures also use “a large set of different lexical bundles, while conversation relies on the extremely frequent use of a smaller set of bundles” (Biber et al., 2004: 379 cited in Nesi and Basturkmen, 2006: 289-290).

Othman (2010) uses a corpus-based approach to investigate the use of three specific DMs ‘*okay*’, ‘*right*’ and ‘*yeah*’ by NS lecturers at Lancaster University. Her study uses both transcribed audio/ video recording of the lectures and subsequent interviews with lecturers to verify Othman’s (2010) interpretation of lecture data from both the lecturer’s (speaker) and her own (hearer) point of view. Othman’s (2010) results support the functional significance of these three DMs in conversational interactions when organizing utterances. Othman (2010) observed, based on interviews, that university lecturers use DMs as signposts on a structural level when taking turns in lecturing as a subconscious behavior.

Bellés-Fortuño's (2006) PhD thesis investigated the uses and functions of DMs in a contrastive North American and Spanish academic lecture corpora from the Social Sciences discipline. Bellés-Fortuño (2006) adopted a corpus-based Hallidayan approach to her research and developed a taxonomy for DM classification (see section 2.4.4) based in part on Halliday’s functional model. The results of her study revealed that generally DMs occurred more often in the North American lectures than in the Spanish lectures.

In reference to Bellés-Fortuño's (2006) three type of DMs (see section 2.4.4), she found that, first, ideational DMs (i.e. 'micro-markers') are widely used in both corpora, possibly due to the nature of the lectures as monologic. Secondly, textual DMs (i.e. 'macro-markers') are the least frequently used category in English and Spanish. Bellés-Fortuño (2006) pointed out that this may be due to the linguistic expressions "which tend to be longer chunks of language, more unsteady and sensitive to changes and prone to form language expressions" and "may vary from one language to another, not having clear corresponding counterparts" (ibid: 256). Thirdly, the last category, interpersonal DMs (i.e. 'operators') are the second highest in in terms of frequency of occurrences in her sub-corpora and indicate "a quite important position" (ibid: 257), as they can be said to illustrate a trend towards a more *open* lecture style.

One critical question Bellés-Fortuño (2006) raises about textual DMs (i.e. 'macro-markers') is whether its position as the least frequent DM in the sub-corpora indicates that it "[has] less relevance in the spoken academic discourse of North-American and Spanish lectures" (ibid: 257). On the other hand, interpersonal DMs (i.e. 'operators') that rhetorically signal the speakers' intentions and affect illocutionary force, are very relevant to the sub-corpora (Bellés-Fortuño, 2006: 258).

Fung and Carter (2007) examined and compared the production of DMs by NS and NNS of English based on a pedagogic sub-corpus from CANCODE and a corpus of interactive classroom discourse of secondary students in Hong Kong. The results of their study indicate that, in both the NS and NNS groups, DMs function as useful interactional tools to structure and organize discourse on interpersonal, referential, structural and cognitive levels. Each group is shown to perform on different functional levels to aid in the development and management of discourse.

The results of the Hong Kong NNS learner corpus indicated a significant use of referentially functional DMs such as, 'and', 'but', 'because', 'ok', 'so', etc. On the other hand, the NNS speakers displayed a relatively restricted use of other markers such as, 'yeah', 'really', 'say', 'sort of', 'I see', 'you see', 'well', 'right', 'actually', 'cos', 'you know', etc. NS are found to use DMs for a wider variety of pragmatic functions and the study, therefore, also discusses some possible pedagogical implications involved in preparing learners to become more interactionally competent speakers.

Rendle-Short (2003), whose research is largely concerned with the areas of spoken interaction and discourse, explored a corpus-based analysis of the DM 'so' in monologic seminar talk within the computer science discipline. She illustrates two distinct functions of 'so' in various contexts based on Schiffrin (1987). First, the DM 'so' connects adjoining clauses on a semantic level to show causal relations and secondly, on a pragmatic level, 'so' marks the potential speaker's discourse transitions which can function as a topic-shifter (Rendle-Short, 2003). These functions of 'so' are analyzed from a corpus consisting of six video-taped and then transcribed computer science seminar talks. The results of Rendle-Short's (2003) analysis shows that 'so' plays different functions in seminar talk depending on its position in the talk (e.g. initial, medial or final). It also finds that monologic talk is not continuous, rather, it is divided into smaller parts or sections which have a finely organized and well-structured discourse pattern (Rendle-Short, 2003).

2.6 Summary

This chapter has discussed the key concepts, theories and studies that inform this research. We have established that academic lectures are a recognized genre of spoken discourse (Chang, 2012; Flowerdew, 1994) and that DMs are a salient feature of this discourse (Biber, 2006). DMs are shown to operate on multiple levels in lecture discourse and fulfill a role in spoken text similar to punctuation in written text (Flowerdew and Tauroza, 1995). Fung and Carter (2007) found that DMs function as useful interactional tools to structure and organize discourse with both NS and NNS. However, they also found that similar to other studies (Fuller, 2003; Müller, 2005) NS generally have a higher frequency of DMs in their spoken discourse than NNS and that NNS use DMs with a limited range of functions in comparison to NS.

CHAPTER 3

CODE-SWITCHING

“[T]o build rapport with individual pupils, create greater personal warmth and encourage greater pupil involvement, the teacher may... when the occasion is suitable, switch to the local language”.

(Ferguson, 2003: 43)

Chapter 3 Code-Switching

3.1 Introduction

In this chapter, I present a brief overview of the extensive literature on code-switching across the world in a variety of languages with respect to the focus of this research. The following sections will discuss the current definitions of code-switching, the varying views of what constitutes code-switching, the theories and stances towards code-switching and lastly, the roles it plays in the Middle East in academic discourse and with DMs.

3.2 Purpose of the Study: The Case of Arabic Code-Switching and Discourse Markers

The discourse of the Saudi-based academic medical lectures include English DMs (e.g., *okay*, *so*, *now*), which are the main focus of the research; however, these DMs do not always occur in isolation from Arabic DMs (e.g., *tayib* {okay}, *'fa'* {so}) or other language choices made throughout each lecture. The addition of these variables to the research will provide valuable insight into English/Arabic code-switching and the functions of English/Arabic DMs in academic medical discourse that are not currently available.

3.3 Background

In an academic community where English is used as the official medium of instruction (EMI), but is otherwise a foreign language in the greater community, it is to be expected that degrees of code-switching will take place. Code-switching (CS) has been defined as “the process whereby bilingual or bi-dialectal speakers switch back and forth between one language or dialect and another within the same conversation” (Trudgill, 1992: 16). Gal (1988: 247) describes CS as a conversational strategy used to establish group boundaries and create, evoke or change interpersonal relations. For example, bilingual communities have been observed using CS as a

mark of solidarity for situations between bi/multilingual friends or colleagues in informal contexts.

The circumstances that prompt the change in linguistic code have long been of interest to the research community and of great interest to myself as a serial code-switcher. Saudi Arabia has an estimated 9.1 million expatriates in the Kingdom (Naffee, 2013) and if Arabic is not an option, English or in some cases, a mixture of both codes, is often the only choice for communication. With the influx of English media and technology in Saudi Arabia, it has become commonplace to hear terms such as *missed call* or *mobile* in Arabic conversations, for example:

SpeakerA: “Ha’a3teeki *missed call*”
 {*I will give you(+female)*}
 SpeakerB: “*Tayib,*”
 {*okay*}

With the presence of English in Saudi Arabia as an unofficial foreign language, most specifically in academic environments, it is not surprising that CS occurs. In the study’s corpus of Saudi-based academic medical lectures, the presence of CS between English and Arabic is observed. The questions that arose during the collection and subsequent transcription of the academic lectures were first and foremost - why code-switch at that moment? And why choose to use a token, phrase or chunk of language in Arabic rather than English and vice-versa? While these questions cannot be definitively answered or the reasons fully understood in this study, it motivated a supplementary aspect, in addition to the use of English DMs in academic lectures, to investigate what has been claimed to be “an inevitable consequence of bilingualism” (Hudson, 1996: 51).

Researchers like Toribio (2004: 78) believe that bilingual speakers are generally aware of the DM equivalents in both languages and therefore their alterations can be classified as CS. Furthermore, these acts of CS are described as a conscious desire to “juxtapose the two codes to achieve some literary effect” (Toribio, 2004: 136). Occurrences of CS by a NS Arabic lecturer are illustrated from the lecture, *Drugs used for Anxiety and Panic Disorders (DAD)*, from the Saudi-based SASME corpus.

“Anxiety without cause. *Tayeb fe fe al* {Okay so so the} demarcation *ya3ni* {i mean} demarcation is not very clear between normal anxiety *ya3ni* {i mean} you are anticipating an exam or bad news...”

The potential motivations and functions of the CS to the L1 in the *DAD* Excerpt “...between normal anxiety *ya3ni* you are...”, are found to act as *attention-getters* and to clarify the purpose of upcoming utterances. The reasons for CS in this excerpt and other NS Arabic lectures will be presented in Results II (see 6.6). The study will present results of the Arabic CS in terms of the Arabic DMs used to initiate (start), connect (combine/link) or close (conclude) utterances in the academic discourse of the native Arabic speakers.

3.4 Definitions of Code-Switching

The phenomenon of CS has no doubt been a matter of interest from the first time language contact resulted in bilingualism. And with this interest, there has been much research with definitions varying extensively depending upon the perspective of the research. In some cases the terminology overlap or are differentiated by labels such as code-switching, code-mixing¹⁸ and borrowing¹⁹. In general, CS, often the umbrella term used, has no unified definitions though many share similarities with the conflict stemming from what the nature of CS could or should entail.

Gumperz (1982) defines CS as “the juxtaposition within the same speech exchange of passages of speech belonging to two different grammatical systems or subsystems” (ibid: 59) and describes the two general types of CS as Conversational (metaphorical) and Situational. *Conversational* CS is used to convey a speaker’s attitude towards the topic of conversation, whereas, *Situational* CS is when the situation determines the choice of language (Gumperz, 1982). Auer (1984) criticizes the distinction from both ends arguing that the relationship between language choices and situation is less rigid and conversationally less independent and thus must be replaced by a continuum including these parameters rather than just two rigid parameters (ibid:

¹⁸ Code-mixing can be defined as applying the grammatical structures on language A to language B. For example, Arabic children raised bilingually use ‘*al3abing*’ a mix of the Arabic word ‘*al3ab*’ meaning play and the English grammatical ‘-ing’ present continuous form.

¹⁹ Borrowing can be defined as adopting a token from language A as language B may or may not have an suitable equivalent. An example of borrowing in Arabic is the word ‘*computer*’, pronounced in Arabic as ‘*cumbuiter*’.

91). Auer (1995: 116) proposes that in order to effectively analyze and theorize the meaning of CS one must interpret each incident in respect to the utterances that precede and follow the code-switch.

Poplack's (1980) definition comes mainly from a grammatical perspective defining CS as a term used within the field of bilingualism or multilingualism to refer to the alternative use of two or more languages in a single discourse, sentence or constituent (ibid: 583) through morphological integration. Poplack's (1980) research, like others (Pfaff, 1976, 1979; Sankoff and Poplack 1981; Timm, 1975), focuses on CS in terms of its systematic linguistic behavior (see 3.5).

Myers-Scotton (1993) defines CS as "the use of two language varieties in the same conversation" (ibid: 239) and approaches her research in terms of viewing language as an interactional negotiation in which speaking is seen as a rational process involving decisions (Myers-Scotton, 1983: 115). Myers-Scotton (1983) developed a model that provides a description of variation in linguistic code choices in terms of markedness and is presented in the next section (see 3.5).

The research definitions are similar in regards to CS; however, as this study mainly focuses on the strategic nature of CS in regards to interactional functions, I adopt an inclusive approach based on Myers-Scotton (1983) and Auer (1995). I follow Myers-Scotton's (1983) definition of CS and furthermore, accept her proposal that variations of CS may have no single motivation that can be marked, unmarked or exploratory. I also accept Auer's (1995) basis for CS analysis that "any theory of conversational code-alternation is bound to fail if it does not take into account that the meaning of code-alternation depends in essential ways on its 'sequential environment'" (ibid: 116). The concept is similar to Schiffrin's (1987) and Fraser's (1999) analysis of DMs with respect to its surrounding context.

3.5 Characteristics and Terminology

Research into the communicative and social implications of CS (Blom and Gumperz, 1972; Gumperz, 1977; Myers-Scotton, 1993a; Romaine, 1995; Auer, 1995) suggests that there are various external and stylistic motivations behind changing languages. Bussman (2000: 78-79) describes the use of CS to contextualize a message of informality in a formal situation or irony in a serious situation. Conversations can sometimes instigate CS when the topic turns to aspects of religion, politics or culture (Timm, 1975) or in other situations, extralinguistic factors such as age, sex, ethnic/sociocultural identity, etc., may play a role (Pfaff, 1976).

Other focuses of research concentrate on the grammatical aspects of CS (Poplack, 1980; Myers-Scotton, 1993b; Myers-Scotton and Jake, 1995; Schmitt 2000). Poplack's (1980) study of Spanish/English bilingual utterances was the first to suggest principles constraining a switch. The first principle is the *Free Morpheme Constraint* in which language codes may be switched after a lexeme, provided that the constituent is not a bound morpheme, otherwise, switches are limited to word boundaries. The second principle is the *Equivalence Constraint* in which code-switches tend to occur at points in discourse where juxtaposition of L1 and L2 elements do not violate a syntactic rule of either language (Poplack, 1980: 586). This complexity suggests that a speaker must have a sophisticated knowledge of both the L1 and L2 grammar and how each language maps onto one another in order to code-switch (Poplack, 1980). Poplack's (1980) research of New York Puerto Rican speakers indicated that low proficiency speakers tended to use less sophisticated CS with single word borrowing and tags, while more proficient bilinguals code-switched at a phrase and clause level.

Myers-Scotton (1993, 1995), who is best known for her research of Swahili/English bilinguals in Eastern Africa, developed her lexically based Matrix Language Frame (MLF) Modal, which in contrast to Poplack (1980), suggests that speakers are not required to have a complete understanding of the complex grammatical structures in both languages. Myers-Scotton (1993) explains her MLF Modal as one in which the languages play an unequal role, the more dominant language described as the Matrix language (ML) while the other language is described as the Embedded language (EL). In other words, the distribution of the languages are asymmetrical with

the ML providing the abstract grammatical frames into which the EL is inserted. The MLF modal is based on an assumption that mixed code utterances will have an identifiable dominant or Matrix Language (ML) and that an asymmetry will always exist between the ML and the Embedded Language (EL). Auer (1998: 3) who accepts the notion of a ML agrees that “balanced proficiency is by no means a prerequisite” for CS and in fact, suggests even participants with “a very limited knowledge of the ‘other’ language” are just as capable of CS. This may be descriptive of NNS learners who use CS in the academic classroom in an attempt to please the lecturer with the L2, but rely on the L1 to clarify their message.

Auer (1998) proposes that code-switching may be divided into two functions:

1. **Discourse-related** CS: “the use of code-switching to organize the conversation by contributing to the interactional meaning of a particular utterance” (Auer 1998: 4).

For example:

Lecturer: *The standard for diagnosing any bacterial infection is what?*

Students: <silence>

Lecturer: ***Ma huwa al*** *world standard to diagnose any bacterial infection?*

{What is the}

Student: *culture*

2. **Participant-related** CS: Takes into account “instances of diverging language preferences” (Auer, 1998: 2) when speaker 2 accepts the language choice of speaker 1 and following that point only language A is used. Auer (1998: 2) proposes the following pattern as an illustration:

A1 B2 A1 B2 A1 // A2 A1 A2 A1.

This participant-related CS pattern exemplifies a pattern of negotiation between participants of which language-of-interaction (Matrix Language) will be established and then accepted following that point (Martin-Jones, 1995: 99). In some contexts, the participant-related CS is particularly salient, for example in communities where the EMI is a foreign language within the academic environment, but, outside this context, students have limited exposure to the foreign language. This presents an enormous challenge for the instructors and the students (Martin-Jones 1995: 100).

3.6 Theory of Code-Switching

A major contribution to CS research has been made by Myers-Scotton (1993, 1998) who proposed the *Markedness Model* based on Grice's (1975) *Cooperative Principle*²⁰. Myers-Scotton (1983) regards speaking, not as situated behavior but instead as interactional behavior in which "code choices are always situated [however] they are not a function of situation, but rather of negotiation, so that speaking is seen as a rational process involving decisions" (ibid: 115). Myers-Scotton (1983) develops her markedness theory as a means to describe the variation in linguistic code choices of speakers that have no single motivation but instead are a process of negotiating their identities based on:

...the premise that participants in conversation interpret all code choices in terms of a natural theory of markedness. That is, part of their communicative competence (Hymes 1972), speakers recognize choices as either unmarked or marked in reference to the norms of their speech community". (Myers-Scotton, 1983:115)

Myers-Scotton's (1983) model proposes that a speaker's motivation in using a language code is dependent on speaker, addressee, situation and the marked vs. unmarked choice. A speaker's negotiations of marked and unmarked code choice are bound by Myers-Scotton's (1983) theory of Rights and Obligations (RO), which are a construct that stands for a set of social features²¹ between a speaker and addressee.

For example, in Saudi Arabia, where the majority of the student population are native Arabs, an Arab bilingual lecturer in a medical college may be greeted by a student in Arabic with "*Asalam u 3alaikum*" {may peace be upon you} as it is the unmarked code choice. However, that same lecturer when visiting the college hospital which retains a high percentage of expatriate workers may choose to greet multilingual peers and co-workers in English with "*Hello, how are you?*" as the marked code choice. However, if this lecturer were to greet the multilingual group of

²⁰ Grice's (1975: 41-58) Cooperative Principle can be described as 'make your conversational contribution only as informative as required, do not say what you believe to be false, be relevant and be brief and avoid ambiguity'.

²¹ Myers-Scotton's (1983) social features are described as 'determined by the social forces in the community' and are not such features as 'age' or 'occupational status'.

staff in Arabic with “*Asalam u 3alaikum*”, this choice of the unmarked code may be seen as symbolic and representative of *we-solidarity* with other Arabic speakers.

Auer (1995) points out that the correlation between particular languages and speech activities predicted by the Markedness Model has failed to be proven by empirical studies. Nevertheless, Myers-Scotton’s Markedness Model is one of the most fully developed and influential models for CS motivations (Nilep, 2006: 12).

3.7 Stance Towards Code-Switching

From a speaker’s point of view the choice of using his or her first language (L1) or their second (L2) can impact or redefine a social situation (e.g., meeting, interview, lecture, etc) from informal to formal, official to personal, serious to humorous and politeness to solidarity (Blom and Gumperz, 1972). The language a speaker chooses to use in a particular situation conveys a meaning that goes beyond the speaker’s actual words (Trudgill, 2006:106).

Languages can be described in terms of identity and members of one or more language communities “attach different rights, identities and obligations to each of their languages” (Stroud, 1998: 322). In Saudi Arabia, while English is viewed as the primary foreign language and its palpable presence in the Saudi educational system is a reflection of the country’s interest and investment in the language (Al-Seghayer, 2012), it is also associated with the West, modernity and technology to name a few. In essence, it is often viewed as that which is not Islamic, Arab or traditional; it is the language of *other*. This study is not here to argue the identity in favor of the use of either language, but merely to describe the stance often expressed regarding the English language in Saudi Arabia as it pertains to CS in the research corpus.

In particular, the stance towards the English language in medical colleges and hospitals may also possess these ideologies regarding English, but it is doubtful that they are as deeply rooted as the role of English has been an integral part of medical training from its inception in 1967 with the first medical college. This is not to say that opposition to a foreign language as the EMI in medicine does not exist as is discussed in chapter 4 (see 4.2.2).

Alternatively, the roles of each language - English and Arabic - within this Saudi context may hold a greater symbolism and functionality in communicative events by virtue of code selection. Trudgill's (2000: 106) research of Cantonese-English CS reports that students at the University of Hong Kong who only spoke English, might be regarded as being disloyal to their community, whereas those who only spoke Cantonese might be regarded as uneducated and unsophisticated.

Similar circumstances in Saudi medical colleges report comparable judgements to the use of L1 vs L2 by Arab students. The construct of the national identity and its connection to language in relation to culture, politics and religion is complex (Suleiman, 2003: 280) and beyond the purview of this research. With respect to the intricacy of the issue, attitudes towards Westernization in the Arab world and all that it entails regarding their national identity and religious commitment may play a role in the frequency and function of CS by native Arabic lecturers. Would NS of Arabic curtail or increase their CS in this Saudi medical community in concern about external perception of their identity? A native Arab Sudanese lecturer (*PSR*) in the NNS corpus warns students that:

*“... information is power. If you have information you are a good doctor, you have power. If you don't the nurses will laugh at you, the technicians will laugh at you **katheer** {a lot} they will laugh at you.”*

The *PSR* excerpt illustrates the response of the lecturer to students walking in late to his class which spurs him to advise them that if they want power it comes from information. We can understand this as an indirect reference to the language medium; lecturers are aware of students' struggle with the L2 as a vehicle to attain the lecture content. The *PSR* lecturer in essence advises students that their power will come from knowing the foreign language which is the tool to information and power and without it, they will be humiliated in the hospital. This may seem a little harsh to students, but it reflects the current role that English plays in this community.

As a researcher, my interest is captured by the single token “*katheer*” {a lot}. Why, in the midst of an utterance meant to chide students, did the lecturer CS a single lexical unit? Considering the discourse is off-topic to the course lecture, why not use Arabic to scold or is it meant as an illustration of the power the lecturer holds with his ability to communicate in English. Granted as a corpus-based research, the *why?* can not be answered by the discourse alone.

However, further exploration of the *when* CS occurs and *what* function it plays in context is investigated and provides evidence which may shed light on potential motivations to pedagogic CS.

3.8 Roles of Code-Switching

It is not always clear what motivates a bilingual speaker's decision to code-switch a specific lexical token, phrase or chunk of discourse. Published research provides us with insight into possible roles CS plays in communication.

Bhatia and Ritchie (2004) reported CS serving an important function in hedging, that is, when participants were reluctant or hesitant to provide a clear answer, they included interjections and sentence fillers as mixed code in their discourse. Other factors, according to Bhatia and Ritchie (2004), included quotations, reiteration, rephrasing, topic, relative clauses, idioms and deep-rooted cultural wisdom (e.g. to repeat direct quotes in its original code for impact with local proverbs, for emphasis and/or attention by repeating or rephrasing information, etc). All of which, they postulate, are influenced by social variables such as age, gender, religion and class (Bhatia and Ritchie, 2004).

Code-switching is reported as being used to emphasize group identity (Grosjean, 1982; Gumperz, 1982) and a change in topic is claimed to work as a contextualization cue²² for side-remarks or the introduction of new topics (Auer, 1992). Grosjean (1982) also classified a number of CS functions utilized by bilinguals in communication to fill a linguistic need for a lexical item, set-phrase, DM or filler, for example:

Lecturer: *It's spread by droplet **razaz** {spray} not more than one meter, **suh** {right}?*

[CTB, SASME corpus]

The inclusion of both a lexical item “razaz” and a DM, ‘suh’ {right} in the utterance allows the lecturer to fill in a gap and continue to hold the floor uninterrupted. Another two functions

²² Gumperz (1982) defines *contextualization cue* as “any feature of linguistic form that contributes to the signaling of contextual presuppositions” (ibid: 131). When participants understand each others' contextual cues, conversations go smoothly; however, when contextual cues are misunderstood, conversations do not run smoothly and participants often label each other as rude, socially awkward, foreign, strange, etc. (ibid: 132).

described by Grosjean (1982) are to specify speaker involvement and mark and emphasize group identity (solidarity). Both of these linguistic functions of CS are representative of the personal and community factors attached to a speaker's languages, in other words, would the speaker express their emotions better in the L1 or the L2? Is the topic of conversation private, sensitive or related to the language of group A or group B? These and other factors are reported as variables in the negotiation to choose the language code for a situation.

Safi's (1992) study of the social identity of educated Saudi native Arabic speakers in the United States who code-switch Arabic/English described Arabic CS used predominately to discuss cultural events or in situations when the code did not provide the "desired religious or national feelings and when the use of some Arabic expressions [better] indicated the speaker's politeness" (Safi, 1992: 73). The motivation for CS included "clarification, avoidance of taboo or embarrassing structures and holding the floor" (Safi, 1992: 41). Similar motivations for CS were reported in studies and the most commonly specified as function of CS in bilingual communities are tabulated in Table 3.1.

Code-Switching Role	Definition
<i>Reiteration</i>	CS used to reinforce, emphasize or clarify a message "transmitted in one code, but not understood" (Eldridge, 1996: 306)
<i>Elaboration / Clarity</i>	CS used to provide more details, give expansions or additional details for other speaker or audience (Sanchez, 1994; Safi, 1992).
<i>Attention / Persuasion</i> (Nerghes, 2011)	CS in order to attract participant's attention and persuade them to listen closer to message, e.g., " <i>ya shabab this is important...</i> " {hey boys}
<i>Topic / Situation</i>	The topic prompts the CS (e.g., personal, cultural, religious topics and/or circumstances of discourse, such as greetings, etc). CS may also be used to avoid taboo words such as in Arabic (Abalhassan and Alshalawi, 2000).
<i>To Show Solidarity</i>	CS used to mark inclusion as part of the in-group; to signal group membership (Holmes, 2000)
<i>To Reflect Social Status</i>	CS used to show a hidden prestige made explicit by attitude (Auer, 2002). Suleiman (1999) describes CS as seen as being prestigious and a sign of education and competence.

Table 3.1 Code-Switching Roles in CS literature

Lawson and Sachdev's (2000) study reports that Tunisian Arabic and French CS is viewed negatively by their participants. However, these participants admitted to using CS and situational features just to maintain in-group relations and exclude outsiders and strangers. This type of CS is illustrative of the function of solidarity in Table 3.1.

3.8.1 Code-Switching in the Middle East

Research of CS in the Middle East are limited and of the available studies, I will present those that focus on Arabic/English CS such as Hussein and Shorrab, 1993; Eid, 1992; Jake, 1994; Abal hassan and Alshalawi, 2000; Al-Enazi, 2002; Bhatia, 2004; and, Taweel and Btoosh, 2010. Furthermore, this sections aims to provide insight into the motivations and functions that play a role in CS.

Abal hassan and Alshalawi's (2000) study attempts to identify the reasons for CS by Arab speakers of English as a second language in the United States and whether participants are even conscious of what they are doing. The researchers found in this cross-sectional study of bilingual US educated Saudis that there are seven major functions based on the primary roles of the code-switching:

Roles of CS by Bilingual Saudi ESL speakers in the United States	
1) Emphasis and contextualization cues	5) Conversation Tags
2) Parallel constructions for emphasis	6) Linguistic repertoire
3) Quotations and random switches	7) Politeness and avoidance of taboo expression
4) Technical terms	

Table 3.2 Primary Roles of Code-Switching (Abal hassan and Alshalawi, 2000: 184)

Some of the reasons stated for CS expressed by participants were “did not know the term in Arabic” (59.2%) and “forgetting the term in Arabic and ease of saying it in English” (22.6%) (Abal hassan and Alshalawi, 2000: 185). In general, the research suggests that participants are not

only aware of their CS use, but additionally, make conscious decisions about when to utilize this linguistic strategy (Abalhassan and Alshalawi, 2000).

Al-Nofaie's (2010) study of Saudi teachers' and students' attitudes towards employing Arabic as a facilitating tool in the English as a foreign language (EFL) classroom were viewed as generally positive. Participants in the study stated that there were specific situations where they preferred using Arabic, for example, to induce "feelings of comfort" (Al-Nofaie, 2010: 78); while students described the specific situations as activities which required them to express their opinions or ask questions (ibid: 77). Though Saudi teachers were aware of potential disadvantages of excessive use of Arabic in the EFL classroom, Al-Nofaie (2010) states that teachers preferred to have L1 inclusion with beginning and low proficiency students (ibid: 78).

Sabbour et al.'s (2010) research of the language barriers in medical education and attitudes towards Arabization of medicine in an Egyptian medical college, showed that 56.3%²³ of students did not view medicine taught in EMI as a problem. However, a similar number of 56.8% claimed difficulty in following lectures that were delivered completely in English and 34% preferred lectures to be delivered in a mixture of Arabic and English. The concept of Arabization, the push towards using the mother tongue (L1) to "free learners from the dualism imposed by thinking in one language and studying in another" (Ali, 2010) (as cited in Sabbour et al., 2010: 1264) is rife with challenges. The first barrier stated in the research regards existing medical terminology which, on the whole is taught solely in English in countries including Egypt and Saudi Arabia and if taught in the L1, would create "difficult Arabic terminology far removed from people's everyday language" as identified by 80.8% of students and 74.6% of staff members (Sabbour et al., 2010: 1267).

In fact, teaching medical terminology in Arabic may not be a feasible option, as Amr (1996) states "there is a backlog of 250,000 terms with no Arabic equivalents in addition to thousands of yearly emerging neologisms" (ibid: 90). With this backlog of existing terminology and the emerging new terminology each year, Hussein (1999) believes that "there is no way for Arabic to catch up with specialized lexicon," and thus CS is the rule rather than the exception (ibid: 283).

²³ The research's response rate is 81.5% (326/400) among students and 73.3% (110/150) among staff members (Sabbour et al., 2010: 1265)

AlEnezi's (2010) similar exploratory study of students' language attitudes towards Arabic and English CS in a Kuwaiti Allied Health Science College reported that 58%²⁴ of students agreed to the desirability of CS during teaching. Students strongly-agreed (35%) / agreed (53%) that CS contributed to better understanding during teaching. In fact, in response to a series of questions regarding the effect of CS on a "teacher's image" students were asked whether they "respect [the] instructor more when teaching in English and Arabic" and 64% agreed that their respect increased (AlEnizi, 2010: 11). The other two questions in this set asked about teaching "only in English" to which 41% agreed that they had more respect and "only in Arabic" to which 29% agreed and 47% disagreed that their level of respect increased in such circumstances (AlEnizi, 2010: 11). This is illustrative of the negative and positive impact linguistic codes have on teacher perception.

These studies are reflective of the attitudes towards and the function of, CS in social and academic contexts with Arab speakers in both the West and in the Middle East.

3.8.2 Code-Switching in Academic Discourse

Inclusion of the L1 in L2 academic classrooms has been a continuous issue. Some scholars believe that the use of the L1 to organize classroom teaching or to instruct language knowledge deprives students of the opportunity to receive input in the target language (Ellis, 1994). Although the current research does not deal with the language classroom, but rather EMI medical lectures, there are aspects of these studies which are applicable to the current research. NS Arabic students are often expected to improve their proficiency of the L2 (English) via the medium while simultaneously learning the content of the subject course.

Alnassar and Dow (2013: 55) state that in Saudi higher education:

...it must be acknowledged that we ask a lot of teachers if we expect them to switch easily back and forth between Arabic and English in their classrooms. That is sometimes assumed, but it is unrealistic in most cases, and there is much controversy amongst linguistics as to the desirability of 'mixed-code' teaching"

²⁴ Exploratory study of one class of seventeen second year students in a science class of *Human Development for Occupational Therapy* (AlEnezi, 2010)

As English in higher education, specifically medical colleges, is only taught explicitly in the initial preparatory year, Alnassar and Dow (2013: 55) highly recommend continued attention to building students' confidence in English in subsequent years. This will require further effort from academic teachers in general, they state and this reinforcement with English will help better prepare students in their academic studies. The commitment to improving student's English skills needs to be ongoing, "rather than entirely focused on the preparatory year" (Alnassar and Dow, 2013: 55)

Atkinson (1987) describes students as sometimes becoming frustrated when expected to carry out oral activities in the language classroom because of their limited vocabulary and teachers often expecting their students to "think in English" (ibid: 245). According to Atkinson (1993: 25-38), there are some functions of L1 inclusion that help students maximize comprehension:

- 1) *Lead-ins*: exploiting the L1 to check whether students have understood the situation
- 2) *Eliciting language*: extracting language from students
- 3) *Giving instructions*: used to clarify written instructions on a worksheet or a textbook
- 4) *Checking comprehension*: verifying whether students understand a word or phrase

These CS functions are often utilized for time-efficiency in situations such as when a medical term is known by students in the L1 and by accessing this knowledge through CS lecture time is used efficiently. This is illustrated further with excerpts from the current research (see 6.6.4) and is believed to help students perform language tasks more effectively (Swain and Lapkin, 2000). Ferguson's (2003) provides an overview of research into CS functions in the classroom and states that a fair degree of similarities in the research allows the functions to be collapsed into three broad categories (ibid: 39).

- 1) *CS for curriculum access*, e.g. essentially, "to help pupils understand the subject matter of their lessons"
- 2) *CS for classroom management discourse*, e.g. "to motivate, discipline and praise pupils..."
- 3) *CS for interpersonal relations*, e.g. "to humanize the affective climate of the classroom and to negotiate different identities"

Each category shows that CS is a valid pedagogic tool and it achieves at least one important purpose in the classroom which is that CS:

“is not as some educational authorities have implied ... a dysfunctional form of speech behavior, but on the contrary, an important, even necessary, communicative resource for learning, especially for pupils with limited proficiency in the official instructional medium”

(Ferguson, 2003: 43).

Üstünel and Seedhouse’s (2005) study of CS in Turkish university EFL classrooms reflects a similar pattern of functional CS and provides evidence of the link of language and pedagogical focus. Their research suggests that the L2 is not always the preferred choice, instead, the preference is whichever language aligns learners with the pedagogic focus at that particular stage in the classroom interaction (Üstünel and Seedhouse, 2005: 321).

Similarly, Zabrodskaja (2007) studied the conversational function and pedagogical aspects of Estonian-Russian CS in university lectures on linguistics taught in the students L2 - Estonian. In this study, students are not being taught Estonian as an L2 per se, rather, the course *Introduction to Linguistics* is taught through the medium of a second language (Zabrodskaja, 2007: 102). The functions of CS in the classroom are viewed as maximizing an opportunity for positive results and the acceptance of the mother tongue. If the teacher were to reject a student’s L1 use in the classroom they may communicate a rejection of the student and lose an opportunity for a meaningful exchange (Zabrodskaja, 2007: 102). Results of Zabrodskaja’s (2007) study shows that there are benefits to involving the L1 in the classroom and students can improve their academic performance when both languages are used in the class.

Chang (2010: 60), who investigated the reactions of undergraduate Taiwanese students in EMI subject courses in private tertiary education, reports that EMI causes some students great difficulties understanding course content²⁵ and often inhibited them from expressing themselves in class. Findings of her study based on student questionnaires and professor interview data

²⁵ Data from six students and six professors in College of Engineering, College of Management and College of Informatics.

showed that the Taiwanese professors often switched from English to Mandarin under certain circumstances (Chang, 2010: 67):

- (1) Looks of confusion from students
- (2) Explanations requested in Mandarin by students
- (3) Introduction of difficult concepts

She also reported that some professors repeated content in Mandarin after every chunk of English in the lecture (Chang, 2010: 67). This CS between the L1 and L2 is stated as reflective of their awareness of student difficulties and a “considerate attitude” meant to help reduce student anxiety and increase satisfaction with the EMI courses (Chang, 2010: 67).

3.8.3 Code-Switching with Discourse Markers

In bilingual speech as previously described, lexical units from an L1 such as DMs are sometimes inserted into L2 linguistic structures. Discourse Markers play an important role in strengthening cohesion and coherence in discourse and Schiffrin’s (1987: 328) definition of DMs states that they can be syntactically detached from a sentence or in this case, utterances. Riehl (2005: 1949) contends that these “kinds of lexemes” are “easily detachable” and the more “gesture-like” these elements appear, the more likely they are to be integrated into the recipients’ language system.

In bilingual communication, DMs attain a role in function orientation that may not express the same “meaning” as they would in “normal” (De Rooij, 1996: 449) discourse. De Rooij (1996: 447-448) describes DMs as a special kind of contextualization cue with signals which speakers rely on to process conversational inference, as such DMs need to be as salient as possible. De Rooij’s (1996: 451) analysis of French DMs in Shaba Swahili conversations found that the French DMs were particularly salient in conversation. As Shaa Swahili is the dominant language or what we would refer to as the ML (cf. Myers-Scotton, 1993) and French as the EL, the salience of the chosen DMs is enhanced by the speakers and its contrast in the linguistic environment attracts more attention.

Auer (1991: 326) states that code-switching DMs in bilingual conversations becomes more significant as a marked choice of a specific code in relation to other linguistic items from other codes.

In an analysis of 3rd-6th generation Texas German speakers, Salmons (1990) found that while speaking German, the participants almost exclusively use North American English DMs and conjunctions such as ‘*now*’, ‘*because*’, ‘*well*’ and ‘*you know*’. Goss and Salmons (2000) and Fuller (2001) argue that English DMs have been borrowed through long term contact with German American dialects (as cited in Boas, 2010: 302).

For example (Auer, 1999: 325):

“*now Kurt, der is ungefähr so hunnert miles from Chicago...*”
 {now Kurt, he is about hundred miles}

Maschler’s (1998) analysis of her Hebrew/English corpus found that the function DMs played in language alteration varied depending on the type of DM involved. Maschler (1998) classifies these types according to the DM function in context: interpersonal, textual and cognitive. Interpersonal DMs are those used to negotiate the relationship between speakers (e.g. *naxon?* {*right?*}). Textual DMs are those which regulate the way conversational actions are structured into a coherent whole (e.g. *biglal* {*because*}). The last type, cognitive DMs, include markers which provide information about the “cognitive processes occurring at frame shifts” (e.g. *ke’ilu* {*like*}) (Maschler, 2000: 536-539).

Maschler (2000: 558) concludes that DMs do function in one particular way when code-switched in bilingual conversation and that is, by framing a signal of “let’s move to metalanguage now.” She reports that, based on her research, this type of framing in monolingual conversation does not occur in the same way.

Mujiono et al.’s (2013) found that the Indonesian participants in their study who were teaching EMI classes would code-switch interchangeably between English, Indonesian and Arabic. The first, English, is the medium and topic of instruction, while the latter two languages, Indonesian and Arabic are code-switched when students do not seem to understand the information being conveyed. The English lecturers code-switched for a variety of pedagogic reasons including,

information clarification, to create humor, repetition or reiteration of a message, to strengthen requests or commands, to make questions and for DMs. Specifically, in the case of DMs, Mujiono et al.'s (2013) results showed that the English lecturers used more Indonesian DMs (e.g. *'sekarang'* {now}, *'makanya'* {so}) than English DMs (e.g. *ok, right, well, so, now*).

The presentation of CS literature in this chapter establishes the uses and functions of the L1 in the L2 academic classroom. It also shows that there is a lack of research into the use of Arabic L1 in English as a foreign medium lectures and the need for additional studies of both students and teaching staff at the tertiary level in Saudi Arabia. This research is intended to address the gap in the research by investigating the use of Arabic DMs and the overall functions of CS in L2 medium academic medical lectures.

CHAPTER 4

METHODOLOGY

*“Corpora have been likened to the invention of telescopes
in the history of astronomy.”*

(Hunston 2002: 20)

Chapter 4 Methodology

4.1 Introduction

This chapter presents additional background information about the context of the NNS academic corpus (also see 1.4.2), introduces the research questions and describes the overall framework of the this study. The chapter describes the corpus-based approach taken, the exploratory study and the corpus selection for this study. It then details the transcription styles for the English and Arabic data and provides a breakdown of the academic lectures selected for the NS and NNS corpora. Lastly, the development and framework for the DMs and CS classifications established for this study are described and illustrated.

4.2 The Study

While the literature suggests that NS generally use DMs at a greater frequency than NNS in spoken discourse (Müller, 2005; Fuller, 2003; Fung and Carter, 2007), I propose that academic discourse in medical lectures will not necessarily follow this pattern. This hypothesis is in light of the fact that NNS lecturers of SASME are highly proficient in the English language having studied all, or part, of their medical or health science degrees in a native English country. Moreover, the theory is based on the audience of these lectures, which in BASME are either native or highly proficient English speakers²⁶, whereas, SASME's audiences are all native Arabic students using English as a foreign language of instruction.

With homogeneous groups, the literature suggests lecturers often have the tendency to simplify their language to a level their students are able to comprehend and may, in fact, speak more slowly, include more redundant information (Nesi, 2001) and use less subordination (i.e., more single clause utterances, repetition and other forms of redundancy including paraphrasing) (Griffiths, 1990) In fact, based on Tauroza and Allison's (1990) findings on speech rates of

²⁶ Based on University requirements that international medical students must achieve an overall IELTS score of 7.5 (see 4.2.1)

lectures in words per minute, the BASME lecturers had an average speed²⁷ of 166 words per minute which is characterized as *moderately fast*. On the other hand, SASME lecturers had an average speed²⁸ of 127 words per minute and is characterized as *average* though with only two points above *moderately slow* (Tauroza and Allison, 1990:102). This may be the case in the SASME lectures due to lecturers modifying their speech for the benefit of the NNS students.

In the case of bilingual speakers, the SASME lecturers consistently include the supplementary use of the L1 by CS to support comprehension of the lecture content, amongst other reasons (see 6.6). NNS have been noted to have an increased frequency in the use of DMs to structure and organize lecture content for greater accessibility.

The highest frequency of discourse structuring among NNS guest lecturers suggests that speaker language background has a strong influence. Their non-native speaker status seemed to prompt them to use discourse structuring expressions ... to help the L2 listeners...

(Camiciottoli, 2004:49)

Therefore, an increased use of ‘simpler’ single-unit DMs might be expected in SASME, to clearly structure and organize the lecture content, while, conceivably, a higher frequency of relatively more ‘complex’ multi-unit DMs (e.g., *We’re going to, as you know, let’s x, etc.*) might be found in BASME.

In undergraduate medical lectures, where the language of instruction is the L1, the discourse of BASME was expected to exceed that of SASME in complexity (e.g., utterance structure, content, etc.) and language choices (e.g., medical jargon). The following are two excerpts from BASME, the native speaker corpus.

²⁷ The average speed of BASME lecturers was found by first calculating total tokens per minute for each lecture (see Table 4.4) and then adding the total number of tokens (per minute) of the nine lectures. This was then divided by total lectures to calculate average.

²⁸ The average speed of SASME lecturers found by first calculating total tokens per minute for each non-Arab lecturer (see Table 4.3). Second, native Arab lecturers, all of whom included both L1 and L2 discourse, were calculated by counting number of tokens spoken during a one minute audio segment of each lecture. Total number of tokens (per minute) of the 11 lectures were divided by total lectures to calculate average.

okay so it's difficult to predict how long stage two is going to be for an individual the key thing about stage two is during stage two although the people are infected they are well they are ge- in general terms healthy now clearly if you are someone that who is infected the thing that you want to know more than anything else is how long am I going to be in stage two

[lslct008, BASME]

...just 'cause we invented it we think we're we're the best at it and have have er rights over it er and er nobody actually nobody knows what Bright's disease is er and some people think it's nephrotic syndrome some people think it's glomerulonephritis er it's better not not to define it 'cause obviously Bright didn't define it I take it to be er synonymous with nephrotic syndrome...

[lslct034, BASME]

In general, the material presented in medical lectures is described as “not conceptually more difficult than many rigorous undergraduate courses, but the volume flow rate of information per hour and per day is much greater” (Apperson, 2014); and lectures have been frequently likened to ‘drinking from a fire-hose’ (Apperson, 2014). My assumption initially with the NNS data was that while the concepts and volume rate may be comparable to the NS, the structure of the NNS data would differ in respect to the use of DMs considering the medium of instruction was a foreign L2.

You know there is a pulmonary edema talk about it today and sign of meningeal location. That's a gradual - not as severe as what you find in the acute yaseer {it happens} gradual. An altered mental status. Remember, the [inaudible] loss might be six, seven years wa al {and the} course can be affected because these are easy to deceive. Maybe they are not easy, but this one, if you have got the facial - or you find that weakness on one side, then what do you have? We teach it in the course.

[CTB, SASME]

Then this way the congenital malformations of the CNS is a huge topic. It really depends on your understanding for the embryology of the CNS. Up to my knowledge we don't really take the embryology in details, we just take major outlines. So it's really beyond your scope to talk about congenital malformations. However, there are common diseases which you will see in the children. May I say today, it's to make sure you know what they mean. What these diseases are.

[CMH, SASME]

The excerpts demonstrate that the general discourse of the corpora lectures are relatively comparable, although the ‘sentence’ structures (e.g., shorter vs. longer utterances) in the corpora do differ as might be expected from NS vs NNS lecturers teaching in EMI.

The aims of this study, as stated in the introductory chapter, relate to the use of spoken DMs in academic medical lectures based on corpus findings. The research focuses primarily on the frequency of occurrences and the uses of English DMs in non-native English medical academic lectures in contrast to its frequency and use by native English medical academic lectures, with a supplementary focus on Arabic DMs used within the NNS English academic discourse.

As presented in the previous two chapters of academic discourse and code-switching (CS), there is limited research available on the uses and functions of DMs in academic discourse or on CS in subject courses where English is the foreign medium of instruction. What the studies do show is that academic lectures are a major part of tertiary education (Flowerdew, 1994; Bellés-Fortuño, 2006; Chang, 2012) (see 2.2.1) and the presence of DMs help create coherence between segments of discourse (Schiffrin, 1987; Fraser, 1999) (see 2.4.2). Furthermore, DMs have been shown to play a positive role in the structuring and level of interaction in lectures (Chaudron and Richards, 1986; Morell, 2000; Aijmer, 2002; Bellés-Fortuño, 2006; Othman, 2010) (see 2.4.4). However, their frequency and functions have also been shown to vary by speech events (Fuller, 2003; Rendle-Short, 2003), academic disciplines (Csomay, 2005), medium of instruction (Bellés-Fortuño, 2006) and native and non-native speaker status (Fung and Carter, 2006; Othman, 2010) (see 2.5.3). This study explores the distribution through a frequency count of occurrences of DMs that NS and NNS academic medical lecturers rely on to convey their pedagogic aims and are intended to improve students’ comprehension of the lecture content. I analyze two sets of DMs which I have classified as *Structural* and *Interactional* DMs based on their functions and taking into consideration the linguistic context (co-text). This classification is presented in detail and discussed in 4.8.

The analysis of the DMs is intended to help gain insight into the similarities and differences between British and Saudi-based academic lecture discourse in the field of medicine in respect to the uses, roles and functions of DMs. We have established from the literature that DMs can be frequent in academic lectures that have a more conversational or interactive style (see 2.4.4); and

that NS are generally found to use a higher frequency and wider range of DM functions than NNSs (cf. Fuller, 2003; Fung and Carter, 2007). The first question and its subparts frame the intention of the analysis to test the theory that the highly proficient Saudi-based NNS lecturers may use more DMs in their lecture discourse. Additionally, it questions whether more single-word DMs (e.g., *okay, so*) are used with NNSs and more multi-word units (e.g., *in other words, any questions?*) are used with NS.

- (1) Do Saudi-based NNS English medical lecturers use more Structural or Interactional discourse markers than British NS English medical lecturers?
 - (ii) Do the NNS English medical lecturers use a higher proportion of Structural discourse markers relative to the use of Interactional discourse markers?
 - (iii) With respect to categories of Structural DMs, do the NNS English medical lecturers use each function with the same proportional frequency as the NS English medical lecturers?
 - (iv) With respect to categories of Interactional DMs, do the NNS English medical lecturers use each function with the same proportional frequency as the NS English medical lecturers?

The second research question is meant to explore whether the CS in the Saudi-based medical lectures is a communicative strategy meant to further the pedagogic aims of the subject course (see 3.8.1). Code-switching has also been shown to strengthen a message given in an L2 by inserting L1 DMs as the marked choice (Auer, 1991) to regulate discourse relations on structural and interpersonal levels (Maschler, 1998) (see 3.8.3). Moreover, the motivations behind CS can be a useful pedagogic communicative strategy aimed at supporting content discourse (Atkinson, 1987; Ferguson, 2003; Zabrodska, 2007; Mujiono et al., 2013) (see 3.8.2).

- (2) Do the native Arabic lecturers code-switch (English L2/Arabic L1) DMs in the Saudi-based medical lectures? If yes, which functions (Structural or Interactional) of DMs are code-switched more often in academic medical lectures?
 - (i) If native Arab lecturers do code-switch, what are the roles of code-switching in academic medical lectures?

Through the data analysis I will attempt to answer the research questions through both qualitative and quantitative means. Quantitatively I will calculate the distribution of frequency of

occurrences (normalized per 1000 tokens) of the DMs identified (see classification framework) across the NS and NNS monologic medical lectures in the corpora. Qualitatively I will evaluate the functions of the DMs used in both the NS and NNS lectures and identify their classification (Structural or Interactional) and (sub)functions in the context of the surrounding spoken text. The qualitative analysis will show the use of DMs by medical lecturers to signal a local and/ or local structural (textual) relationship between and across segments of discourse or signal an interactive Ideational relationship between speaker-hearers.

4.2.1 Context of the Study

King Saud University (KSU) was the first university established in Saudi Arabia in its capital city, Riyadh, in 1957. It is now the largest university in the Kingdom and the Arabian Peninsula as a whole, with over 8,300 faculty members, more than 70,000 students and more than 150 academic programs (Mazi and Altbach, 2013). Instruction in most undergraduate programs at the university is conducted in English with the exception of Arabic and Islamic studies (Smith and Abouammoh, 2013). All universities in Saudi Arabia, with two exceptions²⁹, now have both male and female students; however, based on Article 155 of the Saudi Arabian Education Policy, they are strictly segregated on campus and in most lecture rooms (Smith and Abouammoh, 2013). Most medical schools are an exception to this government policy.

The first medical college in Saudi Arabia was established in 1967 at the King Saud University in Riyadh in affiliation with the University of London until 1979 when the affiliation was terminated (Al-Shehri and Campbell, 2013). Traditionally, the medical education curriculum was built mainly on a content-based model; however, recently, as a result of critical reports, significant developments in medical education have taken place including the adoption of an international trend towards professionalism in education (Al-Shehri and Campbell, 2013).

²⁹ The two exceptions are the King Fahad University for Petroleum and Minerals (KFUPM) in Dhahran - a male-only university and Princess Nora bint Abdulrahman University (PNU) in Riyadh - a female-only university. PNU has recently founded the first female-only Health and Medical Education Colleges in 2012.

4.2.2 English as a Foreign Medium of Instruction

Using English as a Medium of Instruction (EMI) is a common practice in university based health care programs in the Middle East where Arabic is the mother tongue. Students, upon joining university, are obliged to accept the change from Arabic as a medium of instruction in their secondary education, to EMI in their college lectures (Suliman and Tadros, 2010). English, in many cases, is limited to academic domains and not, on average, accessible in students' social lives; thus rendering it a continuous challenge for students with little to no support outside the academic setting.

The English language has been noted to be one of the main obstacles for students, as their secondary education, as it currently exists, does not prepare students with a sufficient command of English to better prepare them for a medical career (Al-Sulaiman, 2000: 16-17). The issue of using a foreign language in higher education is contentious (see Al-Jarf, 2008; Sebai, 1982; Albar and Assuhaimi, 1996) and much debate exists (see Jenkins, 2010; Sabbour et al., 2010; AlEnizi, 2010) on whether medical colleges need to maintain an “all or nothing approach” (Al-Shehri, 2001: 322). Most Saudi medical colleges teach medical terminology in English only (see 3.7.1); however, the use of CS often functions as a supportive linguistic tool when possible (see 3.8). This is important as medical students need to be able to communicate with NS Arabic patients using existing L1 medical terminology or have the ability to explain the L2 terminology in L1 laymen terms.

4.3 Corpus Approach

This study adapts a corpus-based approach to linguistic research which provides a means to investigate academic medical discourse in greater depth and with a scope and reliability of analysis not otherwise possible (Biber and Conrad, 2004). One of the main advantages of using this approach is the frequency information that the corpus can readily provide (McEnery et al., 2006) (see 2.3.1).

The Saudi Academic Spoken Medical English (SASME) corpus was created with respect to criteria, as follows:

Purpose

The purpose of the corpus is to research the academic language of NNS English medical health lecturers in a Saudi medical college with a NNS English student body. A *spoken* corpus is chosen as the means to investigate this specialized academic language, as it assured an insight into *authentic* language usage in the L2 academic classroom.

Representation

A collection of non-native English speakers similar in terms of education, teaching experience and lecturing style were chosen as representative of a main sub-group in the medical college (as opposed to teachers, instructors and graduate assistants, etc.). The research corpus, SASME, is also representative of a *specialized* corpus by containing a very specific set of spoken discourse used in higher education and is similar in concept to such corpora as MICASE (see 2.3.2), BASE (see 2.3.2), amongst others.

Balance

The intent with corpus design is to have a balanced representation of the existing dynamics in the Saudi medical college, which in current years has had mainly native Arab lecturers (e.g., Saudi, Sudanese, etc), with a minor faculty of non-Arab lecturers (e.g., South-East Asian, Western, etc). Therefore, an attempt was made to include both Arab and non-Arab lecturer transcripts in the SASME corpus.

Size

The size of the corpus was subject to the time-frame of the transcriptions. In the course of the transcription process a decision was made to include at least 10 lecture transcripts as a balanced representation of the Saudi-based academic medical discourse. The final corpus size of SASME, while not extensive at 56,250³⁰ total English and Arabic tokens, was not regarded as a

³⁰ The total number of tokens include the English only tokens of 47,574 (see Table 4.3) and the estimated total number of Arabic tokens 8676 (see Table 4.6).

limitation of the target population, as the inclusion of a “range of linguistic distribution” of said population (Biber, 1993: 244) is believed to be represented.

4.4 English, Medicine and Lectures

The topic of the study, EMI lectures in the KSU medical college, is chosen as the focus in large part because the researcher is an English language instructor in the preparatory year for medical health students. Secondary motivations are a desire to increase available research on higher education in Saudi Arabia.

Access to two³¹ of the main medical colleges’ faculty and administration was made possible through colleagues and students.

Ethical Considerations

Lecturers were personally approached by the researcher, when possible³² and invited to take part in the study. Lecturers who agreed to take part in the study were assured that their identities would remain confidential. As privacy is a main concern to lecturers, they were assured their audio recordings and lecture transcripts would not be published in its entirety, though they were informed that excerpts³³ would be used within the thesis. Consequently, the corpus cannot be made publicly accessible³⁴.

Lecture Recordings

Lectures were recorded with a small audio recorder placed as close to the lecturer as possible without it distracting students. In the exploratory study (see 6.5.1), lecturers recorded their own

³¹ For the exploratory and subsequent main study data collection

³² There were issues contacting lecturers in the male-only medical college at KSU; in these cases, an intermediary was used to approach and gain permission to audio record lectures.

³³ Every attempt is made to insure excerpts do not reveal information that may reveal the identity of speakers.

³⁴ Prior to data collection, I was aware lecturers would be reluctant to allow recording of their lectures and that concessions would have to be made to guarantee anonymity. However, in future for the purpose of educational development and research, Saudi Arabia needs corpora of non-native English social and academic language to be made publicly accessible.

classes with the audio recorder given to them by the researcher. This was done so as not to cause any distraction in the classroom with the presence of a non-student and potentially affect the dynamics of the classroom with students and the lecturer, often referred to as *observer's paradox* (Labov, 1972: 209). The audio recorders were collected following each lecture or at an agreed upon time depending upon the lecturers' schedule.

The audio recordings for the final corpus were collected by a few lecturers themselves, but on the whole, a student rep was contacted to assist in each lecture audio recording. Lecture halls varied in size and if no teacher's desk was available, the audio recorders were placed on a student desk in the front row or on an empty desk in the front corner. Audio recordings overall varied in clarity and sound quality.

4.5 Academic Medical Discourse

4.5.1 Exploratory Study

An exploratory study was carried out to explore the linguistic features in academic medical health lectures. A corpus was compiled of non-native speakers of English (NNS English) lectures from two Medical/Health Science Colleges in Saudi Arabia and a native speakers of English (NS English) sample corpus from transcripts in the Life & Medical Sciences discipline in BASE (Nesi and Thompson, 2003).

The first exploratory corpus, Saudi Academic Medical English Exploratory (SAMEX), was created with NNS English medical lectures from two universities in Saudi Arabia - the King Saud University (KSU) and the King Saud bin Abdulaziz University for Health Sciences (KSAU-HS). The SAMEX corpus included four audio lectures from the College of Nursing and College of Medicine, KSAU-HS and eight audio lectures from the College of Medicine, KSU. A total of twelve audio lectures were collected and transcribed with *Soundsciber* (Breck, 2004) and *Dragon NaturallySpeaking*[®] software and had a total count of 30,369 tokens.

Lectures were transcribed by using *Soundsciber* (Breck, 2004) to play and loop audio recordings in 10 to 30 second audio segments for up to 10 loops (i.e. repeated playbacks of

selected segment) while speaking into a microphone on *Dragon NaturallySpeaking*[®]. The transcription software *Dragon NaturallySpeaking*[®] was used for the English discourse by listening to the lecture audio and parroting the recorded text into a microphone which the software recognized and instantly transcribed into written text.

It was not possible to transcribe both the English and Arabic concurrently; and, in fact, at the time of transcription the software was not capable of recognizing or transcribing Arabic discourse. Therefore, the Arabic was transcribed manually using *Soundscriber* (Breck, 2004) following the completion of the English data in the NNS corpus.

The second exploratory sub-corpus, British Academic Spoken English Exploratory (BASEX), was compiled from transcripts in the Life & Medical Sciences group in the BASE holdings that were as similar in topic as possible for the purpose of comparison. The 10 selected transcripts ranged from the Animal & Microbial Sciences department to medicine from the Leicester-Warwick Medical School (LWMS) and were compiled to create an exploratory sub-corpus with a total of 57,069 tokens.

The criteria for academic lecture inclusion to the NNS English exploratory corpus SAMEX were as follows:

- i. *Native Arabic Speaker*: uniformity across lecture transcription data as the majority of teaching staff are bilingual Arabic/English speakers and the student body in the Medical Health Colleges are Saudi³⁵.
- ii. *Background in Science and/or Medicine*: provides relative similarity in language use, terminology and knowledge of medical health field
- iii. *Actively teaching a partial or fully lecture-mode course* (vs. laboratory courses or seminars): provides a larger quantity of spoken academic data
- iv. *Proficient non-native speaker of English*: in an attempt to preclude lecture hour(s) from heavily shifting to or becoming predominately taught in the Arabic language

³⁵ Non-Saudi students make-up 2.63% of all enrollments of which the greatest concentration are at PhD level (Smith and Abouammoh, 2013).

The mind-map (4.1) illustrates the design of the corpora collected for the exploratory study:



Mind-Map 4.1 Lectures included in Exploratory Corpora BASEX & SAMEX

The exploratory study was created to examine the discourse of Saudi medical academic lectures by NNS of English in Saudi Arabia, SAMEX and subsequently, compare the findings to the native English sub-corpus of spoken academic lectures, BASEX. The exploratory corpora

analyses were initiated with a top-down approach investigating similarities and differences between the NNS and NS English corpora using *AntConc 3.2.4* (Anthony, 2012).

AntConc 3.2.4 (Anthony, 2012), a freeware, multiplatform tool, was used to search concordances, concordance plots, clusters / n-grams and investigate collocates. It was also used to count total tokens in the corpora and view the most frequent with word list.

4.5.2 Results of Exploratory Study

Through the Exploratory study, results³⁶ indicated that:

- (1) Infrequent similarities and occurrences with medical terminology as tokens were found to be specialized to the (sub)field or (sub)specialty they occurred in. Nonetheless, an overall similarity was found with discourse markers used as *confirmation*, *starter/closer frame* and *acceptance* functions.
- (2) DMs were highly frequent with both NS and NNS medical health lecture discourse. The exploratory study highlighted the frequency of DM use (e.g. *okay*, *so*, *now*, etc) in the NNS English corpus SAMEX and supported further investigation into its functions in Saudi academic medical discourse.
- (3) More care with selection of lectures was needed - medical terminology and statistics lectures were not comparable to other medical health lectures and thus should be excluded. These lectures were found to be organized and presented in a different style from the other lectures as their content was related to presenting vocabulary and mathematical formulas respectively.
- (4) There were numerous instances of code-switching (CS) with NS Arabic lecturers at KSU and relatively few with KSAU-HS lecturers³⁷. One of the main purposes found for CS was as a contrastive marker which functioned to draw attention and re-direct students' focus to key lecture content and/or rephrasing of content.
- (5) Arabic DMs were found to occur frequently with native Arabic lectures, specifically, '*ya3ni*' {mean; meaning}, '*tamam*' {okay}, '*lakin*' {but; however}, etc.

³⁶ Refer to Appendix B for further results of the exploratory study

³⁷ There is a stricter enforcement, in my experience, of the EMI in KSAU-HS, than there is at KSU medical college

4.5.3 Transcription Style

Originally, a style similar to the BASE³⁸ transcription style was to be used for both the exploratory corpus and for the study's main corpus, the Saudi Academic Spoken Medical English (SASME) corpus audio lectures; however, a series of challenges following the exploratory study compelled a different transcription style. First, the Saudi corpus transcription's spellings were taken from Oxford Advanced *American Dictionary*³⁹, rather than the Oxford English Dictionary. Secondly, while BASE used no punctuation to signal the start/end of discourse, it was included in the SAMEX and SASME as cues to the start and end of utterances. Furthermore, the included punctuation were used in the Saudi corpus to signal rises and falls in a speakers intonation. While intonation was not a focus of my research, it plays an indirectly vital role in identifying functions of DMs. For example, a *comma* indicates milli/second pauses in discourse, *full stops* indicate a complete pause in discourse - in some cases by a fall in intonation (i.e. *okay↓.*), *question marks* indicate a rise in intonation by speakers, which are typically associated with questions (i.e. *okay↑?*) and a *hyphen* indicates a sudden stop in discourse, usually an incomplete word and change in structure/word choice by the speaker (e.g., ‘*wha- where do...*’) (see Table 4.1).

Transcription style	Function	Example
Comma ,	Indicates slight pauses, usually associated with taking a breath before completing an utterance	<i>Okay, so why do they...</i>
Full stop .	Indicates a fall in intonation and/or a long pause	<i>Okay. Now why do the...</i>
Question mark ?	Indicates a rise in intonation, usually associated with asking a question	<i>Okay? Do you understand...</i>
Hyphen -	A sudden stop in discourse, due to interruption or self-initiated by speaker	<i>We will disc- talk about the...</i>
Curly brackets { }	Provides translation of immediately prior Arabic CS	<i>Ya3ni {I mean} the bacteria is...</i>

Table 4.1 Transcription Summary

³⁸ http://www2.warwick.ac.uk/fac/soc/al/research/collect/base/history/base_manual.pdf

³⁹ Researcher is a native American English speaker and transcription software requires speaker to set up a user account and voice profile; one choice is ‘base language’ for software to recognize speech and ‘US English’ was chosen.

Furthermore, unlike BASE, the Saudi corpus includes CS events which were subsequently tagged and transcribed using a sequence of Latin letters (e.g., 'kh' for the Arabic letter 'خ') and Arabic numbers (e.g., '3' for the Arabic letter 'ع'). A detailed description of the Arabic CS transcription is provided in section 4.5.4 in addition to an explanation of reasoning behind romanizing the Arabic CS. Events of CS are subsequently translated by the researcher to English and are identified as translations by curly brackets, (e.g., *Alhamdullilah* {Thank God}).

Other than these aspects of transcription styles, the Saudi files were tagged to mark events and functions in a style comparable to the transcripts of the British Academic Spoken Medical English (BASME) sub-corpus. The following table illustrates a brief view of the tagging method used to mark up occurrences in SASME:

Tag Category	Function	Tag Category Members	Example
<intro>	Introduction	Introduction to start of lecture	<intro> Lecturer: <i>The evening session is starting</i>
<topcon>	Topic Content	Start of lecture topic content	<topcon> <i>So now we will go to the names of the particulars.</i>
<aside>	Aside	Utterances not related to lecture, e.g., off-topic	<aside> <i>Can you wake him up</i> </aside>
<overview>	Overview	Course discussion, reminders related to lecture, etc	<overview> <i>I think you have gone through the signaling through physiology. This I assume...</i>
<conclusion>	Conclusion	Indicating end or wrapping up of lecture topic	<concl> <i>So today what I'll stop here, today what we talk is introduction...</i>
<q>	Confirmation Checks	<okq>, <rq>, <cq>	<i>forms, is kind of more common,</i> <okq> <i>okay</i> </okq>?

Table 4.2 Tagging Method for SASME lectures⁴⁰

The transcription process was difficult as the audio recording were not as clear as expected. This is due to the lecturers moving around as they lectured, mumbling, unclear accents and movement by students shifting, whispering and other noise issues. Moreover, some of the native

⁴⁰ Complete Tagging List provided in Appendix C1

Arabic lecturers would code-switch unexpectedly or rapidly from English to Arabic and vice-versa. The inconsistent and rapid nature of the CS was challenging and time consuming to transcribe.

4.5.4 Arabic Transcription Style

The Arabic discourse from the SASME corpus has been transcribed using a mixture of Latin alphabetical letters (e.g., *a, b, c*) and Arabic numerals (e.g., *1, 2, 3*). The letters of each token are transcribed into characters representative of the Arabic tokens, for example, '*nihna ingool*' {we say}. As Arabic contains sounds which are not found in English, a corresponding character with similarity when possible has been chosen to allow the reader the ability to pronounce the source language clearly (see page ix for Arabic transcription key).

4.6 Corpus Selection

4.6.1 Saudi Academic Spoken Medical English (SASME) Corpus

The data for the SASME corpus was collected from lecturers mainly through assistance from students or from the lecturers themselves. As previously stated, universities in Saudi Arabia are segregated and a female researcher would be prohibited from attending lectures in the male colleges. Additionally, in order to minimize disturbance and retain as much authenticity as possible, lectures by female lecturers were similarly collected by students or the lecturer herself.

The official medium of instruction in medical colleges is English and while lecturers are strongly encouraged to lecture in the foreign language, many instances of CS (see 3.7) native Arabic lecturers do occur. In some cases, the lecturers incorporate more of the L1, Arabic, than English for various reasons which may include the difficulty of the topic matter, the confidence of the lecturer to use English to conduct lectures or, possibly, pressure from students to use the shared language rather than the foreign EMI.

These aspects were considered following the collection of 30 lectures, each of which were analyzed based on a set of five criteria for inclusion in the NNS English corpus.

- a) **Mode of Instruction:** The lectures must be monologic and taught by one lecturer and not a team (two or more lecturers co-teaching). The lecture should ideally be conversational or what would be described as interactive in style and should not feature a reading style in which a lecturer reads aloud from a script or speaks as though reading from a text.
- b) **Academic Audience:** The lectures selected must contain an audience of undergraduate medical students in year two or three and have at least 35 students in attendance during the lecture. Students are grouped in their first year (e.g., Group A, Group B, etc.) into clusters of around 60 to 70 students and continue to take classes together throughout their program though numbers decrease with each subsequent year (i.e. attrition rate).
- c) **Lecture Duration:** Lectures must be no less than 45 minutes and no more than 60 minutes. An academic lecture is officially defined as 50 minutes; however, lectures may end early or run slightly late.
- d) **Department:** The lectures must be within the fields of medicine and/or medical sciences which would include: pathology, biochemistry, pharmacology, physiology and anatomy.
- e) **Language:** Lectures must be in English as the medium of instruction and if Arabic is used, it should ideally not exceed 35%* of total spoken discourse. (*exceptions were made for select lectures)

The following table (see Table 4.3) lists the lectures which matched the criteria for inclusion. An additional variable was to include at least one lecture from each of the departments within the College of Medicine, namely, Pathology, Anatomy, Physiology, Biochemistry and Pharmacology.

#	SASME Audio Transcript	Lecturer Nationality	No. of English Tokens	Duration	Audience	No. of students
1	<i>Anatomy & Physiology of Sympathetic Nervous System (APSNS)</i>	Pakistani	5680	55	UG2	40-50
2	<i>Cerebral TB (CTB)</i>	Sudanese	4756	47	UG2	40-50
3	<i>Vitamin A (VitA)</i>	Indian	6016	45	UG2	40-50
4	<i>Intro to Degenerative Brain Diseases (IDBD)</i>	Saudi	3486	45	UG2	40-50
5	<i>Drugs used in Epilepsy (DE)</i>	Pakistani	7139	50	UG2	40-50
6	<i>Congenital Malformations and Hydrocephalus (CMH)</i>	Saudi	3986	44	UG2	40-50
7	<i>Anatomy of the Ear (AE)</i>	Egyptian	2939	46	UG2	40-50
8	<i>Drugs used in Anxiety Disorders (DAD)</i>	Sudanese	4506	57	UG2	40-50
9	<i>Physiology of Synapses and Receptors (PSR)</i>	Sudanese	2045	54	UG2	40-50
10	<i>Pathology of Brain Tumors (PBT)</i>	Saudi	2806	46	UG2	40-50
11	<i>Drugs used in Balance System (DBS)</i>	Sudanese	4215	51	UG2	40-50
Total no. of English tokens			47574	9:00 (540 min.)		

Table 4.3 Breakdown of SASME Corpus Lectures

Moreover, with the 11 lecturers, an effort was made to include a range of lecturer nationalities, as illustrated in Table 4.3. The SASME corpus used *AntConc 3.2.4* (Anthony, 2012) as the primary concordance program in analyzing the frequency and functions of DM use in the corpus.

Figure 4.1 Screenshot of Concordance Lines ‘right’

Query **right** 53 (901.5 per million)

First | Previous Page 2 of 3 Go Next | Last

file819832 guess where is the disease? Is it on the **right** or the left? Student: The left. Lecturer
file819832 left? Student: The left. Lecturer: At the **right** there, at the right. [Unclear] Student:
file819832 left. Lecturer: At the right there, at the **right** . [Unclear] Student: Right, right. Lecturer
file819832 there, at the right. [Unclear] Student: **Right** , right. Lecturer: [Arabic]. [Unclear] ventricle
file819832 at the right. [Unclear] Student: Right, **right** . Lecturer: [Arabic]. [Unclear] ventricle
file819832 [Unclear] ventricle, that [unclear]. The **right** ventricle, it's lateral ventricle. This
file819832 are comfortable that we have the diagnosis **right** and we can call it pilocytic astrocytomas
file819831 Congenital malformations is a huge chapter. **right** ? So there's actually books called Pediatric
file819831 light is going through his skull, easy, **right** ? This is the trick, be careful. Megalencephaly
file819831 You don't have grey, [unclear] neurones, **right** . So you have a white matter on the outside
file819831 folic, folic acid, which equates foliates, **right** . Folic acid is deficiency - is major risk
file819831 fossa the cerebral is going to be misshaped, **right** . The shape of the cerebral one will be
file819831 has some entry and exits of the [unclear], **right** . So if you have obstructional abnormalities
file819830 But neuronal death can happen anywhere, **right** ? If you have a tumour pushing, if you have
file819830 up to a strong [Arabic]. So also toxins, **right** , and illness and [unclear] yeah, other
file819830 disease, multiple [unclear] familial [pause] **right** ? Familial [Arabic]. Okay, there might be
file819830 in parkin and this [pause] on the slides, **right** ? In parkin- is one of the proteins which
file819829 these channels can propagate the signal. **Right** ? So that's why you need to understand that
file819829 . Seizure, we don't know the mechanisms, **right** , but in some cases you have [unclear] mechanisms
file819829 see the brain activity. Severe excitation, **right** ? And there is a tonic, right? This is a

First | Previous Page 2 of 3 Go Next | Last

Concordance Lines (as shown above) and detailed analyses of data transcripts are also used to distinguish between instances of DM and non-DM use.

4.6.2 British Academic Spoken English (BASE) Corpus

Collecting the SASME data, the intention was to compare it to a suitable NNS English corpus of spoken academic lectures in the medical and health science fields. The comparative corpus was intended to contrast the three academic variables of SASME - the NNS lecturers, the non-native English audience and lastly, the English as a foreign language country. The choices available that fit the criteria and were freely accessible to the researcher were two large NS projects, MICASE, from the United States (see 2.3.1) and BASE, from the United Kingdom (see 2.3.1). The SASME data allows the unique opportunity to freely choose a corpus from either country as historically,

Saudi Arabia has and continues to hold, ties to both countries through its educational practices and curriculum.

Although MICASE contains a wider variety of speech events, my research is focused solely on large monologic lectures and BASE possesses a greater variety of medically related lecture transcripts that optimally suited the criteria for a comparable corpus.

The lecture transcripts selected from BASE were chosen based on the same five criteria used for SASME:

- a) **Mode of Instruction:** The lectures selected were specifically lectures (monologic) and not seminars, as seminars are often more interactive in nature - with smaller audiences - and the aim of the research is to investigate the discourse of lecturers in a traditional lecture context. In addition, similar to SASME, lectures were conducted by one lecturer and not a team.
- b) **Academic Audience:** The lectures selected must contain an audience of undergraduate medical students in year two or three and have more than 35 students in attendance in order to achieve a similar classroom context to SASME. One exception was made for lecture *lslct008*, which had 20 students as indicated by “<20” (see Table 4.4)
- f) **Lecture Duration:** The lectures that were chosen lasted roughly 40 minutes to an hour, with the exception of one lecture (*lslct028*), which fell just short of the minimum duration at 37 minutes.
- g) **Department:** The lectures must be within the fields of medicine and/or medical sciences and as closely related as possible to the courses in SASME.
- h) **Language:** As BASE is being used as a comparative corpus, the selected courses must be taught by NS English lecturers to contrast the NNS of SASME.

Based on the selection criteria, nine lectures from the Life & Medical Sciences disciplinary group were chosen for the British Academic Spoken Medical English (BASME) comparative sub-corpus. As there were no exact matches between courses in the Life & Medical Sciences group, the nine lectures were partially selected based on lecture titles with similar topics and content and

their departments, namely, Biological Sciences under the University of Warwick and Medicine under the Leicester Warwick Medical School (LWMS). The summary (see Table 4.4) clearly outlines the key factors of BASME lectures selected for this comparative research.

#	BASME Transcripts	Lecturer Nationality	No. of Tokens	Duration	Audience	No. of Students
1	<i>Abdominal Aortic Aneurysms. Trial and Error</i> (Islct017)	NS	5752	0:42:43	UG/PG	130
2	<i>Diabetic Nephropathy</i> (Islct032)	NS	9592	0:59:31	UG2/PG	100
3	<i>Glomerulonephritis</i> (Islct034)	NS	8224	0:50:40	UG2/PG	100
4	<i>HIV and AIDS</i> (Islct008)	NS	5894	0:49:06	UG	<20
5	<i>HIV Infection</i> (Islct037)	NS	5444	0:46:22	UG3	35
6	<i>Introduction to Renal Function. The Concept of Clearance</i> (Islct026)	NS	7830	0:43:05	UG2/PG	140
7	<i>Kidney Stones</i> (Islct028)	NS	6157	0:36:57	UG2/PG	140
8	<i>Hepatitis B Endemicity</i> (Islct018)	NS	6473	0:47:40	UG2	45
9	<i>Concentration and Dilution of Urine</i> (Islct029)	NS	8724	0:48:45	UG2/PG	140
Total no. of tokens*			64090	7:05:00		

Table 4.4 Breakdown of BASME Corpus Lectures selected for Comparison

The BASME data was analyzed with the same procedures used to investigate the SASME data. *AntConc* 3.2.4 (Anthony, 2012) was used as the primary concordance program for analyzing the frequency and functions of DM use in both corpora.

4.7 Corpus Description

4.7.1 SASME

The SASME corpus, unlike the exploratory corpus, SAMEX, contains lectures solely from male lecturers. This is not intentional as every effort was made to include female-led lectures in the corpus. Unfortunately, the female lecturers that were recorded in the data collection featured a higher percentage (50-70%) of Arabic use which does not meet one of the basic criteria of the compiled NNS corpus.

Of the total 30 lectures collected, only three were female (non-Saudi) lecturers as others that were approached declined being recorded for various reasons. This is understandable with respect to the culture and traditions to which many subscribe and thus they would be uncomfortable having their voices recorded. The lack of female lecturers is a limitation in the study; however, it can still be regarded as representative as male lecturers teach both male and female students, while, with few exceptions, female lecturers only teach female students.

Originally, the corpus was intended to consist of solely Saudi lecturers; however, the majority of the faculty in the medical college in second year are non-Saudi expatriates. As a result, the corpus represents the existing faculty dynamics within the College of Medicine, which is made up of mainly Arab lecturers (e.g., Saudi, Sudanese, Egyptian, Jordanian, etc.) and a smaller percentage of non-Arab lecturers (e.g., Pakistani, Indian, etc.)

Of the lectures which did meet the criteria, a random selection were chosen in an attempt to take at least one or two from each sub-discipline (e.g., biochemistry, pharmacology, medicine, etc). Furthermore, every effort is made to represent the different nationalities that typically make up the medical college as this would ideally represent the reality students face with accents, dialects and potentially, cultural academic lecturing styles. That is, non-Arab lecturers were included in the corpus to represent another challenge students face when they share no common language with a lecturer. This may not be an issue for the more proficient students; however, for other students who are struggling with the balance between the (foreign) EMI and studying a dense subject the challenge is immense.

4.7.2 BASME

From available BASE holdings information, the nine lectures included in the research from the Life & Medical Sciences disciplinary group were collected from the Biological Sciences under the University of Warwick and Medicine under the Leicester Warwick Medical School between 2002–2003, with the exception of *lslct008* (Biological Sciences) recorded in 1998. The lectures were taught by six male and three female NS English lecturers.

4.8 Framework for Analysis

4.8.1 DM Framework Development

The DM Framework was developed using the work of Brinton's (1998) list of DM functions (see 2.7), Aijmer's (2002) macro-levels of discourse: textual and interpersonal, Fung and Carter's (2007) multifunctional discourse levels (see Table 2.6) and Bellés-Fortuño's (2006) DM classification (see Tables 2.4) presented in Chapter 2 (see 2.4.4).

Through the exploratory study and initial analysis, DMs were identified as common in both the NS and NNS corpora. Frequent DMs that appeared in the corpora word lists were listed under broad categories of functions, mainly using Bellés-Fortuño's (2006) classification as a starting point (e.g. *additional*, *attitudinal*, *rephraser*, *topic-shifter*, etc.), in addition to other identified functions (e.g., *attention-getters*). Once the list was completed, the functions were compared and contrasted and found, in some cases, to be insignificant or extraneous to the needs of the study. For example, the function of *additional* (e.g., 'and'), which Bellés-Fortuño (2006) calls micro-markers were excluded. This was partially due to its relatively excessive frequency and mainly due to its limited function as additive in the spoken discourse; a finding similar to other studies (Schiffrin, 1987). Furthermore, functions such as attention-getter were absorbed into other groups as the DMs in this sub-set were found to simultaneously function as topic starters, developers, closers, etc.

Once the final list was established and using Aijmer (2002) levels of discourse as a starting point, the study's DMs list was found to display two general levels of discourse: textual and interactive. The first level was named *Structural* DMs. I prefer to use the label 'Structural' (cf. Müller, 2005) for this function than 'Textual' (cf. Aijmer, 2002; Brinton, 1996) because these DMs operate on both a local and global level of discourse to signal the relation between and/or across utterances. The groups that were observed to function under this category of Structural DMs are Topic Initiator, Topic Developer, Summarizer and Closer (see 4.8.2). The second level, interactive DMs, are identified as functioning on an interpersonal and active level that indicates speaker interest or lack thereof and aids in shortening the distance between speaker-hearers. This category was then named *Interactional* (cf. Müller, 2005) and the groups observed to function under this category are Confirmation Checks, Elicitors and Rephrasers (see 4.8.3).

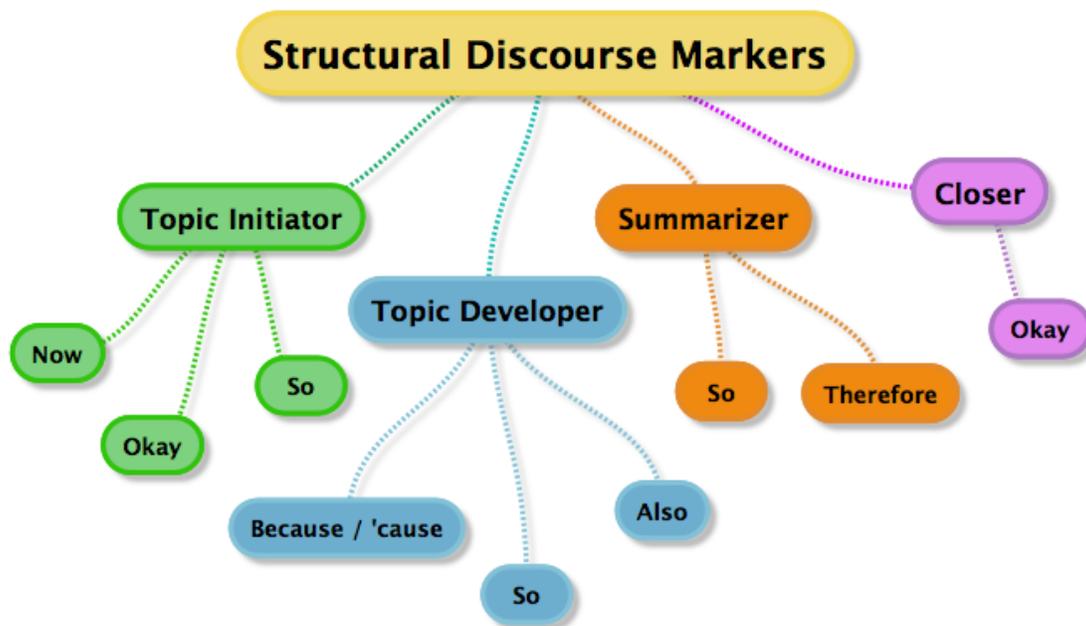
The process of establishing the framework for the DM classification was time consuming and painstaking, but ultimately, allowed for a more organic development of context based DM functions. The work of Schiffrin (1987), Fraser (1999, 2006), Morell (2000), Nattinger and DeCarrico (1988), Fung (2003), Othman (2010) and others presented in the Academic Discourse chapter (see 2.4), all also contributed, in part, to the development of the DM classification established. The following section (see 4.8.2) presents and illustrates the features of the DM Classification Framework used for this study.

4.8.2 English Discourse Markers Classification

The DMs in the corpora are divided into two categories - Structural DMs, which include four functions: Topic Initiators, Summarizers, Topic Developers and Closers; and Interactional DMs, which include three functions: Confirmation Checks, Elicitors and Rephrasers. This section will discuss the English DMs that occur in SASME; Arabic examples are illustrated for context and will be discussed in detail in chapter 6 (see 6.6).

The first category, Structural DMs, are a type of marker that Aijmer (2002: 4) describes as textual and may function on either a "qualifier" or "frame" coherence level. In essence, this category has numerous functions which organize the discourse from shifting topic, framing

discourse, elaborating a subtopic and drawing attention to aspects of an argument. The functions under this category are fundamentally used to signal internal or ideational relations within or across utterances. These markers link ideas and make discourse easier to comprehend and follow. Each function is described briefly and exemplified with excerpts from the SASME data.



Mind-Map 4.2 Functions of Structural Discourse Markers

The first function under the categorization of Structural DMs is *Topic Initiator* which occurs at the beginning of a section of talk, usually after a brief pause indicated by a full stop or comma and functions as a boundary marking the introduction of a new (sub)topic. The DM ‘so’, briefly examined by Howe (1991: 93) is shown to function as a “marker of connection” in its utterance-initial position. Schiffrin (1987: 191-203) also introduced ‘so’ as a marker of main idea units, though her research describes it in terms as complementary to ‘because’ - when the speaker leaves the main argument with ‘because’ to introduce details and explanations, the speaker returns to the main topic with ‘so’. Overall, Topic Initiators function as signals that introduce a sequences of actions (Sacks et al., 1974; Schegloff, 2006).

Function 1A

- [VitA] (a) *So, let's see what is basically vitamin A, okay? Vitamin A, just like any other vitamins, is essential.*
- [APSNS] (b) *Here in the somatic they are myelinated, therefore they conduct impulses. Now, the three main differences with somatic nervous system,*
- [PSR] (c) *What is the transmitter involved in it? [Arabic]⁴¹ glutamate [Arabic]. Okay. What is the cerebral function of LTP? Information of memory.*

The excerpts in 1A illustrate the function of Topic Initiator DMs to signal an introduction to a new aspect of the discourse topic. In example (a) the DM ‘so’ indicates to the listeners (students) that the topic will now shift to a more basic description of vitamin A. With the use of “let’s see” we can assume that the lecturer is inviting the students to join him in a basic revision of what the vitamin is. Also by using a face-saving strategy of inclusive let’s = let *us* see, the lecturer prevents any threat to students’ basic knowledge.

In example (b), the DM ‘now’ indicates a boundary of what was previously discussed and what will subsequently be discussed - the differences with the somatic nervous system. The last example, (c) similarly illustrates an end (boundary) to the previous topic; it functions as a revision in Arabic of the information preceding it, which was discussed in English. Code-switching from English to Arabic often occurs to repeat information in the shared language and is discussed further in 6.6.2.

Topic Developer is the second function of Structural DMs ‘so’, ‘because’ and ‘also’, which often occur utterance-medial and develop and/or elaborate on the previously introduced topic. For example:

Function 2A

- [DE] (a) *Some people are in the habit of playing these video games, so again like watching TV, because the waves are basically interfering with your normal brain activity, so it's not good.*

⁴¹ As the Arabic discourse is a distraction and unnecessary for these illustrations, all instances are labeled [Arabic]

Function 2A (*continued*)

[CTB] (b) *It means that this is very important for the economy of the country, **because** if you say that this disease is airborne manner, I have to build a room for airborne transmission, which costs a lot of money.*

[PCMH] (c) *So, if you have obstructional abnormalities you would have hydrocephalus and remember that these **also** are associated with Arnold-Chiari type 2 with lumbar...*

The excerpts above illustrate the lecturers' use of the DMs to develop the topic such as (a), where 'so' is used to elaborate on the preceding information by connecting the act of playing video games to that of watching TV. In (b), the lecturer's use of 'because' clearly develops on the consequences of declaring a disease as "airborne" by describing its effect on the local countries economy. Lastly, example (c) 'also' is used to develop the links of certain symptoms to a spinal malformation (Arnold-Chiari), which was previously introduced.

The third function is that of *Closer* which occurs at the end of discourse and signals the close of a topic or point in the lecture. The sole English DM to function in this manner is 'okay'. The following examples illustrate its use in lecture discourse:

Function 3A

[PBT] (a) *These [unclear] families [unclear] astrocytomas, oligodendrogliomas and ependymomas, **okay**. Now, there is [unclear] tumors.*

[IDBD] (b) *... the presence of alpha-synuclein in the neuron bodies suggested that there is a thicker degeneration of the protein in the protea [unclear] **okay**. Why do we talk about this? Because it...*

The first example (a) concludes a discussion of primary brain tumors and moves on in the following utterance to discuss prototypes. The second DM (b), similarly functions to close the discourse, describing neuron degeneration and connects it back to the main overall focus of the lecture - brain diseases.

The fourth and last function under this category is *Summarizer*. The DMs ‘*so*’ and ‘*therefore*’ often occur following a dense chunk of lecture discourse and provide a meaning of ‘as a result’ or ‘from the previous information, we can understand that x’. The DMs function as Summarizers illustrated in the following excerpts:

Function 4A

[DRBS] (a) ...when released histamine bonds synaptically on three receptors H1, H2 and H3. **So**, histamine has got three receptors H1, H2, H3 situated postsynaptically but histamine can also act on presynaptic receptor, H3 receptor.

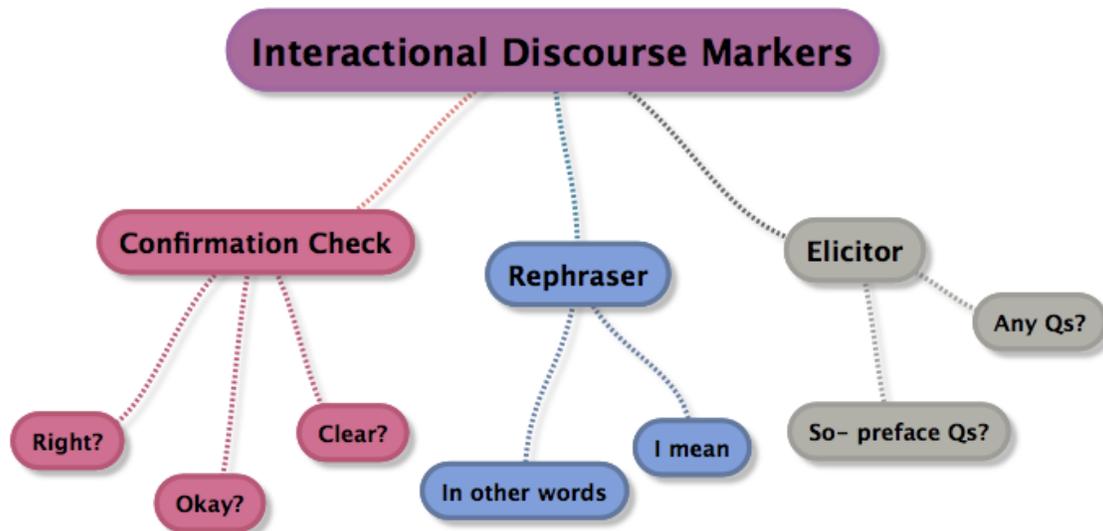
[DAPD] (b) They alleviate the somatic symptoms of anxiety, **therefore** they are useful in case of stage fright or performance anxiety. Like the anti-depressant...

Example (a) illustrates the DM ‘*so*’ signaling a summary of preceding information, which is often informationally dense. The lecturer summarizes histamine’s three receptors names, location and its other actions. The second example (b), further illustrates the function of the category to show the result of information. In this Excerpt (b), the lecturer is describing the use of beta-blockers to alleviate the symptoms of anxiety and uses the DM ‘*therefore*’ to provide a meaning of ‘as a result’. In occurrences alone, the DM ‘*so*’ is used twice as frequently as ‘*therefore*’. The potential reasons for this will be discussed further in Results I (see 5.4.3 and 5.5.3).

4.8.3 Interactional Discourse Markers

The second category is that of Interactional DMs which is based in part on Müller’s (2005) description which pertains to emotions and attitudes involved in the use of these DMs. It is described in broad terms to include various kinds of interactions from conversations to speeches or sermons given to an audience. These DMs play an important role in shortening the distance between speaker-hearer and help foster a level of intimacy between interactants. Elicitation markers, Confirmation Checks and Rephrasers are categorized as Interactional DMs that operate

on an interactive level between the speaker and hearer even when the function is largely rhetorical.



Mind-Map 4.3 Functions of Interactional Discourse Markers

The first function under this category is *Confirmation Checks* (*okay? right? clear?*) that occur in the middle or end of a dense chunk of discourse and function as a means for lecturers to verify students are following the development of the topic and confirm understanding. For example:

Function 1B

[DRBS] (a) ...they produce extra perimeter side effects because they block dopamine receptor in the basal ganglia. Is that **clear**?

[VitA] (b) ...you have difficulty seeing. Because it takes some time for rhodopsin to regenerate in the dark, **right**?

[DE] (c) So these are the types, **okay**? Like partial, just only one part of the brain is affected. This is called partial, not full and then

The examples of Confirmation Checks illustrate its function in the lectures which is to verify student understanding. In example (a), the lecturer uses the DM ‘*clear?*’, one of the most transparent in terms of its function, as a Confirmation Check. The variations in the three DM’s overall occurrences will be discussed in more detail in Results II (see 6.4.1).

The second function is that of *Elicitors* which prefaces sequence-initiating actions (questions); its function is to elicit responses from listeners (students). The Elicitors identified in the data are the multi-word units (MWUs), ‘so -prefacing question’ and ‘*any questions?*’. For example:

Function 2B

[VitA] (a) **So** *what do you mean by light sensitive? When light strikes on this complex it can change its confirmation, okay?*

[DRBS] (b) *Is it clear? So how is balance maintained?*
Student: *Through the cerebellar.*

[APSNS] (c) *...there is no concept, there's no action of the function of that response in the autonomic nervous system. So these are present at various places in the body. Any questions? Yes please?*

These MWUs function as a means for lecturers to draw out information from the students; however, while this is effectively illustrated in examples (b) and (c), the first example (a) functions as a pseudo-Elicitor. Meaning that, the lecturer, after a barely perceptible pause following the MWU, moves on to answer his own question. Elicitors are often used in the data to obtain information and in some cases are used rhetorically as attention-getters (see 6.4.2).

The third and final function, under the Interactional DMs category, *Rephrasers* are used by lecturers to reword information that may have been too dense or complex for students. While it occurs in the English discourse, it is used more frequently by the native Arabic lecturers in SASME with Arabic DMs such as ‘*ya3ni*’ {means; in other words}. This is discussed in more detail in the chapter 6 (see 6.6.3). The Rephraser DMs are generally used with the first person pronoun (I); however, there are some occurrences of second person (you) pronoun use.

Function 3B

[DE] (a) *...this is what I mean neurological science says and the neurologists diagnose...*

[APSNS] (b) *Mainly in physiology you will - it means going to read - in other words, there is a brainstem.*

The two categories of Structural and Interactional DMs and the functions listed under each category, are presented in Table 4.5.

Category/Function		Definition	English DMs
Structural DMs		Are markers used to organize and signal transitions and emphasis in lecture discourse. There are four functions identified as Structural: Topic Initiator, Topic Developer, Summarizer and Closer.	
1	Topic Initiator	Occurs as utterance-initial. Its function is to introduce a new (sub)topic and acts to introduce a sequence of action; it may also act as an attention-getter	<i>So</i> <i>Now</i> <i>Okay</i>
2	Topic Developer	Occurs as utterance-medial and develops or elaborates on previously introduced idea or topic	<i>So</i> <i>Because /</i> <i>'cause</i> <i>Also</i>
3	Summarizer	Occurs as utterance-medial or utterance-final of a dense chunk of discourse and provides a meaning of 'as a result' or 'from the previous information, we can understand that x'. It may also act as an attention-getter	<i>So</i> <i>Therefore</i>
4	Closer	Occurs as utterance-final and indicates closure of a point or topic in the lecture; it may also act as an attention-getter	<i>Okay.</i>
Interactional DMs		Are markers which play an important role in speaker/hearer interaction, even when largely rhetorical. There are three functions identified as Interactional: Confirmation Check, Elicitors and Rephrasers.	
5	Confirmation Check	Occurs as utterance-medial or utterance-final and indicates lecturers' need to check students' understanding of information discussed to that point. The DMs are often marked with a rising intonation (indicated with a '?' mark) and is often in the form of a tag question	<i>Okay?</i> <i>Right?</i> <i>Clear?</i>
6	Elicitor	DM 'so' prefaces sequence-initiating actions (questions); its function is to elicit responses from listeners (i.e. students). It may also act as an attention-getter	<i>So-preface</i> <i>question;</i> <i>Any</i> <i>questions?</i>
7	Rephraser	This DM seems to occur more frequently with first person singular pronoun and is used by the lecturer to indicate a rewording of what was stated in an immediately prior utterance, for a better understanding by the students. It may also act as an attention-getter	<i>I mean</i> <i>In other words</i>

Table 4.5 Discourse Marker Classification Framework for Study

4.8.4 Inter-rater Agreement

In order to test the validity of the Structural and Interactional DM classification, two independent inter-raters⁴² were selected for a statistical measure of inter-rater agreement in order to strengthen the DM Framework. The overall design of the study was explained to coders and each category and their functions were presented and discussed. Sample excerpts with DMs in context were worked out by each coder, with researcher present, as practice before they began their evaluation. Each coder was then given a worksheet with the existing DM framework, a brief description of each category and 30 excerpts (representing 4–5 samples of each function) from the SASME and BASME corpora. Based on their labeling and feedback, the DM classification framework was refined until agreement was achieved overall.

The inter-rater reliability analysis (Geertzen, 2012; Freelon, 2013) using Cohen's Kappa and Krippendorff's Alpha statistics were performed to determine agreement among coders. The inter-rater reliability for the coders was found to be Cohen's Kappa = 0.866 ($p < 0.001$), which is interpreted as *almost perfect agreement* between raters and Krippendorff's Alpha = 0.868, which is considered reliable at $\alpha \geq .800$.

4.8.5 Arabic Discourse Markers Classification

The calculations for the total number of Arabic tokens in each lecture were estimated. The calculations were compromised with a series of linguistic code issues in the lecture transcriptions and thus estimations had to be made. Issues were, for example, occurrences of code-switching which were unclear due in part to audio quality and also to NS Arabic lecturers speaking indistinctly and at a faster pace in the L1. Furthermore, a few occurrences of Arabic discussions that were off-topic were not included in the transcription as participants (i.e. lecturers) spoke in lowered voices and audio was indecipherable. Table 4.6 provides the calculated approximate number of total Arabic tokens used within each of the lectures.

⁴² Coders are both recent PhD graduates in Applied Linguistics, have teaching experience and are non-specialists in the medical health field.

#	SASME Audio Transcript	Lecturer Nationality	No. of English Tokens	No. of Arabic Tokens
1	<i>Anatomy & Physiology of Sympathetic Nervous System (APSNS)</i>	Pakistani	5680	∅
2	<i>Cerebral TB (CTB)</i>	Sudanese	4756	1264
3	<i>Vitamin A (VitA)</i>	Indian	6016	∅
4	<i>Intro to Degenerative Brain Diseases (IDBD)</i>	Saudi	3486	1123
5	<i>Drugs used in Epilepsy (DE)</i>	Pakistani	7139	∅
6	<i>Congenital Malformations and Hydrocephalus (CMH)</i>	Saudi	3986	1050
7	<i>Anatomy of the Ear (AE)</i>	Egyptian	2939	1287
8	<i>Drugs used in Anxiety Disorders (DAD)</i>	Sudanese	4506	1141
9	<i>Physiology of Synapses and Receptors (PSR)</i>	Sudanese	2045	1241
10	<i>Pathology of Brain Tumors (PBT)</i>	Saudi	2806	1145
11	<i>Drugs used in Balance System (DBS)</i>	Sudanese	4215	425
Total no. of tokens			47574	8676

Table 4.6 Breakdown of SASME Corpus Arabic Tokens

The Arabic DMs identified in this research are those with both a high frequency of occurrences across the SASME corpus, as well as functioning in a similar manner to the English DMs analyzed in both the SASME and BASME sub-corpus.

The lectures by native Arabic speakers included numerous short (single tokens or a phrase) and long (propositional utterance) CS events. For the purposes of this research, focus is placed on those DMs which occurred as a link between English utterances or served to initiate a shift in linguistic code. For example:

[PBT] *Al* {the} *location fe al* {in the} *brain jidan muhim mu bas lil* {very important not only for} *staging ...*

While many similarities were identified in regards to the DM functions, their use in Arabic within a predominately English L2 medium of instruction provided an additional effectiveness to utterances. For example, compare the following utterance keeping in mind the audience in this and all other lectures, are homogeneous native Arabic speakers:

[CTB] *I know you don't know, but I want you to think. Think. Someone diagnosed with TB mish kitha?* {is this not right?} *wa ba3dan kan 3ando* {and later he had} *Brucella...*

With the inclusion of code-switching DMs and chunks of discourse, NNS lecturers are reminding students they are part of the same linguistic community (Ferguson, 2003; Lawson and Sashdev, 2000; Holmes, 2000) and providing students with minor breaks from decoding the foreign language of instruction (Ferguson, 2003; Al-Nofaie, 2010). As discussed in the code-switching chapter (see 3.8.2), NNS students often prefer CS to either pure L1 or L2 academic discourse (Zabrodskaia, 2007; AlEnizi, 2010; Sabbour et al., 2010).

The following Tables 4.7 and 4.8 provide a summary of the CS of Arabic DMs that were identified in the NS Arabic lecturers' transcriptions (see 4.5.2) and the identified purposes of CS in the NS Arabic academic medical discourse (see 3.8.1 and 3.8.2).

Category/Function		Definition	Arabic Examples
Structural DMs		Are used to organize and signal transitions and for emphasis in lecture discourse. They are also used as attention-getters through CS. There are four functions identified: Topic Initiator, Topic Developer, Summarizer and Closer.	
1	Topic Initiator	Mostly occurs as utterance-initial. Its use in Arabic is subtle, yet powerful as an attention-getter to signal the introduction of a new (sub)topic.	'fa' {so}, 'khalina' {let's}, 'gool' {say}
2	Topic Developer	Occurs as utterance-medial and develops or elaborates on previously introduced idea or topic. The use of an Arab DM emphasizes its function as an attention-getter.	<i>hata</i> , 'zay', 'aidhan' {also}, 'mithl', 'mathalan' {for example}, 'laan', 'laanu', '3lashan (kitha)' {because}
3	Summarizer	Location varies in respect to code-switching and may be utterance-initial, utterance-medial or less frequency, utterance-final. Provides a meaning of 'as a result' or 'from the previous information, we can understand that x'	<i>3ibara (3an)</i> {meaning}, ' <i>3amooman</i> ' {generally}, 'ghaliban' {in general; for the most part}
4	Closer	Occurs utterance-final and indicates closure of a point or topic in the lecture	'tayib', 'n3am' {okay},
Interactional DMs		Play an important role in speaker/hearer interaction, even when largely rhetorical. They are also used as attention-getters through CS. There are three functions identified: Confirmation Check, Elicitors and Rephrasers.	
5	Confirmation Check	Occurs as utterance-medial or utterance-final of discourse and indicates lecturers' need to check students' understanding of information discussed to that point through the shared L1. The Arabic DMs often follow up utterances or chunks of discourse with dense or complex information.	<i>suh?</i> {right}, 'na3am?' {yes}, 'suh wala ghalat?' {right or wrong} 'mafshoom?', 'mashy?' {understood}, 'fahmeen 3alay?' {do you understand}, 'wadhih?' {clear}, 'tayib?' {okay}
6	Elicitor	Arabic DMs used for elicitation are rarely single tokens and are often an invitation for students to respond in the L1.	'hatha s3ab wala sahal?' {is this hard or easy}, 'ay sewal?' {any questions}, 'ahad y3raf?' {anyone know}, 'laysh...?', 'aysh...?' {why}, 'ma3nâh shino?' {meaning what} 'ma3a ya shabab?' {are you with me boys}
7	Rephraser	Rephrasers are used to signal that prior information will be reworded. However, the use of an Arabic DM does not automatically signal the following utterance will be in the L1.	'ya3ni' {means; that is to say}, 'ida' {what}, 'nahna ingool' {we (can) say}, 'bin gool' {we say}

Table 4.7 Arabic Discourse Marker Classification Framework

The roles and purposes for NS Arabic lecturers' CS are summarized in the following Table 4.8 and are based on the studies discussed in Chapter 3 (see 3.8). The CS classification of roles are discussed and illustrated in Results II (see 6.6.4).

#	Role of Code-Switching	Purpose
1	Solidarity	L1 is used to indicate a shared linguistic bond with listeners and co-membership in community.
2	Reiteration	L1 used to repeat or rephrase immediately prior discourse to ensure listen comprehension.
3	Elaboration	L1 is used to provide addition details, examples and/or analogies that may assist in topic comprehension.
4	Topic	L1 is used in place of the L2 in matters best expressed in the native language, for example, religion, politics, personal matters, etc.
5	Elicitation	L1 used to invite listeners to respond and are an indirect invitation for listeners to respond in the L1.
6	Checking Comprehension	L1 used to confirm whether listeners have understood the prior information and invites listener response, though not required, in the L1 (or L2, if listener feels confident).
7	Classroom Management	L1 used as an attention-getter and for off-topic classroom management.

Table 4.8 Classification of Code-Switching Roles in SASME Lectures

4.9 Conclusion

This chapter gives an overview of the structure and content of the study. I have noted that the study attempts to identify and classify the most frequently spoken DMs in the medical discourse of NS and NNS English lecturers. The classification framework divides the DMs into two distinct categories: Structural and Interactive. Structural DMs are multifunctional at a textual level and can be identified as Topic Initiators, Topic Developers, Summarizers and Closers. Interactional DMs are multifunctional at a speaker-hearer interactive level and can be identified as Confirmation Checks, Elicitors and Rephrasers. These categories are applied to English DMs identified in the native sub-corpus, BASME and the non-native corpus, SASME.

This analysis is complemented by an additional investigation of DMs which function similarly in Arabic and of the role of CS in EMI lectures in the SASME corpus of native Arabic lecturers.

CHAPTER 5
ANALYSIS OF RESULTS I:
STRUCTURAL DISCOURSE MARKERS

“The analysis of discourse markers is part of the more general analysis of discourse coherence—how speakers and hearers jointly integrate forms, meaning, and actions to make overall sense out of what is said.”

(Schiffrin, 1987: 49)

Chapter 5 Analysis of Results I: Structural DMs

5.1 Introduction

This chapter discusses the results of one category of English discourse markers (Structural DMs) to investigate the similarities and differences of their functions in the Saudi Academic Spoken Medical English (SASME) Corpus and the British Academic Spoken Medical English (BASME) corpus. The following chapter (Analysis of Results II) will discuss the results of the Interactional DMs and the Arabic DMs analyzed in the SASME corpus and compare their use and functions with the English counterparts in both the Saudi and British corpora.

The study is intended to investigate and provide insight into the frequency and functions of DMs used in medical based lectures in Saudi Arabia.

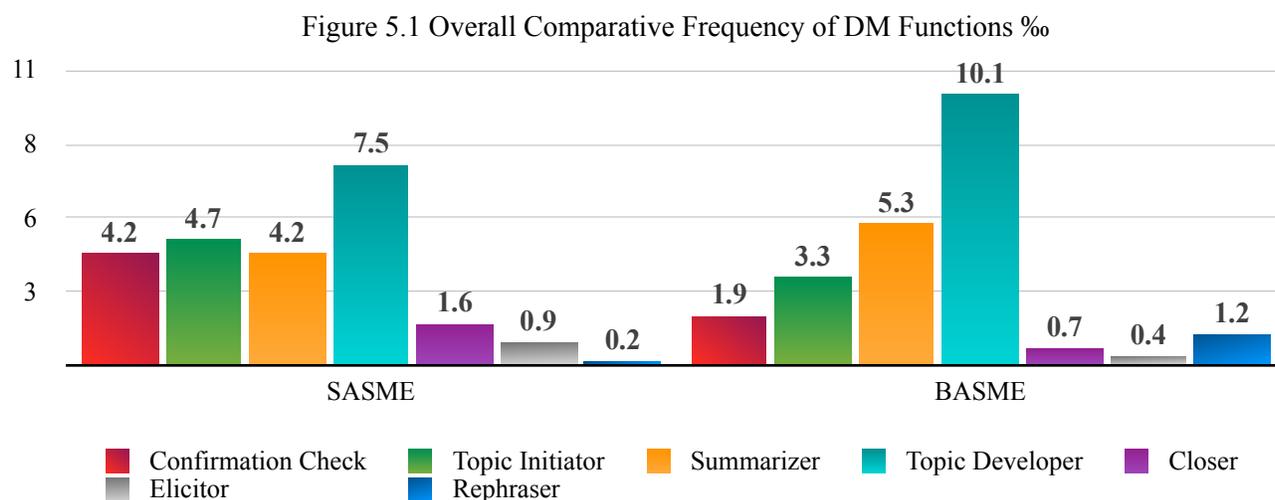
The chapter is organized into three main sections:

1. First, an overall view of the results of the pre-established classification of DMs used in the corpora. I will also briefly discuss the discourse marker breakdown by lecturer, native language status and corpus, for insight into the usage and frequency within each academic lecture.
2. Second, a discussion of the results of the Structural DMs' functions, uses and frequency of occurrences in SASME and BASME.
3. Third, a summary of the results of Structural DMs is given.

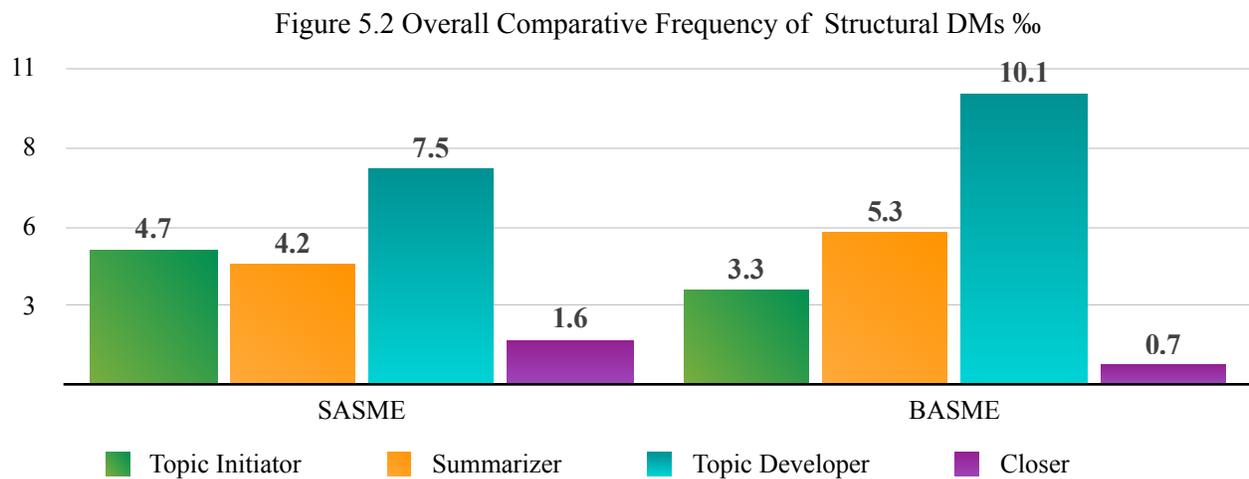
5.2 Results of Discourse Markers in SASME and BASME: An Overview

When analyzing the SASME data, the research questions ask whether a higher frequency of DMs overall would be found in comparison to BASME, a native English sub-corpus. The research questions investigate which category of DM, Structural or Interactional, have more overall occurrences. Additionally, the research investigates whether a higher frequency of single-unit DMs (e.g., *okay*, *so*, *because*, etc) would be found in comparison to BASME, which might contain more MWUs (e.g., *I mean*, *any questions?*). The following figures provide an overview of the comparative frequencies found in the NS and NNS corpora.

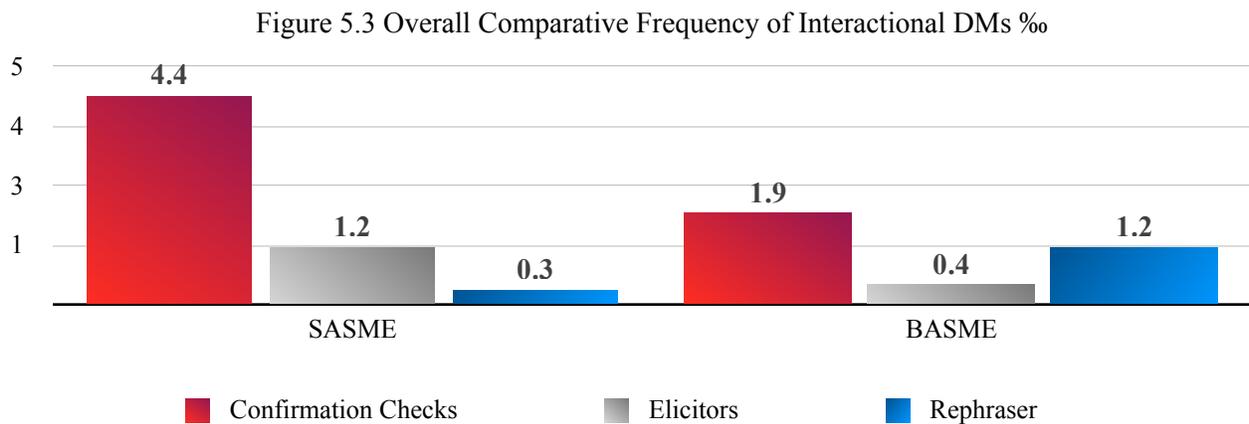
Overall, the number of DM occurrences were relatively equal with BASME at 22.9 occurrences per 1000 (hereafter, %) and SASME only slightly higher at 23.8%. A comparison between the corpora DMs provides evidence that these types of markers found in SASME are also present in the BASME sub-corpus. However, the patterns of frequency for Structural and Interactional DMs do differ between the two corpora.



In BASME, the most frequent DM is ‘*so*’, a Structural marker used to develop topics within the discourse at 5.3% occurrences and is, in fact, a great contributor to the category of Topic Developer (*so*, *because*, *‘cause*, *also*) and ranks highest in frequency at 10.1%. Similarly, Topic Developer DMs have the highest frequency in SASME at 7.5%, but the individual DM that ranked highest in frequency is an Interactional DM within the Confirmation Check category, ‘*okay?*’ with 3.4%.



As illustrated in the two Figures (5.1 and 5.2) neither corpus has a discernible pattern when viewed overall and thus further breakdowns of the patterns by category function and by individual lecturer are needed to investigate whether a perceptible pattern emerges of DM use in academic medical health lectures.

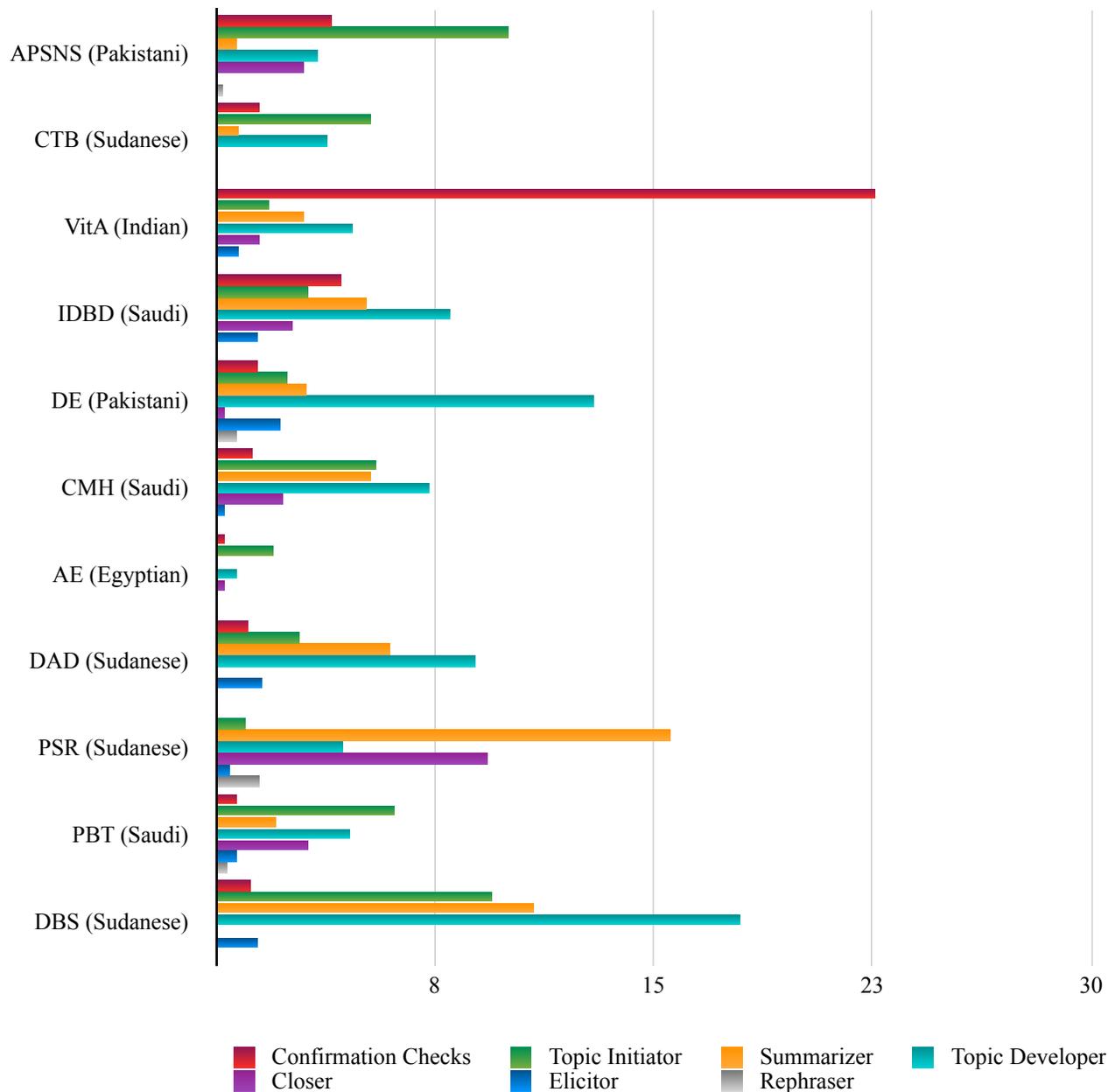


As illustrated, the Interactional DM category functions (see Figure 5.3) are not as similar or as consistently distributed, as the frequencies of the Structural DM markers comparatively (see Figure 5.2). In both SASME and BASME, the Topic Developer DMs are the most frequent with 7.5% and 10.1% overall occurrences respectively.

5.2.1 SASME Lectures

In this chapter, when the distribution of English DMs are discussed, note that I am discussing and comparing only English elements of the lectures which are present in all eleven lectures of SASME (see Figure 5.4).

Figure 5.4 Overall DM Category Distribution across SASME %



However, a total of eight of the eleven lectures have a lower total number of tokens and relatively lower frequencies of DMs than the three non-Arab lecturers' discourse in SASME and

overall when compared with BASME in most cases. This is an important factor to keep in mind as eight of the lectures have the relatively unique option of incorporating two languages (Arabic and English) in the medical discourse. Code-switching has been researched in terms of adult and youth social dynamics and in the classroom (see 3.7), though very little research has been published, to my knowledge, on its effects in higher education and more specifically, in medical colleges which adopt English as the official EMI.

The distribution⁴³ of the DM categories across SASME are shown in Figure 5.4, which illustrates the variation in frequencies between lecturers. One similarity between lecturers, is that at least one function occurs with a considerably higher frequency than other functions in that lecture. For example, in the *AE* lecture, the Topic Initiator marker has a higher number of total occurrences than the other six functions. On the other side of the spectrum, the *VitA* lecture has one function, Confirmation Checks, that is not only considerably higher than any other function in that lecture, but, additionally, in all the lectures in SASME.

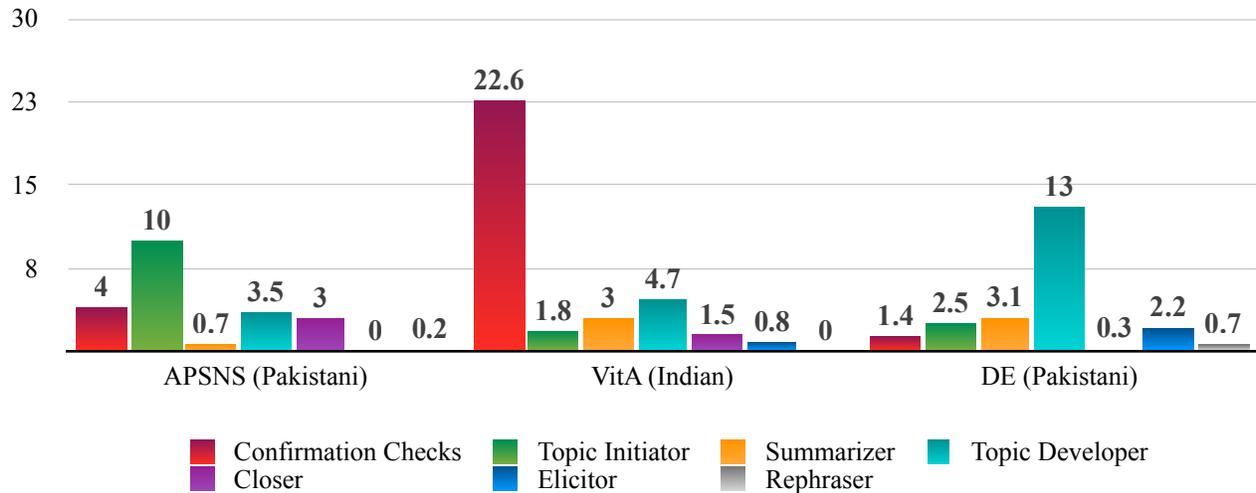
5.2.1.1 Non-native Arabic Lecturers

Three of the SASME lecturers were non-Arabs, as well as non-native English speakers and are representative of the relative percentage of non-Arab staff that currently makes up most medical colleges in Saudi Arabia.

In this chapter, I will examine in more detail the effect that these three lecturers had, especially the *VitA* course lecturer, on the total frequency and distribution of DMs in the SASME corpus. As illustrated in the figure below (see Figure 5.5), the DM results were relatively consistent with their SASME counterparts in that they were generally at or below five occurrences per 1000 (‰). Three separate functions though were considerably higher than most, if not all, other lectures.

⁴³ See Appendix D1 and D3 to view table of SASME overall distribution across lectures per 1000 occurrences

Figure 5.5 Non-Arab Lecturers' DM Distribution %



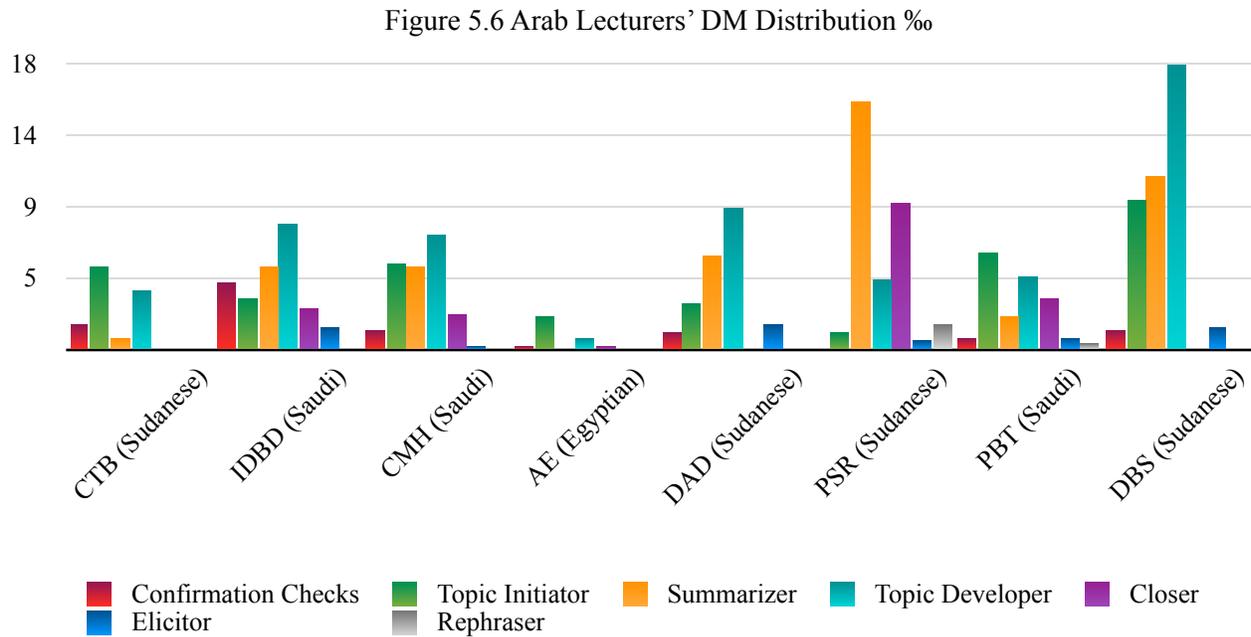
In the case of the *APSNS* course, its frequency of 10% *Topic Initiator* DMs placed it as the highest number of occurrences overall although only by 0.5% occurrences more than the Sudanese *DBS* course lecturer. The *Topic Developer* DM function is the second highest in the *DE* lecture with 13%, but once again, another non-Arab lecturer is surpassed by the Sudanese *DBS* course lecturer with 18%.

In the third noted case, the function of *Confirmation Checks* in the *VitA* course, the lecture is not only the highest in number of occurrences in SASME, but is, in fact, the highest overall in the corpora with 22.6%. That is more than quadruple the next highest number of occurrences in SASME and about three times greater than the highest number of occurrences in BASME. Each of the functions and the instances of exceptional occurrences will be discussed in further detail in the Discussion chapter (see 7.3). These results indicate that, in general, the non-Arab lecturers (see Figure 5.5), have an overall lower occurrence of DMs in contrast to their native Arab peers.

5.2.1.2 Non-Arab Lecturers

The illustrative figure (see 5.6) of the native Arabic lecturers in SASME, shows that the DM functions that occurred at the highest frequencies were not the same as those with the non-Arab lecturers (see Figure 5.5). From Figure 5.6, we see that the highest number of occurrences overall are *Topic Developers* across the lectures. Second, *Summarizers*, followed by *Topic Initiator*. This

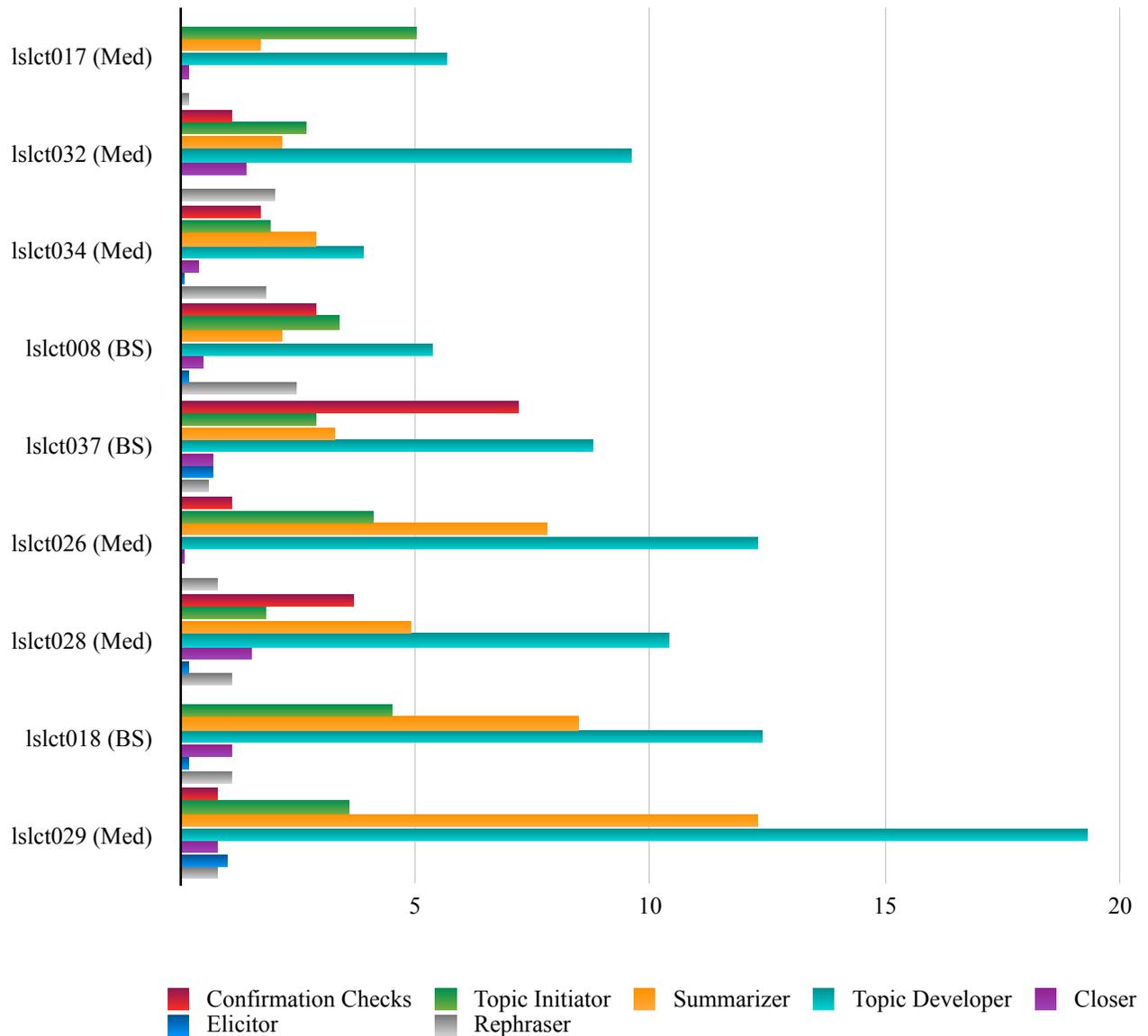
suggests that the native Arab lecturers may share this relative pattern of DM usage with the BASME medical lectures (see Figure 5.8) and to a lesser degree with the Biological Science lecturers (see Figure 5.9).



5.2.2 BASME lecturers

As can be seen in the overall figures of the SASME lecturers (see Figure 5.6) and the BASME lecturers (see Figure 5.7), we can initially conclude that the distribution of DMs throughout the corpora are comparable. We note that the majority of DM functions occur at or less than, 5% occurrences across the BASME lecturers, with exception to three functions - Topic Developer, Summarizer and Confirmation Check. The reasons for this and a more detailed analysis of BASME and its comparability to SASME will be given in the Discussion chapter (see 7.3).

Figure 5.7 BASME Overall DM Distribution across Lectures %

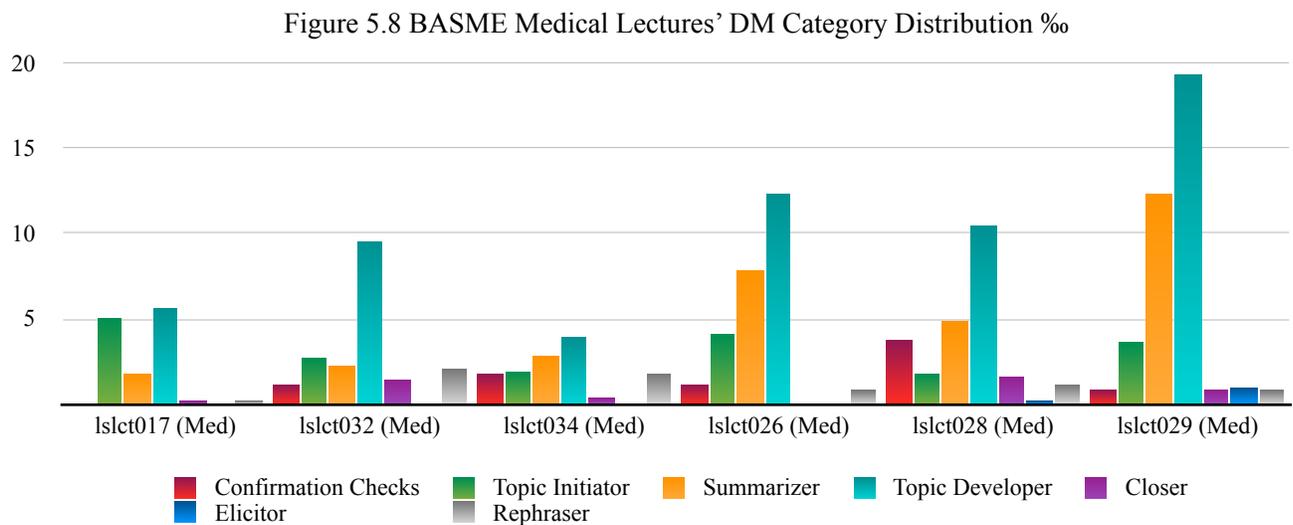


The BASME transcripts compiled for this research are from two academic fields within the Life & Medical Sciences, namely, Medicine and Biological Sciences. The following Figures (5.8 and 5.9) were created to illustrate the breakdown of DM occurrences by function and by academic field⁴⁴.

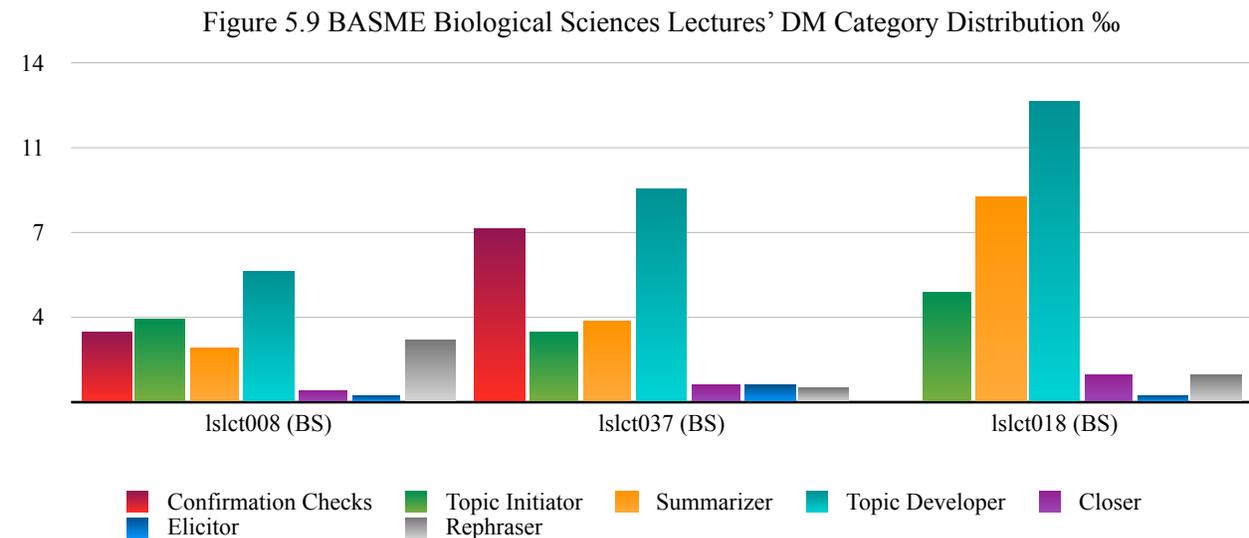
The first figure (5.8), illustrates the DMs by function in the six medical lectures. The three functions, Topic Initiator, Summarizer and Topic Developer, show a higher number of

⁴⁴ See Appendix D2 and D4 to view table of BASME overall distribution across lectures per 1000 occurrences

occurrences in at least five of the BASME medical lectures. The Topic Developer function shows a higher number of occurrences than any other function in each of the medical lectures, followed by Summarizer and then Topic Initiator.



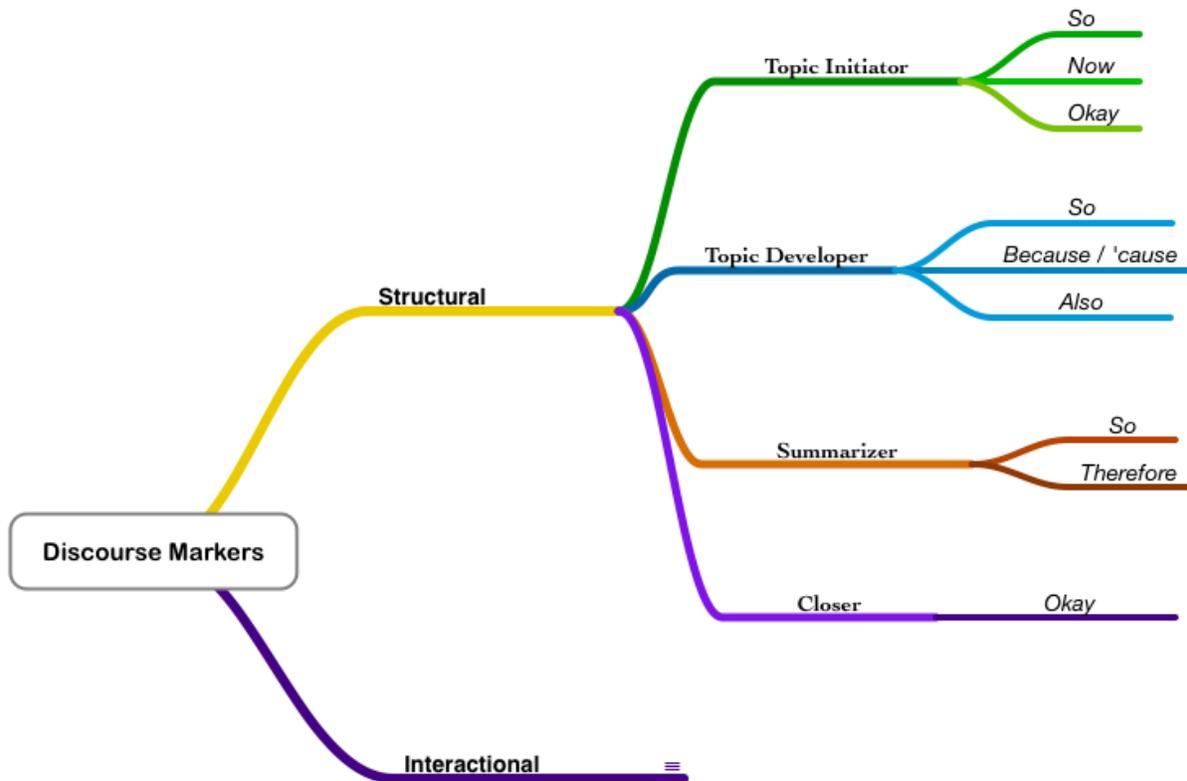
The results vary in Figure 5.9 in contrast to Figure 5.8. In both, they have a similar pattern with Topic Developer occurrences as the highest overall frequency relative to other functions in the Biological Sciences lectures. However, while Summarizers are the second highest in one lecture (*Islct018*), they are third and fourth in total occurrences in the remaining two, *Islct037* and *Islct008* respectively.



Overall, Confirmation Checks have a higher number of overall occurrences across the Biological Sciences lecturers and relatively insignificant numbers across the medical lectures.

5.3 Structural Discourse Markers

Structural DMs are mainly used to signal internal or ideational relations within utterances or propositions. This category of DMs links ideas and is intended to make discourse easier to comprehend and follow, with clear linguistic cues such as ‘*so*’ and ‘*therefore*’. The four functions identified as Structural DMs are Topic Initiators, Topic Developers, Summarizers and Closers (see Mind Map 5.1).



Mind-Map 5.1 Structural Discourse Markers

5.4 Structural Discourse Markers in SASME

5.4.1 Topic Initiator (*so*, *now*, *okay*)

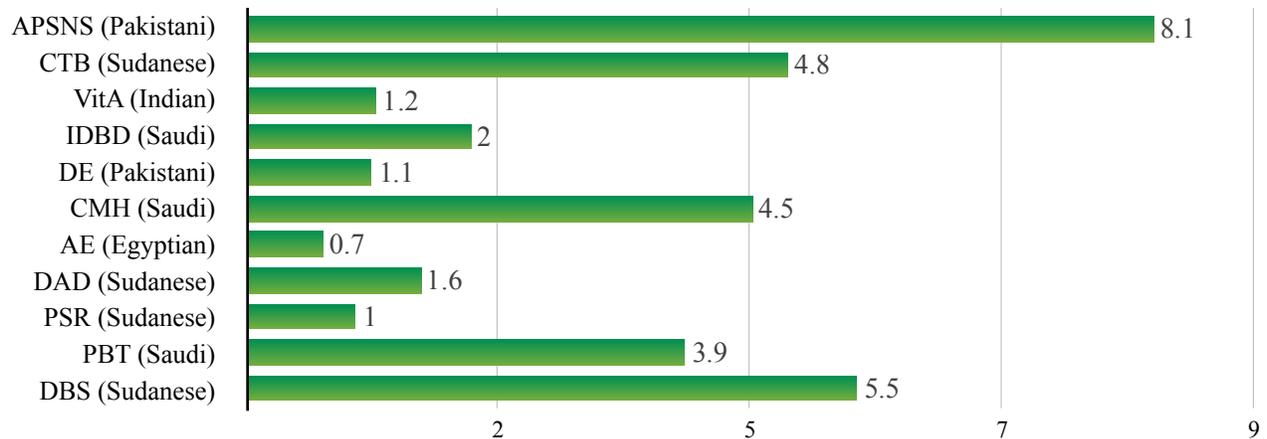
As stated previously, Topic Initiators function under the category of Structural DMs in that they play a role in organizing the discourse of the lecture. As Fraser (1990) states, DMs in and of themselves do not create meaning but rather assist in orienting the listeners and guide them in clarifying the speaker's intention (ibid: 390). In this case, the function of the DMs '*so*', '*now*' and '*okay*' are necessary guides signaling the speaker's communicative intent to mark a new or different (sub)topic within the overall lecture discourse. These DMs are, in respect to this research, always utterance-initial, meaning they are positioned at the start of an utterance and signal that new or previously discussed information will follow.

Like most of the DMs, utterance-initial DMs can also play a secondary role as attention-getter, which, by their nature, are meant to capture the listener's attention and clearly indicate that the preceding information has ended and (sub)topic is about to begin.

5.4.1.1 Topic Initiator '*so*'

As a Topic Initiator, the DM '*so*' functions to introduce a new sequence of action (Sacks et al., 1974; Schegloff, 2006), which is specifically to gain attention and mark the start of a new (sub)topic. The DM connects the immediate discourse to the overall main topic of the lecture. As illustrated in Figure 5.10, the DM '*so*' is used by all of the lecturers in their academic discourse.

Figure 5.10. SASME Frequency % of Topic Initiator ‘so’



In the first Excerpt (1), the Pakistani lecturer of the *Anatomy and Physiology of the Sympathetic Nervous System (APSNS)* lecture is speaking, within Excerpt 1’s context, about the division of the brain - forebrain, midbrain and hindbrain. In (a) the lecturer introduces the collective parts of the brainstem and clearly states that this is called the brainstem. In his next utterance (b), he clarifies this is just one part or region and uses the DM ‘so’ to shift back to the preceding topic prior to (a). In (b) the lecturer’s function of ‘so’ is, as Schiffirin (1987) describes it, a marker of main idea units, which is used to return to the main topic, in this instance, the nervous system.

Excerpt 1

[APSNS] (a) **So**, midbrain, bones and the medulla oblongata, they collectively the one structure, we are calling brainstem. (b) **So**, this is just the region, central nervous system, peripheral nervous system. Central nervous system, brain and spinal cord.

The APSNS lecture contained the highest frequency of the Topic Initiator ‘so’ with 8.1% occurrences. This lecture, in particular, caused concern as the results were out of proportion to the averages of each of the other SASME lecturers. As a result, a second uncoded transcript was analyzed, meaning that each instance of the DM ‘so’ in the transcription, was evaluated based on its surrounding context and labeled according to its function. The original and new version were compared and found to be almost identical.

Based on the analyses and results, the assumption drawn is that the APSNS lecturer applies the DM ‘*so*’ as both an attention-getter for the listeners at the start of a new utterance and to signal the initiation of a new or different (sub)topic. The next Excerpt (2) from the APSNS lecture, illustrates that the Pakistani lecturer’s verbal style of imparting information is relatively short and uncomplicated, by which I mean, considering the nature of the lectures, his utterances are restricted to non-lengthy and in effect, less dense propositions. In lines 1-3, observe that the lecturer uses the DM ‘*or*’, which is not within the scope of this research, to organize and distinctly separate the examples he provides as representative of body functions not in the conscious control of the subject. Line 3 (a) then illustrates the function of the DM ‘*so*’ to initiate a change in (sub)topic from anatomic “increased heart rate” to autonomic with “involvement of skeletal muscles”. The second ‘*so*’ is also labeled as a Topic Initiator based on context as it is still initiating the new (sub)topic; this is in spite of the fact that the utterance seems to be initially repetitive of the first occurrence.

Excerpt 2

[APSNS] 1 *That increased heart rate, or increased respiratory rate, or there is*
 2 *secretions of certain- means secretions from the different brains that is*
 3 *not under control of the subject. (a) **So**, this system, somatic involvement*
 4 *of skeletal muscles. Autonomic, there is enlargement of smooth muscles,*
 5 *cardiac muscle, or glands, or any secretions. (b) **So**, this system, somatic*
 6 *under the control of the subject, anatomic not under the control of the*
 7 *subject. This anatomic nervous system is divided into the two main*
 8 *systems. What is it? Sympathetic nervous system, second is para-*
 9 *sympathetic nervous system.*

Similar to the APSNS lecturer, SASME’s other two non-Arab lecturers, are comparable in their proposition structuring throughout their lectures in straightforward and manageable chunks of information. Both the Pakistani lecturer of the *Drugs used in Epilepsy (DE)* lecture and the Indian lecturer of the *Vitamin A (VitA)* lecture used the DM ‘*so*’ as a Topic Initiator with 1.1‰ and 1.2‰ respectively. Excerpt 3 exemplifies the straightforward structure of the discourse in *DE* and *VitA*. In Excerpt 3(a), the *DE* lecturer initiates a change in (sub)topic with the DM ‘*so*’ but cuts himself off (the *hyphen* - indicates an abrupt break in discourse) to question the listeners about their background knowledge of epilepsy.

Excerpt 3

[DE] ...antidepressant NNT and mood stabilisers. (a) **So**, the disorder of epilepsy to be- I don't know if you have gone through the concept of epilepsy as a disease. Have you gone through this concept? Nobody taught you what epilepsy is?

[VitA] When you can't see, when you sleep, you close your eyes, that cycle is kind of regenerating, okay? (b) **So**, we will talk about how this cycle goes on in our eye. So, I think you have gone through the anatomy of eye, did you?

In Excerpt 3, *VitA*, the lecturer phrases the first utterance as if it was a question, asking what happens when they close their eyes and then provides the answer. In (b), the DM 'so' acts as an explicit marker signaling a new (sub)topic will now begin on the cycle of the eye.

Excerpt 4

[CTB] (a) **So**, most of the cases which are diagnosed as if they are TB meningitis are actually *Brucella meningitis*. (b) I remember when I gave this lecture in the early '80s, one of the students came out and told me that...

In Excerpt 4, the Sudanese lecturer teaching the *Cerebral Tuberculosis and other Chronic Infections (CTB)* course, uses 'so' as a signal to introduce a subtopic (a), which focuses on *Brucella meningitis*, as related to the overall main topic of tuberculosis (hereby referred to as TB). The lecturer applies the DM in order to capture the attention of the listeners and indicate that a shift in the discourse topic is about to occur. It also functions as a cue of new background knowledge before shifting to an anecdote. The anecdote (b), initiated with *I remember*, an Interactional DM, not covered in this research, provides the listeners with anecdotal information about the previous utterance (a). The anecdote, though it begins in English, soon code-switches to Arabic as often occurs with this lecturer; the switch perhaps indicating a building of rapport with listeners by using the shared local language (Ferguson, 2003).

In fact, the Topic Initiator 'so' is one of the few DMs that the *CTB* lecturer uses in English with 4.8%. Overall, the Sudanese lecturer has 11.8% occurrences of DMs in both categories. A comparison of the English and Arabic DMs used within the *CTB* lecture will be discussed in greater detail in the next chapter (see 6.6).

Similarly, the Saudi lecturer of the *Congenital Malformations and Hydrocephalus (CMH)*, a native Arabic speaker, incorporates the L1 into his discourse, though not to the extent of his *CTB* peer. While the frequency of ‘*so*’ in the *CTB* lecture (4.8‰) exceeds those in the *CMH* lecture at 4.5‰; overall, the Saudi lecturer has 22.6‰ to the Sudanese lecturer’s 11.8‰ total occurrences. In Excerpt 3 (a), a topic *forebrain anomalies* is introduced with ‘*so*’ and from the subsequent utterances we surmise that this is important information for the course as a whole and is “seriously a huge topic”.

Excerpt 5

[*CMH*] (a) ***So***, we talk forebrain anomalies, we talk neural tube defects and we have a family called posterior fossa anomalies. I know some of you may have many questions and what is that, what is that. It's really a huge topic that I'm giving you in 25 minutes. It's seriously a huge topic.

The Topic Initiator ‘*so*’ in Excerpt 5 (a) may seem like a subtle means to capture the listeners’ attention and signal the start of a topic within the scope of the overall lecture, however, its effectiveness is illustrated with a larger chunk of transcribed discourse in Excerpt 6. Through a series of steps, the Saudi *CMH* lecturer verbally guides the listeners through his discourse (see Excerpt 6).

Excerpt 6

[*CMH*]

- a. There is a syndrome that can have a huge effect on the brain coming out from the
- b. back, but it could be small. ***So***, this is a general term. [inaudible] of the brain,
- c. wherever it is, it's the brain. Usually it would be the occipital or the posterior fossa
- d. but it's a general term. You can see it in different diseases in - that is what is in
- e. [inaudible]. Done.
- f. Finally, posterior fossa anomalies. ***So***, we talk forebrain anomalies, we talk neural
- g. tube defects and we have [inaudible] called posterior fossa anomalies. I know some
- h. of you may have many questions and what is that, what is that. It's really a huge topic
- i. that I'm giving you in 25 minutes. It's seriously a huge topic. ***So***, what I'm - just trying
- j. to introduce you to how things are divided, that's all. Posterior fossa anomalies of
- k. course is common malformations and usually associated with abnormalities around
- l. the cerebral line. There's the famous Arnold-Chiari malformations...

In lines *a* - *d*, he describes the effects of the abnormal development of a bone in the ear and in line 5 explicitly states that this topic is “done”. In line *f*, the *CMH* lecturer indicates that the next topic is the final topic of the lecture with “finally” and within the next utterance signals that another (sub)topic has begun with the Topic Initiator ‘*so*’. In lines *g* - *i*, the lecturer veers slightly off the topic by engaging with the listeners and demonstrating that he realizes that this is new information - “I know some of you may have many questions...”. He simultaneously emphasizes the importance and depth of the new information by repeating that it is “a huge topic” that he will cover “in 25 minutes”. Subsequently, a second Topic Initiator ‘*so*’ in line *i* functions as a marker to return the listeners’ attention and focus back - prior to the slight tangent off-topic and to initiate the discussion of the new topic “Posterior fossa anomalies”.

The DM ‘*so*’ is the most frequently used single-unit DM throughout SASME with 9.3% overall. It is distributed relatively similarly across the lecture discourse by its functions as Topic Initiator (3.2%), Summarizer (3.1%) and Topic Developer (3.0%).

5.4.1.2 Topic Initiator ‘*now*’

The function of the DM ‘*now*’ has been described to have numerous functions, in addition to its temporal meaning of ‘at the present moment’. Aijmer (2002) argues that on a textual level ‘*now*’ serves as boundary marking functions to shift topic, to frame discourse units, to denote steps in an argument or draw attention to specific units of discourse. Similar to Aijmer’s (2002) findings, the DM ‘*now*’ is analyzed and as an initiating marker, its function is identified as Topic Initiator. As such, the DM ‘*now*’ not only indicates the start or shift in (sub)topic, but also functions as an attention-getter for listeners. The first Excerpt (7), illustrates ‘*now*’ functioning as initial topic markers and drawing the listeners’ attention from the prior discourse to the present (sub)topic as in (a), which initiates discourse of how to classify the sympathetic and parasympathetic nervous system. In (b) the lecturer shifts topic from areas of the brain being affected by TB to the cerebella by using the DM ‘*now*’ as an attention-getter and as a marked topic shift to a question - the answer to which is the new topic.

Excerpt 7

[APSNS] *(a) Now, if we see the sympathetic and the parasympathetic nervous systems I classify first brain then central nervous system, then peripheral nervous system, then somatic and the autonomic nervous system.*

[CTB] *You can get a [inaudible] lesion, depending on what area is being affected. (b) Now, what area of the brain for cerebella is going to ...*

In general, the DM ‘now’ is straightforward in its function as Topic Initiator in the corpus as illustrated in Excerpt 8 which marks the shift to a new (sub)topic as in (a) to vestibular pathway and in (b) the result of neurotransmitters binding with receptors. In (c) the DM functions similar to Excerpt 7(b) as an attention-getter and as a signal of a shift in topic from the development of CNS (central nervous system) to a pseudo elicitation to identify the differences between an abnormality and a mutation in the CNS.

Excerpt 8

[DBS] *So, this is the physiology of balance how is balance maintained. (a) Now these [slide] are the neurotransmitters involved in the vestibular pathway they include, main neurotransmitters, glutamate, acetylcholine, glycine and GABA glutamate is maintained ...*

[DAPD] *(b) Now, when both the GABA and the benzodiazepines bind to the receptors there is greater influx of chloride ions. There is greater influx of chloride ion. What does chloride do?*

[CMH] *... mutations that affect the normal development of the normal CNS and that there are examples which are known. (c) Now, what is the abnormality in the CNS and what is the mutation.*

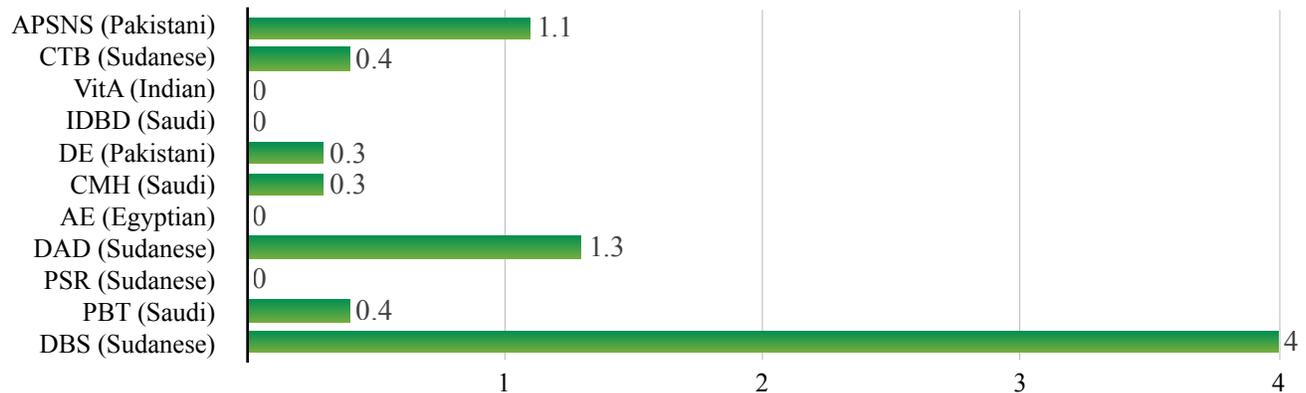
Not all occurrences of the token ‘now’ in the corpus are coded as DMs, for example, the idiom *now and then* in Excerpt 9 (a) of which there is a single occurrence in the SASME data.

Excerpt 9

[DE] *Now, you have to see what drugs are effective in epilepsy. (a) **Now and then** you will see a patient and maybe women, maybe pregnant, or old age or maybe a small child, so, on that basis you have...*

As a Topic Initiator ‘*now*’ is comparably not as frequent as ‘*so*’ (3.2‰), but overall shared a similar number of occurrences to ‘*okay*’ with 0.7‰. Distribution across SASME ranged from 0‰ in four of the eleven lectures to 4‰ in the DBS (see Figure 5.11).

Figure 5.11 SASME Frequency ‰ of Topic Initiator ‘*now*’



From Figure 5.11, we observe that the DM ‘*now*’ is favored by the Sudanese lecturers, all of whom have occurrences of the DM in their lecture discourse. The DBS lecturer’s high frequency of ‘*now*’ is also seen to correlate with his use of the Topic Initiator ‘*so*’.

5.4.1.3 Topic Initiator ‘*okay*’

The third DM that functions as a Topic Initiator is ‘*okay*’, which has 0.7‰ occurrences overall in seven of the eleven lectures in SASME. Comparatively, the lecturers who utilize Topic Initiators in their discourse, seem to prefer the DM ‘*so*’ about four times as much as ‘*okay*’ and ‘*now*’. The DM ‘*okay*’ is defined as a Topic Initiator based on the context and its falling intonation, as indicated by the comma following the DM. As illustrated in Excerpt 10, ‘*okay*’ functions as a signal of a topic shift, as in the first example (b), in which the lecturer diverges from the topic of epileptic patients to inform students they may “break there for this week”. With the DM ‘*okay*’ the speaker draws the listeners back to the main topic previously uttered in (a) with example symptoms an epileptic patient might have.

Excerpt 10

- [DE] Lecturer: (a) You would be able to recognize epileptic patients?
 Student: Yes.
 Lecturer: Easily? Very good. But then you can also know the drugs. I'll just take one minute. We might break there for this week. [inaudible]
 (b) **okay**, like shaking or jerking. As you can see here, he is holding the ball, playing normally.
- [CMH] (c)...Usually you have one, you have two, three or group of different defects, leading to the formation of a syndrome. (d) **Okay**, usually it's not a single disease. As you know, the brain develops in stages. Since formation of the zygote until you have baby, there are different stages, biologic.

In the second example of Excerpt 10, the *CMH* lecturer uses the DM 'okay' to shift the discourse back to the preceding main point prior to (c), which is that congenital malformations are not usually the result of a single disease and continues the shift by describing the development of malformations from the formation of the zygote.

Functionally, the DM 'okay' signals the intended action of the speaker to focus the attention of the listeners (see Excerpt 11) from (a) where the speaker originally assumes the listeners understand the term *cognition* to Excerpt 11(b) where 'okay' draws the listeners' focus to the word *cognition*, which is subsequently defined. Based on context, the DM functioned as a Topic Initiator as it has been defined in the classification. Similarly, in the second example, (*CMH*), the DM 'okay' in Excerpt 11(d) functions as a marker indicating that the introduction of a new (sub)topic, not connected to the previous information (c) on *myelomeningocele*. That is, the context leads us to understand that utterances (d) and (e) indicate that while the previous information is about *myelomeningocele*, the subsequent information is not, as it is now being introduced.

Excerpt 11

- [PBT] Location, location, location. The location. Location for brain [Arabic].
 (a) Like in general cognition. Cognition, why the cognition? (b) **Okay**,
 Cognition. Cognition is the whole...

Excerpt 11 (continued)

[CMH] (c) Also bowel and bladder control abnormalities. (d) **Okay**, so that was an example on neural tube defect which was myelomeningocele. (e) This is [refers to powerpoint slide] of course - it's not myelomeningocele.

The following examples (see Excerpt 12) are taken from the introductory sections of the lectures, as the speaker begins to transition from the background and introductory information to the main topic of the lecture. The Excerpts (a) and (b) illustrate that the DM ‘*okay*’, like the DM ‘*now*’, functions in two ways in the corpus: as a Topic Initiator and as an attention-getter. The second function of ‘*okay*’ respectively has been described as “open[ing]... a new round of talks” within discourse and a means to (re)gain control of the listeners (Othman, 2010: 675). This may be especially on target with the non-Arab lecturer (see Excerpt 12) who may have begun to lose the attention of the listeners in (a) and (b) and implements it to regain attention and preface a new topic of discourse.

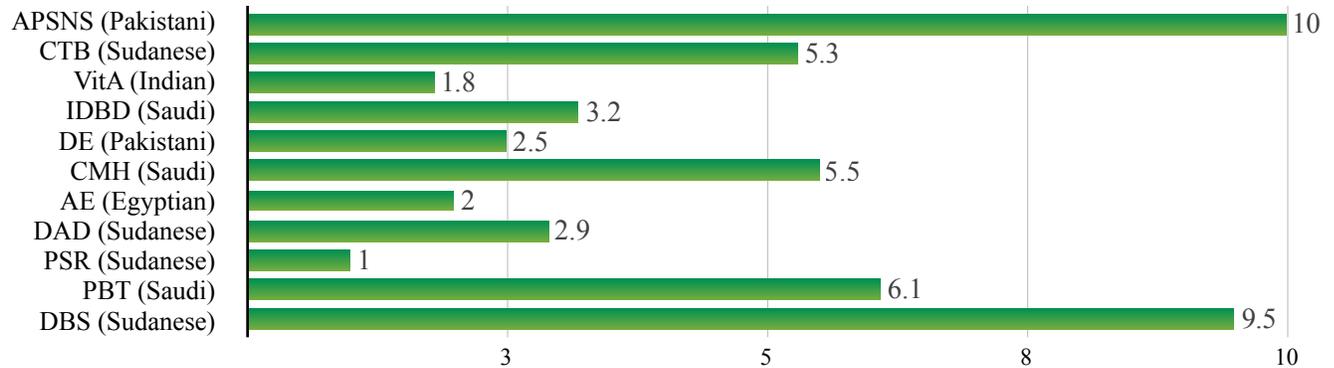
Excerpt 12

[VitA] (a) **Okay**, we'll take it step by step, okay? Alright. So what happens here - so, this is our retina and it contains rod and cone cells...

[DE] Yes the same [inaudible], where it's severe tonic and clonic seizures. Yes, true. (b) **Okay**, we are coming to generalization, when both the spheres of the brain are involved [inaudible]. This is the most common epilepsy in the world.

Overall, the Topic Initiator function is the second highest in frequency at 4.7‰ occurrences following Topic Development at 7.5‰ and Summarizer and Closer at 4.2‰ and 1.6‰ respectively. Figure 5.12 below, illustrates the distribution of Topic Initiators across SASME and comparably, we observe that no patterns or similarities can be drawn by ethnicity (e.g., comparing Saudi or Sudanese lecturers to each other) or by L1 (e.g., native Arabic speakers).

Figure 5.12 Distribution of Topic Initiator DMs ‰ across SASME



5.4.2 Topic Developer (*so, because, 'cause, also*)

The second function of Structural DMs is Topic Developers (*so, because, 'cause, also*), which commonly occur in utterance-medial (i.e. in the middle of an utterance) and function as developers of propositional ideas. That is, the DMs signal that a preceding and established proposition (i.e. the topic has been introduced and discussed in prior proximate discourse) will be elaborated upon in the next utterance through examples, anecdotes, details or definitions.

5.4.2.1 Topic Developer 'so'

In the first example, the Sudanese lecturer of CTB speaks about a PCR test (used to rule out infections related to TB) in Excerpt 13(a) and in (b) the lecturer provides additional information about the PCR test. In the analysis stages, the framework establishes a distinction between the Topic Developer function and that of Summarizer. As you will see in this section and in the subsequent section (see 5.4.3), the functions are differentiated by context and the nature of the information that follows the DMs. Summarizers, as will be discussed, function as signals of recapitulations of preceding information. Topic Developers, in contrast, as illustrated in Excerpt 13, provide additional information about the preceding proposition. We observe in Excerpt 13 (b) as stated, supplementary information is given about the PCR test discussed in (a). Similarly, utterance (d) supplies the listeners with more details about the time frame and speed that the

medication, “Betahistine”, (c) will be “metabolized” and in (e) further details the “metabola” stated in utterance (d).

Excerpt 13

[CTB] (a) PCR is a test that's specific, like [inaudible] is specific. You can look for minor things. (b) **So** a PCR is very helpful, as well, for diagnosis of TB. With Indian ink for *Cryptococcus*, it will indicate a genus of *Cryptococcus neoformans*.

[DBS] (c) Betahistine is taken orally in the form of tablets. It has got the short plasma half-life, (d) **so** this will be taken frequently. It is metabolized partially into active metabola (e) **so** some of the metabolites are pharmacologically active.

Topic Developer ‘so’ provides the listener with verbal cues to observe and note that the information only provides supplementary details related to a main topic and not a topic in and of itself. Previous research (see 2.5.2) suggests that the recognition of these functions is generally believed to be beneficial for retention and recall (Chaudron and Richards, 1986; Jung, 2003) and may thus help listeners to successfully follow lectures (Khuwaileh, 1999). Excerpt 14 illustrates the links in proposition (a) and (b), in that, the first utterance describes the side-effects of taking *phenytoin* which are causing drowsiness and difficulty concentrating; the subsequent utterance (b) relates to the preceding proposition in respect to the side-effects, which the lecturer uses as a springboard to advise the students to stay awake because the information is “very important” for their assessments.

Excerpt 14

[DE] (a)...if a patient is taking one dose of phenytoin, now he is awake, but he will have the same difficulty in concentrating. (b) **So** if you are awake - this is very important for you because some students are frequently asking if the things that you will be discussing in this lecture are included in assessment.

[CMH] (a) There's no absorption of the CSF. That's leading to accumulation of the- communicate. (b) **So**, the commonest is non communicating. Not communicating is the commonest compared to the communicating. So what can cause hydrocephalus, I've said that already, so-

Student: Obstructions, tumors.

In Excerpt 14, the proposition (a) describes the lack of CSF absorption, which leads to over-accumulation of the communicate. Subsequently, the lecturer describes the effects of an obstructed hydrocephalus in (b) and pseudo-elicits the cause from the listeners. Excerpt 13 exemplifies the function of ‘*so*’ in (b) to give new information about the propositional topic and not to outline previously stated information, as it would in a DM ‘*so*’ functioning as a Summarizer.

Excerpt 15 illustrates the Topic Developer ‘*so*’ used slightly more frequently in a chunk of discourse by the Indian lecturer of the *VitA* course. The first DM ‘*so*’ (a) functions as a summary in essence and all immediately previous utterances are summed-up with *that’s the film*. However, the subsequent DM ‘*so*’ varies in function to Topic Developer as it can be observed to provide new information - a description of light creating an image on film. This continues to develop further in (c), gradually helping the listeners understand the similarity of light on the CCD and how information is transmitted in CCD (c) and in the eye (d).

We can observe in Excerpt 15 the extent to which the *VitA* lecturer uses Confirmation Checks; this chunk of discourse contains five Confirmation Checks. The lecturer has more occurrences than any other lecturer in either the NS or NNS corpus. In fact, he has a higher frequency than each corpus combined. This will be discussed under the next category of Interactional DMs (see 6.3.1).

Excerpt 15

[*VitA*] *It used to be films, right, before. No more films. (a) So that's the film. (b) So when the light strikes on the CCD this CCD can merge the light into an image, okay? (c) So similarly this light has information, right? The light [inaudible] has information, okay? I can see you because the light around you is reflecting on my eyes, right? so that I can see your face. (d) So this light has information, okay?*

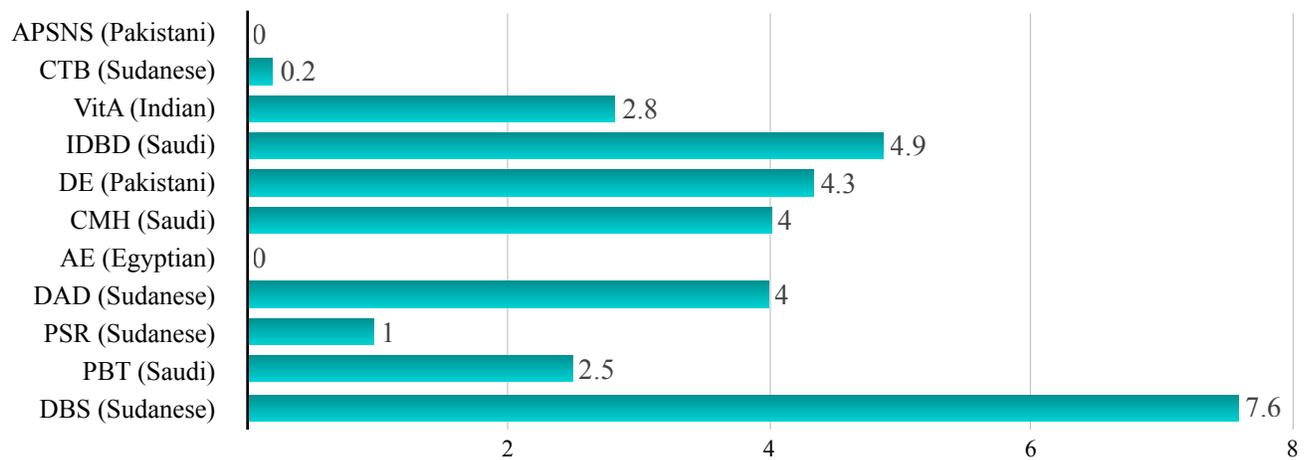
In our last Excerpt (16), an illustrative comparison is given between the function of Topic Developer ‘*so*’ in (b) and (c), which provides examples of anti-depressant drugs; however, utterance (c) sums up the immediately preceding information by stating their categorization within the pharmaceutical family of anti-anxiety drugs.

Excerpt 16

[DAPD] (a) The rest of the drugs are anti-depressant drugs, (b) **so** MAO inhibitors, drugs anti-depressant, 5-hydroxytryptamine reuptake inhibitors, these are all anti-depressant drugs, they are all anti-depressant drugs. (c) **So** we have got benzodiazepines, we have got 5-hydroxytryptamine 1 a receptor, (d) **so**, these are the main anxiolytic or anti-anxiety drugs.

In summary, as a Topic Developer ‘*so*’ had an overall occurrence of 3‰ and is found in all but two of the lectures, one a native Arabic speaker and the other not. In general, the overall distribution of ‘*so*’ is not as skewed as previously seen with the Topic Initiator ‘*now*’, with the exception of the *DBS* Sudanese lecturer who had 7.6‰, followed in frequency by the *IDBD* Saudi lecturer with 4.9‰ (see Figure 5.13).

Figure 5.13 SASME Frequency ‰ of Topic Developer ‘*so*’



5.4.2.2 Topic Developer ‘*because* / ‘*cause*’

As a Topic Developer, the DM ‘*because*’ or its informal token ‘*cause*’, function in the same manner as ‘*so*’ and, more often than not, are utterance-initial in the SASME data. As a DM, ‘*because*’ has often been viewed as functioning solely as a conjunction, a causal marker, linking one utterance to another and its role as a DM has been disputed in that it is not seen to contain a meaning beyond its function as a ‘causal’ conjunction (Chaudron and Richards, 1986). Researchers who do cautiously label ‘*because*’ as a DM, may then categorize them as micro-markers, lower-order markers (Chaudron and Richards, 1986), fundamentally used to link internal

or ideational relations within propositions, clauses or ‘sentences’. Fraser (1999: 939) describes DMs as tokens which “introduce a separate message with its propositional content” and considers marked expressions like ‘*because*’ as DMs since it relates two separate messages (ibid: 940). Within SASME and subsequently, in my analyses of the BASME results, ‘*because*’ is regarded as a DM which signals transitions and emphasis in a lecture and can appear on three planes of discourse (see Table 2.4).

The first example (see Excerpt 17), illustrates the function of the Topic Developer to provide examples, as in (b), where the lecturer is not using ‘*because*’ in its Ideational Structure of cause (Schiffrin, 1987), rather, it functions on the Action Plane, with a meaning of *motive* (of speech act) (see 2.4.2). The DM’s propositional meaning in (b) is that of an ‘inflammation, tumor or other things’ as reasons why “I” have headaches and pain (a). Functionally, the DM ‘*because*’ provides background information - details, examples, elaborations on possible causes of the symptoms, which are Topic Developers.

Excerpt 17

[CTB] (a) *It takes a long time. So, you will have, I have got these headaches for some time, going on for about two or three weeks. I have got neck pain, and I have got change in personality. Why? (b) **Because** there is inflammation, there is tumor, there is a tuberculoma due to TB or other things affecting a brain...*

Similarly, the Sudanese lecturer of the *Drugs used in Anxiety Disorders (DAD)* lecture uses ‘*because*’ to develop the topic of why patients may suffer from anxiety and lists possible causes in Excerpt 18. He uses the DM to elaborate on the causes but then differentiates between normal and pathological fear with another DM ‘*but*’ (not within scope of research), which then shifts focus onto a subtopic.

Excerpt 18

[DAD] ***Because** there is a cause, the exam is a cause, fear from dentist is a cause, but sometimes some people have anxiety without obvious cause, so this is called pathological anxiety. Fear normal anxiety, fear pathological anxiety. So, what are the symptoms of anxiety?*

In the *Drugs for Epilepsy (DE)* lecture (see Excerpt 19), the lecturer digresses slightly to use an opportunity to advise the listeners that exercise is essential for health and not only to reduce the number of epileptic seizures a patient may suffer or have. The Topic Developer ‘*because*’ connects the proposition of lifelong time commitment to the main topic of exercising for health and develops the idea of *time* by giving examples of helpful sports or habits they can take up.

Excerpt 19

[DE] *Try to devote some time for exercise, like doing sports, maybe football or just walking or jogging or gym [inaudible] that much, I mean have that much time for gyms. **Because** you have to commit for lifelong, but if you are in a habit of walking, like running for, 30 minutes or 20 minutes, or you do some like push-ups or things like that, then it would be very helpful for you to be physically active and muscularly strong.*

In virtually all instances of utterance-initial occurrences of ‘*because*’ the DM is preceded by a question as illustrated in the three examples in Excerpt 20. The DM ‘*because*’ establishes the reason why embryology is important in (a) with a personal assurance by the CMH lecturer. In *VitA*, the DM elaborates on the differences between rod and cone cells in (b) and in (c) the lecturer describes the purpose in discussing the genetic makeup of Parkinson’s disease. Each of the Topic Developers immediately follows a question, though in (a) we observe that the lecturer had originally begun his explanation as *that’s why-* but abruptly cuts himself off and instead uses ‘*because*’.

Excerpt 20

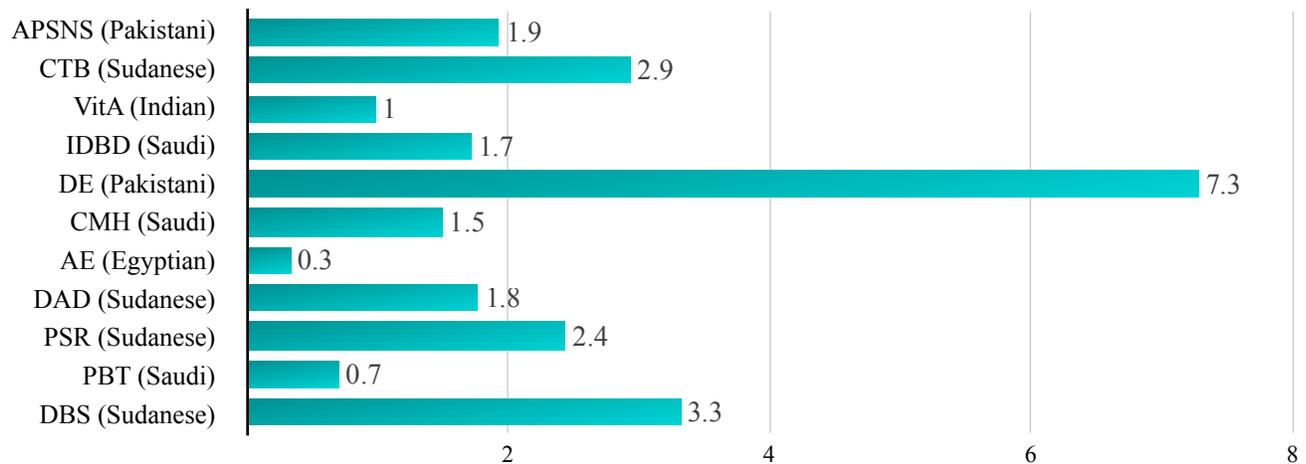
[CMH] *This is the importance of knowing the embryology. That's why- (a) **because** I know the embryology and that is the focus of this in [inaudible]. However, I just want you to understand there is the problems in the development of the neurones...*

[VitA] *...okay? Both have the same function, but since rod and cone cells they are different. Why? (b) **Because** one of them is involved in black and white image and the other is involved in the color image. That's why their proteins are a bit different but they are type of rhodopsins, okay?*

[PDBD] *Why do we talk about this? (c) **Because** it's like two genetic [inaudible] for Parkinson's disease. Two genetic [inaudible]. Genes involved in parkin and this [pause] on the slides, right? In parkin- is one of the proteins which are mutated [inaudible] in Parkinson's disease.*

The Frequency of the Topic Developer ‘*because*’ (‘*cause*’) is relatively comparable in distribution across SASME, though one lecturer (*DE*) does skew the data of this particular DM. The lecturer’s frequency of use is nearly triple that of any other lecturer with 7.3%.

Figure 5.14 SASME Frequency % of Topic Developer ‘*because*’



Analysis of the data suggests that the non-Arab lecturer used the DM ‘*because*’ throughout his discourse to distinctly mark reasons, examples and elaborations for the benefit of the listeners’ comprehension.

5.4.2.3 Topic Developer ‘*also*’

The last DM under Topic Development is ‘*also*’, which functions in a similar manner to ‘*so*’ and ‘*because*’ (or ‘*cause*’). Blakemore (1992: 139) asserts that DMs should be analyzed within their specific context linguistically, which Fraser (1999: 944-945) expands upon to suggest that based on their sentential structure (i.e. S1 and S2) the DMs provide a procedural meaning that specifies how each segment (i.e. utterances), relate to one another in respect to context. He further categorizes DMs like ‘*also*’ as an elaborative marker which “indicate[s] a relationship in which the message of S2 parallels and possibly augments or refines the message of S1” (ibid: 948).

In the SASME data, the DM ‘*also*’, as defined under this function, is found to be mainly utterance-initial and, as described by Fraser (1999), the DM parallels the message of the previous utterance. As in Excerpt 21, in which the *DBS* lecturer lists the characteristics of a “balance

disorder” and utterance (b) expands upon his description of “dizziness” to include “vomiting and nausea”. In contrast, the DM ‘also’ in Excerpt 21(a) is not characterized as a Topic Developer, as in (b), but rather, is described as functioning similar to ‘and’, an additional marker not within the scope of the research.

Excerpt 21

[DBS] *As you will see here we may use histamine agonists, we (a) **also** can use histamine antagonists. You will see that balance disorder is characterized by dizziness but some drugs which are used for treatment of dizziness can induce dizziness. (b) **Also** you will see that balance disorder may result in vomiting and nausea. And you will see...*

The second Excerpt (22) illustrates a further feature of the functions of ‘also’, which is to separate the message in shorter utterances as in *DBS* (a) and (b), in which the utterances could have been produced as a single statement. Instead, the lecturer formed two separate, yet parallel propositions of which a relationship is indicated through ‘also’. This is potentially to illustrate the importance of each detail to the listeners and to support a less content-dense explanation of *benzodiazepines*.

Excerpt 22

[DBS] *(a) They cause addiction therefore in many cases they are avoided but sometimes if you use low doses they may be useful. They may be useful. (b) **Also** benzodiazepines may affect walking, so the patient may fall down.*

[PDBD] *Alpha-synuclein. (a) The mutations for alpha-synuclein can lead to solid domain forms of parkinsonisms. (b) **Also** some gene duplication or triplications. So [inaudible] mutations of alpha-synuclein is known to be a cause for certain groups of Parkinson's disease.*

The second example of ‘also’, in Excerpt 22 (*PDBD*), illustrates the equal importance of the propositions (a) and (b). For the speaker, the function of ‘also’ indicates that “mutations for alpha-synuclein” and “some gene duplication or triplications” can both potentially lead to certain forms of Parkinson’s disease.

Similarly, in Excerpt 23, the transcript samples illustrate the function of the Topic Developer as in (a) and (b), which relate the utterance on the vision cycle (a) to deficiency diseases (b). The DM ‘*also*’ refines the propositional message of (a) to include another subtopic, namely, “diseases”, that will be discussed, though apparently with less focus, as clarified in (c). The second example (d) and (e) also relates the propositional messages in which (e) augments the message of (d).

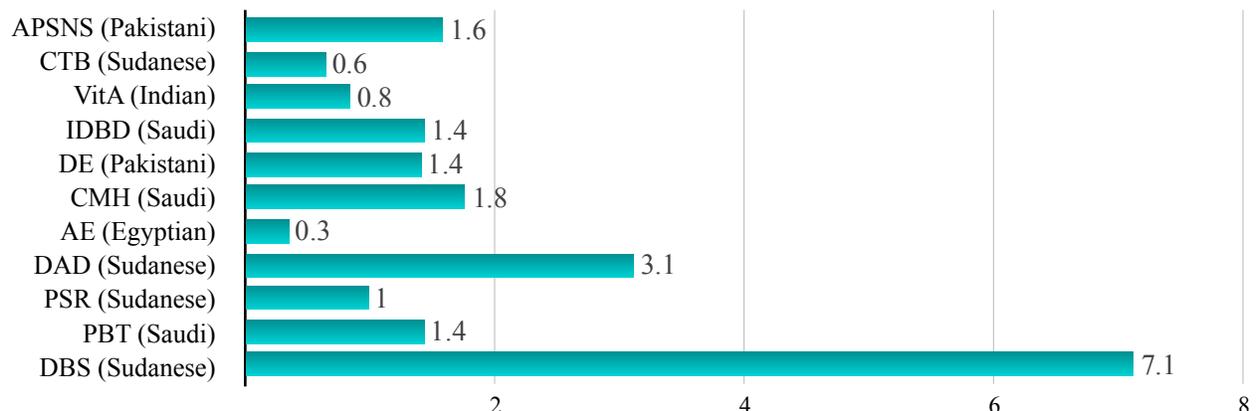
Excerpt 23

[VitA] (a) *Or the visual cycle, whatever you call it. We will go in details about this cycle. (b) **Also**, of course, the diseases, because of the deficiency. (c) Again, the diseases are many but then we will focus mainly on the vision, okay?*

[VitA] (d) *So, that's the normal thing, you can't see, okay? Because you have nothing to see. (e) **Also** when you sleep, when you close your eyes when you sleep, your rhodopsin is regenerated, you know, ample supply of it is regenerated, okay?*

Overall, the function of Topic Developer has the highest overall number of occurrences at 7.5%; however, within the function, the DM ‘*also*’ had the lowest frequency of occurrences at 1.9‰ in comparison to ‘*so*’ (3‰) and ‘*because/cause*’ (2.6‰). Figure 5.15 illustrates that the DM ‘*also*’ seems to be more prevalent in at least one lecturer’s discourse and partially skews the data. In this case, the *DBS* Sudanese lecturer has more than double the occurrences of ‘*also*’ (7.1‰) than the (Sudanese) *DAD* lecturer (3.1‰) and, on average, six times more than the remaining lecturers.

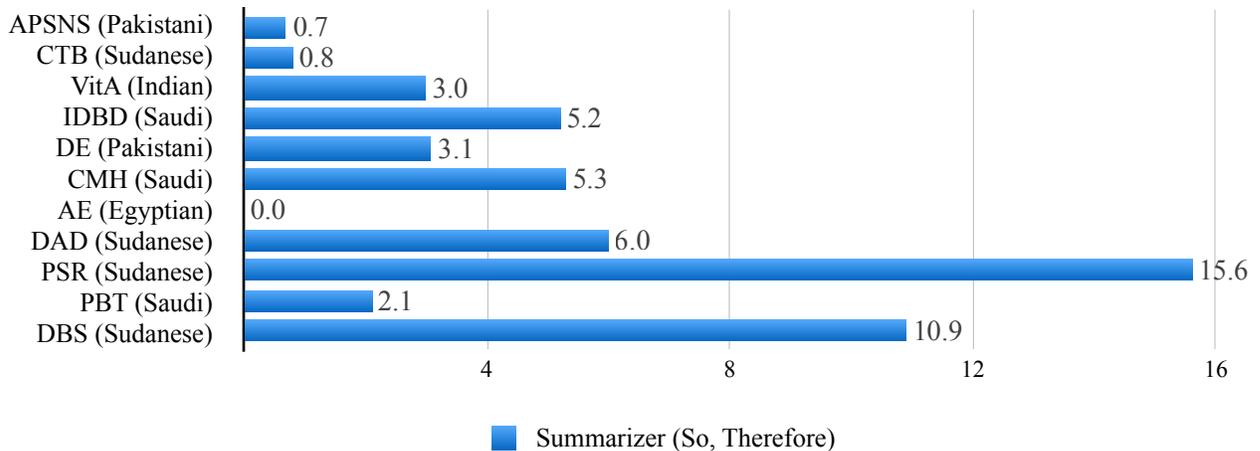
Figure 5.15 SASME Frequency % of Topic Developer ‘also’



5.4.3 Summarizer (*so, therefore*)

The third function, under the category of Structural Markers, is Summarizers, which overall are the third highest in occurrences (4.2‰) after Topic Initiator (4.7‰) and Topic Developers (7.5‰).

Figure 5.16 Summarizer Category Use % across SASME Lecturers'



The DMs '*so*' and '*therefore*' function as markers that signal a summary of prior information. Fraser (1999: 948) labels these DMs as “inferential markers” in which S2 should be interpreted as concluding or summing up the information stated in the S1 (or within the immediate context). Functionally, the DMs '*so*' and '*therefore*' in this group can be seen to express a meaning of *as a result* or *in conclusion* and expresses the same propositional idea as the previous utterance(s).

5.4.3.1 Summarizer '*so*'

These occurrences of '*so*' are identified as DMs that function as Summarizers based on the context in which they appear. In our first group, the DMs function as a Topic Initiator, introducing a proposition that had not been previously discussed in prior discourse. As a Topic Developer, the DM expanded upon a prior propositional idea and is followed by new information related to the previous utterance.

As a Summarizer, the DM '*so*' does not include new information and context illustrates its function is to mark the result or sum up the propositional ideas uttered in the preceding

utterance(s). For example, in Excerpt 24, the first extract shows that the utterance following ‘so’ in (b) expresses the same propositional idea as the previous utterance (a), that is, in spite of the fact that a minor segment of the utterance is inaudible, we are able to clearly assume from the context of (b) that “this” is a result of Meningitis factors presented in (a). In the second example, (see Excerpt 24), the *CMH* lecturer is summing up the basic factors of “embryology” in (d) based on the three factors provided in (c).

Excerpt 24

[CTB] *In some cases, it presents as meningitis. (a) Meningitis was chronic [inaudible] (b) so it had to be this. These are excellent. You can see on the right-hand side here, there is some bacteria of tuberculosis is spread by...*

[CMH] *(c) This is the [inaudible] and you have this - formation of the neural tube and different parts of the CNS. (d) So this is basic embryology. The effects here you can have an open skin and spinal cord coming out...*

The next example (see Excerpt 25), illustrates the function of summarizing the previous proposition as in (b), which sums up the input received by the “vestibular nuclei”, which had been discussed in prior utterances and the input, which is described in (a). The following utterance (c) relates to what Fraser (1999) described as “inferential markers” in which listeners can infer that the step following *input* in (b) would then be *output* in (c).

Excerpt 25

[DBS] *(a) There is an input from the eye and proprioceptors in the neck, in the spine and so on. (b) So, when these inputs are received by the vestibular nuclei they are processed, (c) then an output will take place. To where?*

[DBS] *(a) Now next are the drugs that impair balance. (b) Now this can either impair vestibular function or cause destruction to safety instruction in the vestibular system so they can destroy the hair cells or they can destroy the nerve ending. (c) So, they can either impair the function or destruction.*

The second extract, *DBS*, from Excerpt 25, exemplifies the structural functions of the DMs in the lectures with (a), which establishes the propositional idea - drugs that impair balance and then describes the types of impairment or destruction that can occur as a result of certain drugs. The

last utterance in this example (c) uses the DM ‘*so*’ to mark a summary of the prior utterances and explicitly reiterate the important points.

Similarly, the reiterative nature of the propositional ideas in Excerpt 26 illustrates the key function of ‘*so*’, which is to focus the listeners’ attention on the key factors they need to understand. In (b) and (d) the utterances that follow ‘*so*’ summarize the prior propositional ideas of (a) and those immediately prior to (a). The utterance (c), discussing the link between fear and anxiety in the brain, is a digression within the summary to stress the link between the two manifestations.

Excerpt 26

[DAPD] *(a) Also hoarding ... collect things, collect things for think this is one type of obsessive compulsive disorder. (b) So, these are the types of anxiety [Arabic] fear, the panic disorder, fear and phobia. (c) This is because in the brain they are closely linked and anxiety with fear are closely linked in the brain. (d) So these are the types of, some types, not all types, some types of anxiety.*

Excerpt 27 also exemplifies the function of ‘*so*’ to summarize the prior propositions and in this case, a further purpose in stating the relationship between the disease - *Parkinson's* - and the syndrome in (a), is also to mark a distinction for the listeners as stated in (b).

Excerpt 27

[PDBD] *Excellent. (a) So, Parkinson's disease is an example of Parkinsonism. (b) But Parkinsonism is a clinical picture, it's not a disease... I am presenting a clinical picture.*

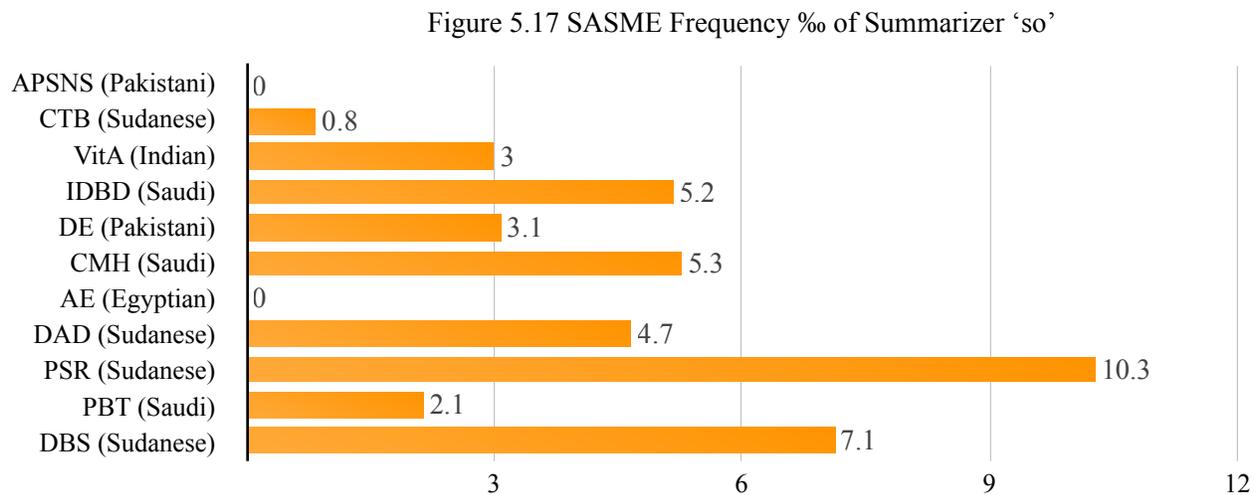
[CMH] *So, these are the complications of - major complications of hydrocephalus. If there is an obstruction in the CSF pathway flow you could call that non-communicating - non-communicating hydrocephalus.*

The lecturer in the *Vitamin A* lecture (*VitA*) (see Excerpt 28), not only uses the DM ‘*so*’ to summarize his previously stated propositions, but prior to (a), he also stresses that these points are summing up the “general points” by explicitly stating this in (b). He continues his summary in (c) again by reiterating the proposition with specific examples.

Excerpt 28

[VitA] (a) **So** these are just the general points. (b) Here I just summarise - you can see that it's the functions and the metabolism. (c) **So**, you can see that beta carotene can be converted to retinol and then retinol can be converted to retinal.

Figure 5.17 summarizes the frequency of occurrences across the SASME lecturers of whom two have zero occurrences of the DM 'so'. Its occurrence is more frequent across the NS Arabic lecturers, with the exception of the *AE* course and has an average of 0.31% in SASME.

**5.4.3.2 Summarizer 'therefore'**

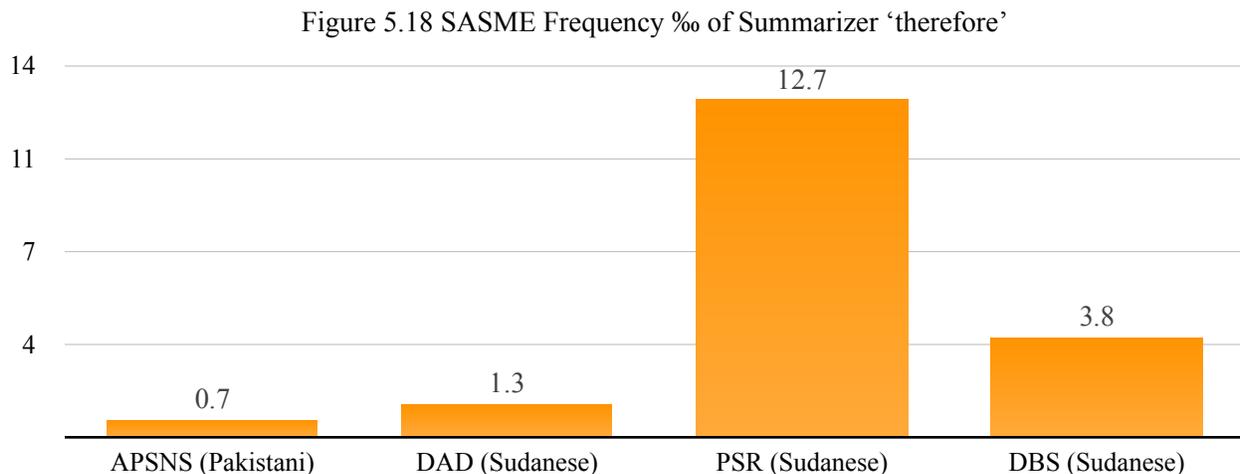
The DM 'therefore' under this function seems interchangeable with its counterpart 'so' within certain contexts in the category of Summarizer. This is exemplified in the transcript extract (see Excerpt 29) in which the DM 'therefore' in utterances (a), (b) and (c) could be replaced with 'so' and provides a similar meaning of *as a result*. The propositional ideas in (a-c) seem to be inferential and linked to background knowledge listeners should possess about the main topic of the lecture *Drugs used in Balance System*.

Excerpt 29

[DBS] *They reduce the blood flow and (a) **therefore** they bring about biochemical changes leading to impairment of electro mechanical transduction. (b) **Therefore** they can prevent the vestibular [inaudible].*

[DBS] *Increase in the motility of the gastro-intestinal tract can also improve nausea or prevent nausea. (c) **Therefore** dopamine antagonists can act centrally on the dopamine D2 receptors to suppress nausea and vomiting and they can also act briefly...*

Overall, the DM ‘*therefore*’ is underused in English in the SASME transcripts. The DM had zero occurrences in seven of the eleven lecturers in the SASME discourse. Of the four lecturers, who used the DM as a Summarizer, one in particular of the *PSR* lecture had a disproportionately high number of occurrences relative to the other lecturers with 12.7‰. This is more than three times higher than the next occurrence at 3.8‰ by the *DBS* lecturer (see Figure 5.18).



The *PSR* lecturer seems to be under the impression that an increased use of the DM ‘*therefore*’ in his lecture adds greater cohesion and structure for the listeners, as well as signaling the importance of the following utterance. Its use is semantically marked and could be viewed as awkward in place of an unmarked ‘*so*’, which could easily replace most of the occurrences of ‘*therefore*’, as in Excerpt 30(a).

Excerpt 30

[PSR] *If you don't, the nurses will laugh at you, the technicians will laugh at you. Ka'theer {a lot} they will laugh at you. (a) **Therefore** nah'na {we} we try to gain power. During- according to the University regulations the lecturer should go ten minutes before.... (b) **Therefore** every lecture the first part is very important.*

Excerpt 31 similarly illustrates the *as a result* meaning of the DM 'therefore' in (b) and (c). Utterance (b), in fact, provides more than just a *as a result* meaning following the DM 'therefore', it also sums up the details provided in utterance (a) and an immediately prior utterance to that, which discusses beta-blockers.

Excerpt 31

[DAD] *(a) Somatic symptoms are manifestations of hyperactivity of the sympathetic nervous system. (b) **Therefore** if we block this hyperactivity we can relieve the somatic symptoms such as the palpitation, the sweating, tremor, this can be relieved by the beta blockers.*

[APSNS] *They are myelinated. Here in the somatic they are myelinated, (c) **therefore** they conduct impulses.*

5.4.4 Topic Closer (okay)

The final function, under the category of Structural Marker, is Closer of which only one DM, namely, 'okay' was identified. It functions as an ending or closing of a (sub)topic and signals to the listener that a propositional idea, presumably having been completed, is now being drawn to a close. Its falling intonation, indicated by the full stop, illustrates that its function is not that of a progress or Confirmation Check, but rather a signal that an action, series of events or the final point in the discourse has drawn to a close.

Excerpt 32

[APSNS] *If simply activated sympathetic and cannot come back to the normal level then it is difficult. (a) Heart rate is 200, who will control it, for example, **okay**. (b) Respiratory rate is more than 18, then who will control it for example, **okay**.*

In Excerpt 32 above, the lecturer provides students with examples of the sympathetic nervous system and signals the closing of each example with the DM ‘okay’ as in (a) and (b). In the next example, Excerpt 33 (a) the discussion on dementia as related to “Huntington” in the first utterance, is closed in (b) and a new subtopic, HIV, as related to dementia, is introduced. The DM ‘okay’ functions as an explicit signal to listeners that the propositional idea being discussed has now been closed.

Excerpt 33

[IDBD] (a) *With Huntington, some patients can have dementia. (b) The dementia is not specific, **okay**. (c) Infections- HIV patients. HIV patients suffer HIV encephalitis.*

[CMH] (d)...*communicating, non-communicating hydrocephalus, general principles or the causes of hydrocephalus and that's it. I didn't give you much really, (e) **okay**. (f) Fine. We'll take 10 minutes break and then we'll talk about the pathology*

The use of ‘okay’ at the end of the utterance, as in Excerpt 33 (e), serves as a cue to signal the proposition is at its end. Similarly, in the next extract (see Excerpt 34) the *PBT* lecturer discusses malignant tumors in (a) and then proceeds to verify whether listeners have any other questions or concerns. This is followed by ‘okay’ (b), which signals the end of the utterance. In the second example, the *VitA* lecturer’s discourse digresses slightly to define the meaning of confirmation for the listeners before closing the digression with the DM ‘okay’ (c). The subsequent utterance begins a new topic about opsin and vitamin A in (d).

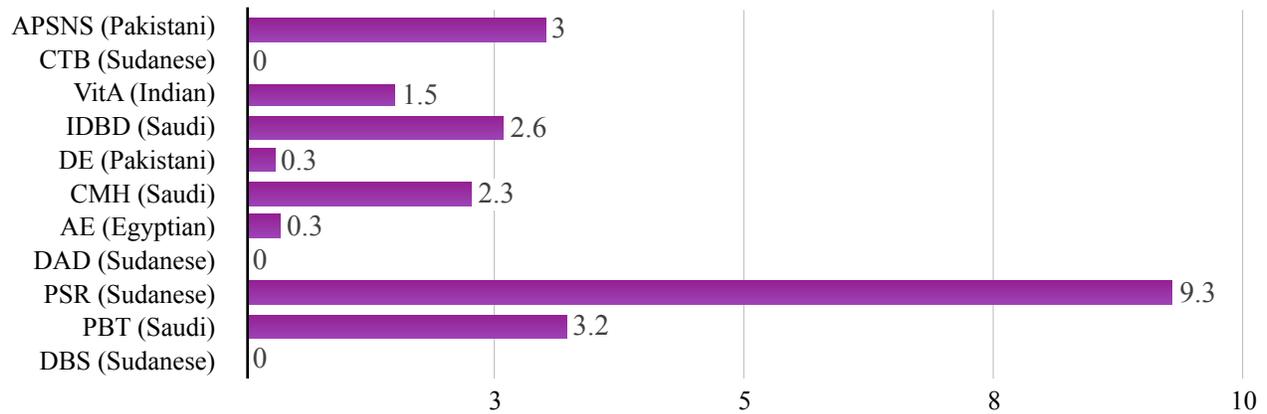
Excerpt 34

[PBT] (a) ... *malignant tumor had [inaudible] glial cells [inaudible] neurons do divide and they do form tumors. (b) Anything else? **Okay**. See you tomorrow.*

[VitA] *For example, the change in the confirmation - (c)confirmation means shape, **okay**. (d) So, let's see here. This is opsin and this is vitamin A, okay?*

The position of the DMs ‘*okay*’ near or at the end of discourse, with a falling intonation (indicated by full stop) and the start of new information in the subsequent utterances, all support the theory that these markers acts as precursors that a (sub)topic is near or at its end.

Figure 5.19 SASME Frequency ‰ of Closer ‘*okay*’

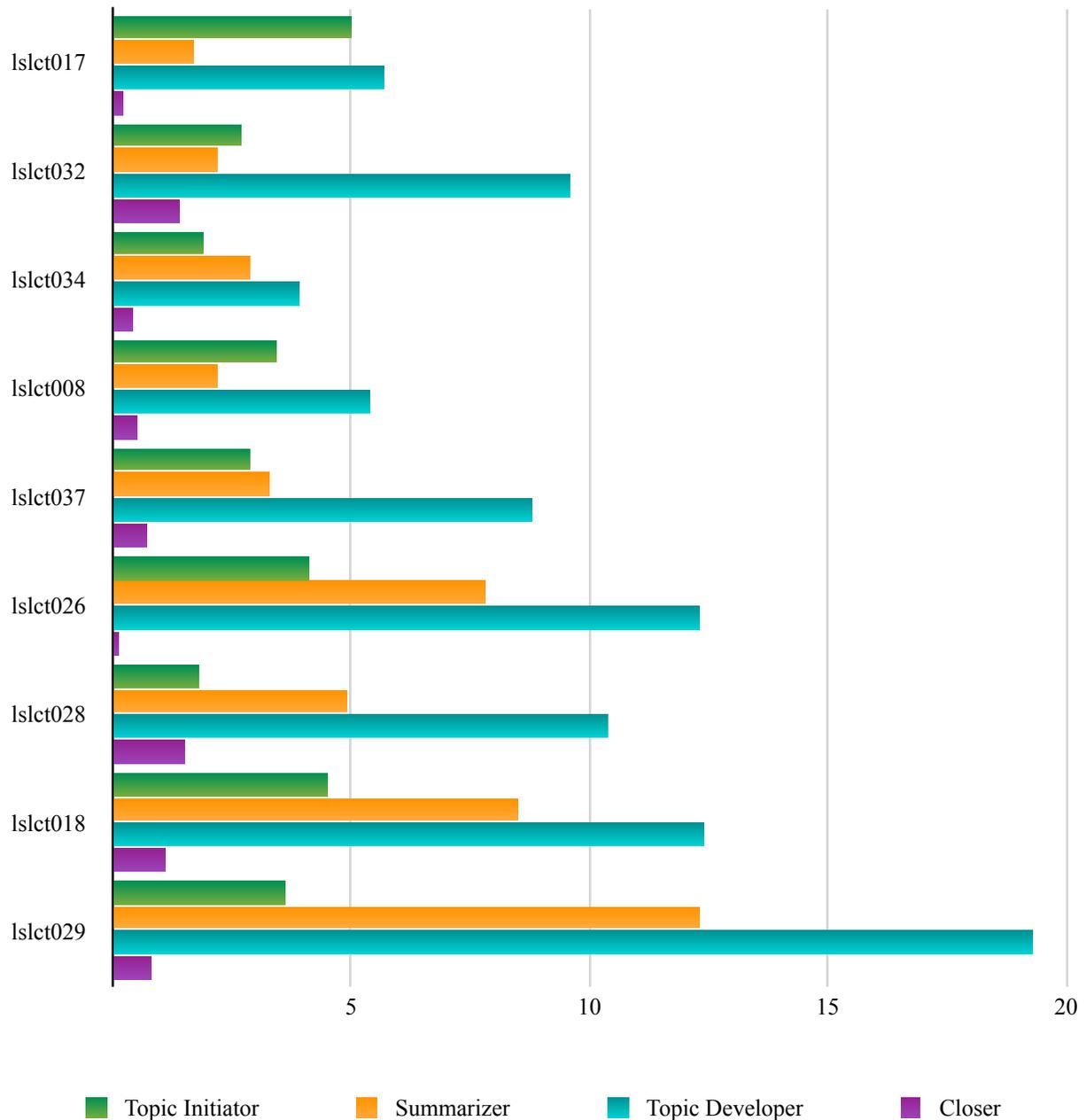


The Closer function is used in eight of the lectures, with the highest number of occurrences in the *PSR* lecture at 9.3‰ (see Figure 5.19). This is three times greater than the next highest frequency in the *PBT* lecture at 3.3‰ occurrences.

5.5 Structural Discourse Markers in BASME

The following section presents the results of the analyses⁴⁵ of the BASME sub-corpus and briefly compares each function, when relevant, to the results of SASME.

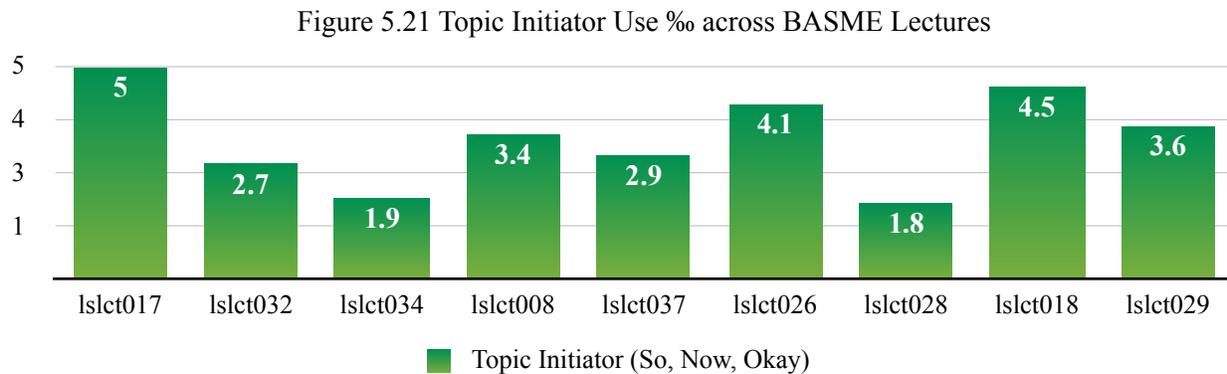
Figure 5.20. BASME Lecturers' Structural DM Use %



⁴⁵ See Appendix D1 to view table of BASME overall distribution of Structural DMs per 1000 occurrences

5.5.1 Topic Initiator (*so, now, okay*)

The DMs that function as Topic Initiators (*so, now, okay*) are comparatively more even in their distribution across BASME lectures ranging from 1.8‰ to 5‰, than those of SASME (see Figure 5.21), which ranged from 1‰ to 10‰.



5.5.1.1 Topic Initiator ‘*so*’

As previously described (see 5.4.1), the DM ‘*so*’ is predominately utterance-initial and functions as an initiator to a new topic or a means to refocus attention back to a topic following a tangent. For example, Excerpt 35 illustrates the function of ‘*so*’ to regain attention and bring back focus to the topic (b) following a tangent discussing the importance of the information in their exams.

Excerpt 35

[ls1ct032] (a) ...that you ha-, know have to know a lot about because it's in your exam but not just not because it's biologically more common or more important (b) **so** diabetes is very common diabetic nephropathy is quite rare why why do you think it's quite rare lady at the end

The second Excerpt (36) similarly provides lecture extracts that illustrate the function of ‘*so*’ to introduce a new topic of *irregular aorta* as in (a). The use of ‘*so*’ functions as a signal to listeners that a new topic has begun that what was discussed in the prior utterance. In (36b) the

utterance is concluded with a “thank you” to listeners and the ‘*okay*’, which acts as an attention-getter and, additionally, as a Topic Initiator, is repeated in function by ‘*so*’ in (c). We can assume, by context, that the lecturer in *lslct029 (Concentration and Dilution of Urine)* attempted not only to regain the attention of the listeners, but also to stress that a new topic is being initiated with the use of both DMs ‘*okay*’ and ‘*so*’ (36b-c).

Excerpt 36

- [*lslct017*] (a) **So** here you've got an aorta slightly irregular but probably about seven centimetres in diameter if you go and do this to the general population
- [*lslct029*] (b) ...thank you okay (c) **So** we're looking about how you concentrate your urine so we're going to look at osmolarity and urine concentration

Excerpt 37 illustrates the DM ‘*so*’ functioning to signal a return to the topic under discussion, *Richard Bright* (37b), following a slight divergence (a). While (c) functions to signal the initiation to diverge off-topic to discuss a set of handouts and the pointer which seems to have been misplaced by the speaker.

Excerpt 37

- [*lslct034*] (a) ...er the story as they say goes back to to eighteen-twenty-seven (b) **so** Richard Bright was a nephrologist perhaps the world's first nephrologist at Guy's Hospital
- [*lslct037*] ...and the infection (c) **so** we have another chunk of handouts and i'll try and get this up while you hand those around if I can find me pointer again i've put the bally thing down

In BASME, the DM ‘*so*’ is not the most frequent in the Topic Initiator category, unlike SASME, with 1.3‰ to 3.2‰ total occurrences respectively.

5.5.1.2 Topic Initiator ‘*now*’

The Topic Initiator ‘*now*’ had the highest frequency of occurrences with 1.5‰, as well as in SASME with 0.7‰. This DM is used in the lecture discourse to initiate a new topic and

additionally, like ‘*so*’, is used as an attention-getter to return focus to the speaker, as in Excerpt 38.

In the following example (Excerpt 38), the speaker, lecturer - *nm0607**⁴⁶, in utterance (38a) indirectly inquires whether the students have the information, discussed in the previous utterance, in their handout. Two students (38b-c) utter a positive response and confirmation that they have received the handout. With this confirmation, the lecturer signals with the DM ‘*now*’ (38d) that he needs their attention re-focused back on him as he introduces new information about the handout.

Excerpt 38

[*lslct037*] *nm0607: ...*(a) *you've got them in your handout*
 sm0608: (b) yeah
 sf0609: (c) u-huh
 nm0607: (d) now I haven't put them on this copy I can't get onto my hard
 disk it's er let's shut this off a minute...

The token ‘*now*’ is used in different contexts with different meanings and functions and the analysis of the corpus data differentiated each use. Only those that functioned as a DM initiating a new topic are included in this category as in (c) and (e) in Excerpt 39. The other uses of the token not included were instances of ‘*now*’ as in, the information is on “the board *now*” (b) and “by *now*” (f), both referring to its temporal meaning ‘at the present moment’.

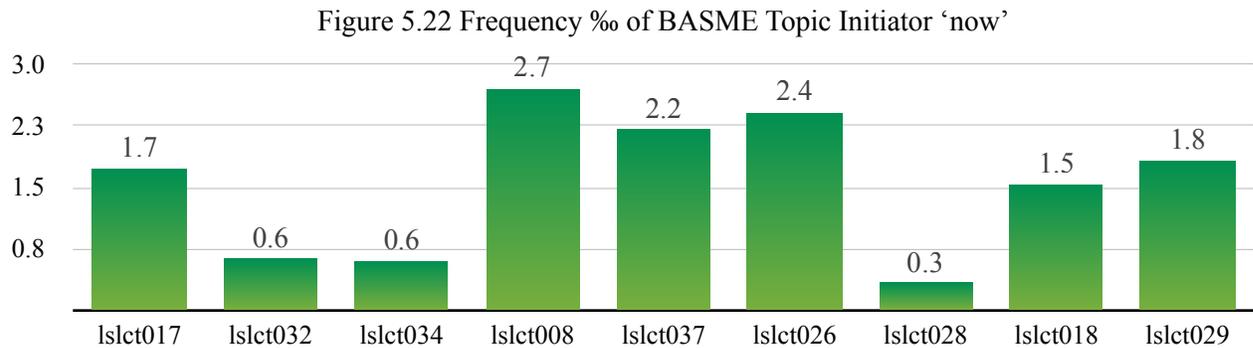
Excerpt 39

[*lslct017*] (a) *...and this is the only equation you're going to see in this lecture and it's*
 on the board there (b) now (c) now what does that tell you does it look
 even vaguely familiar...

[*lslct032*] (d) *okay (e) now why are we interested in diabetes and diabetic*
 nephropathy (f) you know by now my talk's interactive so if you don't
 answer i'll start picking on people

⁴⁶ “n” refers to ‘non-student’, “m” to male and speaker ID 0607. Speakers “sm0608” and “sf0609” refer to ‘s’=student, ‘m’ = male, ‘f’ = female and their speaker ID.

The DM ‘*now*’ varied in total occurrences in BASME from 0.3‰ to 2.7‰ (see Figure 5.22). The Topic Initiator marker is present in each of lecturers unlike in SASME which had occurrences in only two of the lectures with a similar occurrence of 0.2‰.



5.5.1.3 Topic Initiator ‘*okay*’

The DM ‘*okay*’ is the least frequent marker used in this category in BASME with 0.5‰ occurrences. Similar to the previous two markers ‘*so*’ and ‘*now*’, this DM, ‘*okay*’, acts as an attention-getter, but mainly functions to signal to listeners that a new (sub)topic will now begin as in Excerpt 40. In this example, the lecturer, *nm0572*, elicits from students how “proteinuria” is detected in a patient (a), to which the students (ss) respond with the utterance “dipstick” (b). This is clearly the right answer as the lecturer repeats their response (c) and then indicates that the topic will now shift with ‘*okay*’ (d), which introduces the initiation of a question, soon aborted to inquire whether students have experience with this type of test.

Excerpt 40

[lslect034] *nm0572*: (a) ...how do we how do we find out whether somebody has proteinuria
ss: (b) dipstick
nm0572: (c) dipstick (d) **okay** how do you do a dipst-, has any-, who's done a dipstick anybody a few people

In Excerpt 41, the speaker discusses “the notes” for the lecture (a) and although not indicated here, due to varying transcription conventions (see 4.5.3), the utterance is concluded at

“something else to read”. In the analysis, the following utterance is marked as initiated with the DM ‘*okay*’ which, as discussed, functions as an attention-getter, but mainly functions to signal the start of a new (sub)topic - explicitly stated by the lecturer (41c), “what I’m going to do is to talk about hepatitis B virus”.

Excerpt 41

[lslct018] (a) ...that can act as the notes for this lecture if you want something else to read (b) ***okay*** (c) ***so*** what I’m going to do is to talk about hepatitis B virus and I’m going to talk about

With this DM, in its function as Topic Initiator, a pattern or what would be referred to as a collocation, is identified of ‘*okay*’ + ‘*so*’. This co-occurrence of DMs is also identified by Bellés-Fortuño (2006) in her study of academic social science lectures. She reports that there are 87 total occurrences (0.7%) in her NS English corpus which seemed to function as “a summative, resultative evaluation” (Bellés-Fortuño, 2006: 234). Excerpt 41, (b) and (c), illustrate both DMs are used in sequence to capture the audiences’ attention and signal the start of a new (sub)topic.

In at least three examples, the Topic Initiator ‘*okay*’ is used to initiate the lecture discourse as illustrated in Excerpt (42). The first example (lslct026), shows the collocate ‘*okay so*’ as Topic Initiators (a-b), while in the second example, the DM ‘*okay*’ is used alone as the first token uttered (c) at the start of the lecture. The last example, is similar to (c), in that both (d) and (e) function in the same manner and both are utterance-initial. In context, the DM ‘*right*’ (d) could be replaced with ‘*okay*’ with no loss in meaning or function and, as such, is regarded nearly identical to the Topic Initiator ‘*okay*’.

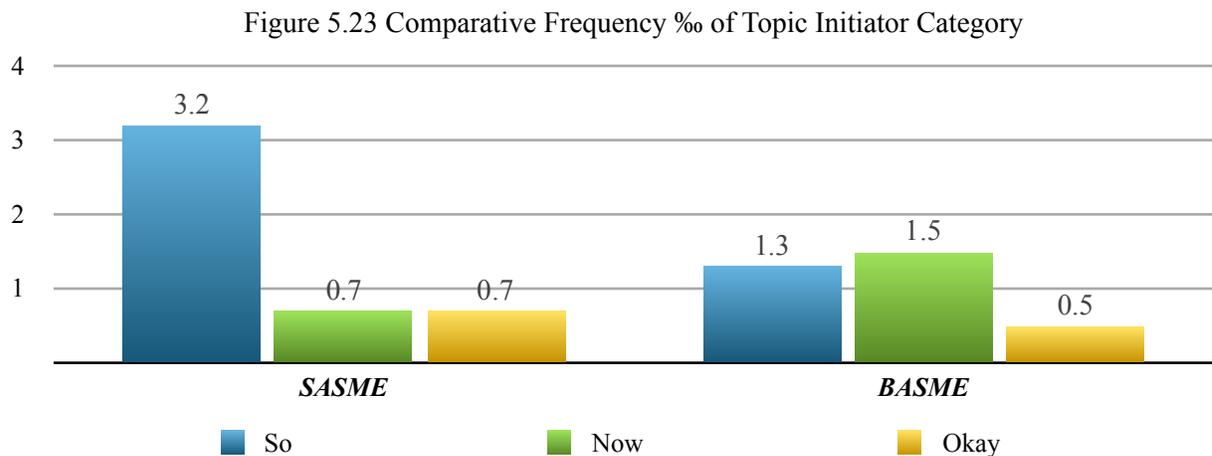
Excerpt 42

[lslct026] nf0368: (a) ***okay*** (b) ***so*** today you’ve got two lectures off me er the first one we’re going to start and look at renal function

[lslct018] nm0283: (c) ***okay*** as promised the lecture er now is looking at hepatitis B virus there are no handouts for the this lecture

[lslct034] nm0572: (d) ***right*** (e) ***okay*** settle down everybody er I think I’m meant to be talking to you about glomerular pathology er I’m not a pathologist

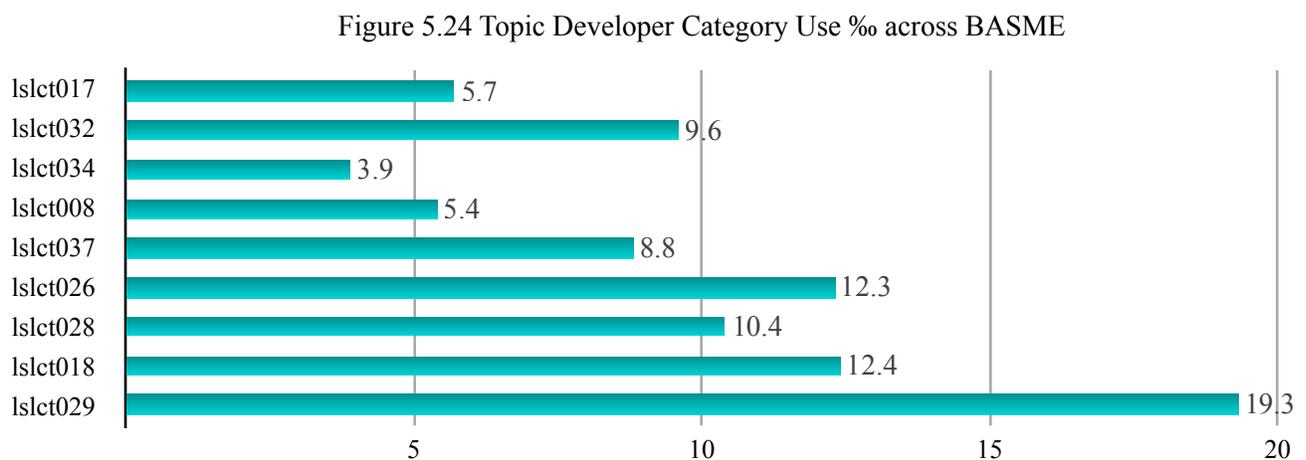
The category of Topic Initiator is the second lowest function in overall frequency to any other under the Structural DM category in BASME with 3.3 ‰ (see Figure 5.23).



The frequency is not significantly different to Topic Initiator in SASME with an overall frequency of 4.7‰; although individual frequencies of the DMs differed in that ‘now’ (1.5‰) is the highest in BASME while ‘so’ (3.2‰) the highest in SASME.

5.5.2 Topic Developer (*so, because, ‘cause, also*)

Topic Developer DMs are used considerably more across the BASME lecturers with occurrences ranging on the lower scale from 3.9‰ to 5.7‰ and on the higher scale, from 8.8‰ to 19.3‰. These occurrences are comparable to results in SASME where lectures have occurrences ranging from 0.7‰ (in one lecture) and from 3.5‰ to 18‰ (in the remaining ten).



5.5.2.1 Topic Developer ‘so’

In some cases, it is difficult to decide whether the incidences of the DM ‘so’ are functioning as a Topic Developer or as a Summarizer (see Excerpt 43). Based on the established framework (see section 4.8.2) a clear distinction is made in the description of each function. First, Topic Developer DMs function to introduce additional information about a topic under discussion and elaborate on the previous utterance. Summarizers, on the other hand, function to conclude information and the DM ‘so’ cannot introduce any new information in the utterance.

In Excerpt 43, the lecturer is discussing the start date of a prior research project and the death of two patients at the start of 2002 (a). The lecturer then expands on this information by explaining explicitly at this point of the utterance (b) that she decided to begin the research at the start of February so that the two deaths would not impact her research statistics (b). The DM ‘so’ in (b) does seem to indicate a meaning of *as a result* which is defined as a characteristic of Summarizers, but, as stated, in this context the utterance (b) is providing additional details and information related to the previous utterance (a).

Excerpt 43

[lslct017] ...*(a) at the beginning of two-thousand-and-two the first two patients operated on died within thirty days after surgery (b) so I thought it would be better to start my series on the first of February two-thousand-and-two*

The defined categories were more challenging to identify unequivocally in BASME than they were in SASME, though, in some incidences, there is no difficulty as in Excerpt 44. The lecturer of *lslct034*, attempts to inject humor with a play on words (a) and, with the DM ‘so’ signals that the lecturer will clarify his attempt at humor by explaining that “there's a bit of a joke in there” (44b).

Excerpt 44

[lslct034] *(a) Marx Groucho Marx okay [laugh] i've got Bright's disease and he's got mine (b) so there's a bit of a joke in there it's obviously not a very funny joke [laughter] it's obviously a er a joke that's only appreciated by kidney doctors*

Excerpt 45 exemplifies the function of ‘so’ to develop (sub)topics as in (b), which expands on the topic of urine production introduced previously in (a). It continues to develop the topic with more information and details about minimum urine production (c). Similarly in (d) the DM ‘so’ introduces additional information to the students that in addition to diabetes, “schizophrenia rheumatoid disease” are two of the most common chronic diseases.

Excerpt 45

[lslct029] (a)...if we have a look at urine production first off then (b) **So** this is your you've got a huge flexibility in how much urine you do or don't produce (c) **so** you've got a minimum of about point-twenty-five er point-two-five mls per minute

[lslct032] it's probably about two per cent of the general population in the U-K five per cent of blacks and Asians (d) **so** it makes it with schizophrenia rheumatoid disease one of the commonest chronic diseases

The last Excerpt 46, illustrates the function of ‘so’ to develop the topic by clarifying information.

Excerpt 46

[lslct034] ...and and most proteins including albumin are made in the liver (a) **so** there's probably liver trouble in nephrotic syndrome (b) but we'll call it nephrotic syndrome for the moment

In this example, the speaker clarifies to listeners that while “nephrotic syndrome” may include liver trouble (a), for the present moment it will just be referred to as “nephrotic syndrome” (b) - a nonspecific kidney disorder.

5.5.2.2 Topic Developer ‘because / ‘cause’

The second DM under the function of Topic Developer is ‘because’ and its informal equivalent ‘cause’. The DMs are not combined as one set under the lemma ‘because’ as they were transcribed as such and the analysis is based on the original unedited data. For this reason, the two

forms of the DM are presented as separate frequencies of occurrence, though they are viewed as functioning in the same manner and maintaining the same essential meaning.

The function of the DMs under Topic Developer, as stated previously, is to signal that additional information, related to the topic under discussion, will be developed or clarified with examples/analogies/facts or with previous background information. For example, following utterance (a) of Excerpt 47, (b) clarifies the importance of holes in pores in relation to size and the syndrome under discussion. The second example, uses '*cause*' (47d) to signal the development of the topic (*Concentration and Dilution of Urine*) with examples “stones urea and sodium” and a reminder “if you remember” that the information is covered previously. The last example, illustrates the function of '*cause*' to develop a topic with additional facts (47f) that the speaker may assume the listeners know but, nonetheless, explicitly states to clarify.

Excerpt 47

[lslct034] (a) ...these holes are are which have to be there there has to be a pore
(b) '*cause* sometimes you may physiologically want to open them up and send things through or not send things

[lslct029] (c) the concentration gradient in the medulla it also does one really other important things (d) '*cause* if you remember from stones urea and sodium and chloride they're all things...

[lslct017] (e) whereas you can get at the end of a research project the date of death and the cause of death of patients entered into this (f) '*cause* they're all recorded centrally in Southport

One point of interest is why the NS in BASME use the informal '*cause*' as opposed to the more formal '*because*'. Statistically, '*cause*' is used 0.12‰ which is less than half the occurrences of the formal use of '*because*' with 0.27‰. There is no clear answer to the question, nor any obvious pattern to when '*cause*' is used in the discourse. However, as English is the L1 of the lecturers, one possibility is that by using a more relaxed informal structure, speakers create a personal environment similar to what NS Arabic lecturers do in SASME by incorporating the L1 in their discourse.

It is also interesting to note, that in comparison to BASME, the NNS English lecturers in SASME had zero occurrences of 'cause'; however, the overall frequency of 'because' had only 0.01% less occurrences in the NNS corpus than in the NS corpus.

As presented previously, the DMs 'because' and 'cause' are located utterance-medial and function to clarify, modify and/or support immediately prior discourse. The utterance following 'because' functions to clarify and support why the number discussed in the prior utterance (48a) is important with an example “if you were producing two litres of urine” and assurance more visual clarification will follow (b). Similarly, the DM 'because' in (d) functions to signal that the subsequent utterance will develop the topic and clarify the importance in (c) with support of *motive* (d).

Excerpt 48

[lslct029] (a) now this is kind of an important number (b) **because** if you were producing two litres of urine a day and you'll see why on the next couple of slides

[lslct032] (c) ...full-blown diabetic nephropathy but it's more important than that (d) **because** it's possibly the forerunner of all the small and large vessel complications of diabetes and ...

The lecturer in *lslct037 - HIV Infection* (see Excerpt 49), supports his statement that medical professions were under “an enormous amount of pressure” in the 80s (a) with an assertion that “that's what people were worried about” and imitating what a patient might state “I am sick doctor you know I am infected with something doctor” (b).

Excerpt 49

[lslct037] (a) ...an enormous amount of pressure in the mid-eighties to try and come up with prognostic indicators (b) **because** that's what people were worried about I am sick doctor you know I am infected with something doctor...

The DM 'because' seems to be one of the clearest forms of DMs in regards to its function as a Topic Developer to the listeners. It provides an explicit signal that conveys a meaning of *why* and develops the previously uttered information with support.

5.5.2.3 Topic Developer ‘also’

The final DM, under the function of Topic Developer, is ‘also’ which in BASME varies in location from utterance-initial to utterance-medial. In both cases, the DM ‘also’ functions to signal to the listeners that the following utterance is not only a continuation of the previous topic but also that a connection will be established, clarified or supported. For example, the following two transcript extracts (see Excerpt 50) are illustrative of the speakers supporting their statements through indirect anecdotes such as (b) describing “some hospital doctors” as “sit[ting] on their arses” and “pontificate[ing]” all week. Or, through an analogy (c) comparing the “dull[ness]” of train-spotting to how people feel about “kidney stone disease” (d).

Excerpt 50

[lslct034] (a) ...er yeah G-Ps don't want to get called out at weekends. (b) **Also** some hospital doctors er sit on their arses all week go to the mess er sit around pontificate now bloody hell Thursday all the oh bad blood tests

[lslct028] (c) The reason why it's similar to trainspotting is that most people think trainspotting is rather dull and er most people (d) **also** think kidney stone disease is rather dull

Advice is also another form of development of a topic that follows the DM ‘also’. For example, in the following first extract (see Excerpt 51) the speaker reminds the students to “get excretion and secretion the right way round” (a) and signals with ‘also’ that the subsequent utterance is still related to the previous topic and providing additional advice (b) on the topic - *Introduction to Renal Function: the Concept of Clearance*. The second extract, *lslct018 (Hepatitis B Endemicity)* discusses the transference of Hepatitis B through sexual contact (c) and develops on the topic of transference with (d) another example - intravenous means including blood transfusions.

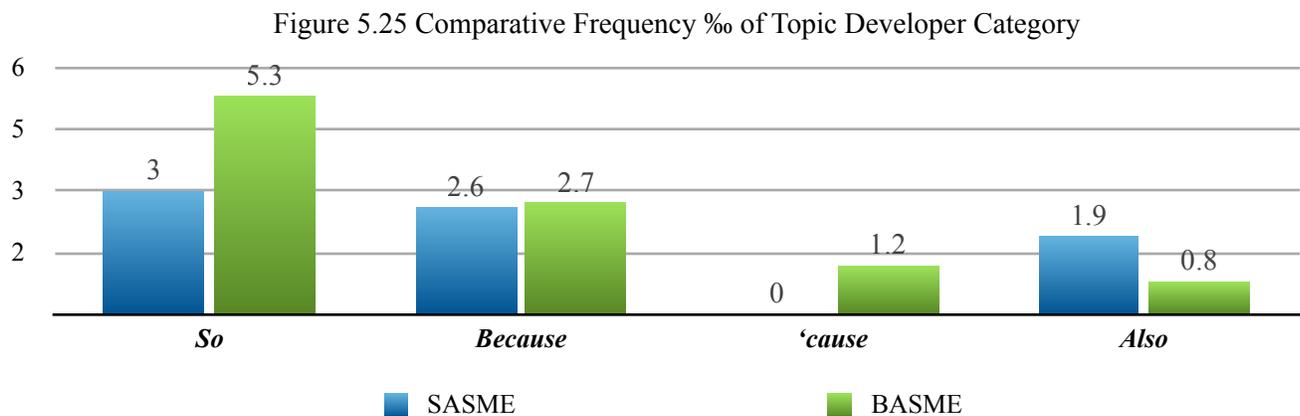
Excerpt 51

[lslct026] (a) that this is comes in remember I said remember to get excretion and secretion the right way round (b) **also** remember to start getting in the habit of using urine and filtrate in the right er context urine is only urine

Excerpt 51 (continued)

[lslct018] (c) *sexual contact has an important role in in developed countries and*
 (d) **also** *intravenous and particularly used to be very bunch of blood borne diseases associated with tranfusions until all tranfusions started to be screened*

The category of Topic Developer is the highest in overall frequency to any other category in BASME with 10.5% overall occurrences. This is similar to SASME which also has the highest number of overall occurrences in Topic Developer with 7.1%. Individual frequencies of the DMs such as, ‘so’ which is the highest DM in both BASME and SASME, though differ slightly with 5.3% and 3% respectively (see Figure 5.25).



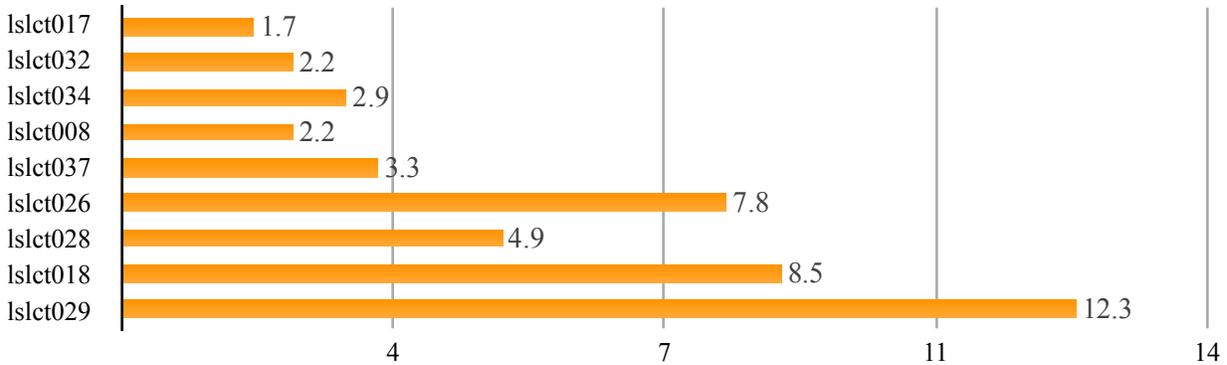
The DM ‘because’ is relatively equal in overall frequency with 2.7% and 2.6% respectively. However, if you include the informal token ‘cause’ to the total, BASME has a higher overall frequency with 3.9%. The last DM ‘also’ is used less frequently in BASME with 0.8% than SASME (1.9%), indicating that the DM is not as favored by the NS as by the NNS in SASME.

5.5.3 Summarizer (so, therefore)

The DMs under the Summarizer function were used in a comparably similar manner in BASME (5.3%) as they were in SASME (4.2%). In Figure 5.26, we see that the distribution of the DMs ‘so’ and ‘therefore’ were comparable in the first five lectures (*lslct017 - lslct037*) and

(*lslct028*). The remaining three (*lslct026*, *lslct018* - *lslct029*) illustrate a higher frequency of occurrences.

Figure 5.26 Summarizer Category Use % across BASME Lecturers



5.5.3.1 Summarizer 'so'

The summary function of 'so' is illustrated with the following Excerpts (52) in which 'so' is used by the lecturer to signal the point - a summary of what students should understand from his previous utterance(s). In the first Excerpt (52), the lecturer wraps up his utterance describing the "growth rate" of about "two-point-four millini-, metres a year" for smokers (a), by stating in what could be described as more everyday language (b), as opposed to 'medical speak' (a). The second example (52c-d), illustrates the function of 'so' to summarize information already covered prior to reminding the students of the information and clarifying the relationship between utterance (c) and (d). The third example (e-f), is similar to the first example in that the lecturer chooses to use everyday language to make the point that the information in utterance (e) regarding "immigration" and "endemicity" are not caused by "host genetics" (f). The DM in (f) is marked as a Summarizer in this context as the lecturer is not providing new information but is, in fact, summing up what the listeners should conclude from the previous information.

Excerpt 52

[*lslct017*] (a) *growth rate about two-point-four millini-, metres a year those who are actively smoking two-point-nine millimetres a year big difference (b) so if you keep smoking you're going to grow faster you're going to come to surgery faster*

Excerpt 52 (continued)

[lslct029] (c) *er three and four are also in the collecting duct er but these are in the basolateral membrane (d) so that's the interstitial tissue er border with the tubule lumen remember*

[lslct018] (e) *...if you transfer people through immigration from a high endemicity area to a low endemicity area then after a generation they've actually got low endemicity infection (f) so it's not not something to do with host genetics that makes people more susceptible*

The next two examples (see Excerpt 53) illustrate the DM 'so' functioning as a Summarizer, (53a-c), but also illustrate instances when the lecturers presumably felt a need to follow up their initial utterance with an explanation of why "it's bad for you" (b) or differentiate between previously used terms, e.g., "false positive" vs "false negative" (d) used in the summary.

Excerpt 53

[lslct032] *...they're getting the complications there (a) so that's bad for you (b) because if you wait long enough you'll get the complication of diabetes then something else might happen*

[lslct034] *which will give you a trace of proteinuria (c) so that's a false positive er (d) a false negative is when you miss important biological levels of proteinuria which a dipstick said was a trace*

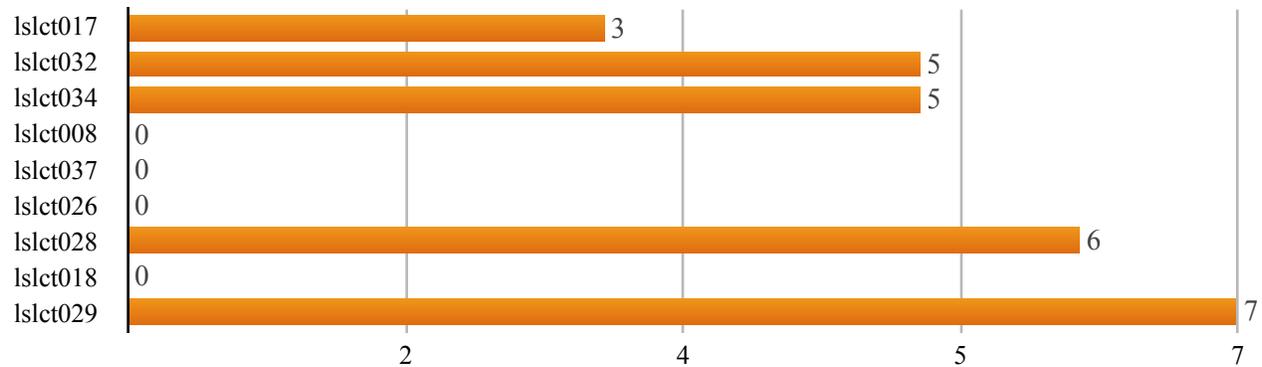
It is again not surprising that lecturers, NS and NNS, supplement their speech following summaries (and other DM functions) with examples and details as the topics are dense and content heavy for students. This is discussed further in the discussion chapter (see 7.2.3.2).

5.5.3.2 Summarizer 'therefore'

The DM 'therefore' as a Summarizer functions similarly to its DM counterpart in SASME. Moreover, it is similarly underused in the BASME lectures appearing in five of the total nine lectures, with 0.4‰ overall occurrences. However, it is more evenly distributed across lectures with three of the lectures at 3‰ or 5‰ and two at 6‰ and 7‰ (see Figure 5.27). SASME

distribution is skewed across four lectures of which three ranged from 0.7% to 3.8% and one lecture had 12.7%.

Figure 5.27 Frequency of BASME Summarizer ‘therefore’



Based on the analysis and the lack of any other form of summarizing DMs, we can speculate that the ‘so’ marker, which can function interchangeably with ‘therefore’, is used as the preferred marker for summarizing.

The function of ‘therefore’ used in BASME is illustrated in the first Excerpt (54) in which the DM signals a summary of the previous utterance which describes two treatments in terms of its “quality of life” (a). With the addition of ‘so’, which functions as an attention-getter and Summarizer signal, the DM ‘therefore’ is supported and stresses the importance of utterance (b) to the listeners of summarizing the main points.

Excerpt 54

[lslect017] (a)...but there was almost no different in quality of life between the two treatment arms (b) So **therefore** with a similar survival similar quality of life but one treatment costing much more than the other

[lslect029] So you've got changes here of plasma osmolarity if you get alterations in your extracellular volume (c) **Therefore** you're becoming dehydrated that will also change its release

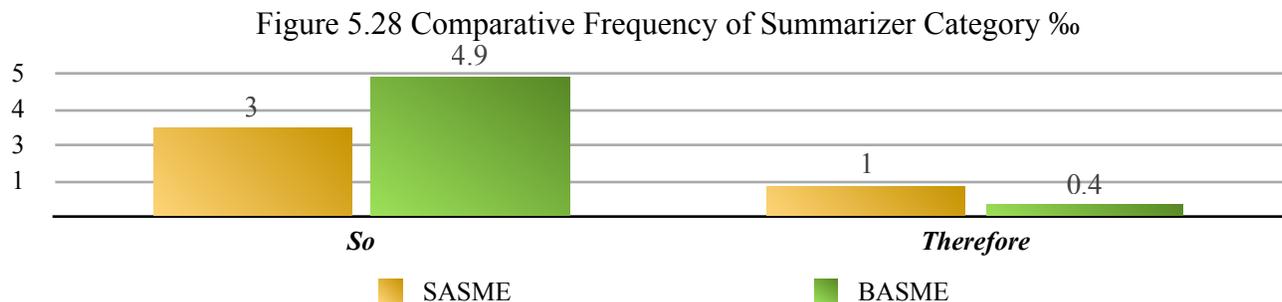
The second example (c) illustrates the DM ‘therefore’ functioning to summarize the information - “you're becoming dehydrated” - that listeners, presumably, should understand at that point in the lecture.

Excerpt 55

[lslct032] *it's much more common in IDDMs but IDDM is a much rarer disease than NIDDM (a) So by definition (b) **therefore** far more people with NIDDM and that's where the problem lies get diabetic nephropathy*

[lslct034] *and er sorry I have to look this up 'cause I always forget er yeah (c) most proteins are very negatively charged And they (d) **therefore** repel each other and also the glomerular basement membrane is...*

In Excerpt 55, the deliberate use of ‘*therefore*’ (b) is illustrated to mark a summary. In the prior utterance, the lecturer describes “IDDMs” and begins to utter “So by definition”, but then seems to reconsider and instead uses the Summarizer ‘*therefore*’ (b) to signal to the listeners that the following utterance will clarify and sum up the main points of the previous utterance. The next example (*lslct034*) illustrates a brief background summary point for listeners as the lecturer describes the proteins as “very negatively charged” and thus indicates the main point students should understand (d).



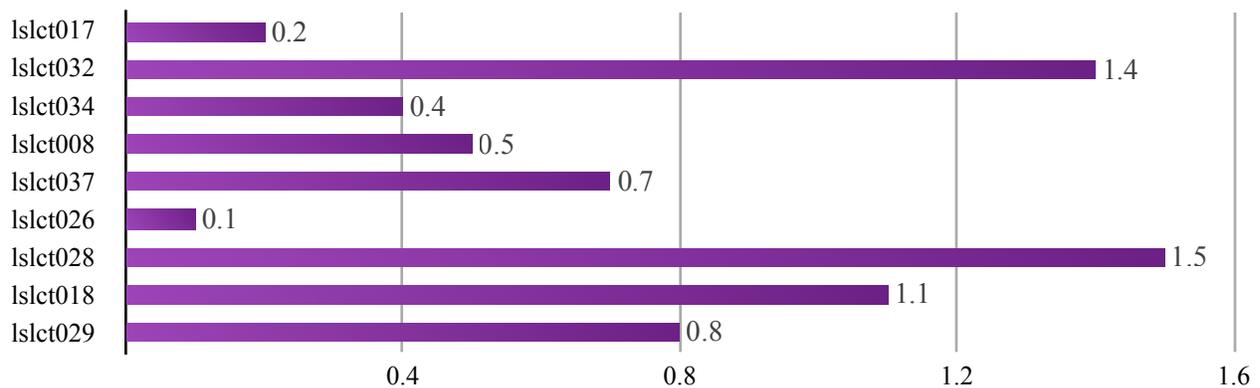
In contrast to SASME, as previously described, the Summarizer ‘*therefore*’ is not used significantly in BASME. However, viewing Figure 5.28, we see that the use of its functional counterpart is more significant in BASME with 4.9‰ than in SASME with 3.1‰.

5.5.4 Topic Closer (*okay*)

The last function in the Structural DM Category is Closer, with one DM, ‘*okay*’. The DM is the only one identified during the analysis of common features in the NS and NNS discourse. The DM ‘*okay*’ is marked with a full stop to signify its function, as does the falling intonation.

The Closer ‘*okay*’ signals to listeners that a (sub)topic has closed and in some cases, functions as an attention-getter. As an attention-getter, the DM is used when discourse has moved off-topic or attention has shifted away from the lecturer and functions as a signal to regain the floor. In Figure 5.29, the Closer is used across BASME with occurrences ranging from 0.2‰ to 1.5‰. Although, the number of occurrences are lower in BASME than the average occurrences in SASME, the NS lecturers show a relatively more similar distribution in their use than NNS.

Figure 5.29 Closer Category Use ‰ across BASME Lecturers



As a Closer the lecturer may also use the DM to signal the close of a digression as in (56a) where the lecturer seems to be talking to himself. The lecturer uses the Closer, to not only indicate he ended the prior utterance, but additionally to regain the students’ attention. This is followed with ‘*right*’, functioning as reiteration for the closer and ‘*fine*’ (b) as a possible filler or attention-getter. The additional markers are used either in support of regaining attention, as the first may not have been successful or more likely, provided the lecturer a moment to gather her thoughts before beginning the next utterance (c).

The next example in Excerpt 56, illustrates a pattern found in SASME, though not as extensively in BASME, of ‘*okay so*’. Each DM is used as a different function in context, with ‘*okay*’ (d) signaling the close of the prior utterance describing diabetes. The sentence following (d) is a summary of information as indicated by ‘*so*’ and the lack of new information in the utterance. The Closer (d) is an example of the defined function of the DM ‘*okay*’ to close one utterance in a pre-emptive move that indicates a new (sub)topic will begin - in this case a summary of information.

Excerpt 56

[lslct017] *er well that's great 'cause the screen's gone off so I er oh (a) okay. (b) Right. Fine (c) i've got you this is a cartoons of the aorta*

[lslct032] *...it's all wrong and it's too complicated so just think about it simply (d) okay. (e) So diabetes is difficult to define and therefore diabetic nephropathy is difficult to define but...*

In both examples in Excerpt 57, the lecturers signal the end of the previous utterances (a) and (b). Each of the DMs, ‘okay’, are followed with brief pauses similar to those in SASME that emphasize the end of an utterance. Unlike the Closers in Excerpt 56, the DMs in (a) and (b) are not followed by other DMs and instead, are followed by casual discourse phrasing such as, “Bit of pathology...” (lslct032) and “Let me go back...” (lslct018).

Excerpt 57

[lslct032] *some of the complications of diabetes if namex can crack it in her in in in her lifetime (a) okay. Bit of pathology er I know your course doesn't emphasize pathology*

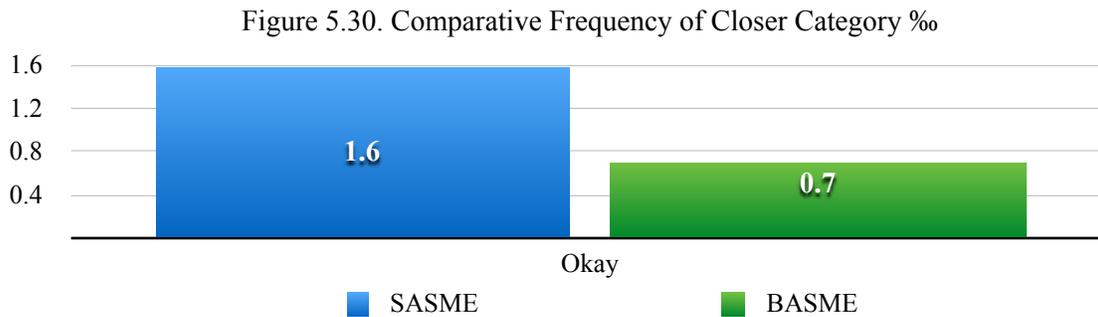
[lslct018] *what I'm doing is I'm reducing the basic reproduction number (b) okay. Let me go back I'm actually starting on here yeah So I'm starting at a value here*

The following extract (see Excerpt 58), illustrates multiple functions of the Closer ‘okay’, including the use of humor to engage with students. Humor was similarly identified as a pedagogic function in BASE lectures by Nesi (2012). The Closer ‘okay’ signals the lecturer’s attempt to regain control of the floor and continue with the previous lecture topic or a new topic of discussion.

Extract 58

[lslct028] *(a) nm0370: 'cause they're completely non-thinking people who couldn't be physicians [laughter] is that what you said
(b) sf0414: yeah
(c) nm0370: [laughter] Okay. Okay sorry I don't I don't don't let me put words in your mouth though*

In (a) the lecturer (*nm0370*) teases the students by rephrasing a previous response and is rewarded with laughter (a) and confirmation by a female student (b). Following more laughter, the lecturer tries to bring the laughter to a close with ‘*okay*’ (c), repeating it twice before apologizing and then moving on with the lecture topic.



Overall, the Closer function has a higher total frequency (1.6%) in the NNS lectures than in the NS lectures. This may be indicative of a need by NNS lecturers to use the DM to signal the close of a (sub)topic for the benefit of the NNS students; whereas, the NS lecturers may not feel this is necessary for their heterogenous study body.

5.6 Overview of Structural DMs in SASME and BASME

The following tables are representative of the total number of raw occurrences and occurrences per thousand (‰) in each of the functions under the Structural category in each corpus.

Topic Initiator	#DMs raw	‰	Topic Developer	#DMs raw	‰	Summarizer	#DMs raw	‰	Closer	#DMs raw	‰
<i>So</i>	154	3.2	<i>So</i>	141	3	<i>So</i>	146	3.1	<i>Okay</i>	75	1.6
<i>Now</i>	35	0.7	<i>Because</i>	125	2.6						
<i>Okay</i>	33	0.7	<i>'cause</i>	0	0	<i>Therefore</i>	52	1			
			<i>Also</i>	90	1.9						
Total	222	4.7	Total	356	7.5	Total	198	4.2	Total	75	1.6

Table 5.1 SASME Structural DMs Summary

Topic Initiator	#DMs raw	‰	Topic Developer	#DMs raw	‰	Summarizer	#DMs raw	‰	Closer	#DMs raw	‰
<i>So</i>	81	1.3	<i>So</i>	342	5.3	<i>So</i>	313	4.9	<i>Okay</i>	48	0.7
<i>Now</i>	96	1.5	<i>Because</i>	171	2.7						
<i>Okay</i>	33	0.5	<i>'cause</i>	78	1.2	<i>Therefore</i>	26	0.4			
			<i>Also</i>	54	0.8						
Total	210	3.3	Total	645	10.1	Total	339	5.3	Total	48	0.7

Table 5.2 BASME Structural DMs Summary

	<i>Total DMs #</i>	<i>Total DMs ‰</i>
SASME Structural DM Categories Total	865	18.2
BASME Structural DM Categories Total	1466	22.9

Table 5.3 Structural DMs Overview

5.7 Results I: Structural Discourse Markers Chapter Summary

This results are summarized as points in the following:

- (1) The Topic Initiator DMs are the second highest used function in the SASME corpus with 4.7‰ total occurrences and are used as both attention-getters and to signal a shift to a new (sub)topic. The DM ‘*so*’ is the preferred DM with 3.2‰ occurrences, in contrast to ‘*now*’ (0.7‰) and ‘*okay*’ (0.7‰).
- (2) The Topic Initiator DMs are the third highest used function in the BASME corpus with 3.3‰ total occurrences and are used mainly to signal a shift in discourse to a new (sub)topic. Similarly, the DM ‘*so*’ is the preferred DM with 1.3‰ occurrences, in contrast to ‘*now*’ (1.5‰) and ‘*okay*’ (0.5‰).
- (3) The SASME Topic Developer DMs are the highest used function with 7.5‰ total occurrences and are mainly used as a verbal cue to signal the continuation of a prior initiated topic. Topic Developer DMs signal new details, examples or analogies, in either English or Arabic functioning as supplementary information of the main topic.
- (4) The preferred Topic Developer DM in SASME is ‘*so*’ with 3‰ total occurrences, with ‘*because*’ a close second at 2.6‰ and ‘*also*’ a distant third with 1.9‰. There are no occurrences of the informal ‘*cause*’ in SASME which may be due to the nature of the NNS English corpus.
- (5) The BASME Topic Developer is, similarly, the highest occurring function in the corpus with 10.1‰, surpassing the total number of occurrences in SASME. The Topic Developer DMs in BASME are used to signal the development and/or clarification of the main or (sub)topic with details, examples or analogies similar to SASME.
- (6) The preferred Topic Developer DM in BASME is similar to SASME with the exception of ‘*so*’, which at 5.3‰ is the most frequently used DM within its own function group, as well as all other functions in BASME and SASME. The remaining DMs under Topic Developer are not significantly different in frequency (to SASME) with ‘*because*’ at 2.7‰, ‘*cause*’ at 1.2‰ and ‘*also*’ at 0.8‰. Though no occurrences of the informal ‘*cause*’ are identified in SASME, it is not surprising to find it in a NS English corpus.

- (7) A challenge in both the SASME and BASME sub-corpora involved distinguishing the functions of Topic Developer and Summarizer. A distinction is made regarding the content of utterances in that *Topic Developers* include new information while *Summarizers* function as recapitulations of previously introduced information.
- (8) The SASME Summarizers are the third most frequently used function with 4.2‰ occurrences overall, as stated in (7), these DMs function as recapitulations of information introduced in (usually) the immediately prior utterance(s). Summarizer DMs often function as a means of signaling a summing up of information through less formal language and/or grammatical structure of the utterance.
- (9) The preferred DM under Summarizer in SASME is ‘*so*’ with 3.1‰ occurrences overall; with a similar and more frequent, preference in BASME with 4.9‰. The second DM, ‘*therefore*’ is not preferred in the SASME corpus, with a total 1‰ of occurrences and 0.4‰ in BASME. Potentially, the DM ‘*therefore*’ may be too formal in spoken academic discourse and its marked use in discourse, in contrast to the unmarked use of ‘*so*’, may be responsible for its low frequency in both the NS and NNS English corpus.
- (10) Summarizers in BASME are the second most frequently used function with 5.3‰ occurrences overall and are used as recapitulations of previously introduced information.
- (11) Closers, which include only one identified DM, ‘*okay*’, are the least used function in SASME with 1.6‰ total occurrences and are used to signal the end of a (sub)topic. Furthermore, they are also used as a verbal cue to draw back attention to the speaker and allow the lecturer to regain the floor, following, for example, off-topic discussions or laughter.
- (12) Similarly, Closers in BASME are the least used function with 0.7‰ occurrences overall and are used to draw topic, subtopics or off-topic discourse to a close. Closers function to regain focus on the topic and allow the speaker to signal a topic has or will soon be, drawn to a close.

CHAPTER 6
ANALYSIS OF RESULTS II:
INTERACTIONAL & ARABIC DISCOURSE MARKERS

*“Topics are broached, commented on, developed, extended, replaced, retrieved...
and all this conversational flux is continuously shaped and negotiated by
interactants. Crucial to this... is the inserting of discourse markers and other
interactional signals into the stream of talk.”*

(Thornbury and Slade, 2006: 57)

Chapter 6 Analysis of Results II: Interactional & Arabic DMs

6.1 Introduction

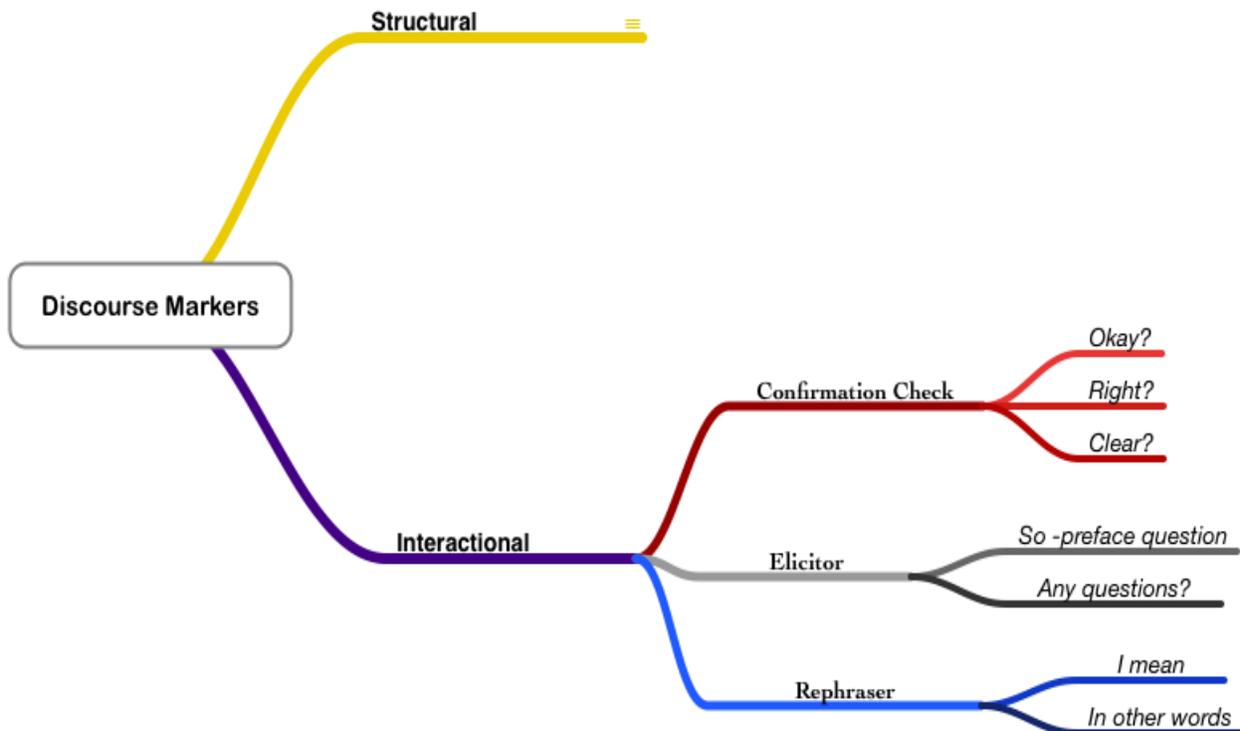
This chapter contains two sub-sections, the first discusses the results of one category of English *Interactional* DMs to investigate the similarities and differences of their functions in the Saudi Academic Spoken Medical English (SASME) corpus and the British Academic Spoken Medical English (BASME) sub-corpus. The second subsection discusses the results of the Arabic DMs analyzed in the SASME corpus and compares their use and functions with their English counterparts in both the Saudi and British corpora.

The chapter is organized into three main sections:

1. First, a discussion of the results of the Interactional discourse markers' functions, uses and frequency of occurrences in SASME and BASME.
2. Second, results of the Arabic DMs, both Structural and Interactional (e.g., *ya3ni* {it means}; *fa* {so}; *tamam* {okay}, etc.), are presented. They are discussed in relation to the use of English DMs by the Saudi-based lecturers. In addition, the Arabic discourse markers (ADMs) will be compared to the NS English lecturers' use of DMs in BASME.
3. Lastly, a summary of the results presented in this chapter.

6.2 Interactional Discourse Markers

Interactional DMs play an important role in speaker/hearer interaction, even when largely rhetorical. These DMs function at an interactive level that invites otherwise passive listeners to interact in the discourse and play a more active role. The three functions identified as interactional are: Confirmation Checks, Elicitors and Rephrasers (see Mind-Map 6.1).



Mind-Map 6.1 Interactional Discourse Marker Functions

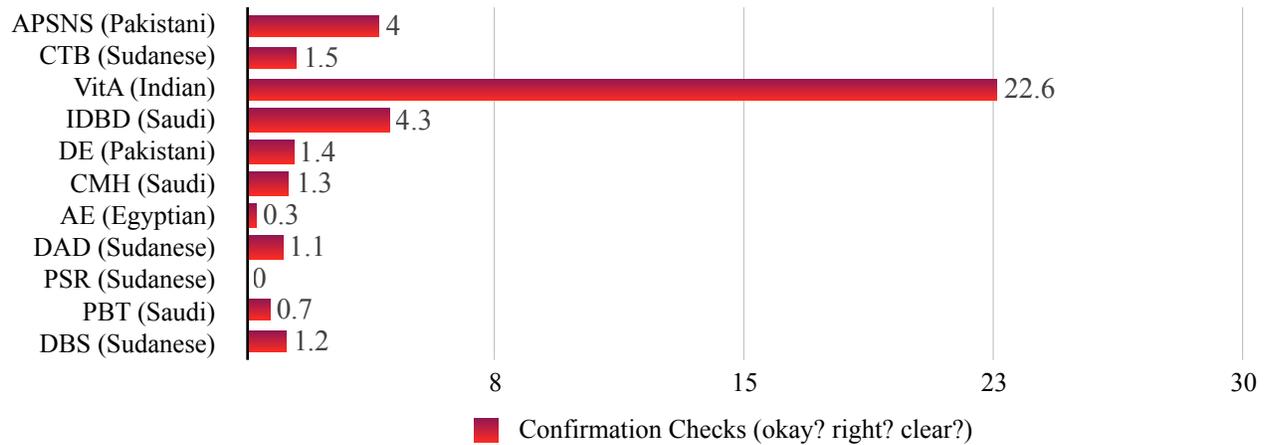
6.3 Interactional DMs in SASME

6.3.1 Confirmation Checks (*okay? right? clear?*)

Confirmation Checks are structured in a similar manner to tag questions; however, their function is, on average, not to elicit information or even to invite responses from the listeners. These types of Confirmation Checks have been described as “pseudo-turn offers” (Schleef, 2005)

as they project an image of eliciting a response from listeners, but in most cases no verbal response is required or expected. Othman (2010) reports that the DMs ‘*okay?*’, ‘*right?*’ and ‘*so?*’ suggest a sense of pre-closure and signal the speakers readiness to move on to the next point; this is similar to the DMs in this study identified as Confirmations Checks.

Figure 6.1 Use of Confirmation Check % across SASME



The overall frequency of Confirmation Check markers distributed across SASME ranged from zero in the *PBT* lecture, 0.3‰ to 1.5‰ over seven lectures, 4‰ to 4.3‰ in two others and in one lecture, *VitA*, a significantly higher 20.8‰. Further details about the Confirmation Checks distribution across the corpus will be discussed in the subsequent sections.

6.3.1.1 Confirmation Checks ‘*okay?*’

As a Confirmation Check, the DM ‘*okay?*’ indicates the end of a topic or point in the lecture and the lecturers’ need to check that the students have understood the information discussed thus far. The DM ‘*okay?*’ is often marked with a rising intonation- as denoted by the question mark and is often in the form of a (pseudo-) tag question. Its overall frequency of occurrences is 3.5‰ words, which is at least seven times higher than the other two Confirmation Checks (*right?* and *clear?*) in this group. This indicates a strong preference for the DM ‘*okay?*’ by the NNS as the marker of choice to check students’ comprehension.

‘*Okay?*’ is sometimes viewed as an elicitation request from a speaker to hearers; however, in this group its function does not act to elicit information or invite the hearer to take the floor. Rather, as a Confirmation Check, the listeners are not expected to provide a verbal response. Instead, as illustrated below in Excerpt 59 (a) and (b), lecturers tend to move on following the Confirmation Check ‘*okay?*’. Based on audio recording, the pauses following ‘*okay?*’ are equivalent to a breath between utterances and therefore, we can assume that the speakers have no intention of eliciting a response from the listeners. However, in some cases, while speakers may not expect or require a response, there seems to be an implicit agreement that, if needed, listeners are permitted to interrupt with a question or request clarification following the Comprehension Check ‘*okay?*’ as in (c).

Excerpt 59

- [VitA] (a) *The body can take, the body can handle, **okay?** (b) But if you take more than that, it can cause toxic effects on your liver, **okay?***
 (c) *Student: Can I take more than that - but you'll excrete it normally?*

It is noted that a considerably high number of DMs, mainly, ‘*okay?*’, were used in the Vitamin A lecture taught by a non-Arab speaker. The lecturer, an Indian national, had a frequency of 22.6‰ ‘*okay?*’ occurrences, which is at least six times greater than the next highest occurrence of the DM. The lack of a common language between speaker and hearers, may have driven the NNS lecturer to structure his discourse in a manner which enables him to interact with the listeners and verify comprehension throughout lecture discourse. Excerpt 60 illustrates the use of ‘*okay?*’ following consecutive utterances as in (a), (b) and (d).

Excerpt 60

- [VitA] (a) *Here we will discuss mainly the biochemistry of the vision, **okay?** (b) How the vision cycle or the visual cycle takes place in your eye, **okay?***
 (c) *The vitamin A has a very important role in your vision, **okay?***

The lecturer’s repeated use of ‘*okay?*’ (see Excerpt 60) following each utterance marks his progress check of listeners’ comprehension of the propositional ideas. As described previously,

the minor pauses (equal to a breath or heartbeat) following each comprehension check does not lend itself to a true function of eliciting a response of understanding. Yet unanimously, the lectures of SASME apply these DMs as pseudo-elicitors that may, but are not required to, evoke a response from listeners.

Excerpt 61 further illustrates the use of ‘*okay?*’ as a DM, functioning not only as a progress check in (b) and (d), but additionally, as a break in discourse. The break seems linked to the speakers’ knowledge that the discourse may be informationally dense (a) and (c) for the listeners and provides an opportunity for confirmation or interruption, if points have not been viewed as clearly stated.

Excerpt 61

[IDBD] (a) ... *the [inaudible] plaques and plaques have a core of amyloid beta and we said that there's a collection of DAU, so silver staining is to detect neuritis. (b) [Inaudible] neuritis of the abnormal cortices, okay? Abnormal axons, abnormal dendrites.*

[DE] (c) ...*you cannot use lamotrigine in absence seizure but you can use valproic acid, because it combines the whole thing; sodium, GABA, glutamate, potassium. (d) Okay? And you can see how the activation, the sodium channels are frozen; they are active and inactive.*

In some cases, the Confirmation Check ‘*okay?*’ is apparently not viewed as sufficient by the speaker and consequently, the lecturers (see Excerpt 62) follow up with a clarification (b) or a reiteration of the points (e).

Excerpt 62

[APSNS] *That one - only one motor neuron. Very clear from cell body one is this one. I'm going to say just like this one, between me and you. (a) One is here, okay? (b) Transmission here changes, then signal's going here.*

[CMH] (c) *This is the trick, be careful. Megalencephaly is large volume of brain substance. (d) It is not large skull, okay? (e) It's not large skull. Anencephaly again, forebrain anomalies.*

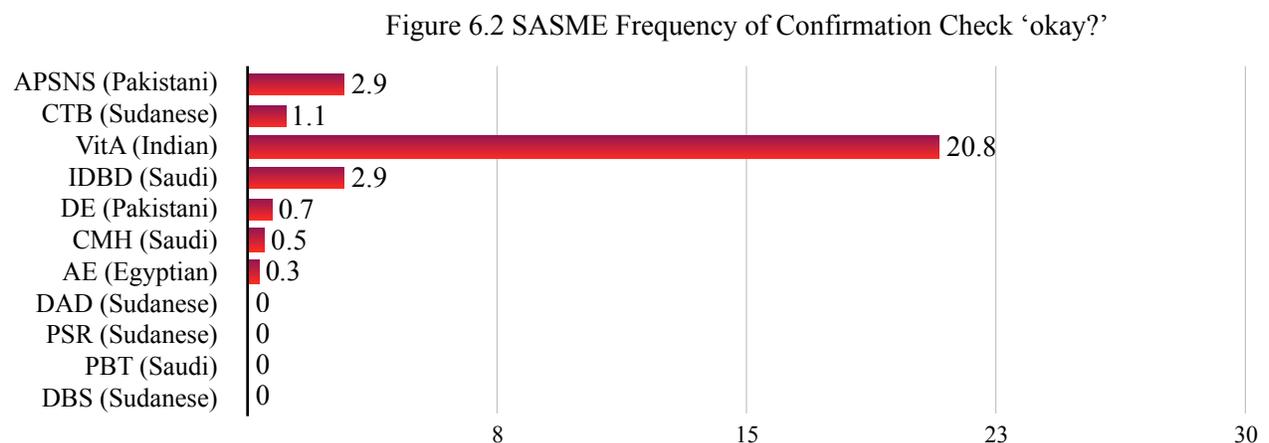
Although the function of ‘*okay?*’ as a Confirmation Check (a) is assumed to be clear to the listeners - based on the lack of follow up subsequent to its use, in one instance in the *CTB* lecture (see Excerpt 63) it is followed by a second progress check (b) “is good?”.

Excerpt 63

[CTB] *TB Brucella can be confused, one. Number two, people with TB, instead of TB they have Brucella laan yajehum al {because they get the} pharmacene, part of the chemotherapy of TB, (a) okay? (b) Is good? So, the chronic cerebral infection...*

The reiteration of a Confirmation Check in English only occurs in this lecture and in only this instance. This is a unique occurrence as other instances involve the reiteration or repetition of an English DM followed by an Arabic DM for clarity or as a stronger elicitation. In section 6.6, the frequency and context of the Arabic DMs within the SASME data and its similarity to Excerpt 63 (a) and (b), is discussed in more detail.

The DM, ‘okay?’, varied in frequency with the Arab and non-Arab lecturers. It is noted that two of the non-Arab lecturers, Pakistani (APSNS) and Indian (VitA), have the highest frequency of use with 3‰ and 20.8‰ respectively. The Indian VitA lecturer had nearly six times as many Confirmation Checks, overall, than any other lecturer. Initially, it is presumed that the discrepancy is due to the lack of a shared language with the homogenous NNS Arabic students; however, the additional data of two non-Arabs did not support this hypothesis. The substantial gap in the number of occurrences per 1000 (‰) is illustrated in Figure 6.2.



Based on the results, it can be surmised that the high number of Confirmation Checks used by the Indian lecturer is a matter of personal preference and a heightened need to verify students’ progress throughout the lecture. Similarly, personal preference seems to play a role in four of the

eleven lectures taught by native Arabic speakers which have no occurrences of the English DM ‘*okay?*’. Results show us that these lecturers prefer to use equivalent DMs in Arabic, rather than English. There are a considerable number of occurrences of Arabic DMs (e.g., *suh?* {right}) which function in a similar manner to English DMs (e.g., *okay?*, *right?*). This is illustrated in Excerpt 64.

Excerpt 64 [Arabic DM]

[PSR] (a) ... *time misinformation and information is money. (b) Information is power; suh?* {right?} *information is power...*

Excerpt 64 exemplifies the use of the Arabic word ‘*suh*’, which translates to ‘*right*’ or ‘*correct*’ within this context; the lecturer uses the DM to confirm students are following his point. The functions of this and the other Arabic DMs within the scope of this research are discussed in more detail in section 6.6.

6.3.1.2 Confirmation Checks ‘*right?*’

The second DM with the function of Confirmation Check is ‘*right?*’ which had the second highest number of occurrences at 0.5‰, in contrast to ‘*okay?*’ (3.5‰) and ‘*clear?*’ (0.4‰). It functions in the same manner as ‘*okay?*’, though it is preferred less in contrast to both the English DM ‘*okay?*’ and the equivalent DM in Arabic (see 6.6.3).

The meaning of the DM ‘*right?*’ can be interchangeable in some contexts, as in Excerpt 65, where the meaning is not ‘*right?*’ as in *correct* or a verification of facts. In context, like (a), the speaker is confirming that the propositional points of transmitters, channels and the resulting propagation of signals is clear to the listeners. The DM can, in context, be replaced with ‘*okay?*’.

Excerpt 65

[DE] ...*sodium and potassium channels, because these transmitters are basically working through channels, opening sodium channels, because the opening of these channels can propagate the signal. (a) Right? So that's why you need to understand that there is a balance-*

Excerpt 65 (*continued*)

[DE] (b) *Severe excitation, **right?*** (c) *And there is a tonic, **right?** This is a tonic and this is a clonic. You can see contraction, relaxed. Contract, relaxed and it goes like this and finally ...*

In the second example (see Excerpt 65 *DE*), the speaker concludes each utterance (b) and (c) with ‘*right?*’, which functions similarly to a pre-closer, signaling the proposition is drawing to an end and also functioning as a Confirmation Check. Similarly, the DMs (b) and (c) act as pseudo-Confirmation Checks like ‘*okay?*’ and could replace the DMs ‘*right?*’ without affecting the meaning or function.

In the next example (see Excerpt 66), the DM ‘*right?*’ functions as a Confirmation Check rather than as a pseudo-checker. Like most occurrences of ‘*okay?*’, in this context (a), also functions as an elicitor of agreement. Congenital malformations is a huge chapter in the students’ textbook and the lecturer seeks to confirm that students agree with his statement by using ‘*right?*’. It could be replaced with ‘*okay?*’ but I believe some subtle meaning would be lost in the exchange.

Excerpt 66

[CMH] *Also the congenital malformations of hydrocephalus. Congenital malformations is a huge chapter. (a) **Right?** So there's actually books called Pediatric Neuropathology book, huge book.*

The *VitA* lecturer, as previously mentioned, had the highest number of overall occurrences in confirmation markers, most notably with ‘*okay?*’ (20.8%) and ‘*right?*’ (1.7%). In Excerpt 67, the first transcript sample illustrates the lecturer’s use of both ‘*okay?*’ and ‘*right?*’ in all four utterances, alternating from one DM to the other. Looking closer at the extract, the lecturer seems to choose to use ‘*right?*’ in (a) and (c) following the first mention of the proposition about light containing information and reflecting it. In the subsequent reiteration, he switches to use ‘*okay?*’ (b) and it could be concluded that in the first instance (a) the speaker is eliciting confirmation that the listeners agree with and understand this first point, before reiterating it in (b) and (d).

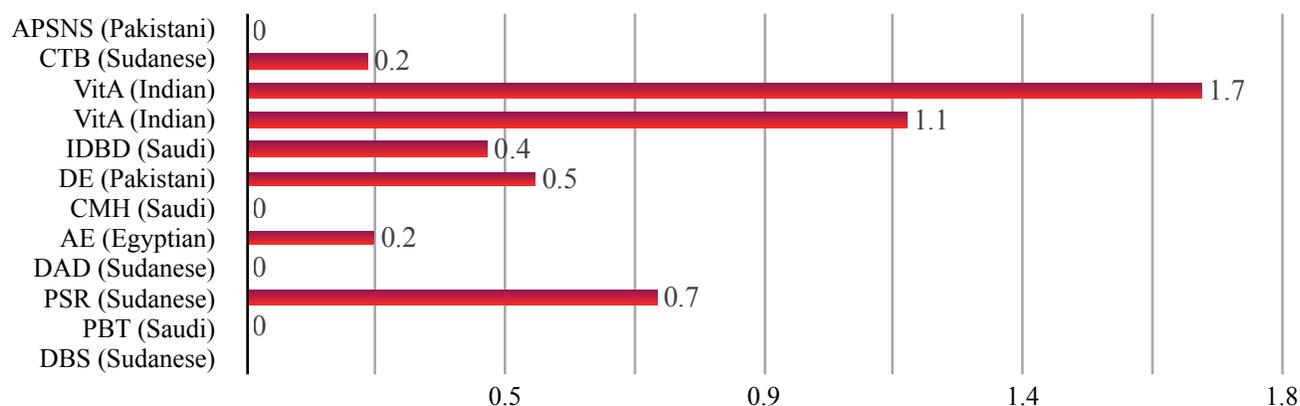
Excerpt 67

[VitA] ... okay? (a) So, similarly this light has information, **right?** (b) The light [inaudible] has information, **okay?** (c) I can see you because the light around you is reflecting on my eyes, **right?** so that I can see your face. (d) So this light has information, **okay?**

[VitA] (e) Okay, but you've gone through it **right?** You can take any CNE anatomy book and see where it's conjunctiva, it will tell you. (f) **Yes?** (g) Student: Are these diseases corrected only by giving vitamin A or sometimes the damage is irreversible?

In the next extract (see Excerpt 67e), the *VitA* lecturer implements the DM ‘*right?*’, similar to ‘*okay?*’ as a pseudo-Confirmation Check, with no expectation of eliciting a response from the listeners. However, following the subsequent utterance, the speaker opts to use ‘*yes?*’ as opposed to the DMs he has previously favored. What prompted the change of DM cannot be confirmed due to lack of video recording, though based on the audio recording we could speculate that a movement or facial expression of the audience instigated the change. The lecturer concludes his utterance (f) as a distinct break in discourse to elicit questions or concerns from the listeners as in (g).

Figure 6.3 SASME Frequency ‰ of Confirmation Check ‘*right?*’



Overall, the Confirmation Check ‘*right?*’ is not the preferred method of eliciting confirmation, nor is the DM ‘*clear?*’, which has an occurrence only 0.1‰ less than ‘*right?*’. Neither DM challenged the preferred DM ‘*okay?*’ which has an overall occurrence about seven times higher than either of them. There were no occurrences of ‘*right?*’ in four of the SASME lectures, three of

which are by native Arabic speakers. The highest number of occurrences are with the *VitA* lecturer, previously discussed, with 1.7% and a Saudi lecturer of the IDBD lecture with 1.1%.

6.3.1.3 Confirmation Check ‘*clear?*’

The last DM identified as a Confirmation Check is ‘*clear?*’, which is the least preferred marker by NNS lecturers, though as previously stated, by a mere fraction of a point from the DM ‘*right?*’. The DM functioned similarly to the other confirmation markers ‘*okay?*’ and ‘*right?*’ though it seems to have a subtle difference in meaning, in that it seems more explicit in its role as a Confirmation Check. That is to say, the DM ‘*clear?*’ seems explicit in its meaning to the listener that the speaker is seeking to clarify whether the propositional ideas are clearly understood or whether clarification may be needed. This is highlighted in the second extract (68b), in which the Confirmation Check ‘*clear?*’ follows a direct elicitation for confirmation by the speaker “Do you have any questions about any part of the lecture?”.

Excerpt 68

[APSNS] ...or it is inhibited. There are inceptors. They are accepting that neurotransmitter, yes or no. (a) **Clear?** It's specific receptors.

[PDDB] That will show them [inaudible] (b) Do you have any questions about any part of the lecture? (b) **Clear?**

The utterance 43(a) functions to indirectly elicit questions from the listeners and (b) seem to stress that if the listeners have no questions, then to confirm whether the propositional points are clear or not. In one case, the lecturer of *CTB* phrases the Confirmation Check as a more explicit elicitation “Are you all clear?” (see Excerpt 69a). This type of Confirmation Check, with a propositional meaning similar to *do you understand?* and *is it clear?*, also occurred as Arabic DMs with native Arabic lecturers such as, the *CTB* speaker (e.g. ‘*wadhih*’ {clear}). This will be discussed in more detail in the next section (see 6.6.3).

Excerpt 69

[CTB] *You have got the patients who have got [inaudible]. The patient is ill, but you cannot grow the organism...and I put them by their importance. (a) Are you all clear?*

In most instances, the DM appears as part of a question structure (e.g., MWU) and not as a single unit, as in Excerpt 69 (a). This did not take away from its propositional meaning or function, but in other instances in SASME, the DM forms a pattern of BE + PRONOUN + *clear* (see Excerpt 70a-c).

Excerpt 70

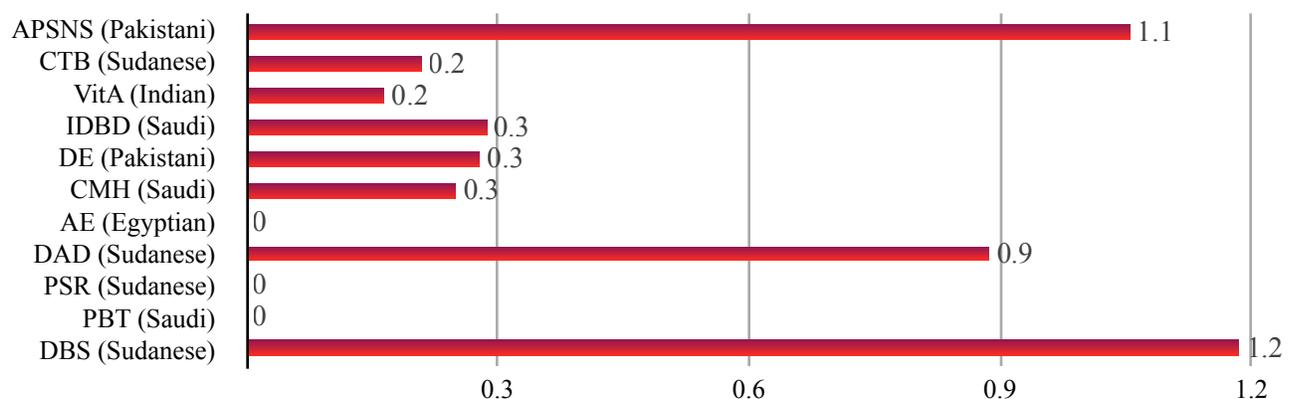
[DBS] *So, they cause vasodilatation in the middle ear and this will relieve the pressure in the inner ear. (a) Is that clear? These three classes are used for prophylaxis. The postulated mechanics could be as follows.*

[DAPD] *The monoamines with the dopamine, serotonin or noradrenaline, they are all catabolised by MAO enzymes. So, depression is treated by elevating the level of monoamines, (b) is it clear?*

[DAD] *Buspirone is preferred to benzodiazepine, the benzodiazepines in old people. (c) Is it clear? Buspirone is preferred to benzodiazepines in old people.*

The DM ‘*clear?*’ is distributed across the corpus with a comparably higher use in three of the lectures - *APSNS*, *DAD* and *DBS* - the latter having the highest number of occurrences overall. With the remaining lectures, three have 0‰ occurrences and five have a relatively similar number of occurrences ranging from 0.2‰ to 0.3‰ (see Figure 6.4).

Figure 6.4 SASME Frequency of Confirmation Check ‘clear?’



6.3.2 Elicitors (*So-prefaced question; Any questions?*)

The second function, under the category of Interactional DMs, is Elicitors and this group includes two sets of multi-word units (MWUs) which function to elicit a response from listeners: ‘*so-prefacing question*’ and ‘*any questions*’.

Although the main function of elicitation markers is to invite listeners to participate in the lecture discourse, the elicitations are not always successful and “sometimes there is no answer producing default interaction” (Bellés-Fortuño, 2006: 268). In SASME, with the MWU ‘*so-prefacing question*’, the default interaction, following no immediate response or even pre-empting the possibility, is the rephrasing, restructuring or vocalization of more explicit and detailed questions (see Excerpt 71 c-f). In select cases, the rephrasing, restructuring or repetition of questions are in the L1 - Arabic.

In contrast to the Confirmation Checks (see Section 6.3.1), the function of Elicitors is to draw out information from the listeners and invite them to take the floor. In English, this category is the second least popular function in SASME, with more occurrences overall in Arabic with NS Arabic lecturers (see 6.6.3).

Elicitors have a total frequency of 0.09‰ occurrences which is the second lowest English Category in SASME. This is not unexpected as the lectures are monologic in style and thus the floor is held by one single speaker, the lecturer and listeners are not generally encouraged to interrupt or intervene in discourse unless invited. Within this corpus, the Elicitation MWUs in English, though relatively few, are occasions when listeners are invited to take the floor.

6.3.2.1 Elicitors ‘*So-preface question*’

The Elicitor ‘*so-preface question*’ functions at an Interactional level and marks the speech act of questioning listeners’ understanding of the discourse. It functions at an Interactional level as it does not primarily act to structure the discourse at a textual level. The Elicitor works to organize and elicit the participation of the listeners - who are ideally active participants in the lecture (Müller, 2005). The *so-prefaced* construction units are recipient (listener) attentive, rather than

self-attentive and the ‘*so-preface question*’ speech acts are related to the overall topic of the discourse and function to elicit information from the listeners, as well as highlight, in most cases, the speakers' interest and concern for the listeners, i.e., students (Bolden, 2006).

In the first Excerpt (71), the DM ‘*so*’ prefaces question sequences which function to elicit a response from listeners, as in utterances (a), (c) and (c). Each sequence has the pattern of *SO* + *WHAT* + *VERB* + *CLAUSE* (see Excerpt 71) and in each of the elicitations (a-c) they are immediately followed by another question as either a repetition (b), elaboration with examples (d and g) or as a clarification of an immediately previous question (f). The occurrence of repeated questions is similar across lecturers in SASME and can be viewed on the basis that lectures are in a foreign EMI and listeners may need questions and intent of elicitation clarified.

Excerpt 71

[DBS] (a) ***So***, *what occurs in the vestibular apparatus?* (b) *What occurs?*

[DE] (c) ***So***, *what else have you observed?* (d) *Any salivation or facial [inaudible] movements, have you seen that? Lip-smacking or rolling of eyes; like shaking his legs everything else?* (e) ***So***, *what do you think;* (f) *what is the abnormality in epilepsy?* (g) *Why is this patient having uncontrolled shaking of hands? Any mechanism? Yes?*

One type of ‘*so-preface question*’ elicitation used in the lectures functioned as rhetorical, meaning that the speaker did not expect or require a response from the listeners. Chang (2012: 106) classified this type of elicitation as a "class management/engagement" question, which is designed not to elicit a response like other questions, but rather to "enhance the teaching flow" which can "build rapport or increase the level of interaction" with the students. Excerpt 72 illustrates this type of *enhancement* in each speaker’s ‘*so-prefacing questions*’ (a) and (b), which seems designed to capture the listeners’ attention and engage with them on a rhetorical level.

Excerpt 72

[PDDBD] *They are trying to understand what causes the disease. They look at the risk factors and how we can fix it. (a) So*, *what causes dementia?* *Dementia is a topic that is not specific, although it's commonly associated or it's a main feature for Alzheimer's disease.*

[VitA] (b) ***So***, *what happens?* *there's no more film. It used to be films, right, before. No more films. So, that's the film.*

There are select occurrences of ‘*so-prefacing question*’ which functioned not only as Elicitors but additionally as indirect chiding of students’ actions. In this extract (see Excerpt 73), the lecturer is annoyed with students walking into his lecture late and reminds them (a) that lectures are only meant to be 50 minutes and they, as students, have the right to a 10 minute break to complete the hour. The lecturer uses the ‘*so-prefacing question*’ structure to build upon his previous utterance that they have rights and must demand them (b). The function of (b) seems intended to elicit agreement from the students, as well as place responsibility for their tardiness on their shoulders and not their previous lecturer.

Excerpt 73

[PSR] (a) *Am I correct wala {or} - you have told - you have been told this. (b) So, why do you leave your rights? okay. Because your brain - you will finish and know that it's...*

Johnson’s (2002) research into language in the legal process found that police interviews use a statistically significant number of ‘*so-preface questions*’ and suggests that it is likely that it also extends into other types of interviews. In police interviews, it is found that the ‘*so-preface question*’ is exclusively found in the domain of the interviewer and were a particular feature of child interviews. Johnson (2002: 100) believes the interviewer compensates for the child’s limited ability to produce complex and extensive narratives by making the ‘*so-preface questions*’ “highly content and substance-concentrated”. By using this type of question style, the interviewer supports the interviewee and allows them to “agree with, adapt or expand the interviewer's utterance” (ibid: 100).

Johnson’s (2002) observations could be applied to SASME on the presumption that, while the students are not children nor is there any doubt about their intelligence, their challenges with English as a foreign EMI may be comparable. Johnson (2002) suggests ‘*so-preface questions*’ are highly concentrated in content and substance in order for listeners to more easily formulate responses. As the language of instruction is not the L1 of the lecturers or the students, lecturers may consciously or subconsciously feel a need to formulate questions as clearly as possible.

The first extract (Excerpt 74) illustrates the lecturer's choice to phrase his '*so-preface question*' as "So, what's happening in the picture?" (a) rather than 'So, what's happening here?'. Moreover, the lecturer chooses to follow up his question with a second that clarifies the answer he is attempting to elicit and instead provides the answer and concludes with a Confirmation Check, '*right?*'.

Excerpt 74

[CMH] (a) **So**, what's happening in the picture? The light is going through his skull, easy, right? This is the trick, be careful. Megalencephaly is large volume of brain substance.

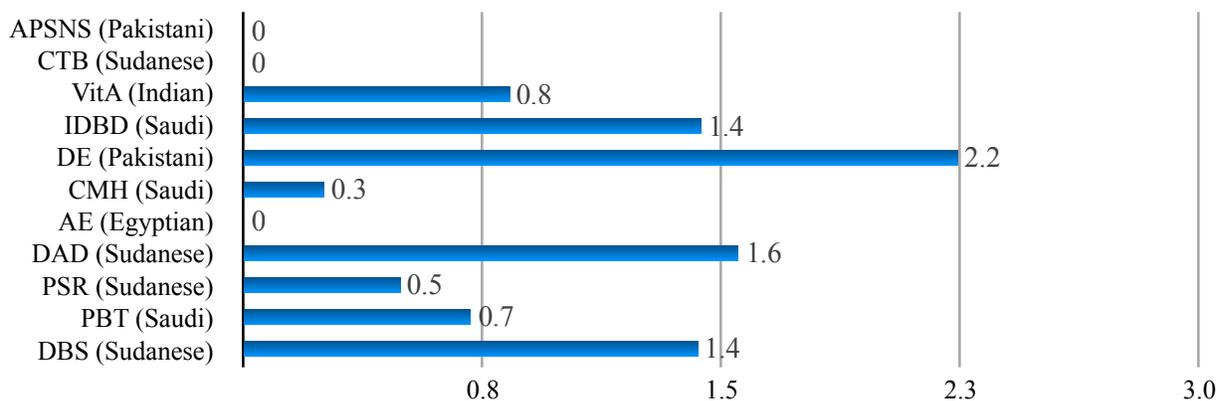
[DAPD] Next are the beta blockers. (b) **So**, what is the rationale of using beta blockers in case of anxiety? [student(s) respond]

[DE] Excitation. Very good. (c) **So**, what is the excitation? What else transmitters? [student(s) respond]

[PBT] We have a tumor, malignant tumor; we try to stage and grade, right, (d) **so**, what is the difference between staging and grading? Are the basic foundation blocks.

The next two extracts (b-c) illustrate '*so-preface questions*' eliciting a response from listeners. The structure, described by Johnson (2002) as providing high content, can be seen in the Excerpt 74d. The lecturer repeats the tokens *stage* and *grade* that were uttered immediately prior to the '*so-preface question*' to make the purpose of the Elicitation clear to the students.

Figure 6.5 Frequency % of SASME Elicitation So-Preface Question



The MWU ‘*so-prefacing question*’ occurs overall in SASME 0.09‰ in English and is a pattern that does occur frequently in Arabic. The Elicitation ‘*so-prefacing question*’ occurs in eight of the eleven lectures ranging from 0.3‰ to 2.2‰ as illustrated in Figure 6.5.

6.3.2.2 Elicitors ‘*any questions?*’

The second MWU functioning as an Elicitor is ‘*any questions*’ which can be viewed as one of the most direct and clearly stated DMs under either category. Although ‘*any questions*’ is not used at a significant level, relative to the other DMs within the Interactional category, it is noted during analysis and is used in six of the eleven lectures. Considering the low frequency per thousand (‰) occurrences, repeated analyses focused on investigating how lecturers elicit information from the student listeners. One other Elicitation MWU noted in English is ‘*You understand?*’ (0.1‰); however, this phrase has a lower frequency than the Elicitor DMs included in this study and more importantly, for the comparability of the research, this MWU did not appear in BASME. Moreover, the results indicate that in respect to SASME, most Elicitors of native Arabic lecturers are in Arabic (see to 6.6.3).

The Elicitor ‘*any questions*’ is successful in the majority of occurrences in eliciting a response from students, though not immediately in some occurrences (see Excerpt 75). In Excerpt 75, the Elicitors are successful at immediately drawing out a response from students following ‘*any questions*’ utterances (a-b).

Excerpt 75

[APSNS] *Second thing is myelinated and this is unmyelinated. Unmyelinated means conduction is slow. (a) Any questions from your side? No?*

[PBT] *...pilocytic with circumscribed hair like processes, Rosenthal fibres, eosinophilic granular bodies, [inaudible] pilocytic astrocytes, okay. (b) Do you have any questions?*

In Excerpt 76, though these examples of Elicitors (a-b) are successful in drawing out questions from students, the responses are delayed as the lecturers do not provide an adequate pause

following either occurrence (a-b). In (a) the lecturer explains to students that he “tried really to go very basic, very easy” which he may view as offering students an opportunity to refute such a claim. No response is forthcoming until he utters “Do *we* have anything?” which does finally prompt a response from students. In the second example, though students do not respond to the initial Elicitor (76b) they are heard on the audio tape in the background murmuring in agreement. The *PDBD* example also illustrates the CS from English to Arabic, which occurs during the NS Arabic lecturers’ discourse.

Excerpt 76

[CMH] *There is no real treatment for such disease. So, that's it, I guess. (a) Do you have **any questions** for congenital malformations? I tried really to go very basic, very easy, few terms and few examples...So is it clear? Do we have anything?*

[PDBD] *... hydrocephalus was not hydrocephalus. This inta 'aj fe {results in} atrophy of the brain. (b) **Any other questions?** izan al {so the} plan wathha? {is clear?} Dementia ash hiya al- {what is the-} dementia ...*

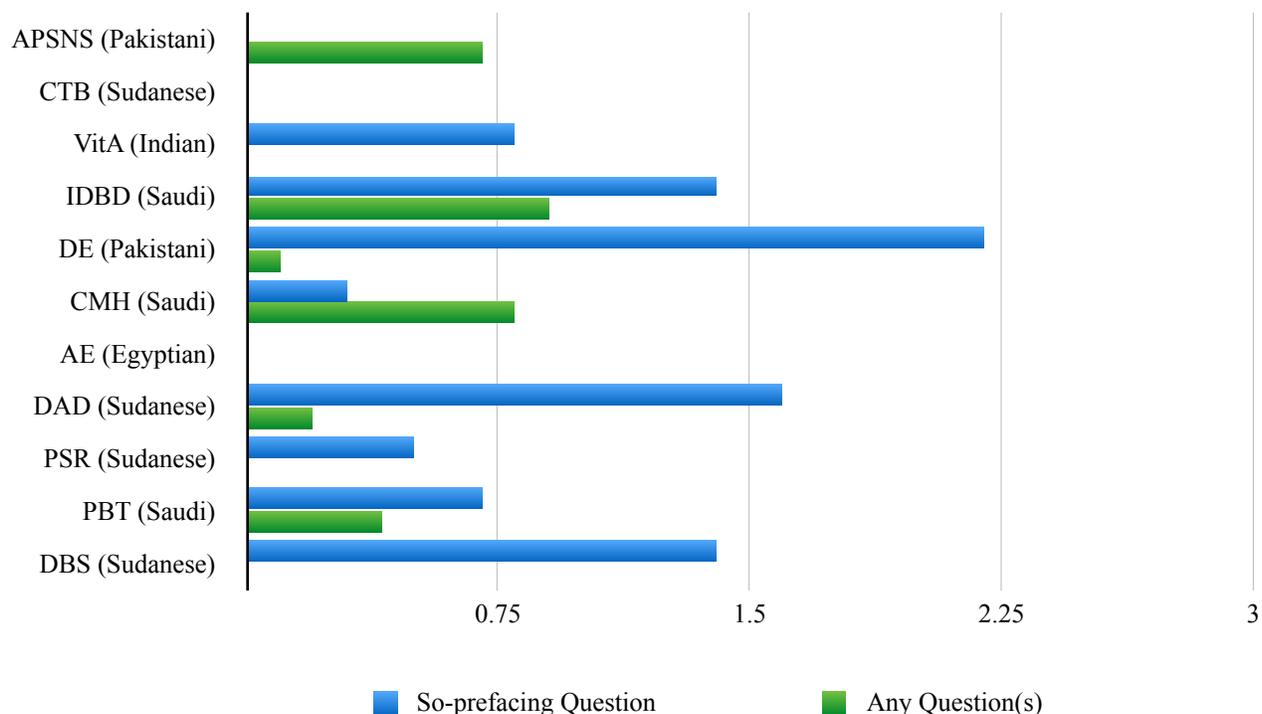
The last extract (see Excerpt 77) illustrates the function of ‘*any questions*’ as an Elicitation phrased as an open invitation to listeners that welcomes them to ask questions, even those beyond the limitations of the classroom.

Excerpt 77

[VitA] *That was my last lecture for you in (CNS). I will come back in [inaudible]. (a) If you have **any questions** you know my office. You can come anytime.*

This occurrence, (77a), is included in the Elicitation category as it functions in generally the same manner as other incidences, with the exception of expected outcome (i.e. immediate response) which is delayed by *if-clause* and token “anytime”.

Figure 6.6 Frequency of SASME Category Elicitors



The Elicitation function (see Figure 6.6) is the second highest in overall occurrences (1.2%) following the Confirmation Checks function (4.4%).

6.3.3 Rephrasers (*I mean; in other words*)

The third and final function, under Interactional DMs, are Rephrasers which are used by lecturers to rephrase information that may be too dense, complex or unclear for students. The Rephraser DMs are usually used with the first person pronoun (*I, we*), though there are some occurrences of second person (*you*) and third person (*it*) pronoun use.

While the DMs under this function do occur 0.3% in English, they are used more frequently by the NS Arabic lecturers in the L1 (e.g., ‘*ya3ni*’, translated as ‘*I mean*’ or, ‘*that is to say*’).

Excerpt 78

[PSR] *Because the university will tell me to start in time. So, we have - if we had any problem - (a) ya3ni {I mean} it is a lecturer problem. It is finish ten minutes before...*

Excerpt 78 illustrates the use of the Rephraser in Arabic, ‘*ya3ni*’ {*I mean*}, by the lecturer to clarify that the issue he’s discussing is not a student problem but rather one with lecturers (a).

6.3.3.1 Rephraser ‘*I mean*’

The MWU ‘*I mean*’ functions as a Rephraser which acts to modify the speakers ideas or intentions (Schiffrin, 1987: 267) and maintains the speaker’s focus on prior material (Schiffrin, 1987: 297). It also signals to listeners that previous utterance(s) will be reiterated or rephrased, in some cases with additional examples or details. The MWU ‘*I mean*’ is also a representation of the lemma or headword for the set of MWUs which occur in the corpus (e.g. *I mean*, *we mean* and *it means*). Similar to Fuller’s (2003) findings, the MWU ‘*I mean*’ is often positioned utterance-medial in discourse as illustrated in the excerpts and rephrased information uttered immediately prior to the lemma, ‘*I mean*’.

In the first example (see Excerpt 79) the lecturer of the Cerebral Tuberculosis (*CTB*) lecture uses the Rephraser DM to make certain that listeners understand his reference to “partially treated meningitis”. Without challenging the students' knowledge, the lecturer rephrases his previous expression (a) by explicitly stating what the phrase is referring to (b).

Excerpt 79

[*CTB*] (a) *So, partially treated meningitis, (b) I mean someone with chronic meningeal [inaudible], please think first. Depending on their age, depending on knowing whether he has been having - the meningitis has been partially treated or not.*

In the first example, Excerpt 80(b), the lecturer rephrases his point to make it clear that "that much" (a) refers back to spending time in the gym. This illustrates the function of the Rephraser to rephrase prior utterance(s) such as, the listing of examples “doing sports, maybe football”, which may have confused students of the main point the lecturer is attempting to relay. The second example also rephrases the immediately prior utterance (c) to clarify what - "neurological science" and who - “neurologists”, diagnose two seizures as epilepsy (d).

Excerpt 80

[DE] *Try to devote some time for exercise, like doing sports, maybe football or just walking or jogging or gym [inaudible] (a) that much, (b) **I mean** have that much time for gyms. Because you have to commit for lifelong,*

[DE] *(c) If he experiences one seizure episode, no episode at all in two years, he is not epileptic, but he appeared - this is what (d) **I mean** neurological science says and the neurologists diagnose as epilepsy - two seizures in six to 12 months.*

In most cases, the Rephraser function is used in a positive sense for '*I mean*' rather than the negative '*I don't mean*' or in the following extract, "it doesn't mean" (see Excerpt 81). In the few incidences when the multi-word unit is used in a negative sense, it functioned to rephrase and clarify information that the speaker believes the students will potentially misunderstand. Excerpt 81 illustrates the Rephraser function (c) to clarify to students that the term alcohol in utterance (a) and (b) is not referring to the social drink but to the natural chemical compound.

Excerpt 81

[VitA] *(a) So, retinol, as the name suggests also is an alcoholic form, okay? There's an alcohol with this vitamin A. (b) When I say alcohol (c) **it doesn't mean** beer or wine, it's a chemistry it's an alcoholic retinol, CH₂OH, okay?*

The clarification, made by a non-Arab lecturer, may be in light of the cultural and religious sensitivities in the country and to verify that students have correctly understood the reference to alcohol in the context of Vitamin A.

6.3.3.2 Rephraser 'In other words'

The second MWU under the category of Rephraser is '*in other words*' which is similar to '*I mean*' and functions as a signal that prior information will be rephrased. The Rephraser simplifies the information by making the proposition more explicit or by describing it with more details and examples for the benefit of listeners.

Considering the density of the information lecturers are presenting to the students, it is not surprising to find Rephrasers, especially those which by their pattern explicitly illustrates to listeners that the following information is being repeated in, ideally, a more comprehensible manner. This MWU is not found to a significant degree in the corpus as an English DM. Excerpt 82 illustrates this Rephraser in (b), which signals to listeners that the vital information described thus far (a) are all related to the "brainstem".

Excerpt 82

[APSNS] (a) *This part consists of bones, medulla oblongata and the cerebral. Mainly in physiology you will - it means going to read - (b) **in other words**, there is a brainstem. The brainstem, the brain is there...*

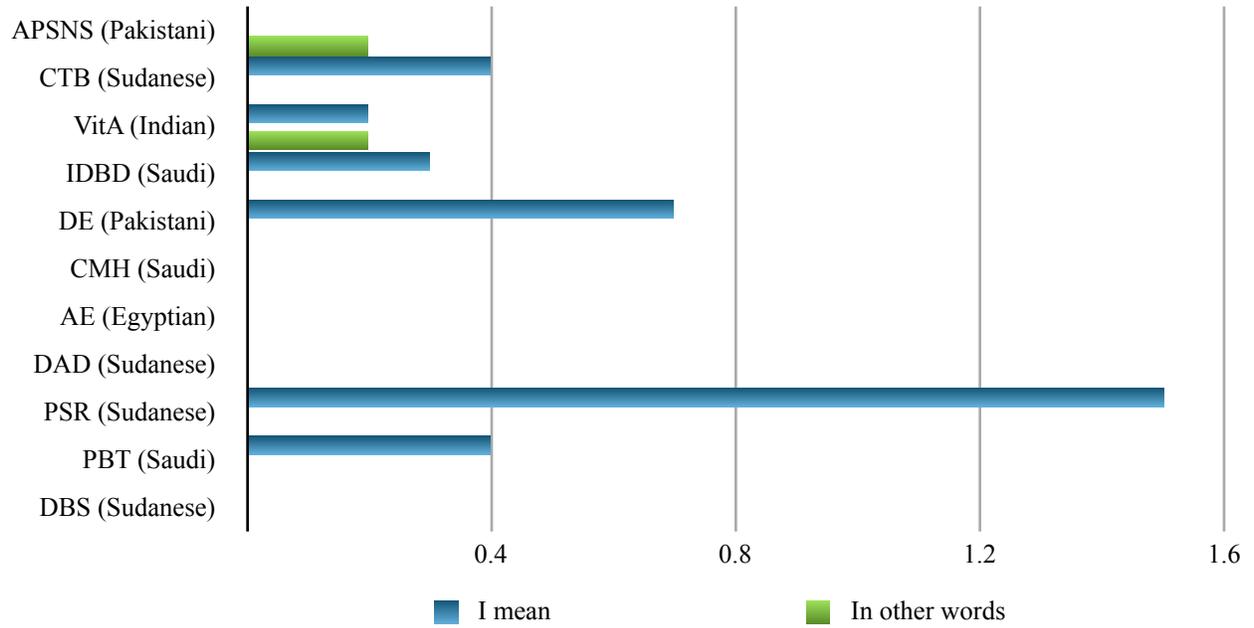
The non-Arab lecturer of *VitA*, provides a series of Rephrasers (see Excerpt 83) for the students as he seems to suspect that they might not understand the meaning of "confirmation" (83a). He attempts to redefine the token by describing it as "chang[ing] structure" (b) and then in a third attempt describes it as a change in shape (b). Lastly, the lecturer uses the DM 'okay?' to check whether the students have understood that "opsin will change its confirmation". The repeated use of re-structuring and rephrasing is expected with this lecturer as he illustrated the highest frequencies of DM use to interact and verify information with students.

Excerpt 83

[VitA] *...this becomes photosensitive. (a) So, opsin will change its confirmation. Confirmation means it will change its structure. (b) It will change its shape **in other words**, (c) okay?*

Overall, Rephrasers in English are not frequent in SASME, with no occurrences in four of the eleven lectures. Their total occurrences are the lowest at 0.3‰, in contract to Confirmation Checks (4.4‰) and Elicitors (0.9‰) and as all three non-Arab lecturers have occurrences of one or both DMs, we can speculate that Arab lecturers preferred to use Rephrasers in the L1 rather than in the L2.

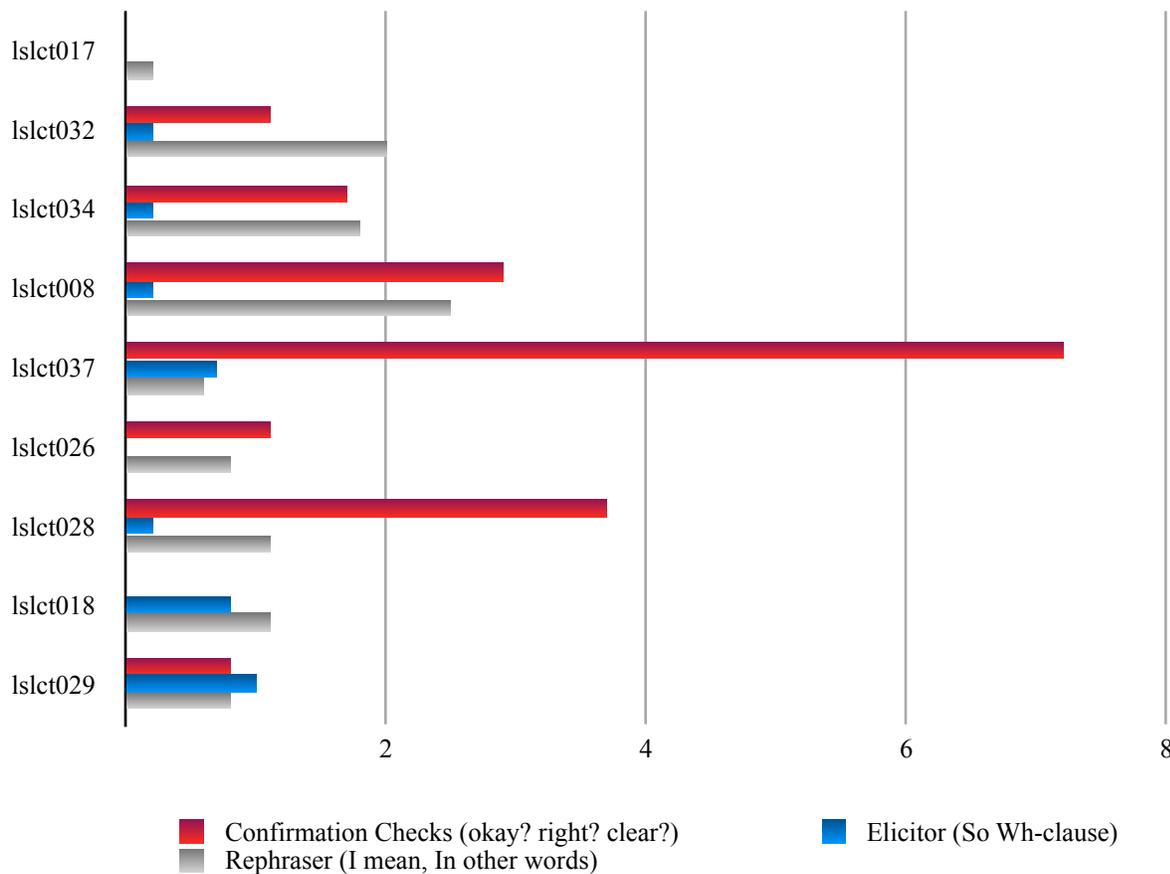
Figure 6.7 Frequency % of SASME Category Rephrasers



6.4 Interactional Discourse Markers in BASME

The next section presents the data results of the BASME sub-corpus analysis and the three functions that fall under the Interactional DM category. As illustrated in Figure 6.8, the Interactional DMs are used throughout BASME and have more occurrences overall with 2.2%, than in SASME with 0.6%.

Figure 6.8 BASME Lecturers' Interactive DM Use %

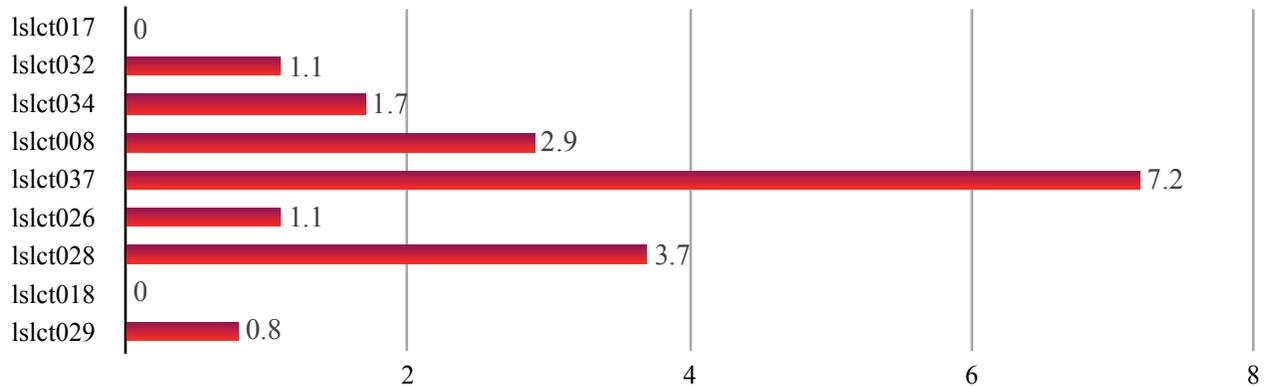


6.4.1 Confirmation Checks (*okay? right? clear?*)

The Confirmation Check functions in BASME are structured in a similar manner to SASME, with each DM ending in a rising intonation denoting its function as a question. The DMs function as pseudo-elicitors, which are not designed to elicit information or invite listeners to take the

floor. Instead, Confirmation Checks are moves by the speakers (lecturers) to seek confirmation of the listeners (students) understanding of the previous utterance(s).

Figure 6.9 Confirmation Check Category Use % across BASME Lecturers



The overall frequency of Confirmation Checks distributed across BASME ranged from zero occurrences in two lectures and 0.8% to 1.7% in four lecturers, 2.9% and 3.7% in two lectures (*lslect008*, *lslect028* respectively) and 7.2% in one lecture (*lslect037*). Details of the use of Confirmation Checks across BASME, in contrast to SASME, will be discussed in the subsequent sections.

6.4.1.1 Confirmation Checks ‘okay?’

The Confirmation Check ‘okay?’ has the highest number of occurrences within this function at 1.8%. It is the preferred DM with BASME lecturers and, when used in discourse, functions as a Confirmation Check which can be identified as such through context. In Excerpt 84, note that there is a similar pattern in SASME of ‘okay?’ + ‘so’ co-occurring in discourse. This co-occurrence is found to occur 0.5% in BASME and 0.9% in SASME. In Excerpt 84, ‘okay?’ is illustrated as located in final-utterance and following a dense chunk of information as in (a) and (c). Both are followed by a new utterance initiated with ‘so’ which functions as a Summarizer in (b) and a Topic Initiator in (d).

Extract 84

- [lslct029] (a) ...*this part is permeable to water er this part isn't permeable to water but has sodium pumps okay?* (b) *So this is what would happen if nothing happened...*
- [lslct034] (c) *they they are coming you will have them at the end and maybe it's a good idea that you pay a bit of attention okay?* (d) *So who said i've got Bright's disease and he's got mine*

Another pattern identified as co-occurring is ‘*okay?*’ + ‘*now*’, which occurs twice in the BASME sub-corpus with 0.03‰ and nine times in SASME with 0.2‰. This co-occurrence is also identified in Bellés-Fortuño’s (2006: 239) study of social science lectures, with 19 occurrences overall (0.1‰) in the NS corpus, mostly functioning as topic shifters. In Excerpt 85, the DM ‘*okay?*’ functions as checks of confirmation in utterance-final positions which are summing up information prior to coming to a close with the DM as in (a) and (d).

Excerpt 85

- [lslct008] (a) ...*move on then from a description of the disease to say something about the current situation globally okay?* (b) **Now** *remember i've handed you out all these handouts*
- [lslct026] (c) ...*and again it's a bit less in females and it decreases with age as you get older okay?* (d) **Now** *your G-F-R actually remains constant your blood pressure goes up and down dur-, out the day goes up and down*

Although the Confirmation Checks are not designed to elicit information from students, they do often act in concert with questions that follow the DM. In Excerpt 86, the utterance (a) is initiated with ‘*so*’ which functions as a Summarizer and is concluded with ‘*okay?*’ to confirm that listeners have understood the main points thus far (b). The second example, too, illustrates a successful elicitation to a question which immediately follows the confirmation marker (c).

Excerpt 86

- [lslct037] (a) **So** *they have the greatest chance at this point of passing on the infection* (b) **okay?** *As far as seroconversion is concerned okay What is the first antibody you would expect to detect after a seroconversion?*

Excerpt 86 (continued)

[lslct028] (c) *Pain is the hallmark of kidney stone disease **okay**?* (d) *Tell me about the pain of kidney stone disease?*

In both cases, there is a response from students suggesting that when Confirmation Checks are used in BASME, in general they are used in preparation for elicitation questions.

6.4.1.2 Confirmation Checks ‘right?’

There are few instances of *right?* as a Confirmation Check in BASME with 0.06‰ occurrences, in contrast to SASME with 0.4‰. Neither corpus has a high occurrence of ‘*right?*’; however, in both, the DM ‘*okay?*’ has the highest number of occurrences.

Excerpt 87

[lslct037] (a) *you're not likely to go out of this period of clinical latency for fifty years You're not going to worry about it too much **right?*** (b) *'cause you're probably going to be killed by something else before you get to seventy*

[lslct034] (c) *What I suggest you do is don't read the back of that okay? er 'cause it's it's all about diabetes* (d) *Is that **right?** Yeah right It's all about diabetes and ...*

In the few occurrences of ‘*right?*’, the DM is used to function more as rapport than to confirm information. In (a) the speaker describes the “period of clinical latency” and then uses humor to bond with students and tells them not to worry about it because they’ll probably be dead long before the disease could kill them (b). Similarly, utterance (c) uses ‘*right?*’ to confirm information (d) and, again, as a means of rapport after a mistake with the handouts is discovered.

6.4.1.3 Confirmation Checks ‘clear?’

The least used Confirmation Check marker in BASME is ‘*clear?*’ which occurs at 0.04‰, though in each instance, it holds the pattern “Is that clear?”. In SASME, the total number of

occurrences are 0.5‰, not a high number of occurrences but, in contrast, used more frequently in the NNS corpus than in the NS corpus.

Excerpt 88

[lslct029] *which will make this more concentrated (a) Is that clear? which is what happens here So then you end up with sort of one cycle of urine*

[lslct029] *because they're still not drinking the A-D-H will cause them to retain water okay (b) Is that clear on my second time round? okay so this is ...*

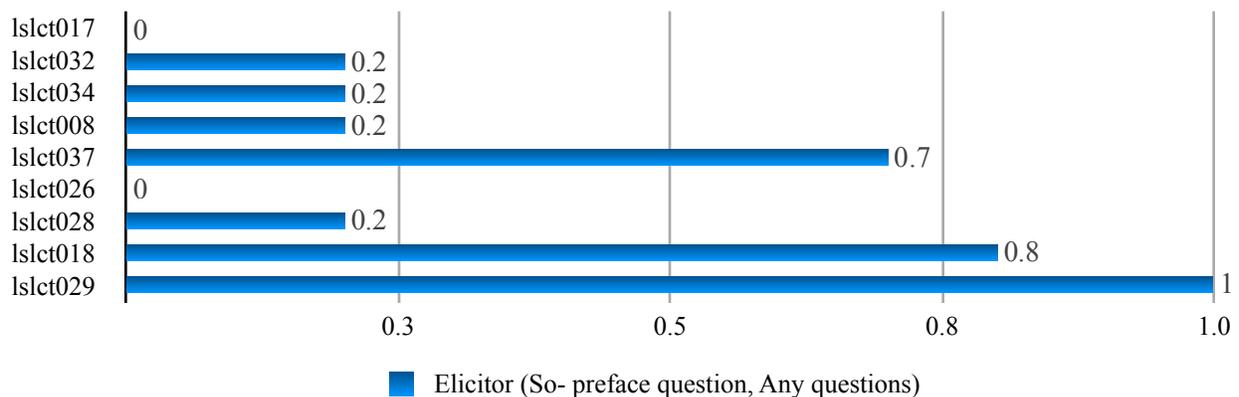
[lslct026] *but what you clear from the blood stays the same (c) Is that clear? So the urine concentration has nothing to do with your renal clearance...*

Excerpt 88 illustrates the function of the Confirmation Check, in context, in the few instances it is used in BASME. In each example, the Confirmation Check is used in the phrase “Is that clear?” (a-c) following content dense utterances.

6.4.2 Elicitors (so-prefaced question?; any questions?)

Elicitors are the least used function in both categories of the BASME sub-corpus, with a total of 0.4‰ occurrences. In contrast, SASME has three times as many overall Elicitor occurrences (0.9‰). The function is used in seven of the nine NS lectures with positive occurrences ranging from 0.2‰ to 1‰ (see Figure 6.10).

Figure 6.10. Use of Elicitor Function ‰ across BASME



6.4.2.1 Elicitors ‘*So-preface question*’

The Elicitor ‘*so-preface question*’ functions in the same manner as in SASME, with some occurrences successfully eliciting responses from the listeners, while others function rhetorically.

Excerpt 89

[lslct028] (a) *So would anybody like to be brave and tell me er about the pathogenesis of kidney stone disease?* (b) *In other words the common causes of the common stones we've been talked about who's going to be brave and tell me about some risk factors?*

[lslct034] nm0572: ...which we call the glomerular basement membrane and that's really what a kidney is (c) *So what comes in at this end?* ss: (d) *blood*

Excerpt 89 illustrates the use of ‘*so-preface question*’ to elicit responses from the students, in each case successfully. However, in the first example, prior to the response(s) and following the MWU (a) are further details that the lecturer seems to deem necessary. In (89a) the lecturer asks if any student is brave enough to answer the question - which is surprising considering the students up to that point in the lecture have been quite interactive and responsive. Following the Elicitor (a), the lecturer, giving no time for a response, moves on to a rephrasing of the question before repeating it (b). This is perhaps to clarify the type of response being elicited by the lecturer and potentially to provide additional time for students to formulate a response which they do following utterance (b). In the second example, the Elicitor (89c) is a clear illustration of a successful elicitation from students (d).

However, Excerpt (90) illustrates unsuccessful elicitation which seem to function intentionally as rhetorical Elicitors aimed at capturing the listeners attention and engaging with them. The first two examples illustrate this function as lecturers ask rhetorical questions about the meaning of figures (a) and regulated water absorption (c). They each provide the answers (b) and (d) with no expectation of listener response.

Excerpt 90

[lslct008] you're going to die in about a year just under a year (a) *So what do these figures mean?* (b) *These figures mean that in sub-Saharan Africa...*

Excerpt 90 (continued)

[lslct029] *that actually gets to this part of the nephron (c) **So** how does the er collecting duct regulate water reabsorption? (d) So it's regulated by this hormone...*

[lslct018] *yeah. (e) **So** does that make sense? remember not to go backwards again I've never done it with loudspeakers on...*

This supports the idea that these types of Elicitors can be described as “class management/engagement” (Chang, 2012: 106) questions. The last example (e), illustrates the ‘*so-preface question*’ functioning as a rhetorical comprehension check which may elicit a response but does not require one.

6.4.2.2 Elicitors ‘Any questions?’

The second MWU functioning as an Elicitor is ‘*any questions*’. Its occurrence overall is quite low in both corpora, though even more significantly in BASME with 0.1%. The results are surprising as I expected results somewhat similar to Bellés-Fortuño's (2006) findings with her MICASE sub-corpus which had 0.6%. The reason for the inconsistent results may lie in the NS corpus as Bellés-Fortuño's (2006) consists of transcripts from the academic division of Humanities & Social Sciences, while SASME are of Life & Medical Sciences discipline.

The MWU Elicitor, as described previously (see section 6.3.3), functions quite explicitly as an Elicitor and provides listeners with a clear invitation to take the floor. The first example, in Extract 91, is the sole occurrence of the DM ‘*so*’ (a) used utterance-initial with the Elicitor ‘*any questions*’ (b). In context, the DM ‘*so*’ (a) functions as an attention-getter to refocus the students’ attention back to the lecturer and marks the start of the Elicitor (b).

Excerpt 91

[lslct018] *Remember not to go backwards again. I've never done it with loudspeakers on (a) **So** (b) **any questions?***

The second example illustrates the successful elicitation of a student response (b) following two repeated uses of ‘*any questions*’ (a). The second repetition seems to indirectly indicate, with

the immediately preceding utterance, that the topic following after the break will be about a different (sub)topic and this is their opportunity to ask questions. This lecturer seems to be strongly encouraging students to respond with repeated use of the Elicitor; moreover, these are the only two occurrences in this lecture (*lslct032*). It is possible to speculate that his reiterated urging may be related to the density of the subject and his experience as a lecturer with this subject matter.

Excerpt 92

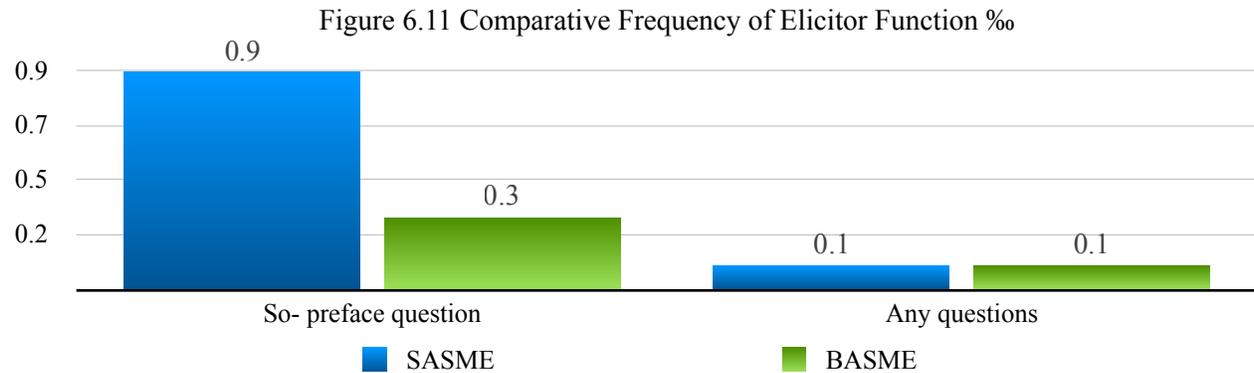
- [*lslct032*] (a) *nm0454: that's it on diabetes have a ten minute break Before you go **any questions** on diabetes? and then w-, when you come back it'll be glomerular disease **Any questions** diabetes*
 (b) *sm0499: yes*
 (c) *nm0454: diabetic nephropathy If you're shy come down the front...*

The Elicitor is also used in the concluding section of a subject matter and as a lecture draws to a close. In the first example, the Elicitor ‘*any questions*’ is presented as though a final opportunity for listeners to ask questions. This is supported by the use of the closer ‘*okay*’ (a), the phrase *that's it* (b) and the phrasing of the Elicitor to include numerous (sub)topics and “or anything?” (c) in the first example. In the second, the phrasing “come and see me” supports the view of the Elicitor as an open invitation.

Excerpt 93

- [*lslct034*] *...immune system and autoimmunity if you understand crescentric G-N*
 (a) *Okay.* (b) *That's it* (c) ***Any questions** on nephrotic syndrome glomerular disease glomerulonephritis or anything?*
- [*lslct029*] *...about quarter to eleven ten to eleven okay? And (d) if you've got **any questions** at all come and see me.*

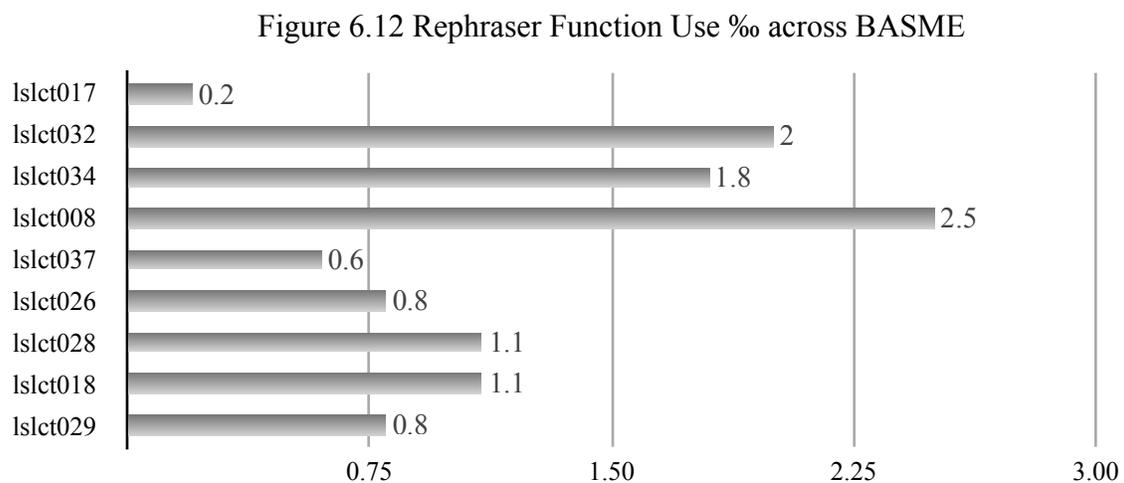
As illustrated in Figure 6.11, the distribution of the Elicitor ‘*any questions*’ is equal in the corpora with 0.1%. However, ‘*so-prefacing question*’ has three times more occurrences in the SASME corpus (0.9%) than in BASME (0.3%). The higher frequency of ‘*so-prefacing question*’ with NNS discourse suggests that the lecturers feel a greater need to use Elicitors with the NNS study body to regularly check understanding and coherence.



On the other hand, it is possible to postulate that the low number of Elicitor occurrences are a reflection of a more interactive style in the UK. Students may feel more confident participating and thus need less invitation to interact during lectures. There is no support to confirm this theory and while the lecturer of *lslct028* would seem to contradict this by stating “would anybody like to be brave and tell me...” (see Excerpt 89), the overall occurrences of the Confirmation Checks (1.9‰) and Elicitors (0.4‰) seem to suggest otherwise.

6.4.3 Rephasers (*I mean; in other words*)

Rephasers are the one Interactional function used more frequently in English in the BASME sub-corpus (1.2‰) than in the SASME corpus (0.3‰). The MWUs are used throughout BASME lectures, varying from as 0.2‰ to 2.5‰ and function similar to their use in SASME.



6.4.3.1 Rephrasers '*I mean*'

The Rephraser '*I mean*' is always used in the first person singular subject pronoun in BASME and functions to signpost the start of a rewording of what the speaker has uttered in the immediately prior utterance(s). Its purpose is to rephrase information for the benefit of the student listeners' comprehension of the dense subject matter.

The first extract (94) are examples of the DM '*I mean*' used to signal the rephrasing of the prior utterance, with its repetition illustrating the lecturer's possible need to make the intended point clear to students and to re-orient himself within his lecture discourse (a-c).

Excerpt 94

[lslct032] *that's all rubbish though isn't it (a) **I mean I mean** it doesn't take a a rocket scientist to work out that big blood vessels get smaller don't they*

[lslct008] *it's not solely due to changes in the infection rate Well this is just another (b) **I mean** you can **I mean** I won't go through this this is just impact of AIDS on under child*

[lslct026] *...and remember as mentioned get your secretion and excretion right (c) **I mean** it's difficult to say **I mean** even I'm not foolproof I you'll find I sometimes talk about the wrong one*

The next extract (Excerpt 95) illustrates the use of the Rephraser '*I mean*' to clarify the previous utterance as a whole or a token used within. For example, in (a) the lecturer discusses the average amount of protein an individual eats but then qualifies the utterance with the MWU (b) and describes different forms of protein. Similarly, the next example illustrates the qualification of the token "younger" (c) by the lecturer who defines it and signals his intention by explicitly uttering "by that I mean" (d).

Excerpt 95

[lslct034] *(a) how much protein do we eat a day eighty grams roughly (b) **I mean** it depends on whether you eat ten steaks a day or two chickpeas or whatever*

[lslct018] *who develop carriage is highly age-dependent yeah (c) if you are young and (d) by that **I mean** the younger you are the say months old*

6.4.3.2 Rephrasers '*In other words*'

The MWU Rephraser is used typically in three patterns in BASME. The first pattern involved the use of a preceding discourse marker 'so' and an if-clause following the MWU '*in other words*', the second is preceded with only the DM 'so' and the third uses the MWU on its own.

In Excerpt 96, the lecturer uses the DM 'so' which functions in context as a Summarizer, in addition to the Rephraser '*in other words*', to signal to listeners that the following utterance (a-b) is rephrasing prior information in ideally a more comprehensible manner. The if-clauses allow the lecturers to rephrase the previous information into hypothetical scenarios. Research in the use of if-clauses in spoken medical discourse suggests that, "[t]he if-operator creates maneuvering room within some type of suppositional or speculative world, enabling for example the promotion of claims and the confrontation of different viewpoints" (Carter-Thomas and Rowley-Jolivet, 2008: 194).

Excerpt 96

[lslct032] ...do a kidney biopsy on somebody with suspected diabetic nephropathy (a) ***So in other words*** if they have renal problems and diabetes we assume by and large that it is diabetic nephropathy

[lslct008] ...involuntary weight loss at greater than ten per cent of baseline (b) ***So in other words*** if you weighed ten stone you've got to have lost more than a stone in a short period of time

The second Excerpt illustrates the addition of the Summarizer 'so' to Rephraser '*in other words*' in the same manner as the previous Excerpt (see 97). The initial 'so' signals to listeners that a result will follow in the utterance related to the information being discussed and the MWU indicates that the information will be rephrased and summed up in more accessible language.

Excerpt 97

[lslct034] which are missed by a urinary dipstick ***So in other words*** you get false positives and you get false negatives er with urinary dipsticks

[lslct028] ...and most of those salts are at a near supersaturation level ***So in other words*** you don't have to do much to the urine for the salts to come out of solution

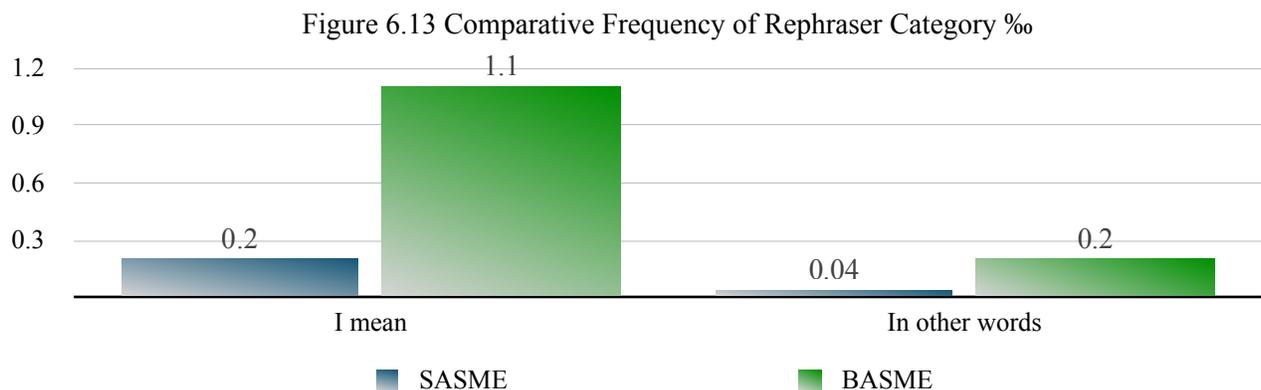
The last Excerpt (98) functions similarly to signal to listeners that the information following the MWU will not be new, rather a rephrasing of information in the immediately prior utterance(s). For example, "atheroma" is more directly defined (b), to clarify the explanation (a) for students.

Excerpt 98

[lslct032] (a) *the complications the large vessel complications of diabetes in a non-diabetic* (b) **In other words** *atheroma in somebody who doesn't have diabetes and there is evidence from...*

[lslct034] (c) *hence it's called crescentric er and the reason we like it so much is because it's a paradigm* (d) **In other words** *it's a disease which helps you understand other diseases*

In (c), the *lslct034* lecturer describes the disease as being a paradigm and signals with the MWU ‘*in other words*’ (d) that the prior “paradigm” will be subsequently reworded for coherence. This clarity of the rephrased utterance is illustrated in (d) and helps students understand normal physiology and other diseases.



The Rephraser function is used more frequently in BASME than in SASME (see Figure 6.13) and is used across the board by the NS lecturers. BASME’s overall frequency is 1.2%, in contrast to SASME with 0.3%.

6.5 Overview of Interactional DMs in SASME and BASME

The following tables are representative of the total number of raw occurrences and occurrences per thousand (‰) in each of the three functions under the Interactional Category in each corpus.

Confirmation Checks	#DMs raw	‰	Elicitors	#DMs raw	‰	Rephrasers	#DMs raw	‰
<i>Okay?</i>	165	3.5	<i>So prefacing Q</i>	43	0.9	<i>I mean</i>	13	0.3
<i>Clear?</i>	23	0.5						
<i>Right?</i>	21	0.4	<i>Any Qs</i>	13	0.3	<i>In other words</i>	1	0.02
Total	209	4.4	Total	56	1.2	Total	14	0.3

Table 6.1 SASME Interactional DMs Summary

Confirmation Checks	#DMs raw	‰	Elicitors	#DMs raw	‰	Rephrasers	#DMs raw	‰
<i>Okay?</i>	114	1.8	<i>So prefacing Q</i>	17	0.3	<i>I mean</i>	68	1.1
<i>Clear?</i>	3	0.04						
<i>Right?</i>	4	0.06	<i>Any Qs</i>	7	0.1	<i>In other words</i>	12	0.2
Total	120	1.9	Total	24	0.4	Total	80	1.2

Table 6.2 BASME Interactional DMs Summary

	<i>Total DMs raw #</i>	<i>Total DMs ‰</i>
SASME Interactional DM Categories Total	265	5.6
BASME Interactional DM Categories Total	144	2.3

Table 6.3 Interactional DMs Overview

6.6 Arabic Discourse Markers

6.6.1 Introductions

In the third part of this study, I present the results of the CS analysis of the Arabic SASME sub-corpus. This analysis approaches the Arabic segments of SASME as a separate subsection of the NNS English corpus as a whole. Meaning that the Arabic DM occurrences and CS events are viewed and calculated in respect to the overall estimated frequency of Arabic tokens, i.e. 8676 tokens.

As stated previously, the medical lectures in Saudi Arabia use English as the language of instruction which, as such, functions as the dominant language or what Myers-Scotton (1983) calls the Matrix language (ML) (see 3.4), in the SASME academic lectures. English holds the role of Matrix language as it is the official code of instruction, the majority of terminology in the discipline is in the L2 and all textbooks/references are also in English. The L1, Arabic, on the other hand, adopts the role of Embedded language (EL) and its use in CS is reliant on the syntactic patterns of the English ML. Certain grammatical frames, in both the ML and EL, share features that support the smooth switch of linguistic codes, for example:

[CTB] *It means that **inu** {that} this is very important for the economy of the country...*

The repetition of the English ‘*that*’ and Arabic ‘*inu*’ is used by the lecturer for a number of possible reasons, including attention-getter (of students), as emphasis for students to take note of the subsequent information and as a momentary pause for the lecturer to reformulate his words. This form of CS can occur from English to Arabic; however, CS, especially DMs, from Arabic to English within a chunk of discourse is not as easy to do in that it may sound awkward and unnatural (Abalhassan and Alshalawi, 2000). The embedded language (Myers-Scotton, 1993) in the EMI lectures take the form of single L1 lexical units, such as DMs or chunks of L1 discourse referred to as “embedded language islands” (Myers-Scotton, 2002) which follow the structure of the ML.

Another aspect of the CS analysis is dialects. The native Arabic lecturers in the corpus varied in dialectal background and this played a part with the CS of DMs, in that with English we can expect to see British, Australian, American and NNS English learners use similar markers such as 'so' and 'okay'. However, in Arabic, the use of tokens in colloquial Arabic is dependent on dialect, for example, a Sudanese speaker may ask '*ma3náh shinoo*', an Egyptian '*bin'ool di aay?*' and a Saudi '*ash ma3náh hazi?*'. Each of the phrases have a similar meaning in English - *what does this mean?*

In light of the focus of the study, which mainly observes the functions of the DMs and not necessarily the individual tokens, the dialectal versions of the Arabic DMs were listed as separate tokens, but treated as one meaning where relevant. For example, '*laan*' and '*3ashan*' both function in context as Topic Developers and can both be translated as '*because*'.

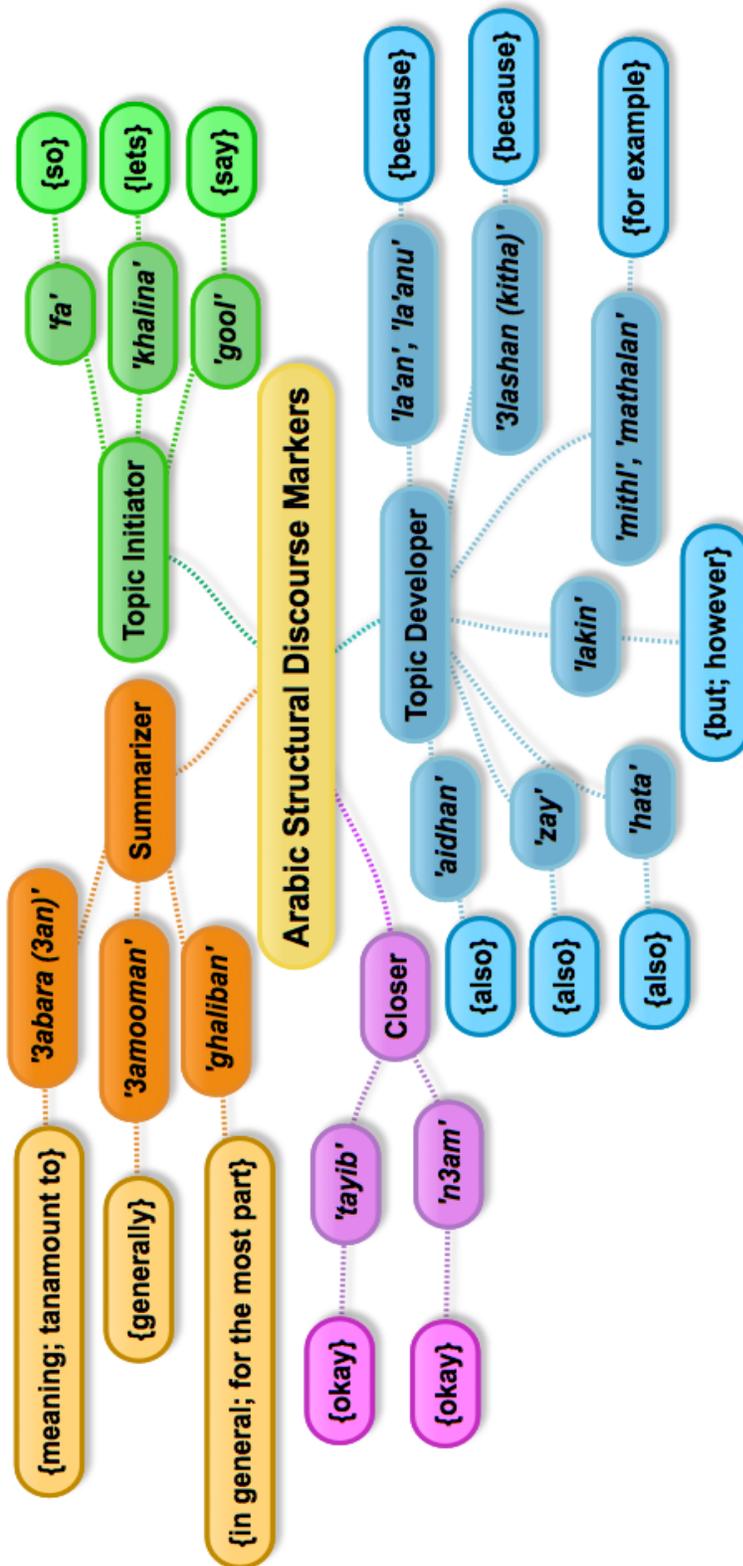
The following subsections present the results in four parts:

- (1) Arabic Structural discourse markers in the SASME lectures
- (2) Arabic Interactional discourse markers in the SASME lectures
- (3) The Roles of code-switching in the SASME lectures
- (4) An overview of the results observed in the SASME lectures

6.6.2 Code-switching and ADMs: Structural Discourse Markers

The results of the Structural DMs showed that the two most important groups, in terms of variety of DMs and overall frequency, are Topic Developers with 15.5‰ and Topic Initiators with 10‰. The Summarizer 1.2‰ and Closer 1.2‰ functions have a similar number of relatively low overall occurrences in Arabic CS.

The Arabic DMs (ADMs) (see Mind-map 6.2) that functioned as Topic Initiators were located utterance-initial or medial and had an overall frequency of 10‰ occurrences.



Mind-Map 6.2. Arabic Structural Discourse Markers by Function

In terms of preference, the ADM ‘*fa*’ is observed to be the most recurrent as a Topic Initiator and is used, in context, as a marker meaning ‘*so*’. The ADM is often used to initiate utterances such as those in Excerpt 98 which illustrate the initiation of information in both the L1 and L2.

Excerpt 98

[PBT] *So sometimes you have a tumor which is located in a sensitive area. Fa’iza ja al {so if came the} tumor bi’mikan hasas b’aldamagh {in a sensitive area in the brain}*

[DAD] *weeping for this psychological ... symptoms... Fa 3andaha {so has (+female)} emotional symptoms wa 3andaha {and (+female) has} somatic symptom.*

[PSR] *Fa {so} how can we determine inhadid {determine} the number of synapses between two neurons...*

The use of ‘*fa*’ {*so*} in Excerpt 98 draws the listeners attention and acts as a cue that the information in the prior utterance(s) will be continued and developed such as in the first example where prior information is reiterated. In the second example, it draws the students attention to the “emotional symptoms” for emphasis and to exemplify the psychological factors.

The ADM ‘*fa*’ {*so*} in the identified contexts is not an easy DM to observe; its presence as a monosyllabic Arabic sound amidst the English context can be perceived as subtle (i.e. comprehended on an unconscious level) to L1 listeners. Furthermore, its location as a Topic Initiator allowed the smooth (sometimes undetectable) transfer from the L1 to the L2. Other ADMs such as ‘*gool*’ meaning ‘*say*’ were used to signal the initiation of a new (sub)topic in the L2 or the initiation of reiterated information in the L1.

Excerpt 99

[AE] *Zay ma goolna {as we said; say} the ear is formed from three parts, external ear, middle ear and internal ear min hina lee hina {from here to here}*

[PBT] *Fa {so} hata low 3andi {even if (I) have} number of vessels...*

The Topic Initiators in Arabic are similar in function to those used in English, with the exception of the location and newness of following utterance. With ADMs, the structure of phrasing is sometimes not as clear cut as with single English DM tokens like ‘*so*’ or ‘*now*’. In the case of ‘*now*’, there are no occurrences of the Arabic equivalent ‘*al’an*’ in the SASME corpus, which indicates, either it is not used as a Topic Initiator in Arabic academic medical discourse or that for some reason, this ADM is not a preferred initiator by the NS Arabic lecturers.

The second, and preferred function by lecturers, is Topic Developer with 15.5% total occurrences. The Arabic Discourse markers are used in the discourse to signal an elaboration, clarification, example and/or provide analogies in support of the (sub)topic immediately prior. The ADMs are used as signals to introduce both English and Arabic utterances, for example:

Excerpt 100

- [CTB] *You cannot grow in artificial media, why? **laanu** rubana khuluk albacteria {because God created the bacteria} to grow on the human body not on the blood culture*
- [AE] *...the floor is called the jugular wall **laanu** {because (it is)} separated - it separates between the inner ear cavity and the bulb of the internal jugular vein...*
- [DAD] *The site of injection act as a reservoir wa min al {and from the} site biyihsal al {occurs the} slow release the diazepam. **3ashan** kitha {because of this} orally there is more rapid onset of action...*

The first Excerpt illustrates the use of the ADMs ‘*laanu*’ {*because*} and ‘*3ashan*’ {*because*} which both function as a signal that the utterance following the ADM contains details related to the prior utterance. In Excerpt 100, the lecturers link a preceding and subsequent L2 utterance with an ADM. In its role as CS, the DM functions to attract the attention of the students through the L1 and emphasizes that ADM ‘*because*’ will contain a result or explanation in the following utterance.

Other ADMs which functioned to signal the development of information in the lectures were ‘*mathalan*’ {*for example*} and ‘*ma3náh*’ {*this means; as a result*} and are illustrated in Excerpt 102.

Excerpt 101

[PBT] *Al {the} frontal lobe masoul **mathalan** 3an {is responsible **for example** of} cognition ...*

[CTB] *...because if you say that this disease is airborne manner; **ma3náh** {this means; as a result} I have to build a room for airborne transmission, which costs a lot of money.*

The first example from *PBT*, illustrates the ADM ‘*mathalan*’ {for example} which is used across the corpus and in this Excerpt (see 101) provides an example of what the frontal lobe is responsible for. The second extract from *CTB*, signals the result of an airborne disease with the ADM ‘*ma3náh*’ which can be translated to *this means* or *as a result*. In both cases, the ADMs do not function any differently from their English counterparts which strongly supports the view that these DMs were used to attract listener attention and to emphasize the importance of the information that followed.

In the last example, the CS is especially noteworthy in that the lecturer of the *DBS* spoke solely in the L2 throughout the class until this point nearly two thirds into the lecture. Prior to this interaction, students had asked questions in a mix of English and Arabic; however, all the lecturer’s questions and responses were solely in English. Excerpt 102 illustrates the point the lecturer chooses to CS to respond to a student’s question in the affirmative with ‘*na3am*’ {yes}. Following this, the lecturer continues with ‘*la fe*’ {no there are} to indicate that part of the response was correct, but clarifies that while the “nystagmus”, is “vertical” there are also other types of vertigo. The use of the ADM ‘*lakin*’ {however} seems to function, not only develop the topic, but also to emphasize and clarify the topic presented prior to the students’ responses.

Excerpt 102

[DBS] *ss: [Arabic unclear]
Lecturer: na3am. la fe {yes. no there are} other types, **lakin** {however} the nystagmus is vertical. These are the symptoms of vertigo.*

In respect of the two functions, Topic Developers and Summarizers, the definition of each had to be redefined in respect to the Arabic CS. The functions of each are not necessarily used with the same intentions as the English DMs. For example, the Topic Developers in the English

discourse were used to clarify, exemplify and/or elaborate previously introduced information with new, additional facts. However, in the Arabic code-switched discourse, most cases of CS rarely included new information and instead were used to clarify, exemplify and/or elaborate previously introduced information in the native L1 (see Excerpts 100-102). Meaning, that at their core, the English functions were similar with exception of one factor, namely, the giving of new information. With Arabic CS, this variant is not and could not be used as a defining factor because the L1 is not used, in general, to present new information. Instead, it is used to repeat or elaborate information for assurance of students' comprehension. In this case, Summarizers are then defined as functions that clearly provided a summing-up of information from prior discourse and were limited to three ADMs (see Mind-Map 6.2).

Neither the third nor fourth functions of the Structural ADMs, Summarizers (1.2‰) and Closers (1.2‰), have a high number of occurrences in the overall discourse. The Summarizer function, as previously described, is similar to the Topic Developer as both provide further details on (sub)topics, the main difference being that Summarizers sum-up information. For example, in the following Excerpt, *DAD*, illustrates the *sum* of the topic discussing the metabolic rate and varying absorption of certain drugs through muscles or orally. The ADM '*3amooman*' {generally} signals that a summary of points will follow in the subsequent utterance.

Excerpt 103

- [*DAD*] ***3amooman*** {generally} *the onset of action from intramuscular route is more rapid than the oral route, except for diazepam and chlorodiazepoxide*
- [*PBT*] ...*heya 3ibara 3an kitla min alkhalaya alshaza ghair altabi3i tizhar fe jiza min aldamagh wa twa'dy illa...*
 {It is **generally** a mass of abnormal cells not natural appearing in any part of the brain and leads to...}

The second example, taken from the *Pathology of Brain Tumors (PBT)* lecture, illustrates a summarization of the key points presented in the prior utterance. By CS the summary to Arabic, the lecturer ensures that the students have understood the key points and helps support a less tense learning atmosphere. This lecturer specifically, tended to speak using larger chunks of the L2 medium. However, when he would summarize dense information or in a lead up to feedback

elicitation from students, the rate of CS would summarily increase until a point when the lecturer is assured students were following the information. The cycle would then begin again with EMI discourse and then gradually an increase of CS incidences with both DMs and linguistic chunks of information for summarization and elicitation purposes.

The identified ADMs used as Closers are not used frequently in the corpus and of those which did occur, are located in the utterance-final position. The ADMs functioned as signals that the (sub)topic has ended or is coming to a close. For example, in the first Excerpt (104), the *DBD* lecturer signals both the close of the prior topic, but also uses the ADM as a signal that the previous topic has now drawn to a close and a new topic will soon begin. The use of the L1 also draws attention back to the speaker as he begins his presentation on Parkinson's disease.

Excerpt 104

- [*DBD*] Lecturer: ***Tayib***. *hina fe nazariyya* {Okay. here in theory} *we are talking about Parkinson disease in...*
 Student: *hyper excitation.*
- [*DAD*] Lecturer: *Increased excitation tani?* {also; and}
 <silence>
Tayib. {Okay} *mish- fe alsabah ma kan 3andakoom* {Didn't- in the morning didn't you (+plural) have} *exam?*
 Students: *Yes.*
 Lecturer: ***Tayib***. {okay} *q'bl al* {before the} *exam.*

The second lecturer of the *Drugs used for Anxiety and Panic Disorders (DAD)* (see Excerpt 104), uses the ADM '*tayib*' {okay}, with a falling intonation, to signal a close to prior actions, in this instance, the silence as the lecturer awaits a response from the students about their exam. As the lecturer is obviously fishing for a specific type of response, the students seem confused and thus silent. By using the ADM '*tayib*' {okay} the lecturer indicates that the question is closed and then proceeds to rephrase the prior question inquiring if students had an exam that morning and indirectly trying to elicit how they felt about the exam. As the lecturer proceeds to draw out the response of feeling anxious by the students, the discourse is observed to shift from the L1 gradually into the L2. With most speaker-hearer interactions in the lecture, successful elicitation of responses are in part due to the CS from the L2 to the native L1.

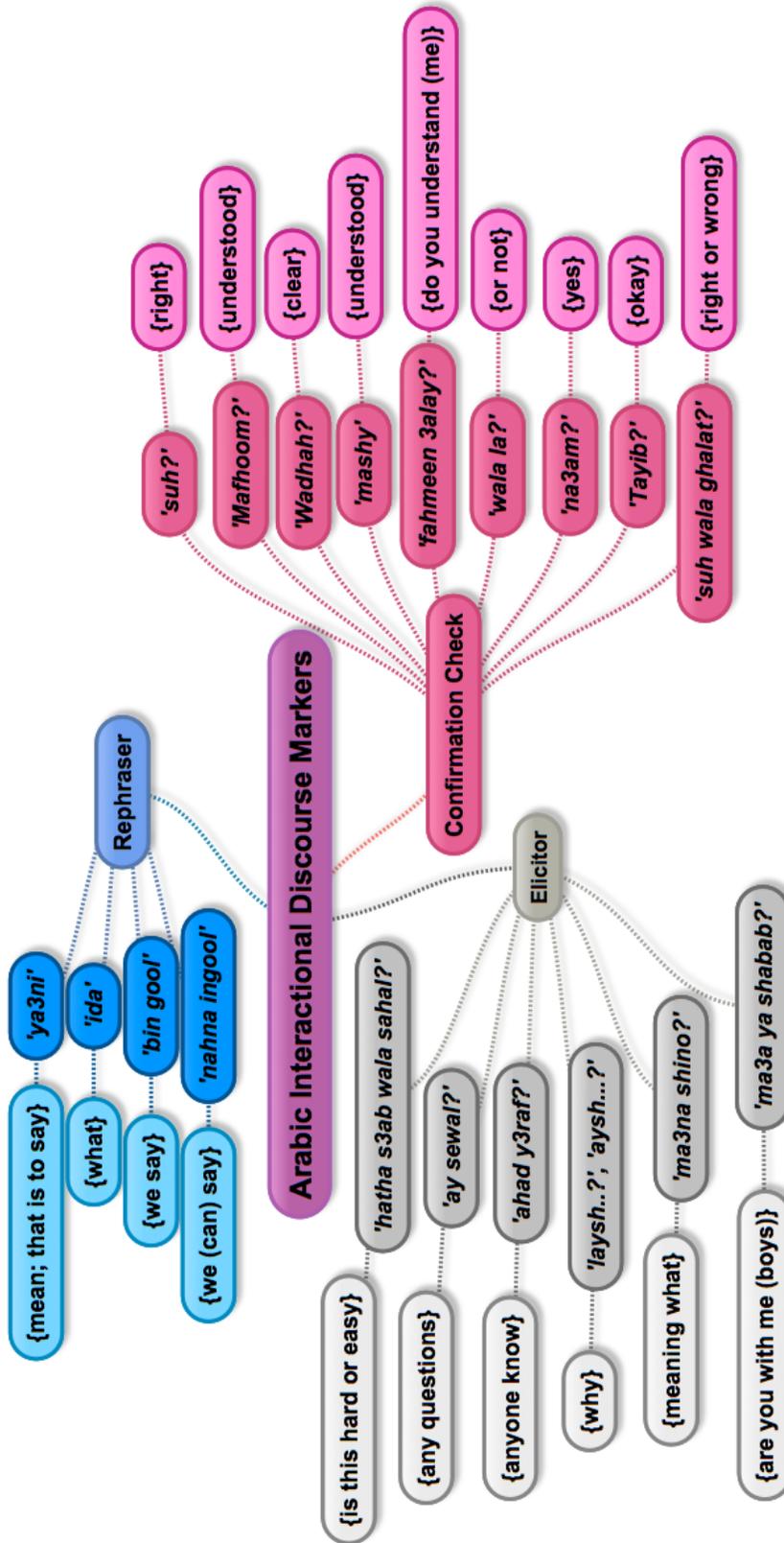
6.6.3 Code-switching and ADMs: Interactional Arabic Discourse Markers

Interactional DMs (10.2%) are the most frequent in terms of total ADMs in the SASME corpus by mere fractions of a point, 0.2%. This is not altogether unexpected, as interactions between speaker-hearer are often more personal and successful interactions in the classroom are highly dependent on language cooperation of the students to take an active role in the lectures.

The results of the Interactional ADMs show that the most recurrent functions, in order of overall occurrences and variety of DMs used, are Rephrasers, which have the highest number of overall occurrences with 10.1%. Elicitors had the second highest number of occurrences with 7% overall and lastly, Confirmations Checks, which have the highest number of overall occurrences with the English DMs, are the lowest in terms of overall occurrences in the Arabic discourse with 6.3%.

Confirmation checks are the least used Interactional function in the Arabic discourse in terms of overall occurrences. The Arabic DMs (ADMs) were used between speaker-hearers to check whether students had understood and/or were following the lecture up to that point. Arabic Confirmation Checks seem to function like pseudo-Elicitors; however, unlike the latter, they are not always meant to elicit a verbal response. Furthermore, in some cases, the momentary pause following the Confirmation Check does not provide time for even a non-verbal response to be noted. Other research, namely Waer (2012) who studied CS of Egyptian EFL teachers in Egypt, found similar results with the use of Confirmation Checks in English (*okay, right*) and in Arabic (*Tayyib* {okay}, *moutakid* {are you sure}); though, overall, more student responses are achieved in respect to Confirmation Checks in her study (ibid: 116).

The Mind-Map 6.3 illustrates the Interactional DM functions used by the NS Arabic lecturers in the SASME corpus.



Mind-Map 6.3. Arabic Interactional Discourse Markers by Function

What I have observed from the analysis is that Confirmation Checks are utterance-final and as pseudo-elicitors are not always successful in eliciting verbal responses from the students, assuming that is the intention. For example, in the first Excerpt (105), the examples illustrate the lecturer's CS from the L2 to the L1 to verify that students have understood the information up to that point. In *CTB*, the Sudanese lecturer uses an English DM 'okay' and follows this with an ADM 'mafهوم' which translates to 'understood'.

Excerpt 105

- [CTB] *You have a patient who has white blood cells but you cannot kill the organasim, okay? mafهوم? {understood}*
- [PBT] *In adults the commonest location of CNS tumors is the brain, it's the supratentorial. Wadhīh wala la? {clear or not}*
- [AE] *where we see a part of the external [unclear] temporal nerve, mandibular division of the- ma3ay ya shabab? {are you with me boys}*

The second example illustrates a segment from the Saudi, *Pathology of Brain Tumors (PBT)*, lecturer's presentation of a large chunk of dense information (equivalent to a 4–5 sentence paragraph) about the anatomical locations of brain tumors, spoken mainly in the L2 with embedded language islands of L1. He concludes this segment by asking students whether this chunk of discourse is 'wadhih' {clear} or 'la' {not}. Neither the first, nor the second lecturer seem to expect a response, though in both cases murmurs of agreement are heard in the background of the audio recordings, though no response is actually given.

The third example, by the Egyptian, *Anatomy of the Ear (AE)*, lecturer, illustrates the use of CS to interrupt the presentation of dense facts (by the speaker himself) in order to verify that the students are with him. This means that the lecturer seeks to gain confirmation that the students are following the lecture and understand the information. The use of the token 'shabab' {boys} is symbolic of a solidarity marker (Holmes, 2000); an attempt through CS to create a more personal environment between speaker-hearers.

Excerpt 106

[CTB] ...but you cannot grow the organism. **mafhoom** al-kalam di? {Are these words **understood**}, and I put them by their importance. Are you all clear? **mufhoom**? {**understood**} akeed? {for sure} hata antum al [unclear] illi wara? {even those of you (unclear) in the back.}

In the English DM analysis of SASME, the lecturer of CTB is observed as having a relatively low number of L2 Confirmation Checks in his discourse - numbering less than seven raw instances in the 47 minute lecture. However, as illustrated above, Excerpt 106 indicates a preference on the part of this lecturer to use ADMs rather than those in the L2. This may also be representative of the idiosyncratic style of the lecturer and knowledge of the benefit of using the L1 for interactions, as opposed to L2 which may create a formal distance between teacher and student.

The Rephraser function is the highest overall with 10.1% occurrences. This is not unexpected, as the total number of English Rephraser occurrences in SASME is insignificant with 0.3%. In the initial analysis, it was observed that one token ‘ya3ni’ {means; that is to say} is used significantly across the NS Arabic speakers lectures. The Rephraser function is used as a means to repeat information in a more comprehensible manner in the L1, though there were instances when the Rephraser DMs are used in the L1 to signal the function of the subsequent utterance which is in the L2. The following Excerpt 107 illustrates the use of the ADM ‘ya3ni’ {means; that is to say} to rephrase the prior utterance into clearer information for the benefit of student comprehension.

Excerpt 107

[DAD] *Second type, panic disorder, this is overwhelming fear, it is overwhelming fear linked with somatic activity. **Ya3ni**, alshakhs {**Meaning**, the person} the person may urinate or defecate, has got excessive palpitations, excessive sweat...*

[CTB] *Tuberculosis is one of the only three infectious diseases known which are transmitted by airborne. Airborne ma3nâh shinoo? {meaning what?} <student response>
Through al alhawa {the air}. **Ya3ni** low {**Meaning** if} NameX kah wa samah Allah 3andoo {coughed and God forbid he has} TB in this room, the bacteria would be hanging until the room [unclear].*

In the first example, DAD, the lecturer uses the ADM ‘*ya3ni*’ {meaning} to signal that the prior utterance will be rephrased in the following utterance. Specifically in this case, that the panic disorder with somatic activity is presented in the subsequent utterance with concrete examples that illustrate the meaning of the initial utterance. The second example, *CTB*, similarly illustrates the function of the Rephraser ADM ‘*ya3ni*’ by reiterating the information in the initial utterance through a scenario in both the L1 and L2.

Excerpt 108

[*IDBD*] *Fa ya3ni lama bi'tikra haja hawal iza ma mant 3arifha. Ana a3rif fe nas yugoolo ya akhy aqra wa binifham ba3dan lakin iza hafuzt 3umrik ma tifham.*
 {So **meaning** when you read something try if you don't know it. I know there are people who say brother I will read and understand later but if you memorize your whole life you won't learn}

The Rephraser function is also used in what would be referred to as classroom management (see 6.6.4). That is, as not only a means to signal the importance of the information that follows the ADM, but also to help create an atmosphere of learning by decreasing language tension with L2 overuse. In Excerpt 108, the Saudi lecturer of *IDBD* had been describing to students in the L2 the importance of understanding the topic and investing time into reading the textbooks of the course. To emphasize the importance of the information, the lecture code-switches to an embedded language island of L1 and signals with the ADM ‘*ya3ni*’ {meaning} that the information is related to the immediately prior utterance.

The last function under the Interactional category are Elicitors which function as ADMs that signal the elicitation of information. Similar to the English DMs, the Arabic Elicitors were not always successful; however, overall they were more successful in the L1 than in the L2. This is believed to be due to the use of the mother tongue by the lecturer to elicit information which can be understood as an invitation to respond in the L1 rather than the L2 - which some students may not feel confident in.

Excerpt 109

- [DAD] *It potentiates the action of GABA and potentiates the binding of GABA to this receptor. More chloride influx, none. **biya3mal ay al** {What does it do the} chloride influx?*
- [DB] *La'an almukh mojud dakhil qafus {because the brain is located inside a cage}. **fa nsammihaysh** {so we call this what}? <students unclear> The skull.*
- [PBT] *What is incidents? Ya3ni {(I) mean}, **fe ashya fahmaha sa3b alwahid yitazakar ha bas** {there are things that are hard to remember but} **shinoo** {what (is)} incidents?*

Elicitor ADMs are the second highest function used in terms of overall occurrences in the Arabic discourse. Often lecturers would code-switch to elicit the meaning of an L2 term to confirm that students comprehended the meaning before continuing with the topic. NS Arabic lecturers used Elicitors as a source to test understanding and in some instances, when students' responses were wrong, lecturers would code-switch to evaluate the level of comprehension at that point in the lecture. The examples in Excerpt 109 illustrate the use of the ADMs such as 'shinoo' {what} to elicit responses from the students in either the L1 or L2.

6.6.4 Code-switching: Its Role in Academic Medical Discourse

Analysis of the NNS English academic medical discourse CS illustrated that there are reasons, whether conscious or unconscious, behind lecturers use of CS from English to Arabic and vice versa. These reasons are identified as seven roles, namely, *Solidarity*, *Reiteration*, *Elaboration*, *Topic*, *Elicitation*, *Checking Comprehension* and *Classroom Management*. Each role is presented with excerpts from the SASME corpus and discussed.

One way to promote the relationship between teachers and students and create a bond of *Solidarity* between them in a situation where the teacher, with knowledge of the L2, maintains all the power dynamics in the classroom is through CS. Ferguson (2003) points out that the use of L2 represents a formal relation between teachers and students, while the mother tongue represents a

on most occasions repetition is generally to elaborate in detail or emphasize prior information (Sert, 2005). With CS to the L1, the lecturers stress the importance of the information and save time by efficiently transferring the important points and by avoiding lengthy explanations in the L2 (English). Eldridge (1996) asserts that when discourse has already been “transmitted” but not understood by hearers in that code, then by CS the “messages are reinforced; emphasized or clarified” (ibid: 303)

Excerpt 112

[DBD] *We cannot really find an etiology and there is a pattern. So, for example, Alzheimers aw alkharaf {or Alzheimers}. Fe {there are} patterns...*

In order to ensure students have understood the term *Alzheimers* the first time it is used in the lecture, the Saudi lecturer repeats the name of the disease in the L1, *alkharaf*, rather than risk students not comprehending the subsequent information. This is an illustration of single token repetitions which were numerous throughout the SASME corpus. So, too, were reiterations of chunks of information such as definitions and explanations. In the *PTB* lecture, the lecturer introduces four cell types (astrocytes, oligodendroglia, ependymal cells and neurons) in the prior utterance and describes changes in one of the cell types to serve as a prototype for the other. The following (see Excerpt 113) illustrates a repetition of information as support for students to follow the dense information that will subsequently continue in the L2.

Excerpt 113

[PBT] *... this is a prototype ya3ni haza mithal 3la al3a'ila ya3ni mithla haza alshakhs yimathel mithlan 3asheertu {This means that an example of family means, for example, that person represents, for example, his clan}*

Elaboration is similar to reiteration, but in this role, CS is used to provide new details, examples and/or analogies that will illustrate the topic more clearly in the students' L1. By CS elaborative details, a students' previous L1 learning experiences and background knowledge can be utilized by the teacher (Cole, 1998). Code-switching saves lecture time and avoids waste due to repeated explanations in the L2 and the failure of students' understanding. In Excerpt 114, the *CTB* lecturer has been describing the importance of taking the history of patients in order to accurately diagnose them and prevent misdiagnoses. To illustrate, he provides an example of a

misdiagnosis of a coworker who had traveled to America for a conference and had gotten sick upon her return to Saudi Arabia.

Excerpt 114

[CTB] *The lyme disease wa alhamdulillah da maradh fe Amrica fe jabal alrowky ma 3andana {and Thank God this is an illness in America in the Rocky Mountains we don't have (it here)} Do not forget and remember you have to ask about history if you are consulted with someone with this disease.*

The Elaboration by use of an analogy is meant to strengthen the importance of taking a history which the lecturer previously stresses is vital for patient health and for the students as future doctors. The role of CS supports the retention of information and elaborative details such as imagery are used by lecturers such as in the *Anatomy of the Ear*:

Excerpt 115

[AE] *Identify the middle ear. The middle ear bizubt 3amla zay 3ilba alkabreet al sageera {exactly acting like a small matchbox}*

By CS, the lecturer provides the students with an image in the L1 which may help students more confidently connect the L1 and L2 information, in this case, identification of the middle ear.

The *Topic* of conversation plays the fourth role in CS as linguistic codes are strongly tied in terms of identity, rights and obligations (Stroud, 1998: 322). The use of religious tokens, for example, can determine the linguistic code. In SASME, the shared religious identity of Islam by the lecturers, both Arab and non-Arab, is identified in the corpus. For example, two of the non-Arab lecturers, both Pakistani, used religious phrases in their lecture discourse and while neither are NS of Arabic, both used the terms in their Arabic forms (Excerpt 116) when discussing future plans or goals. For example:

Excerpt 116

[DE] *Today **inshallah** {God Willing} we shall take this topic anti-epileptic drugs and I hope you all are aware of epilepsy.*

[SNS] *In teaching and understanding autonomic nervous system I consider it a little bit hard but I assure you, **inshallah** {God Willing} you will understand and each every point of the system.*

The lecturers preferred to CS the religious phrase to its Arabic form, *inshallah*, rather than continue uninterrupted in the L2. While we can say that the literal meaning of the phrase remains the same whether used in L2 or L1; we must assume there is a conscious choice to CS to, what may be viewed as its original form, Arabic. Similarly, the NS Arabic lecturers would code-switch when including religious phrases, for example, in the *Cerebral Tuberculosis* lecture, the Sudanese lecturer states:

Excerpt 117

[CTB] *Ya3ni* {Meaning} you can never imitate what *subhanahu wa ta3ala* {Glorified and Exalted be He} has created.

The lecturer explains that bacteria grown as cultures in a lab are not the same as those that can grow in your body and uses the religious phrase “*subhanahu wa ta3ala*” {Glorified and Exalted be He} to illustrate the limitless power of *God* - L1, versus the limitations of *science* - L2.

Another topic that initiated CS is that of greetings. For example, on a few audio recordings, there were occurrences of greetings in the L1 as students entered the lecture room. The greeting (see Excerpt 118) is a typical one between Muslims around the world. While this greeting is not isolated to native Arabic speakers, there were no recorded occurrences of this greeting with the non-Arab lecturers in SASME.

Excerpt 118

[DBS] Student: *Assalamu alaikum* {Peace be with you}
Lecturer: *Wa alaikum assalam* {And upon you be peace}

The fifth role of CS is *Elicitation*, which is used to ease the classroom atmosphere and makes the lecture more welcoming of the students’ native identity. Lecturers will often use CS to the L1 in the classroom to encourage students and decrease the build up of tension which may occur in a lecture with only L2 usage. Students often feel uncomfortable with the L2 and this may prevent them from speaking up and asking questions or responding to the lecturer (Atkinson, 1987; Zabrodskaia, 2007; Al-Nofaie, 2010). When students are asked to answer questions, especially in cases when their answers are incorrect, students’ anxiety levels may be so serious that it affects the students’ confidence and hampers their ability to learn or participate in class (Krashen, 1987,

1988). By incorporating the mother tongue in the lecture with CS, the linguistic gap between students and lecturers is shortened (Holmes, 2000) and the pressure is relieved making students more willing to take part in the lecture (Atkinson, 1987; Ferguson 2003). Ideally, the use of the L1 encourages students to actively take part in the lecture. However, in the *PSR* lecture, students seemed reluctant, even with the inclusion of the L1 and in at least two separate incidents, near the start and the middle of the lecture, the lecturer expresses his frustration (see Excerpt 121).

Excerpt 119

[*PSR*] *Ya shabab, laysh sakteen?*
{Oh boys, why are you silent}

[*PSR*] *Ya shabab! Titkalamu 3arabi wala hindi?*
{Oh boys! Do you speak Arabic or Indian?} <laughter>

Similar to the function of Confirmation Checks with DMs (see 6.6.3), the sixth role of CS is for *Checking Comprehension*; a linguistic tool used to verify that students are following the lecturer's presentation and explanations of the (sub)topics. By implementing the L1, lecturers are able to use the shared linguistic code to confirm that students understand, comprehend and/or are following the information being given. For example, lecturers often used Arabic phrases such as '*anta shayif?*' {(can) you see}, '*mish kida?*' {is this not so} and '*wadhih?*' {clear} as a means of incorporating the L1, easing the tension in students and allowing a platform for students to respond, though it is not necessarily required by the speaker. This role additionally, allows lecturers to illustrate to the students that they have indirect permission to respond in the native L1, as they do with the role of Elicitors which also creates a welcome invitation for CS.

One of the most important roles of CS in the academic classroom is that of *Classroom Management*. A change of the linguistic code at any point during an utterance may be done not because of the subject matter but as a means to recapture students' attention which has wandered or been lost completely. Teachers seem to deliberately CS to the L1 to arouse students' attention and once captured, may then return back to the L2 used prior to the switch. Classroom Management also included CS for the purpose of decreasing or minimizing students' anxiety and overall classroom tension. Tien and Liu (2006) claim CS has socializing effects and that in order to obtain students' recognition and to maintain interest, CS is used to win the students' hearts.

And while, in some cases, no direct reason can be attributed to the use of CS, the motivation seems to be a reaction to the social dynamics in the lecture hall or with the students.

With one NS Arabic lecturer's discourse, his style of speaking could be described as a monotone, unchanging in pitch and remaining steadily in the L2 for the first half of the lecture. This type of discourse can cause students to drift off and lose focus in the lecture; for this reason, the lecturer in the second half of the class, realizing he is losing the students' attention, begins to incorporate L1 phrases such as '*lahthoo*' {focus} and '*intibhu*' {pay attention}.

Myers-Scotton (2006) cites Krashen's Monitor Model - the *Input Hypothesis* (1981, 1985) - to explain the effect boredom, stress and frustration have on a student's acquisition of language (i.e. in this study, adapted to mean acquiring lecture content through the medium of language). For Krashen (1981, 1985), intake (and, thus acquisition) is dependent upon a learner receiving enough comprehensible input inside and outside the classroom. Krashen (1981, 1985) claims that acquisition is an unconscious process, helped or hindered by what he terms an *Affective Filter* - the learner's state of anxiety towards the target language. The higher the filter, the more likely a learner's input will be blocked from becoming intake. Krashen's (1981, 1985) *Input Hypothesis*, is what Ellis (1997) has termed an *implicit* view of language learning. For Krashen (1981, 1985), procedural knowledge of a language is unconscious, while conscious learning involves declarative knowledge and is only seen to play a role in monitoring what has been previously acquired; it does not add to acquisition in its own right.

The academic lectures in SASME are not intended to be language classes, but rather medical lectures. However, the assumption by curriculum designers and the administrators of these English EMI medical programs is that learners will improve their proficiency in the L2 through the course lectures. Mendelsohn (1984) cynically refers to this as the *osmosis approach*, the belief that if students listen to the target language all day, they will, as a result, improve their listening skills and overall comprehension through the experience. The literature describes challenges with L1 and L2 conflict in the language classroom and this can be taken as reflective of events occurring in EMI subject classrooms to a certain degree (see 3.7). Students may sometimes become resentful towards the learning environments exclusion of the L1 and may feel abandoned in the L2 environment (Krashen, 1981, 1985).

The Egyptian lecturer seems to deliberately incorporate the L1 into his explanations for clarity and coherence. This is illustrated in Excerpt 120 when he describes a powerpoint illustration of the anatomy of the head in relation to the ears.

Excerpt 120

[AE] ...*sa3at bingool 3alayha* {sometimes we say it is} *Superior wa sa3at bingool 3alayha* {and sometimes we say it is} *Anterior lakin dee bingool 3alayha* {but, this we say it is} *Posterior wa dee* {and this} *Lateral ...*

The lecturer is capable, based on his discourse in the lecture, of describing the anatomical terms of location using only the L2, instead, he chooses to use the L1 as embedded language islands to guide the students through the illustration. In Excerpt 121, the lecturer uses the L1 to encourage students to take advantage of an opportunity to learn the topic - by drawing.

Excerpt 121

[PBT] *Arsam kharita li'nafsik* {draw a map for yourself}. *Oorsoomi al* {draw the} *function li'nafsik* {for yourself}. *ya3ni hawal inik tistafeedi min alkorse* {i mean try to benefit from the course}.

The roles of CS are not always clear; however, their importance in the academic classroom cannot be underestimated. It is a powerful tool that speakers who share the L1 have at hand to increase the productivity and comprehension in a foreign EMI environment.

6.6.5 Code-switching and ADMs: Summary

The findings discussed above show that, on the whole, there are more identified occurrences of Arabic Interactional DMs than Arabic Structural DMs in SASME.

Lecturers showed a preference towards using Interactional DMs, especially Rephrasers (10.1‰) and Elicitors (7‰) in the native language of Arabic. Considering the students are under extreme stress (Almoallim et al., 2010: 1276) to learn academic medical content through a foreign

EMI, it is not unexpected to observe lecturers CS in an attempt to create a more relaxed learning environment.

One function that has the highest number of occurrences in either category is Topic Developer (15.5%). This Structural DM, again, is not unexpected in terms of its significant use in the SASME Arabic sub-corpus. Academic medical discourse is dense and content-heavy, with a great deal of information to cover in a relatively short 50± minute lecture and thus Topic Developers which help maintain student focus and support comprehension will be favored.

6.7 Results II: Interactional & Arabic DMs Chapter Summary

The results are summarized as points in the following:

- (1) Confirmation Checks in SASME are the highest in frequency with 4.4‰ occurrences overall and are used as pseudo-elicitors, in that no response is required in most cases, rather the DM functions as a signal for listeners to verbally or non-verbally indicate their understanding of the main or subtopic. Confirmation Checks are more frequent when information is dense and/or complex, but can additionally be found following relatively less complex chunks of information. The latter may indicate a need by the lecturers to verify that the student listeners are following and equally comprehending the lecture discourse up to the point of the DM.
- (2) The DM ‘*okay?*’ is the preferred Confirmation Check in SASME with 3.5‰ occurrences, a frequency four times higher than the DMs ‘*clear?*’ (0.5‰) and ‘*right?*’ (0.4‰).
- (3) A non-Arab lecturer (*VitA*) in SASME had the highest frequency of DMs with 22.6‰, a frequency four and a half times higher than the next higher frequency. This single source of such a high frequency skews the overall data frequency results. A hypothesis of the use of Confirmation Checks by non-Arab lecturers and native Arab lecturers will be explored in the Discussion (see 7.3).
- (4) Confirmation Checks in BASME are also the highest in frequency with 1.9‰ occurrences overall and are used to verify that listeners have understood the lecture discourse up to the point of the DM. One hypothesis which will be discussed further in the discussion (see 7.3.2), is the link between the lower frequency in the NS corpus and the nature of interaction in British academic medical lecture.
- (5) The preferred Confirmation Check in BASME is ‘*okay?*’ with 1.8‰ occurrences - the majority of all occurrences. The other two DMs, ‘*clear?*’ (0.03‰) and ‘*right?*’ (0.1‰) have significantly fewer occurrences in BASME and do not seem to represent a type of Confirmation Check used in native English academic medical lectures.
- (6) Elicitors, the second function under Interactional DMs, are the second highest in frequency in SASME with 1.2‰. Elicitor DMs are used in most cases to draw a response from listeners, though they can also be used as pseudo-elicitors, in which case no response is necessary.

- (7) The preferred Elicitor in SASME is the MWU ‘*so-prefacing question*’ with 0.9‰ occurrences overall. This is three times higher than the second identified MWU ‘*any questions*’ with 0.3‰ total occurrences.
- (8) Elicitors in BASME are the least frequent function under this category with 0.4‰ occurrences and are used, in most cases, successfully to draw a response from listeners. The preferred MWU in Elicitors, is the ‘*so-prefacing question*’ with 0.3‰ occurrences overall, a total only slightly higher than the second Elicitor ‘*any questions*’ with 0.1‰.
- (9) Rephrasers are the least frequent function used in SASME with 0.3‰ occurrences and are used to signal the rewording or restructuring of immediately prior information in a more comprehensible structure for the benefit of listener comprehension. The preferred DM is the MWU ‘*I mean*’ with 0.3‰ and following this, ‘*in other words*’ with 0.02‰. Rephrasers are the least used function in SASME in respect to English DMs; however, if we consider the frequency of Arabic Rephraser DMs, this function exceeds that of BASME (see 6.4.3).
- (10) Rephrasers are the second highest in frequency in BASME with 1.2‰ occurrences, a frequency four times higher than that of SASME with regards to only English DMs. Rephrasers are used similarly in both the NS and NNS corpora, as signals indicating prior information will be reworded or restructured following the DM. Similarly, the preferred Rephraser MWU is ‘*I mean*’ with 1.1‰ and the second, ‘*in other words*’ with 0.2‰.
- (11) Arabic DMs are used to attract the attention of the native Arabic students and to draw their attention to the speaker’s utterance. They are followed by utterances in the L2, supporting the theory of attention-getter. In other cases the lecturers continued in the L1 for reiteration, repetition, clarification, exemplification, development, emphasis, elicitation and attention-drawing discourse cues.
- (12) Arabic Structural DMs showed that two of the most important functions used in the L1 in terms of variety and overall frequency are Topic Developers with 15.5‰ and Topic Initiators with 10‰. Summarizers 1.2‰ and Closers 1.2‰ are found to be less frequent in contrast.
- (13) Arabic Topic Initiators occurred in both utterance-initial or medial positions to initiate topics or subtopics in the discourse. The most frequent ADM used is ‘*fa*’ {so}, followed by ADMs such as ‘*gool*’ {say} to signal the initiation of new (sub)topics or the intention to reiterate or repeat information in the L1.

- (14) Arabic Topic Developers, which usually occurred utterance-medial, are the highest in frequency overall with 15.5‰. The most frequently used ADMs are 'laanu' {because} and '3ashan' {because}.
- (15) Arabic Summarizers and Closers, both of which had a frequency of 1.2‰ overall occurrences were not as popular in their Arabic form indicating a preference towards using EDMs. The most frequent ADM used as a Summarizer is '3amooman' {generally} and as a Closer, 'Tayib' {okay}.
- (16) Interactional ADMs are frequent in the native Arabic lecturers' discourse. The most frequent ADM functions are Rephrasers, 10.1‰, followed by Elicitors, 7‰ and lastly, Confirmation Checks, 6.3‰.
- (17) Rephraser ADMs did not have a high frequency when used as English DMs; however, they were quite popular as ADMs, such as, 'ya3ni' {means; that is to say}.
- (18) Elicitor ADMs are used by the lectures in the L1 to encourage student participation and possibly indicate permission for student response in the L1.
- (19) Confirmation Checks are the least frequent of the Interactional ADMs and are used to check whether students understand the content presented and are following the lecturers explanation. Frequent ADMs include 'mafhoom?' {understood} and 'wadhih' {clear}.
- (20) Seven roles for code-switching are identified in the SASME corpus, namely, *Solidarity*, *Reiteration*, *Elaboration*, *Topic*, *Elicitation*, *Checking Comprehension* and *Classroom Management*.

CHAPTER 7
DISCUSSION

*“One way to evaluate how information is processed and transferred in talk is to
rely on discourse markers.”*

(Yang, 2011: 97, cf. Jucker and Smith, 1998)

Chapter 7 Discussion

7.1 Introduction

In this chapter, I will identify noteworthy differences and similarities in the corpora. The scope of this study focuses on the use and function of DMs and CS in academic medical lectures. The main focus of the research is the analysis of the NNS lecturers who teach medicine through a foreign EMI to homogeneously NS Arabic student bodies in Saudi Arabia.

The objectives of this study are to analyze and compare the uses of the previously classified categories of DMs (see Table 4.5 and Table 4.7) across two different academic medical corpora: the NNS English corpus, SASME and the NS English sub-corpus, BASME.

The aims of the study are threefold:

- (a) To investigate the roles of DMs in academic medical lecture discourse using data from a NNS English academic corpus and a native English academic sub-corpus.
- (b) To compare and contrast qualitative and quantitative differences in the use of spoken DMs in academic medical settings between Saudi-based NNS English speakers and British NS English speakers.
- (c) To identify the functions and roles of speaker CS in the non-native English medium lectures.

This chapter reviews the results data in three sections: Structural DMs, Interactional DMs and Arabic DMs and the role of CS in academic lectures. The first subsection comments on the use and functions of DMs in the Structural DM category across the NS and NNS data which highlights significant results from the data. The second subsection discusses the uses and functions of DMs in the Interactional DM category and presents new information on the level of interaction in the NS and NNS English lectures. The last subsection comments on the use, functions and roles of ADMs and CS in the NNS English foreign medium lectures. The chapter is concluded with a summary of the study's findings as related to Structural DMs.

7.2 Comments on Structural Discourse Markers

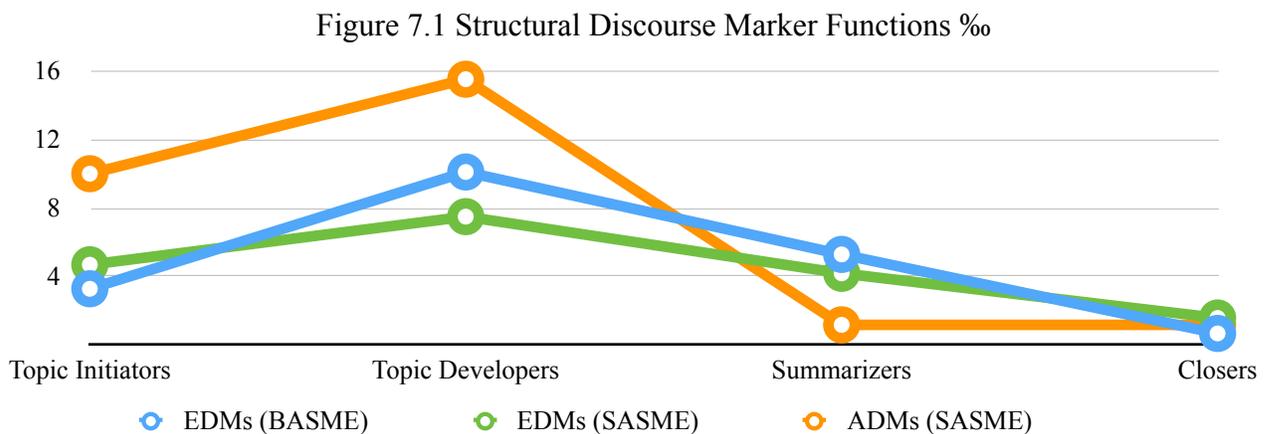
7.2.1 Structure in Medical Lecture Discourse

Structural discourse markers (SDMs) in the two corpora were those DMs identified as playing a clear role in the organization of the discourse on a structural or textual level (Aijmer 2002: 40) by marking transitions from one information stage to another. Discourse markers such as those identified in this study, ‘*so*’, ‘*okay*’, ‘*now*’, ‘*also*’, ‘*because/‘cause*’ and ‘*therefore*’, mark transitions and are signals to the hearer of the relationships of an utterance to the previous utterance(s) and/or what to expect from upcoming utterance(s). These SDMs are sub-divided into their various functions in context (Topic Initiator, Topic Developer, Summarizer and Closer) and compared across the NS and NNS English lecturers’ discourse (see 5.4 and 5.5).

The results of the study revealed that the NS English lecturers of BASME had a higher overall frequency of DM use at 22.9‰ occurrences, in contrast to the NNS English lecturers with 18.2‰. This would seem to indicate that the British NS lectures had a greater degree of organization and structure in their discourse than the Saudi-based NNS lecturers. Meaning that, NS lecturers who incorporated a greater degree of DMs such as Topic Developers (see 5.5.2) and Summarizers (see 5.5.3) in their discourse, could be viewed as helping listeners establish the relationship between utterances on local and global coherence level (Schiffrin, 1987) (see 2.4.2). Previous research (Romero-Trillo, 2002; Müller, 2005; Fuller, 2003; Hellermann and Vergun, 2006) suggests that while NNS of English are able to use the same DMs as NS and do so correctly, they tend to employ them less frequently and the results here corroborate this finding. Additionally, it has been observed that the frequency of use by NNS English lecturers seems to depend on the individual’s level of proficiency and exposure to the language, as well as the nature of the DM (Buyse, 2010). For example, some NNS English lecturers seem to favor the DMs ‘*so*’ and ‘*okay*’, both of which had relatively high frequencies in the overall corpora, but were especially favored in the NS corpus in respect to the other DMs. This popularity correlates with existing literature that identifies these DMs as two of the most frequently used and researched in pedagogic contexts (Levin and Gray, 1983; Rendle-Short, 2000; Swales and Malczewski, 2001; Bellés-Fortuño, 2006; Schlee, 2008).

Another aspect relating to the popularity of the SDMs is whether it is used by the NNS English lecturers in the English L2 medium or code-switched to the L1 during discourse. As the ADMs and the chunks of code-switched discourse are not comparable to the NS English sub-corpus, BASME, the results of the corpora cannot be accurately contrasted. Note in Figure 7.1, a contrast of the DMs used in the NS sub-corpus, BASME and the English and Arabic DMs used in the NNS corpus, SASME.

In Figure 7.1, we note that the overall frequency of ADMs as Topic Developers (15.5%) exceeds those in the BASME corpus as Topic Developers (10.1%) and suggests that the cause behind this lies in part due to the code-switching of the NS Arabic lecturers in SASME.



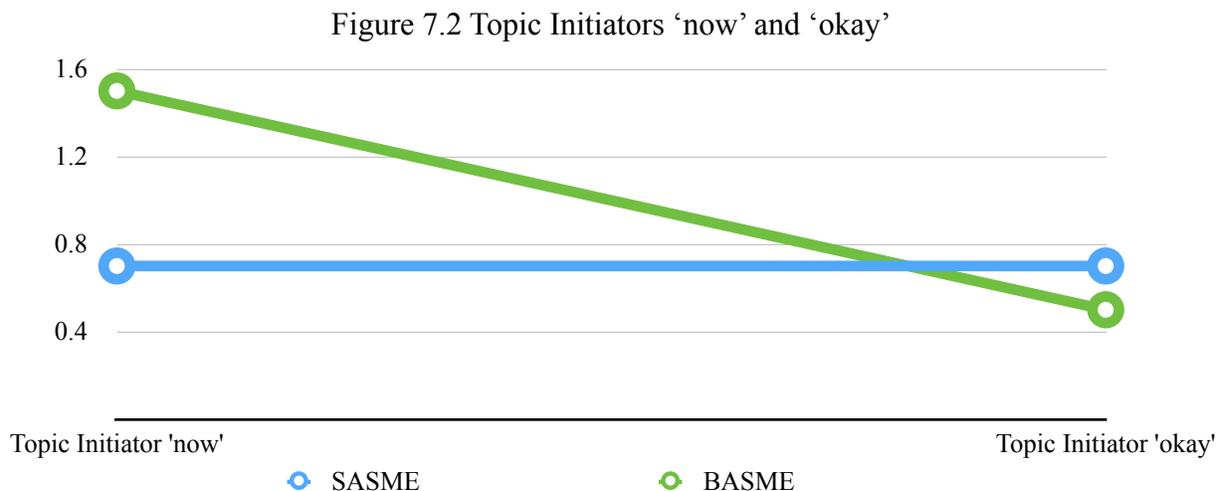
As presented in the literature (see 6.6), NS Arabic lecturers use ADMs for two main purposes, the first is to attract student attention with the use of the L1 and the second is to signal an elaboration, clarification, example and/or provide analogies in support of the (sub)topic in the prior utterance(s). Based on this, I hypothesize that the high frequency of ADMs is a result of NS Arabic lecturers reiterating and rephrasing the prior utterance(s) for the benefit of student comprehension. The ADMs are used to develop the previous utterance(s) and may follow with either a continuation of the topic in the L1 or may, at this point, be code-switched if a topic is highly dense or complex.

Figure 7.1 illustrates the overall frequency per thousand occurrences (‰) of the English discourse markers (EDMs) across the NS and NNS English lectures and the Arabic discourse

markers (ADMs) in the NS lectures. A slight pattern can be detected in the uses of the EDMs and ADMs in that there is a sharp contrast between the frequency as Topic Initiators and as Topic Developers. Moreover, all display a drop in frequency in their functions as Summarizers and Closers, though in the ADMs these two functions are equally infrequent.

7.2.2 The Lecturers' 'okay' vs. 'now'

Levin and Gray (1983) link the structural 'okay' not to the discourse setting, but to the task of lecturing itself where an "agenda involves a number of items to be covered seriatim" (ibid: 199). The DM 'okay', like 'now', marks transitions from one information stage to another and in this data, their particular function is to initiate new topics or subtopics within discourse. Figure 7.2 shows the relative frequencies of use of 'now' and 'okay' as a Topic Initiator in the two corpora.



The frequency of these two DMs as Topic Initiators differed in the NS and NNS English lectures. In the native English corpus, the DM 'now' had a higher frequency at 1.5% than in the NNS English corpus with less than half at 0.7%. On the other hand, the DM 'okay' is shown to have a higher frequency in the NNS English corpus with 0.7% occurrences to the 0.5% in the NS corpus. This is only a small difference, but the question remained of why the differences might have occurred and what difference the use of 'okay' and 'now' might make in the NS and NNS English discourse.

Schleef (2008) contends that most native speakers of English would agree that the DM ‘*now*’ seems more formal based on its conventional use which is often in formal written texts such as academic research articles. This is rarely the case with the DM ‘*okay*’ which most NS speakers might view as informal, used more often in casual conversation. This shift in the increased use of ‘*okay*’ and potentially decreased use of ‘*now*’ may stem from two factors (Schleef, 2008: 80):

- (1) An increase in interactive academic lectures (see 2.2.2), what Goffman (1981) referred to as *fresh talk* and a decrease of read-out speech (cf. Dudley-Evans and Johns, 1981)
- (2) A general reduction of formality in academic life⁴⁷ and the DMs characteristic of fresh talk in lectures.

However, what we need to note from these studies is that both are based on American lectures and we can assume based on the present data, that British lectures may involve more formality than their North American counterparts. Schleef’s (2008) factors may clarify the differences in the NS and NNS English lectures and allow us to suggest that British lectures, while still interactive and classified as fresh talk (Goffman, 1981), may retain a relatively greater formality than that found in the SASME corpus. Moreover, the small differences in the frequency of ‘*okay*’ may indicate a level of informality in the lectures and may even be tied to the academic discipline itself, an aspect both Levin and Gray (1983) and Schleef (2008) consider. That these medical lectures involve informationally dense, fact-based content must, in some part, play a role in determining the type of DMs used as it has with ‘*okay*’ and ‘*now*’.

Nattinger and DeCarrico (1992) claim that the DM ‘*now*’ would be out of place in a casual conversation because it implies the speaker knows more about the topic than the listener. Webber’s (2005) analysis of speakers at medical conferences finds that ‘*now*’ is used more in transactional than social discourse because the DM is speaker-focused and frames a unit in the upcoming utterance. When a conference speaker is entitled to a long turn, it is found that the DM ‘*now*’ invites the listener “to follow the speaker’s reasoning in a cumulative series of stages” (ibid: 174). This makes it particularly suited for lectures and conferences (Webber, 2005). Swales and Malczewski (1995) find that ‘*now*’, along with other DMs such as ‘*well*’, were among what they

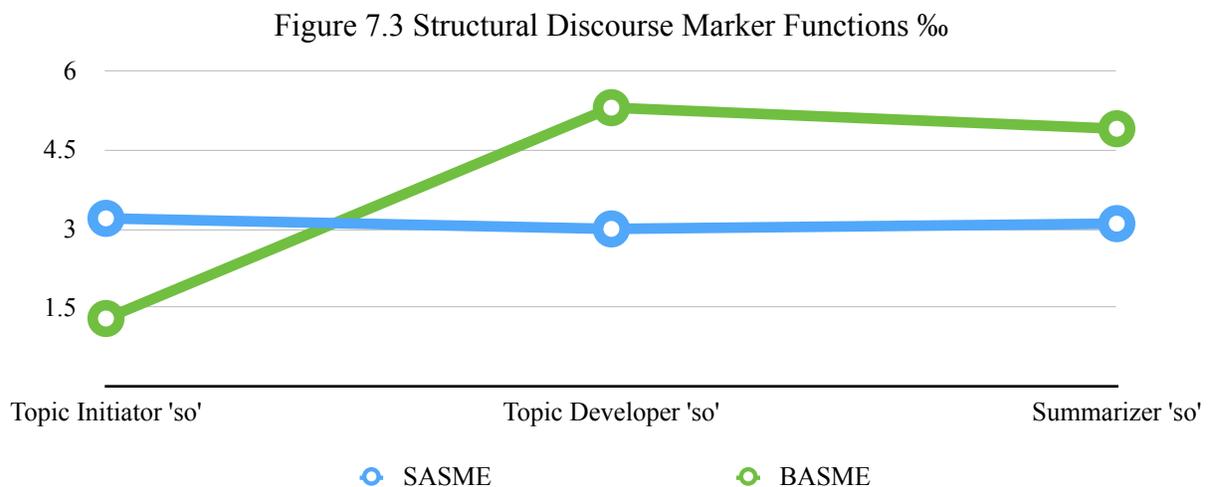
⁴⁷ Schleef (2008: 80) bases his hypothesis on anecdotal evidence from older professors in the United States which suggests that academic culture has generally become more “relaxed” over the last few decades.

defined as semantically empty markers used in academic lectures. However, this view is not shared by other researchers such as Aijmer (2002) who argues that DMs like ‘*now*’ function on a textual level to shift topics, frame discourse and draw attention to units of discourse to name a few. Schiffrin (1987) claims that ‘*now*’ functions to indicate upcoming shifts in talk ordered in a sequence of subordinating parts. Moreover, both Fraser (1999) and Schiffrin (1987) assert that DMs such as ‘*now*’ and ‘*well*’ have core meanings (see 2.4.2).

7.2.3 The High Frequency Structural DM ‘*so*’

The DM ‘*so*’ is highly versatile and can function in a number of ways in a relatively short chunk of context from Topic Initiator, Topic Developer or Summarizer (see also, ‘*so-prefaced questions*’ in 7.3.3). It is also by far the most popular and most frequent, DM in this study, in both the native and the non-native corpora.

In the NS lectures, the DM ‘*so*’ is used more frequently in Topic Developers (5.3‰) and Summarizers (4.9‰) than in the NNS English lectures with 3‰ and 3.1‰ overall occurrences respectively. However, as a Topic Initiator, the NNS English lecturers use the DM ‘*so*’ with more than double the frequency than the NS lecturers with 3.2‰ to 1.3‰ respectively.



In general, Figure 7.3 illustrates a somewhat more equal distribution of the DM ‘*so*’ in its various functions across lectures in the NNS English corpus. However, its distribution in the NS lecture varies extensively and is discussed further in the next section (7.2.3.1).

7.2.3.1 Use of ‘*so*’ as Topic Initiators and Topic Developers’

In the analysis of the Topic Initiator DMs, which overall are significantly less frequent in the NS corpus than in the NNS English corpus, it is found that NS lecturers in BASME are less likely than the NNS English SASME lecturers to initiate a topic with a single unit DM. In most cases, NS lecturers used MWUs such as, ‘*I’m/we’re going to talk about*’ and ‘*as you can see*’ which are discussed further in section 7.3.5.

The point in this case is that the NS lecturers, to whom the L1 is the EMI, use explicit markers of Topic Initiation less frequently than NNS English lecturers. However, they do use significantly more of the DM ‘*so*’ to develop their topics and signal recaps of the main points. One possible explanation for this disparity in DM ‘*so*’ frequency is the lengths of the NS and NNS English utterances. In general, the NNS English lecturers use relatively shorter and less complex (i.e. less than one, if any, dependent clause) utterances, while the NS English lecturer often use longer, more complex utterances (i.e. with one or more dependent and/or independent clauses). For example, note the following extracts taken from a NS and NNS English lecture in the corpora.

Excerpt 122

[*Islct018*] *A proportion of individuals however will fail to clear that infection and will become carriers and that's where the serious consequences of infection occur because they induce a high turnover of liver cells. **So** the the hepatitis is killing the liver cells produces a high turnover and **so** cirrhosis and primary hepatic cancer or P-H-C are actually consequences of long term carriage of hepatitis B but people remain infectious as carriers for many decades many years. Yeah, **so** there is a contrast between the acute phase that's months and the carriage phase which is years.*

[*VitA*] ***So** we divide them in two classes. One is fat soluble and the other is water soluble. **So** we have a big list of water soluble vitamins. We have B1, B2, B6,*

B12 and all of these. And these are four fat soluble vitamins. So we will consider most of these vitamins in GIT the nutritional parts, okay?

The first example, taken from BASME, illustrates the longer, more complex structure of the NS lecturer, in contrast to the shorter, less complex structure of the NNS English lecture from SASME. The contrast between the discourse styles of the NS and NNS is profound and exemplifies the nature of discourse spoken by a NS in a native English country versus a NNS in a non-English country. NS may not find it as necessary in their lecture discourse to explicitly signal the initiation of a new topic which would account for the lower occurrences of ‘*so*’ (1.3‰). However, NNS who understand the difficulties students face with the foreign EMI may feel it necessary to explicitly signal Topic Initiations with the DM ‘*so*’ (3.2‰).

7.2.3.2 Use of ‘*so*’ vs. ‘*therefore*’

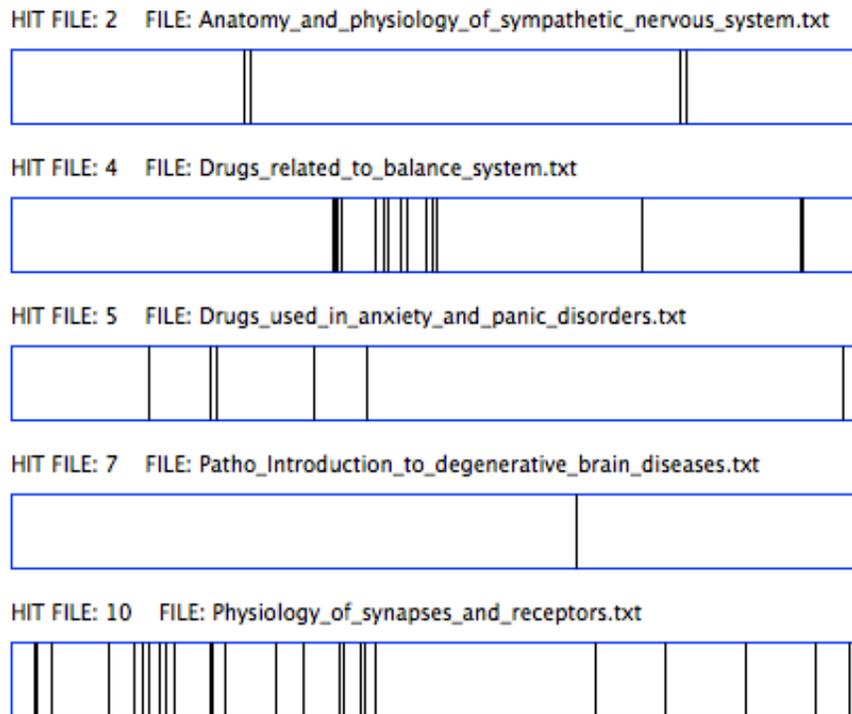
In its function as a Summarizer, the DM ‘*so*’ is one of two markers used to signal a summation of prior information. The second DM ‘*therefore*’ while not as frequent as ‘*so*’ in either corpus, is used more frequently in the NNS lectures than in the NS lectures. Similar to the concept of ‘*now*’ previously discussed, the DM ‘*therefore*’ can be viewed as a more formal marker (than ‘*so*’) and associated to a greater extent to written texts than spoken.

Its distribution across the NS and NNS lecture is relatively infrequent with an overall occurrence of 0.4‰ in five of the nine BASME lectures, while its overall occurrence in SASME is 1‰ across four of the eleven lectures. Its dispersion is illustrated in the plot below (see Figure 7.4) and can be seen to occur more frequently in one particular lecture, *PSR*, than any other. Moreover, the DM ‘*therefore*’ has no discernible pattern in the dispersion plot, other than to generally occur in the first half of the lectures, with exceptions.

The *PSR*, *Physiology of Synapses and Receptors*, lecture, which skews the data with a frequency of 12.7‰ for the Summarizer ‘*therefore*’, is unique in the number of occurrences of ‘*therefore*’ in contrast to other lecturers in both the NS and NNS English lectures. The function of this marker is as both an attention-getter and as a signal that the previous points would be

summarized in the following utterance; however, the *PSR* lecturer, based on his style and lecture discourse, uses ‘*therefore*’ as a means of instilling formality in his lecture. As noted from the dispersion plot (see Figure 7.4), the *PSR* lecturer shows a preference to the DM ‘*therefore*’ in contrast to other lecturers in SASME.

Figure 7.4 Dispersion plot of Summarizer ‘*therefore*’ in SASME corpus



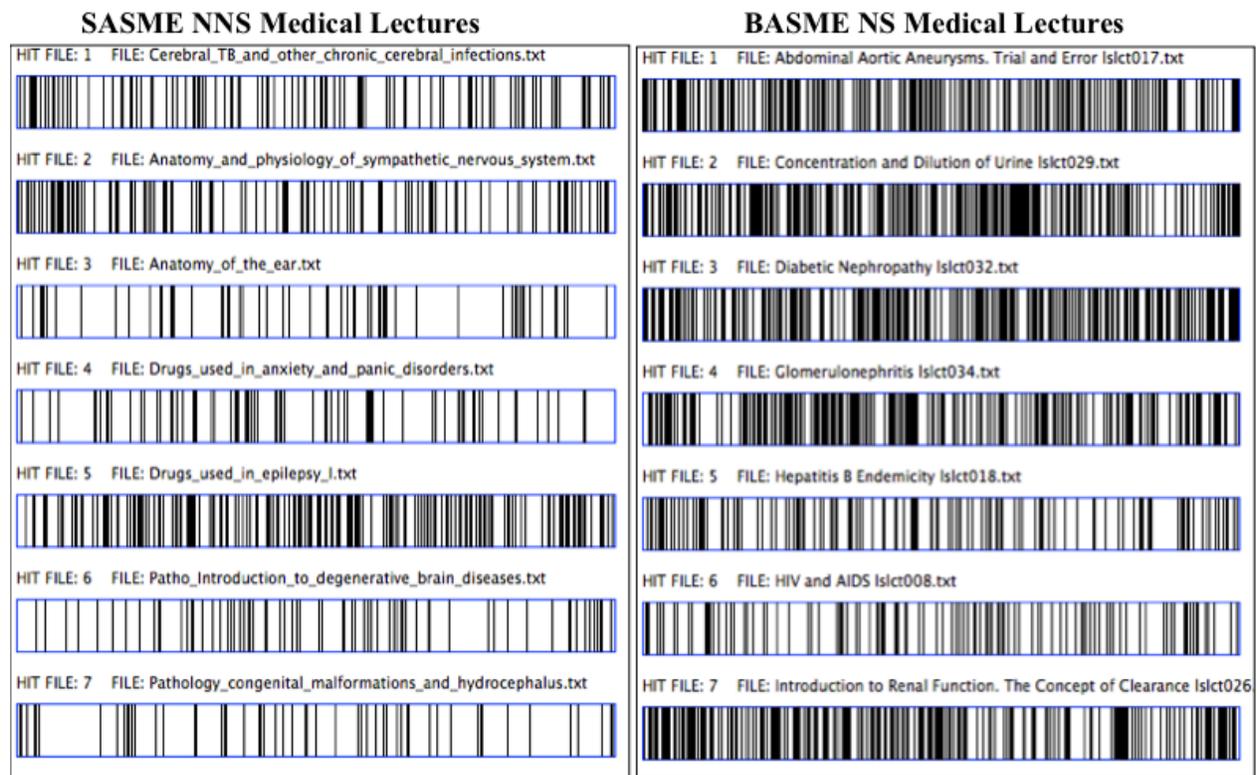
Dispersion plots help indicate whether DMs are evenly distributed throughout the corpus and its salience across lecturers’ discourse (Baker, 2010: 27). The *PSR* lecturer’s idiosyncratic use of ‘*therefore*’ as the preferred choice of Summarizer, continues in the first half of the lecture before abruptly ceasing and then being used intermittently in the remaining third of the lecture. As an observer, I noted the abrupt end to the DM coincided with growing grumbling and shifting from students (i.e., increased whispers between students and movement from their seats), which may have alerted the lecturer to the dissatisfaction of students at that point. Following this, the lecturer began to use more Arabic CS and incorporated humor in his discourse and students seemed more satisfied in contrast to their previous state.

7.2.4 Topic Developer ‘also’

The DM ‘also’ functions as an elaborative marker in the data which “augments or refines the message of S1” (Fraser, 1999: 948). It occurred more frequently in the NNS English lectures (1.9‰) than in the NS lectures (0.8‰), which may be due to the simpler structure of NNS utterances in respect to the more complex NS utterances previously discussed. The marker ‘also’ functions clearly as a DM signaling the addition of related information to prior discourse and in most cases in the NNS English lectures, the S2 utterance provides clearer, less content dense summaries of the prior information.

There are other DMs used in the academic corpora and while NS lectures have a lower frequency of the Topic Developer ‘also’, it is believed that this is due to other DMs and MWUs used in its place and unique to the NS corpus. For example, markers such as ‘*actually*’, ‘*likewise*’, ‘*as well (as)*’ and ‘*including*’ were used to develop topics in the NS lectures. Additionally, another DM ‘*and*’ (Bellés-Fortuño, 2006) that is not included in the scope of the research, mainly due to its extremely high frequency in both corpora, is also thought to be a principal factor in the low frequency of ‘also’.

Figure 7.5 Dispersion Plots of DM ‘and’ in the SASME and BASME corpus



The Dispersion plots above (Figure 7.5) are screen shots from *AntConc* (Anthony, 2012) and illustrate the dispersion of the DM ‘and’ in the first seven lectures of each corpus. Contrasting the dispersion plots makes it evident that ‘and’ is used overall at a much higher frequency and would support the hypothesis that it is used as a Topic Developer much more than the DM ‘also’.

7.2.5 Formal ‘because’ vs. Non-Formal ‘cause’

Schleppegrell (1991: 323) describes the two distinct roles of non-subordinating markers such as ‘because’, as “an expressive, non-causal link which introduces elaboration of a prior proposition” or “as a DM which indicates continuation and response in conversational interaction”. These roles are supportive of the study’s description of the DM ‘because’ or its form ‘cause’, as Topic Developers in the NS and NNS lecture discourse. The DM, in either form, can be used as an elaborator, be “involved in information management” (Passot, 2007: 130) and maintain an utterance-initial position in discourse.

Schiffrin's (1987) analysis of *'because'* and *'so'*, which similarly functions in the study as an elaborator, indicates that these DMs are not restricted to a linking function within subordination, but rather, take part in the construction of an argument's position and its support.

Overall, this marker of elaboration is used with a higher frequency in the BASME sub-corpus at 3.9‰ total occurrences, in contrast to 2.6‰ in the SASME corpus. Moreover, its presence in the relaxed form - *'cause'* - is only found in the NS corpus which is similar to other research such as Bellés-Fortuño's (2006) study of Spanish and North American lectures. Bellés-Fortuño (2006) combines both forms into one in her analysis and does not view either form as distinct in function. Similarly, while both forms were classed as functioning in an identical manner in this study, they were still listed as individual occurrences to highlight their formal and relaxed structures within context of each utterance. In this respect, the NS English lecturers used the formal *'because'* (2.7‰) more than twice as often as its shortened form *'cause'* (1.2‰).

In a recent study, Aşık and Cephe (2013) report that NNS only use *'because'* for referential purposes, while NS use both *'because'* and its relaxed form *'cuz'*. They further note that *'cuz'* is not used in just any instance, but functions in a more overtly discourse marking role in spoken discourse to signal both the continuation of topics as well as link causal relationships between utterances.

While no overt distinctions were found in this study between the forms *'because'* and *'cause'*, it is an avenue of further research that could be taken, especially its impact in pedagogical contexts.

7.3 Comments on Interactional Discourse Markers

Interactional discourse markers (IDMs) in the SASME and BASME corpora were those DMs identified as playing a clear role in speaker-hearer interaction, even when largely rhetorical and marking student (pseudo) involvement in the lecture discourse. These IDMs are sub-divided by their various functions in context (Confirmation Checks, Elicitors and Rephrasers) and compared across the NS and NNS lecturers' discourse (see 6.3 and 6.4).

The results of the study reveal that the NNS lecturers of SASME have a higher overall frequency at 5.6% occurrences, than the NS lecturers of BASME with 2.2%. This indicates that the NNS English lecturers have a greater degree of interaction in their lectures than the NS English lecturers.

7.3.1 Interaction in Medical Lectures

The concept of *interaction* in academic lectures is a complex issue and can be interpreted to be representative of various academic functions. In this study, interaction involves the increased use of linguistic expressions, namely, DMs (single and MWUs) and the decreased distance between speaker and hearers which in effect personalizes the lecture discourse (Morell, 2004: 335). The degree to which a lecture in this study is deemed interactive is based on the number of speaker-hearer interactions that occur in speech and relate to the lecture topic. For example, in Excerpt 123, two examples of speaker-hearer interactive exchanges are illustrated. The first excerpt is from a NS lecturer in BASME and the second is a NNS English lecturer in SASME. Each utterance by the speaker and hearer are counted individually, thus in Excerpt 123 there are three incidences in *lslct037* and similarly, three incidences in *DBS*.

Excerpt 123

[*lslct037*] *Lecturer: ... so which group in the Baltimore classification scheme is this in?*
 Student: Six

...

Lecturer: Six that's right six

[*DBS*] *Lecturer: So, how is balance maintained?*
 Student: Through the cerebellar.

...

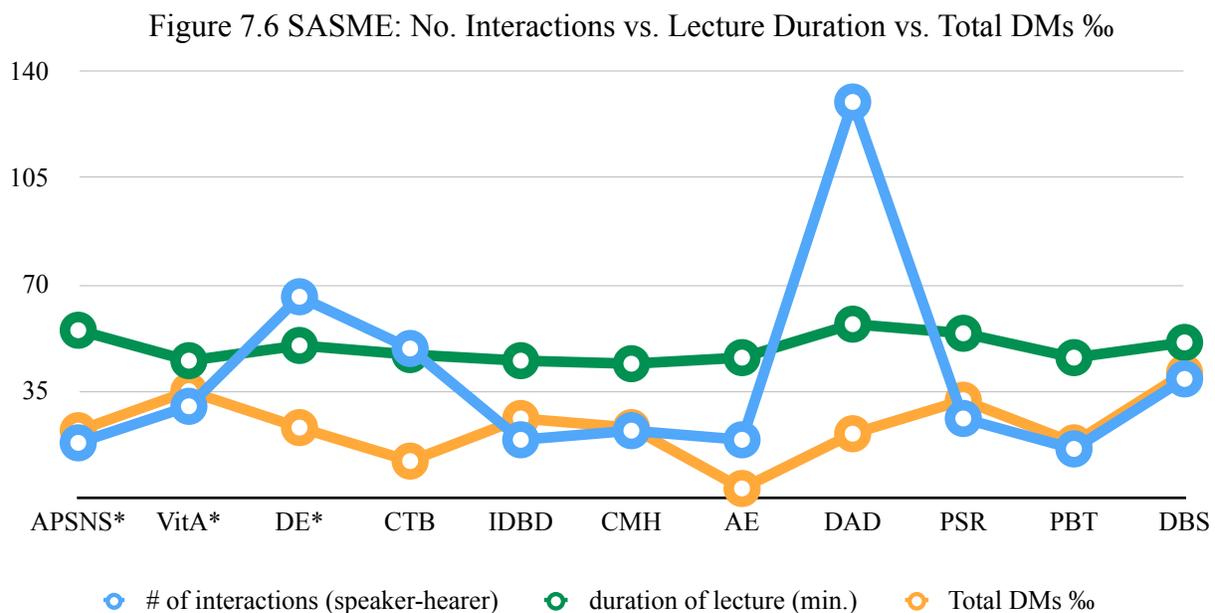
Lecturer: Through the cerebellar. Yeah.

Total incidences varied by lecture, for example, the *lslct008* lecture, *HIV & Aids*, in BASME, had zero incidences of speaker-hearer interaction, while a similar NS lecture, *lslct037*, *HIV Infection*, had 19 incidences. In the NNS English SASME corpus, the *DBS* lecture, *Drugs used in the Balance System*, had 39 incidents, while a similar lecture, *DE*, *Drugs used in Epilepsy*, had 66 incidents of speaker-hearer interaction.

The value of interactive lecturing rests on the premise that active participation and involvement by students is a prerequisite for learning beyond the recall of facts. Moreover, that students must be attentive and motivated in order for learning to occur (Steinert and Snell, 1999: 37). This is a particularly important feature in medical education where the application and use of information is as important as the retention and recall of facts (Steinert and Snell, 1999: 38).

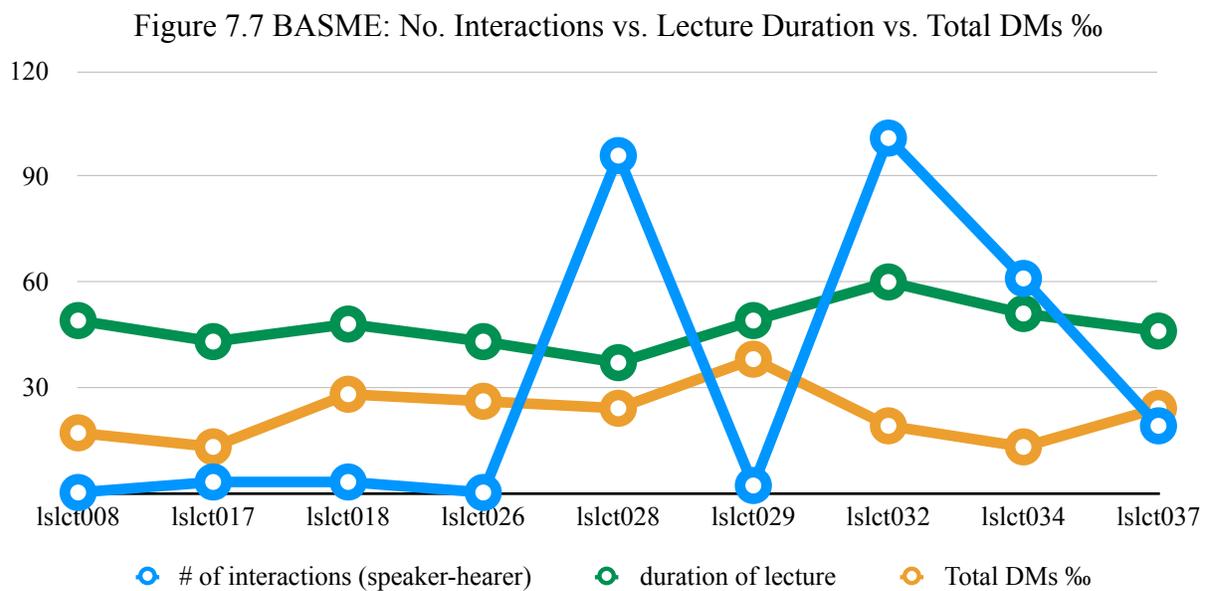
The following two figures (7.6 and 7.7) are used to comparatively illustrate the number of speaker-hearer interactions per lecture, the duration of the lecture (i.e. total minutes) and the total number of DMs used, normalized per thousand (%). Neither figure illustrates a similar pattern with the exception of the duration of lectures. One aspect in relation to number of interactions, is that all of the NNS lecturers have some degree of speaker-hearer interactions, while two of the NS English lecturers have zero and three other lecturers have between two to three speaker-hearer interactions. The number of interactions does not appear to play a role in the overall frequency of DMs that lecturers use.

In Figure 7.6, a similar relative upwards and downwards pattern appears between total number of English DMs and the number of interactions in seven of the eleven SASME lectures; the remaining four lectures (*DE*, *IDBD*, *CMH*, *PSR*) show no such pattern. Figure 7.6 also illustrates that, with exception of the *AE* lecture, there is a significant difference between the total number of DMs in contrast to the number of speaker-hearer interactions, as seen with *DE*, *CTB* and most exceptionally, by the *DAD* lecture.



Additionally, no pattern could be detected between the duration of the lectures and the remaining two variables, which we can assume indicates that the idiosyncrasies of each lecturer plays a much greater role in the number of interactions and frequency of DMs in a lecture regardless of the length of the lecture. These idiosyncrasies, moreover, cannot be linked to the Arab or non-Arab status of lecturers as the two lowest overall occurrences of speaker-hearer interactions were a native Arabic Speaker (Sudanese) and a non-Arab lecturer (Pakistani) and the two highest were a Pakistani lecturer and a Sudanese lecturer (see Table 7.1).

In the second Figure (7.7), with the exception of lecture *lslct032*, we can observe a similar pattern in an upwards and downwards movement between the duration of the lecture and the total number of DMs. This upwards and downwards pattern is visually observed to be more steady and consistent in the BASME lectures than with the NNS lectures. However, the number of interactions is not found to display any similar patterns to the other two variables. In order to investigate this further, I noted that even class size had no visible effect on the number of interactions in each lecture. For example, the *lslct017* lecture had a 130 students and only three speaker-hearer interactions; the *lslct028* lecture had 140 students and 96 speaker-hearer interactions; and the *lslct026* lecture had a 140 students and zero cases of speaker-hearer interaction.



The dramatic spikes in Figure 7.7 led us to investigate what the two relatively consistent lectures (*lslct028* and *lslct032*) and a third relatively less frequent interaction exchanges (*lslct034*)

might have in common. In order to do this, a table was created with the relevant speaker-hearer exchanges data of both the NS and NNS lectures (see Table 7.1). The three BASME lectures are similar in that all are medical and each are taught by a male lecturer. Moreover, the topics of each lecture are related with *Islct028 - Kidney Stones*, *Islct032 - Diabetic Nephropathy* (kidney disease among diabetics) and lastly, *Islct034 - Glomerulonephritis* (a term used to refer to several renal diseases which usually affects both kidneys). This correlation is challenged, however, when the following lectures, namely, *Islct026 - Introduction to Renal Function: The Concept of Clearance* which similarly covers the Kidney and *Islct029 - Concentration and Dilution of Urine* which is a product of kidney function, both reveal zero and two occurrences respectively of speaker-hearer interaction. My conclusion, therefore, is that while there may be a correlation between the number of interactions and lecture topic, these results force me to conclude that the similarities are idiosyncratic teaching styles of the NS lecturers.

SASME Lectures (All male)	No. of Speaker- Hearer exchanges	BASME Lectures (All British)	No. of Speaker- Hearer exchanges
APSNS (Pakistani)	18	Islct008 (male) (BS)	0
VitA (Indian)	30	Islct017 (female)	3
DE (Pakistani)	66	Islct018 (male) (BS)	3
CTB (Sudanese)	49	Islct026 (female)	0
IDBD (Saudi)	19	Islct028 (male)	96
CMH (Saudi)	22	Islct029 (female)	2
AE (Egyptian)	19	Islct032 (male)	101
DAD (Sudanese)	130	Islct034 (male)	61
PSR (Sudanese)	26	Islct037 (male) (BS)	19
PBT (Saudi)	16		
DBS (Sudanese)	39		
<i>Exchanges Median:</i>	<i>39.5</i>	<i>Median:</i>	<i>31.6</i>

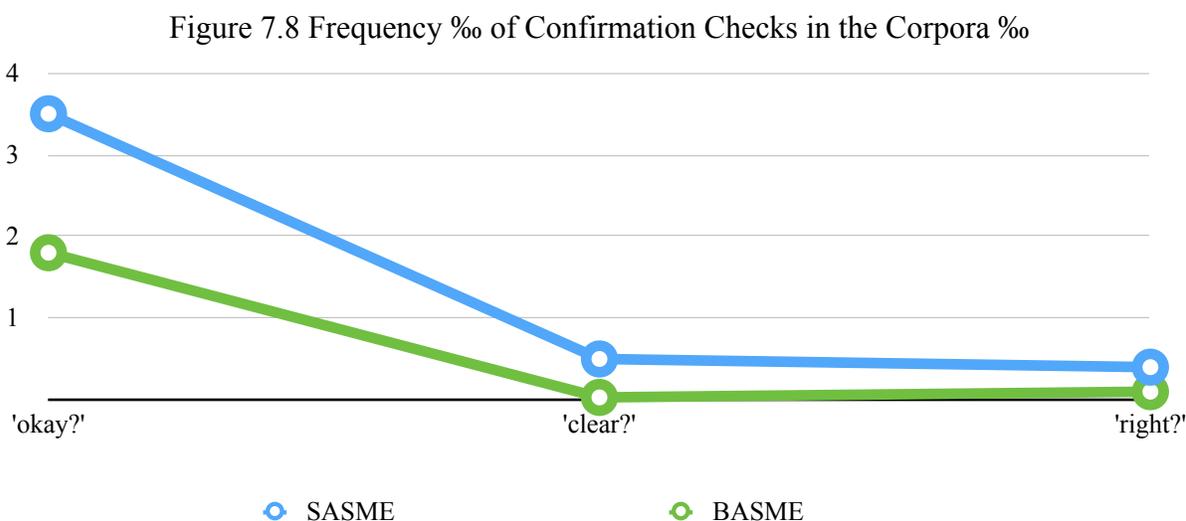
(BS) - Basic science lectures (remaining are medical)

Table 7.1 Speaker-Hearer Exchanges in the NS and NNS corpora

Additionally, other factors revealed in the NS sub-corpus are that the three basic science lectures have between 0–19 speaker-hearer exchanges in contrast to the medical lectures which varied from 0—110 exchanges. This raises an interesting question of whether basic science lectures in general have a lower rate of lecturer - student interaction than medical lectures or whether this is merely a random result of the study. Another point relates to gender which was not focused upon within this study mainly considering the NNS corpus are all male lecturers (see 4.7.1). Based on these findings, it seems gender could be a minor variable in speaker-hearer exchanges as all three female lecturers in BASME had some of the lowest number of exchanges, namely, *lslct017* (3), *lslct026* (0) and *lslct029* (2). However, we must also note that two male lecturers, *lslct008* (0) and *lslct008* (3), had few, if any, speaker-hearer exchanges as well.

7.3.2 Lecturers' Use of Confirmation Checks

The results of the study show that Confirmation Checks are used much more frequently in the NNS English lectures of SASME lectures in contrast to the NS English lectures of BASME. These results are similar to Bellés-Fortuño's (2006) study, which also show that Confirmation Checks (e.g., *okay?*, *right?* and *alright?*) are used more by the Spanish lecturers than by the North American lecturers. Bellés-Fortuño (2006) contends that this data could be interpreted to show that North American students are more used to participating in class “thus providing feedback to the lecturers in ways that do not occur in Spanish lectures” (ibid: 267).



As previously mentioned in the results (see 6.4.1), the Indian lecturer in SASME used the Confirmation Check ‘*okay?*’ more than any other lecturer in either the NNS or NS English corpus. The lecturer’s use of Confirmation Checks in a single lecture, with 22.6‰ occurrences, is significantly greater than the overall use of the Confirmation Checks in the SASME corpus with 4.4‰ or in the BASME sub-corpus with 1.9‰. This led to the question of why there was such significant overuse of these DMs.

The original hypothesis was that this overuse might be shared with other non-Arabic speakers (i.e., non-Arabs) who have no common language with the students; however, subsequent analysis showed that the other two non-Arab lecturers, both Pakistanis, had total occurrences of 1.4‰ (*DE*) and 4‰ (*APSNS*). These figures are considerably lower than those for the *VitA* lecturer and closer in similarity to the other lecturers, who are native Arabic speakers. Thus the overuse by the Indian lecturer is assumed to be more indicative of the lecturer’s personal style and linguistic preferences in discourse. Another potential supposition is that the overuse is a form of defense, meaning that with the increased implementation of student course reviews and possible student complaints, the lecturer may feel he is shielding himself by providing numerous opportunities for students to interact and question the lecture content. Gallagher (1989: 578) describes “non-Saudi Arabic speakers” as more aware of their “vulnerability” in their employment as “expatriate ‘guest’ faculty” and this may be a factor in the *VitA* lecturer’s overuse of Confirmation Checks.

7.3.3 The Use of Questions to Elicit Responses

The use of questions is probably one of the most frequently used interactive techniques and one of the easiest to implement. This is not to suggest that by asking questions in lectures a response is guaranteed or even necessary. On the contrary, in some cases, rhetorical questions are used to stimulate interest, draw attention and serve as open invitations for students to take the floor.

This study focuses on two linguistic MWUs that function as Elicitor DMs for the lecturers: ‘*so-prefaced questions?*’ and ‘*any questions?*’. These Elicitors occurred more frequently in the

NNS English lectures (1.2‰) than in the NS lectures (0.4‰) and may be reflective of the interactivity in the NNS and NS English lectures.

The use of L1 questions to elicit responses from students is evident to a greater degree in the native Arabic lecturers' discourse in SASME, with 7‰ occurrences overall and includes DMs such as, '*ay sewal?*' {any questions} and '*ma3náh aysh?*' {meaning what}.

7.3.4 Rephrasers: Native vs. Non-Native and the Use of ADMs

Fox Tree and Schrock (2002) describe the DM '*I mean*' as a signal indicating "upcoming adjustments" (2002: 741). Schiffrin (1987: 267) states that '*I mean*' can function with a literal meaning, in which case it marks orientation of speakers own talk (i.e. modifies ideas and intentions) and when not used in a literal sense, it maintains attention on the speaker. In this study, the Rephraser '*I mean*' is observed as a function signaling speaker correction, reformulation or clarification of preceding utterance(s). This is consistent with Schiffrin's (1987) proposal that its primary use is to function as a cue for students of an upcoming adjustment in the discourse.

The use of Rephrasers such as, '*I mean*' and '*in other words*', are used in both the NS and NNS corpora and as described in Results II (see 6.4.3), both identified Rephrasers, '*I mean*' and '*in other words*', were used more frequently in the NS sub-corpus, BASME, than in the NNS corpus, SASME. Their frequency overall in the NS sub-corpus is 1.2‰ in contrast to 0.3‰ in the NNS corpus.

The low frequency of Rephrasers in English in the NNS English corpus seems to correspond to the high frequency of these MWUs in Arabic with 10.1‰ total occurrences. This sharp contrast of the English and Arabic usage of Rephrasers in SASME is indicative of the native Arabic lecturers' preference for these linguistic expressions in the L1, rather than in the foreign medium, English. As a Rephraser, DMs such as '*ya3ni*' {means; that is to say}, are conceivably used more in the L1 as a more effective means to gain the attention of the NS Arabic student audience, to signal the subsequent function and to highlight the purpose of the following utterance.

7.3.5 IDMs in Native and Non-Native Medical Lectures

The Interactional English DMs identified in the study may be classified as increasing accessibility to lecture content for NNS English learners and this may account for why they were more frequent in the SASME corpus than in the BASME sub-corpus data. Overtly explicit Confirmation Checks and Elicitors are used more frequently by the NNS English lectures in a foreign English context to promote interaction between speaker-hearers and, as stated, a means to verify student understanding of L2 content-heavy discourse. The added cognitive load of a second or foreign, language on students while learning medicine should not be underestimated.

The EMI lectures are delivered by NS in a native English context, in which NNS English learners are expected to interact in the L1 within and outside of the educational context. Thus, with such expectations, NS lecturers would not necessarily alter their speech to accommodate NNS English students in the lecture hall. This is evident, in part, from the use of colloquial language by the lecturers such as, “*the upshot of that*”, “*to tease out*” (lslct026); “*a biblical plague*”, “*whizzing straight past*” (lslct008); “*the knock-on effect of*” (lslct029); and “*it's all claptrap isn't it*” (lslct032). These phrases are more often than not used to create a bonding effect through humor and informal (colloquial) language between speaker-hearers.

It is possible to speculate that the NS lecturers believed that explicit markers of confirmation or elicitors are not necessary in the British classroom, as were found in the NNS English lecturer discourse. To investigate this further, an examination was made of what might have been used in place of the identified Confirmation Checks and Elicitors (see Table 7.2).

MWUs		
<i>well</i>	<i>let's</i>	<i>i think</i>
<i>as you can see</i>	<i>you can see</i>	<i>you might think/ask</i>
<i>you remember</i>	<i>(just) remember</i>	<i>you should know/remember</i>
<i>remember I said</i>	<i>bear (that) in mind</i>	<i>you might know</i>
<i>you all know</i>	<i>you know</i>	-

Table 7.2 Multi-Word Units used by NS Lecturers of BASME

What were identified are MWUs that are rarely or never used in the NNS English lecturers' data. The MWUs in Table 7.2, were found to have a minimum of two occurrences in more than a single lecture in the NS English lecture transcripts.

In a comparative analysis, using *AntConc* (Anthony, 2012), a search for similar or parallel MWUs to those in Table 7.3 were made in the SASME Corpus. Each MWU is required to occur a minimum of two times in more than a single lecture in the NNS English lectures to be included.

MWUs		
<i>well*</i>	<i>let's</i>	<i>i think</i>
<i>as you can see</i>	<i>you can see</i>	<i>you might think/ask</i>
<i>you remember</i>	<i>(just) remember</i>	<i>you should know*/remember*</i>
<i>remember I said</i>	<i>bear (that) in mind</i>	<i>you might know</i>
<i>you all know*</i>	<i>you know</i>	-

* Used by only one lecturer in corpus

Table 7.3 Multi-Word Units used by NNS Lecturers of SASME

Table 7.3 shows the similar MWUs used by both the NS and NNS in the corpora. A strikethrough is placed through those MWUs which did not occur and an asterisk is used to show that the occurrence(s) of the marker were limited to a single lecturer and thus not representative. The second Table 7.4, illustrates MWUs that were used solely by the NNS in the SASME corpus.

MWUs in SASME		
<i>we can say</i>	<i>as you know</i>	<i>as I told you</i>
<i>first we are going to define/review</i>	<i>we are going to discuss/study/define/review</i>	<i>we'll talk about/discuss</i>
<i>we are calling it</i>	-	-

Table 7.4 Multi-Word Units used only by NNS Lecturers of SASME

These MWUs in Tables 7.2 - 7.4, which were not within the scope of the research, are similar to the lexical phrases (multi-word chunks) found by Nattinger and DeCarrico, (1988) (see 2.4.4).

These MWUs provide us with a closer look at what is occurring in the NNS and NS lectures and allow insight into why the NS sub-corpus had a lower overall frequency of Interactive DMs (2.2‰) to the NNS corpus (5.6‰). The BASME lecturers may have used a lower frequency of the single DMs (e.g. *okay? right?*), but in general, they used a higher frequency of the Rephraser MWUs (*I mean; in other words*) and other MWUs (Table 7.2). The MWUs are similar to those found in other studies which have been reported as DMs that improve retention and recall in post-lecture tests (Jung, 2003), are generally beneficial for activating content schemata (Nattinger and DeCarrico, 1988) and help students successfully follow the lecture discourse (Khuwaileh, 1999).

As a non-specialist observer of the NS lectures, I would agree that the MWUs are beneficial in guiding the listener during the lecturer; however, whether this is a serious gap in the NNS lectures is a difficult call to make for a number of reasons. First, Bellés-Fortuño (2006) reports that while single unit DMs such as ‘*so*’, ‘*well*’, ‘*because*’, etc. are more “fixed and invariably signaling cues” (ibid: 128), MWUs such as ‘*I want to discuss*’, ‘*to end/finish up/with*’, etc. can vary according to “discourse, disciplinary variations or even lecturers’ personal style[s]” (ibid: 128).

This idiosyncratic nature of lecturer discourse may make it quite difficult to teach NNS students to recognize MWUs and identify their function in a foreign EMI. Secondly, NNS lecturers, as observed in the SASME corpus, seem to organize their lecture content to be as direct and less complex as possible considering the density of the medical subject and the challenge of a foreign EMI. NNS lecturers seem sympathetic to the difficult task students face in lectures with the foreign medium and limited time to cover a great deal of information. As such, is it feasible in the current Saudi medical education system to incorporate more meaningful MWUs in contrast to single unit DMs in the discourse of NNS lecturers? I would state that it may not be currently realistic as NNS students may find it an additional challenge, but its inclusion is necessary in respect to the overall structure of the lecture discourse. Given the steady growth and development of the Saudi EMI medical education system I would suspect that in time NNS discourse will include more complex MWU structures as NNS students’ proficiency in the L2 increases.

7.4 Comments on Arabic Discourse Markers and Lecturers' Code-Switching

Homogeneous student bodies, such as those in Saudi Arabia, are a gift and a curse to lecturers and to the students themselves. The curse is the constant pressure to create an adequate linguistic environment for students to acquire the disciplinary knowledge through the foreign language medium. The English foreign medium which students may rarely or never use outside the educational context is nonetheless a crucial skill students must to acquire for academic success. Having both NS Arabic lecturers and NS Arabic students who share a language and culture creates a challenge for all to implement an “officially sanctioned [English] immersion approach to education” (Hyland, 1997: 79).

The gift in this situation is through code-switching (CS), which can be viewed as a useful classroom communication strategy. As I previously discussed (see 3.7.2), CS can be a benefit in foreign language instruction if used strategically (Atkinson, 1993; Ferguson, 2003). By sharing the same language and in part, culture, NS Arabic lecturers are able to apply CS and use linguistic devices such as DMs, to support students' knowledge acquisition in lectures.

7.4.1 ADMs in English Medium Lectures

The impacts of lecturers' use of ADMs in the medical lectures are discernible, in that they function similarly to their English counterparts; however, their use in the shared L1 heighten their linguistic impact in context. The use of ADMs functions strategically to draw students' attention, to highlight the linguistic signal of the upcoming utterance (e.g., new topic, shift in topic, summary, etc), to elicit information from students, to promote classroom teaching and management and to promote a closer speaker-hearer relationship.

Waer 's (2012) study of Egyptian EFL teachers reported similar findings in regards to ADMs which “... attract the learners' attention and so facilitate [student] participation as well as their understanding” (ibid: 139). She reports that ADMs such as ‘*ya3ni*’ {means; that is to say} and ‘*maslan*’ {for example}, amongst others, can function to organize the discourse or shift the hearer

from one topic to another (similar to Topic Developer) and “guide the learners” in much the same way as similar English DMs (Waer, 2012: 139).

Native Arabic lecturers mainly used two Structural functions in their lecture discourse: Topic Initiators (10%) and Topic Developers (15.5%). The first, Topic Initiators, are the second highest used function under the Structural category and involved ADM such as, ‘*fa*’ {so}, ‘*gool*’ {say} and ‘*khalina*’ {lets}. These ADMs are often positioned utterance-initial and in many cases were isolated incidences of Arabic code-switching in the discourse. As an observer, the impact of the ADMs in the utterances seemed to heighten their functionality as initiators by standing out in the midst of English discourse and thus drew the attention of students who may at that point have become overwhelmed with information.

Excerpt 124

[DAD] *Fa* {so} drugs which are metabolized by oxidation should be avoided in old people [unclear] metabolize deteriorates with age.

Excerpt 124 illustrates the strategy of implementing a single ADM as an attention getter and a signal of upcoming new information. Whether this is a conscious or unconscious choice on the part of the lecturer cannot be determined with the existing data; however, it is clear that the use of a student’s mother tongue in the midst of foreign discourse will draw their attention. This function alone illustrates the beneficial impact of CS DMs when initiating a (sub)topic during a lecture.

This beneficial impact is more apparent with Topic Developers (15.5%), which have the highest overall frequency. NS Arabic lecturers seem to increase the use of these ADMs to attract student attention and highlight the development and elaboration of a (sub)topic. ADMs such as ‘*laanu*’ {because}, ‘*ma3näh*’ {meaning}, ‘*lakin*’ {but; however} were quite frequent and the function itself is the highest overall in the Structural category. The ADMs are used in this context to mark developments in closely related moves from the previous (sub)topic and the impact of an *Arabic* DM, rather than an *English* DM seems deliberate, though it, in fact, may not be. In Excerpt 125, note that the lecturer could feasibly replace the ADM with an EDM without overtly affecting the discourse content or function of the existing marker:

Excerpt 125

[CMH] *As you may know - I'm sure you know - if hydrocephalus developed before the closure of the sutures of the skull you'd have a large skull. **Lakin** {but; however} if it develops later you would have-*

The use of ‘*lakin*’ {but; however} to elaborate on the topic is striking in context and draws the attention of students as apparently intended with a student response following the ADM utterance. The function of these ADMs remains similar to EDMs as signals of how “the speaker intends the present contribution to be related to preceding and/or following parts of the discourse” (Lenk, 1998: 52). Having said this, the use of these linguistic features in Arabic may be an effort by lecturers to lighten the cognitive load on students. Mantras (2000: 505) states that:

“[B]ilingual speakers, in an unconscious effort to reduce the mental effort which is necessary to monitor and direct the hearer’s responses and reactions to the speaker’s utterances, can simplify monitoring-and-directing operations by eliminating the language-specific options available to them, thereby automaticizing the choice of expressions.”

Why lecturers choose to use English DMs in some instances and Arabic DMs in others is not always apparent and the context does not usually explain the linguistic choices. It may be that the motivations in CS at these points are a matter of bridging a gap in discourse. Lecturers speaking in a foreign medium in real time may have difficulty accessing their foreign lexicon and instead utilize their L1 to fill the gap. Conversely, as “many bilinguals have an intuitive sense of linguistic well-formedness in their language use, including codeswitching” (Toribio, 2004: 138) it seems likely it is a conscious choice by NNS lecturers who are aware that ADMs will attract student focus and potentially minimize anxiety. Whatever the reasons, it does seem clear that the CS choices of lecturers have a positive impact on students based on their responses in the lectures.

7.4.2 Code-Switching in English Medium Lectures

The inclusion of CS in this study is viewed as necessary and a part of the authentic foreign EMI context. With NS lectures such as those in BASME, a monolingual lecture is expected to adhere strictly to the EMI, as it has a heterogeneous audience and a shared focus of acquiring the subject content. However, to expect the same in lectures such as those in SASME, with homogeneous audiences and in the majority of cases, NS Arabic lecturers, would create in my view, unrealistic expectations and unauthentic classroom environments. Particularly in the contexts where the foreign EMI is the vehicle for the lecture's content, medicine. Academic lectures like these are usually cognitively demanding and context-reduced (Abriam-Yago et al., 1999: 145), which causes multiple challenges for NNS English students and thus the judicious inclusion of CS can be highly beneficial.

The uses of CS by speakers are not necessarily influenced by any one motivation, nor are they necessarily based on the functions of situation, rather, they are an interactive negotiation involving decisions (Myers-Scotton, 1983). This is particularly apt in describing the motivations of academic lecturers who must play multiple roles before an audience in order to effectively deliver the lecture content as the expert and capture and maintain their attention as entertainer. In the case of a foreign medium classroom, lecturers must also continuously negotiate the effectiveness of the L1 and L2 throughout the duration of the lecture.

Waer's (2012) study of CS in EFL classes in Egypt had similar results to those found in this study. Egyptian teachers were shown to use the L1 to provide students with "an Arabic translation, to comment on a text, to highlight an important point and to resume reading a text" (ibid: 33).

Code-switching to the mother tongue in the SASME lectures is used on the whole by NS Arabic lecturers to enable students to control and strengthen instructional content, arouse student attention, save time and as a tool in class management (Merritt et al., 1992), among other roles. Each of these functional roles behind code-switching are discussed in respect to the data (see 6.6.4):

- (1) **Repetition:** Lecturers often use the L1 as a *repeat and restore* function to illustrate information, either terms or concepts, more clearly and is rarely used to introduce new information. On most occasions, repetition is used to elaborate in detail or emphasize information - a means of transferring academic content (Creese, 2005: 2-5). In one lecture, from the Exploratory Study, SAMEX, a Sudanese lecturer used the L1 to create a mental picture for students of two rather complex forms of bacteria by likening each form to soldiers in war: one acting as foot soldiers on the ground fighting and the other as navy soldiers out at sea on a ship using missiles. The repetition of information through an analogy drew appreciation from the audience and as an observer I judged that the use of the L1 was effective for the knowledge transfer.
- (2) **Arouse Attention:** Lecturers quite frequently implement the L1 to arouse student attention and often convert back to the L2 once that has been achieved. It is used in the midst of an utterance to refocus student attention back on to the topic under discussion (e.g., the use of an ADM in an L1 utterance) and to revert attention back to the teacher after it has been drawn away through a distraction (e.g., students entering late, shuffling, etc).
- (3) **Efficiency:** Lecturers use CS to avoid the waste of classroom time. By switching to the L1, the lecturers allow students to access their background knowledge, rather than use precious time to explain information in the L2 that students may already have in the L1. Excessive attention to information only wastes time and results in low efficiency. For example, when discussing dementia in the *PBT* lecture, the Saudi speaker code-switches to repeat the term in the L1 *alkharf* {dementia} and establishes shared knowledge, rather than risk students not recognizing the L2 term.
- (4) **Classroom Management:** Lecturers use of CS promotes a closer relationship between speaker and hearer in terms of increasing or decreasing the social distance (Myers-Scotton, 2006). Ferguson (2003) describes the L2 as representative of the formal relationship between teachers and students, while the L1 represents a more intimate and private relationship that can create a sense of solidarity. The use of CS encourages students' participation by fostering a more relaxed learning environment. For example, the cordial addresses that lecturers used such as, '*ya shabab*' {oh guys; boys} and Confirmation Checks such as, '*wadhah*' {clear} and '*mafهوم*' {understood}, play a role in classroom management.

Another function that has not been previously discussed is the use of CS for humor; a function that is evident with student laughter. This interpersonal function decreases the distance between speaker-hearer and is often found to lower anxiety in students, thus creating a more relaxed teaching environment. Waer (2012) found that Egyptian teachers in her study would use CS to insert humorous comments in their discourse. This is done, I suspect, in a similar manner to the native Arabic lecturers in the SASME corpus, who mainly used humor to attract student attention, bond with students (solidarity), provide a means to relieve student anxiety (e.g., laughter) and help encourage a more relaxed learning environment.

The inclusion of CS in the academic lecture can be viewed as a double edged sword. Its benefits can be attested to in its roles as described above, in addition to countless other benefits, for example, there is the psychological affect of CS for students who are often under stress and anxious as they decode a foreign language in real time to access the lecture content (Ferguson, 2003). Krashen's (1981) *Input Hypothesis* describes the acquisition of information as an unconscious process which can be greatly affected by a learner's stress level. Therefore, it seems reasonable to assume that the students in the SASME corpus may feel high levels of stress and anxiety while listening to medical lectures in a foreign medium. Furthermore, the knowledge that the lecturer shares a common language, in this case Arabic and fails to implement it as a teaching tool may increase the learners' *Affective Filter* (Krashen, 1981) and prevent learning acquisition, in addition to potentially building resentment from students.

However, it can also be said that detractors to CS can be observed if it is used inefficiently. For example, the overuse of CS in the lecture can result in an over-reliance by students on the L1 and manifest in poor L2 language skills (Ellis, 1994). Secondly, in order to accomplish learning goals, teachers must be conscious of their use of CS to support that goal (Jacobsson, 2010; Sert, 2005). In Khuwaileh's (1999) study, he states that the use of CS by the Jordanian lecturer in his study, could have facilitated the students' knowledge in the EMI Civil Engineering lecture. However, his observations were that the use of CS "did not seem a tidy manner of lecturing" (ibid: 259), in that there was no clear purpose or benefit to the use of Arabic "words" and "chunks" in the lecture.

Any potential negative aspects of CS though, in my opinion, are minimal in light of the advantages of incorporating CS in English medium lectures considering at least 30% of students must have difficulty following their L2 curriculum even when CS is employed (Kwo, 1987). Cook (2001: 105-106) describes the ability to go from one language to another as highly desirable and states that there is not much point being multi-competent if one is restricted to a single language.

The following chapter is the conclusion of this study and presents the closing remarks, implications, limitations and suggestions for further research in light of this study.

CHAPTER 8

CONCLUSION

“For Saudi medical students, the English/Arabic issue has everyday practical significance in the task of learning; during examinations, the daily pressure escalates into a frantic struggle for academic survival.”

(Gallagher, 1995: 12)

Chapter 8 Conclusion

8.1 Introduction

Graddol (1997: 45) pointed out more than 10 years ago that “one of the most significant educational trends world-wide is the teaching of a growing number of courses in universities through the medium of English”. This trend is now a reality in Saudi medical colleges, at least for the foreseeable future and its dominant presence as *the* medium of instruction provides an opportunity to research the impact of this in relation to EMI in Arabic L1 higher education settings.

In this thesis, I have reviewed the relevant studies on DMs (see chapter 2) and the phenomenon of CS (see chapter 3), particularly as it relates to the genre of academic lectures in higher education. The corpus-based methodology was described (see chapter 4) and the combination of both qualitative and quantitative research methods provides strength to this study. Following this, two chapters (see chapter 5 and 6) presented in detail, the findings of the DMs identified in the NS and NNS academic medical lectures.

The first results chapter (see chapter 5) discussed Structural DMs (SDMs), which organize and signal the relationship between utterances, that function as Topic Initiators, Topic Developers, Summarizers and Closers. The second results chapter (see chapter 6), discussed Interactional DMs (IDMs) which rhetorically signal the speaker’s intentions and goals (Bellés-Fortuño, 2006) and play a role in speaker-hearer interaction. The noteworthy differences and similarities in the corpora are explored in the discussions chapter (see chapter 7).

With the data results analyzed and discussed, the potential implications of the study's findings are presented in respect to NS and NNS academic medical discourse, English and Arabic DMs and the role of CS in EMI lectures. The chapter begins with a summary of the thesis in relation to the research questions, then moves on to the implications, followed by the limitations and concludes with possible future directions and the closing remarks for this study.

8.2 Thesis Summary

This summary seeks to provide definitive answers to the Research Questions based on the findings of this study. Each question is answered in brief as a summary of the information discussed in the discussion chapter and in the previous results chapters (see 5 and 6).

1. Do Saudi-based NNS English medical lecturers use more Structural or Interactional discourse markers than British NS English medical lecturers?

The NNS English medical lecturers used a higher overall frequency of IDMs than the NS medical lecturers and this is viewed as reflective of the greater degree of speaker-hearer interaction in the NNS English lectures than the NS lectures (see 7.3.1).

However, a higher overall frequency of SDMs were used by the NS lecturers, 22.9%, (see 5.5) and this is viewed as indicative of the greater degree of organization and structure in these lectures in contrast to the NNS English lectures, 18.2% (see 5.4).

(i) Do the NNS English medical lecturers use a higher proportion of Structural discourse markers relative to the use of Interactional discourse markers?

The results indicate that English IDMs are used more frequently in NNS medical lectures with 5.6% total occurrences and less than half that, 2.3% occurrences, in BASME (see 7.3). However, within this category of IDMs, the results illustrate that one type of function, Rephrasers, are used more frequently in the BASME sub-corpus with 1.2% occurrences overall (see 6.4.2), in contrast to SASME with 0.3% total occurrences (see 6.3.2). On a related note, the bilingual lecturers of SASME were found to have a higher overall frequency of Rephraser ADMs with 10.1% (see 6.6.3).

(ii) With respect to categories of Structural DMs, do the NNS English medical lecturers use each function with the same proportional frequency as the NS English medical lecturers?

It was found that the proportional frequency of SDMs differed between the NNS English and NS lecturers (see 5.4 & 5.5 respectively).

Specifically, Topic Initiators are used more frequently by NNS lecturers (4.7‰) than by NS lecturers (3.3‰); as are Closers with NNS lecturers (1.6‰) and NS lecturers (0.7‰).

Topic Developers, on the other hand, are used more by NS lecturers (10.1‰) than by NNS lecturers (7.5‰); and similarly, Summarizers more with NS lecturers (5.3‰) than by NNS lecturers (4.2‰).

This indicates a preference by NS lecturers to use SDMs when developing (see 5.5.2) and summarizing (see 5.5.3) information in the lecture. NNS lecturers in contrast showed a higher preference for using SDMs to signal the initiation (see 5.4.1) and the close of topics (see 5.4.4).

(iii) With respect to categories of Interactional DMs, do the NNS English medical lecturers use each function with the same proportional frequency as the NS English medical lecturers?

It was found that the proportional frequency of IDMs differs between the NNS and NS English lecturers.

Specifically, Confirmation Checks are used more frequently by NNS lecturers (4.4‰) than by NS lecturers (1.9‰); as are Elicitors with NNS lecturers (1.2‰) and NS lecturers (0.4‰). The last function, Rephrasers are used more by NS lecturers (1.2‰) than by NNS lecturers (0.3‰).

The findings indicate a greater preference by NNS lecturers to use IDMs when checking student comprehension (see 6.3.1) and evoking or drawing out student responses (see 6.3.2) in the lecture. NNS lecturers in contrast showed a higher preference for using IDMs to reword or paraphrase information (see 6.4.3).

2. Do native Arabic lecturers code-switch (English L2/Arabic L1) DMs in the Saudi based medical lectures?

The findings indicate that the NS Arabic lecturers do code-switch (see 7.4.1) and that one of the main motives for CS is as an attention-getter to draw students' focus. A second motivation is to clarify the purpose of a shift from one utterance to another. For example, using the Rephraser 'ya3ni' {means; that is to say} to indicate a rewording of prior information (see 6.6.3).

(i) If yes, which categories (Structural or Interactional) of DMs are code-switched more often in academic medical lectures?

The overall frequency of Interactional ADMs (see 6.6.3) occurrences are 10.2‰, which is a small fraction higher than those of SDMs (see 6.6.2) with 10‰ overall, indicating a generally similar preference overall towards both categories of code-switching. However, certain functions were observed to be used more frequently by the NS Arabic lecturers, specifically, Topic Developer (15.5‰), Rephrasers (10.1‰) and Topic Initiators (10‰).

(ii) If Native Arab lecturers do code-switch, what are the roles of code-switching in academic medical lectures?

The main roles of code-switching in the SASME academic lectures are identified as *Solidarity*, *Reiteration*, *Elaboration*, *Topic*, *Elicitation*, *Checking Comprehension* and *Classroom Management* (see 6.6.4). Code-switching is used as a supplementary means of communication to deepen academic teaching of the highly dense subject content and to ease pressure on students, making them more willing to actively participate in their learning.

8.3 Contribution in Relation to Research Findings

8.3.1 English DMs in Academic Medical Lectures

The results of this study correspond to the findings of other research, namely that academic DMs have a considerably higher frequency in contrast to other word forms (Fung and Carter, 2007), are important elements for the structure and organization of discourse (Nattinger and

DeCarrico, 1988; Olsen and Huckin, 1990) and function as effective interactional features (Schiffrin, 1987, Fraser, 1999, Aijmer, 2002; Fung and Carter, 2007). DMs are also valuable for listeners' formation of coherent mental maps which allows them to view how segments of discourse are interrelated (Thompson, 2003).

Regarding the first category, SDMs and observing the general differences in audience language proficiency and context, the results show that the BASME lecturers (22.9%) use more SDMs overall in their discourse than the SASME lecturers (18.2%). It is possible to assume that the NS lecturers use more SDMs because their discourse involves more complex grammatical scaffolding and developing of arguments in the lectures (see 7.2.3.1). Alternatively, I believe that the SASME lecturers may have used less English SDMs because they (native Arabic speakers) realized the value of CS similar DMs in terms of overall student comprehension and engagement (see 7.4.1).

A closer look at the four functions under this category, shows that the identified SDMs are used in both BASME and SASME in a similar manner with Topic Developers having the highest frequency in both. Topic Initiators and Summarizers are the second and third most frequent though their rank order differs in each corpus. We can conclude from this that SDMs have an important place in both the NS and NNS discourse by framing and structuring the medical academic discourse. From our knowledge, we know that medical students need to connect information on a primary level (e.g., symptoms) to a more in-depth understanding of how it relates to the body as a whole. SDMs play a role in this by signaling the relationship between propositions on a local and global coherence level.

IDMs, on the other hand, are shown to have a higher frequency in the SASME corpus than with BASME. This correlates with the higher number of speaker-hearer interactions found with the NNS lectures and based on this, we could suggest that the SASME lectures are more interactive than the NS lectures. However, it is also possible to theorize that the higher frequency of SDMs in BASME organized the subject content in a way that presented it in a more comprehensible manner to listeners. Thus the comparatively higher frequency of SDMs may have helped develop a rich mental map and thus the BASME audience had less need to ask questions or elicit clarifications.

In either case, it is significant and illustrative that the NNS of SASME have more overall occurrences of IDMs that promote a speaker-hearer relationship and encourage the students to take on a more participatory role. The significantly higher number of Confirmation Check and Elicitor functions in SASME can be said to be a result of the foreign EMI which led the NNS lecturers to continuously check student understanding. The low frequency of IDMs in BASME, on the other hand, could again be linked to the higher use of SDMs. On the other hand, as noted in the introduction (see 1.5.1), lectures are not the primary source of information for medical students in the BASME lectures. As such, students may not need to ask questions during the lecture as they are aware they will have other group-related academic opportunities to discuss the subject(s). The lectures in SASME, however, are the main vehicle of information for the NNS medical students and this is further complicated for them as the EMI is a foreign language (see 1.5.2).

8.3.2 Arabic DMs in Academic Medical Lectures

The use of ADMs in the native Arabic lecturers discourse is illustrative of the reality of Saudi-based academic lectures. The findings support that ADMs are used selectively by the lecturers mainly as attention-getters and Rephrasers (e.g. *ya3ni* {means; that is to say}) to signal the upcoming rephrasing of a prior utterance.

The power of using single-unit DMs (e.g. *ma3nâh* {this means; as a result}) or multi-word DMs (e.g. '*ma3ay ya shabab*' {are you with me boys}) in Arabic is shown to draw the attention of students and invites them to take an active part in the lecture. It also illustrates a shared language bond between the lecturer and the students and can function as a means to relax students and lower their *Affective Filter* (Krashen, 1981).

I have discussed previously (see 7.2.1) that the NS English lecturers in BASME have a higher overall frequency of SDMs in their lectures which may indicate that there is a higher degree of organization or scaffolding (see 8.3.1) in the NS lectures. However, I propose that the findings of the ADMs suggests that the lecturers incorporate more SDMs in Arabic to support the students' development of a more cohesive understanding of how the information relates within chunks of

discourse and with the topic overall. Medical education is about teaching students how anatomy, pathology, pharmacology, etc, all relate to form a broader understanding of medicine as a whole and not as individual unrelated parts. As such, NS Arabic lecturers' use of the mother tongue, even as single DMs, is not to be underestimated in its power to guide students' understanding of how segments of information relate.

8.3.3 Implications for Lectures

There are two lecturers in particular from each corpus that I observed to be quite good in their delivery and ability to engage the audience. In the NS BASME corpus, the female lecturer of *lslct017 - Abdominal Aortic Aneurysms: Trials and Errors* - speaks quite effectively, in a clear voice, implementing pauses as she presents on the topic. The *lslct017* lecturer has the highest number of Topic Initiator occurrences (5%) overall; however, is in the bottom three with the other SDM functions. I believe the relatively high use of Topic Initiators contributed to the clarity of the subject content. In fact, what makes this particular lecture interesting is that it had the lowest overall occurrences of IDMs, only three speaker-hearer interactions and was the only lecture in BASME to receive applause at the end of the lecture.

Similarly, the Saudi male lecturer in the SASME corpus, *PBT - Pathology of Brain Tumors* - was an effective speaker in his lecture. Like *lslct017*, the *PBT* lecture had a clear introduction, developed the topic gradually with examples and descriptions and provided a closing summary of the lecture subject. The Saudi lecturer has the third highest overall occurrences of Topic Initiators and the second highest number of Closers. The former is a similarity both lecturers share, as well as the fact that the *PBT* lecturer has the lowest number of speaker-hearer interactions in the NNS corpus. Furthermore, this Saudi lecturer has one of the lowest occurrences of code-switching in the lecture. At the end of the lecture, similar to the applause in *lslct017*, the students are heard commenting, in Arabic, to the lecturer and each other that the lecture was excellent. This was not noted in any other recorded lecture in the NNS corpus.

The implications of this are that academic medical lectures, while content dense, can be made more accessible to learners by explicitly organizing the lecture into coherent parts (i.e. introduction, body, conclusion). Moreover, the use of Topic Initiators seems to play a vital role in helping learners identify the start of (sub)topics and allows them to process the development of ideas without the explicit need of speaker-hearer interaction. While I believe interaction in academic lectures is important, based on the results, medical lectures, which are primarily fact-based, may not require a high number of interactive exchanges if the subject content is presented in an organized, well structured and engaging manner.

8.4 Implications of Code-Switching in Academic Lectures

English is the official medium of instruction in Saudi medical colleges. However, the reality of the situation, as shown in the results, is that CS is, in fact, used as a pedagogical tool in the lectures. Previous research has shown that even when existing policy or rules ban the use of the mother tongue in the classroom, CS does not occur “due to laziness or inattention to detail,” but of the lectures analyzed in this study due to sound pedagogical reasons (Turnbull and Dailey-O’Cain, 2009:183). This was also true of my finding. The NS Arabic lecturers were observed as deliberately incorporating the L1 in their academic discourse for the purpose of improving and supporting student comprehension.

This study has identified the most frequent roles of CS in the academic medical lectures and illustrated the pedagogical purposes of incorporating the L1 in EMI lectures. First and foremost, it saves valuable time for lecturers and students by clarifying a point through the shared language. For example, when students already understand a concept in the L1, lecturers have an obligation to tap into that knowledge base and move on to the next point, rather than use class time to try to explain a concept where the only obstacle is the medium. Likewise, if students seem confused about information, especially content-dense, complex concepts, the reiteration or repetition of information in the L1 can illustrate it more clearly.

The results of this study correspond to other researcher's findings, in that the use of the Arabic L1, as a facilitating tool in the classroom, is beneficial when used effectively (Al-Nofaie, 2010; Waer, 2012). The use of CS provides opportunities for meaningful exchanges (see 6.6.4) and can improve students' academic performance when both languages are used (Zabrodskaia, 2007). However, as with most things, moderation is key and the overuse of the L1 can become a potential disadvantage that detracts from the positive impact of CS (Al-Nofaie, 2010).

8.5 Real World Implications of Research Findings

8.5.1 Pedagogic Implications

The findings of this research could be used as a contribution to English for Academic and English for Specific Purposes (EAP/ESP) training for lecturers and students. Both NNS and NS lecturers would benefit from an explicit awareness of DM use in academic lectures and its roles in organizing and structuring academic discourse. A training course could include a component where lecturers evaluate video recordings of NS and NNS speakers teaching with EMI in their discipline and comment on the effectiveness of the lecture(s). The added knowledge of a wider range of single unit and multi-unit DMs, allows the lecturers the option to consciously manage their speech for maximum impact by pacing their utterances, stressing intonation and using more effective cues for major points in the lecture.

The findings suggest that NNS English lecturers of SASME must re-evaluate the existing frequency and multifunctionality of their use of DMs. This is particularly important with the multifunctionality of DMs (e.g., *so*) which have been reported as be problematic for NNS students (Aijmer, 2002: 3). Considering that NNS English students are under pressure to follow and comprehend the dense information of EMI lectures in real time, it is unlikely they will recognize the multiple functions that DMs such as '*so*' play in medical discourse (as reported in the study, see 5.4 and 5.5). For example,

Topic Initiator: [CTB] *So, most of the cases which are diagnosed as if they are TB meningitis are actually Brucella meningitis.*

Topic Developer: [DBS] *Betahistine is taken orally in the form of tablets. It has got the short plasma half-life, **so** this will be taken frequently.*

Summarizer: [CMH] ***So**, these are the complications of - major complications of hydrocephalus. If there is an obstruction ...*

NNS English students may find it difficult to interpret the multiple functions of 'so' illustrated in the example above. Based on the research findings and on the two 'ideal' lectures discussed in 8.3.3, I suggest that NNS English lecturers attempt to implement a greater range of DMs and include more MWUs (e.g., *(as) you can see, you should know/remember*, see 7.3.5 for further examples) in their discourse. For example, the following illustrates the DMs NNS English lecturers could use in place of 'so':

Topic Initiator: [CTB] ***It is important to know (that)** most of the cases which are diagnosed as if they are TB meningitis are actually Brucella meningitis.*

Topic Developer: [DBS] *Betahistine is taken orally in the form of tablets. It has got the short plasma half-life, **meaning that** this will be taken frequently.*

Summarizer: [CMH] ***You should remember (that)** these are the complications of - major complications of hydrocephalus. If there is an obstruction ...*

By implementing a greater variety of DMs in academic medical discourse, I believe the functions used to frame and bridge discourse will become clearer for students and may consequently improve overall EMI lecture comprehension.

NNS English students, too, would benefit from the inclusion of explicit teaching about the range of DM roles and functions in discourse, especially, those most frequent in their discipline. This discovery of linguistic cues which previously may have passed unnoticed, will heighten their awareness of the significance that these “little words” contribute to academic lectures (Andersson and Trudgill, 1990: 94). The identification of the most frequent DMs in academic medical lectures could be used as a starting point for the development of a more comprehensive list that could heighten students’ awareness of how DMs are used in academic discourse to signal the initiation, development, summary and close of topics.

In regards to CS, it is both inevitable and necessary that CS will occur in Saudi-based academic classrooms and existing regulations in Saudi universities regarding a *no Arabic* policy must be re-examined (Jenkins, 2010). There is a strong need for teachers, administrators and others to move away from the deficit notions of code-switching in the classroom and, instead, explore how the L1 can contribute to the accomplishment of teaching and learning in the classroom.

At present, the issue of CS between the L1 and the official language medium in academic lectures remains largely un-researched. As Turnbull and Dailey-O’Cain (2009: 183) put it, “sweeping this complex topic under the carpet, so to speak, can lead to teacher and student guilt and anxiety”. It is important that this beneficial pedagogic tool not be dismissed by policy makers; instead, training components need to be developed to show teachers how and when to effectively incorporate CS in their discourse and what roles it can play (e.g., *classroom management, reiteration, solidarity*, etc.) (see 6.6.4).

8.5.2 Post-Academia Implications

An explicit awareness of DMs is a valuable linguistic tool for NNS professionals (e.g. doctors, nurses, administrative staff, etc) who wish to engage with other English speakers. Brinton (1996: 64) notes that while DMs may be syntactically independent they serve a variety of “interpersonal functions” and language learners who wish to engage with others in a socially meaningful and acceptable way must be aware of the roles DMs have in discourse. On the whole, the NNS lectures used the DMs in an engaging and meaningful manner in both the L1 and L2. However, there were instances when the NNS lectures’ overuse of English DMs could be perceived as unnatural and awkward, for example, the non-Arab *VitA* lecturer who used Confirmation Checks six times more (20.8‰) than any other NS or NNS lecturer in the corpora. And, the native Arab *PSR* lecture who had a disproportionately high number of Topic Developer occurrences relative to the other lecturers with 12.7‰. Both cases illustrate that for successful communication with students, DMs must be used in moderation and subject content constructed in a coherent manner.

The omission or overuse of these linguistic cues in spoken discourse may cause NNS speakers to be judged as *unnatural*, *awkward*, *disjointed*, *impolite*, *unfriendly* or *dogmatic* within the communicative context (Brinton, 1996: 35-36). In professional contexts, this judgement may be to their disadvantage and could potentially result in problematic miscommunications.

On a broader scale, the students at Saudi medical colleges will work not only within an expatriate non-Arabic speaking community, but will be required to continuously update their knowledge through journals, conferences, symposiums, as well as other venues and most, if not all, will be presented through English as the medium of communication. Moreover, as discussed in the introduction, less than 20% of physicians in the country are Saudi (see 1.5.2), about 50% of nurses in the government sector and less than 5% in the private sector are Saudi nationals (MOH, 2009), which is why most hospitals and clinics use English as the official medium. Suliman and Tadros (2010: 2) state that “[t]he difficulties facing the students in communicating verbally and in writing in the English language, particularly with staff in a multicultural and multinational clinical training environment, are ongoing major concerns”.

8.6 Limitations of the Study

Having presented the overall results and potential implications of this study, I must acknowledge the limitations of the study:

- (1) Audio Recordings: The availability of only audio and not video, recordings in the Saudi-based lectures was a drawback. While video was not an option due to cultural and administrative policy in Saudi Arabia, its absence prevented observations about lecturers' body language, the use of powerpoint presentations, white board, etc. and visual support to help with the transcription of sections with weak audio quality.
- (2) Corpus software: Limitations of corpus software with the Arabic transcripts compelled me to analyze the data manually and may have affected the calculations of frequency and overall functions. Software limitations⁴⁸, at the time of the analysis, prevented

⁴⁸ On Mac with Arabic language recognition

- concordancing, generated word lists and other tools that were available for the English texts (see 4.5.1).
- (3) Gender: Although the intention was to include both male and female lecturers in the SASME corpus, unfortunately, the female lecturer recordings did not fulfill the study's selection variables. While this is a limitation, it must be noted that even the inclusion of female lecturers would still retain the same issue in regard to gender as Saudi medical colleges are still segregated by gender (see 4.2.1). Furthermore, a study by Schleef (2005) investigating how academic instructors' gender may influence the frequency of DMs, such as *'okay'* and *'right'*, reported that no distinct correlation could be detected. In fact, academic discipline and conversational role of the instructor were shown to have a greater influence than gender.
 - (4) Supplementary Data: Though time constraints would not make this feasible, the addition of interview data with the lecturers or a questionnaire of students' perceptions of the EMI could have enriched the data's conclusions. For example, interviews with NNS lecturers would provide insight into their DM usage in the academic discourse, as similarly researched by Othman (2010). Student interviews or questionnaires would also provide insight of how NNS perceive the use of English and Arabic DMs, assuming they are even aware of their presence in discourse.
 - (5) Comparability of Corpora: By focusing mainly on the EMI aspects of the NS and NNS corpora, it was hoped that limitations would be minimized with the English DM findings. However, with the ADMs, findings could not be accurately compared or contrasted to the NS corpus as it had no cases of CS.

8.7 Further Research

There is great potential in this area for further research. Foremost, a broader corpus with similar lectures from other medical colleges in Riyadh or in other Saudi cities (e.g. Jeddah, Dammam) or even Gulf countries (e.g. Kuwait, UAE) would provide richer data and extended generalizations about the uses, functions and roles of DMs and CS in EMI lectures. Similar research can be used to check whether the results of the study can be verified or refuted, by using

schools different from King Saud University and a different sub-corpus of NS medical lectures. Additionally, further research could investigate how academic lectures are received by students by testing their comprehension or through interviews and questionnaires.

Feedback from academicians and policy makers should also be encouraged in order to develop a better understanding of how CS contributes to teacher education and student learning in EMI academic classrooms. Research can also be undertaken to investigate how effective CS in academic lectures are by testing the comprehension of students in groups taught with and absent of L1 CS.

8.8 Concluding Remarks

Discourse markers are one small piece in a larger foundation that supports the development of medical education in Saudi Arabia's higher education. The steps that government and policy makers have taken to improve and strengthen medical education is commendable; however, let us not forget that with all structures, the strength of the foundation is integral. Realistically, for medical education in Saudi Arabia to continue on its path and to become a contender on the global market, there must be additional focus on developing a more successful English for Medical Purposes (EMP) program in the preparatory year and continuation of the EMP program for at least the first few years of medical schools for students.

Moreover, consideration should be given for the creation of mandatory training sessions for academic staff to help raise awareness of the linguistic tools available to them which can improve the effectiveness of their lectures. A conscious knowledge of the value of DMs and CS in structuring and developing the content of academic lectures can only help to further enhance their pedagogic goals.

English is the official foreign medium of instruction and communication in medical colleges and hospitals in Saudi Arabia and as such, we must continuously strive to strengthen the educational system. This research is aimed at helping fill this research gap and contribute to the

development of a stronger and more efficient EMI program for the benefit of the country and its society.



LIST OF REFERENCES

- Abalhassan, K., and Al Shalawi, H. (2000). Code-switching behavior of Arab speakers of English as a second language in the United States*. *Intercultural Communication Studies*, X (1): 179-188.
- Abriam-Yago, K., Yoder, M., and Kataoka-Yahiro, M. (1999). The Cummins model: A framework for teaching nursing students for whom English is a second language. *Journal of Transcultural Nursing*, 10: 143-149.
- Aijmer, K. (1996). *Conversational routines in English*. Longman, London
- Aijmer, K. (2002). *English discourse particles: evidence from a corpus*. Amsterdam: Benjamins.
- Aijmer, K., Foolen, A. & Simon-Vanderbergen, A. (2005). Discourse particles in the perspective of heteroglossia. In Fischer, A. (Ed.), *Approaches to discourse particles*. Amsterdam & Philadelphia: John Benjamins.
- Al-Ayed, I. and Sheik, S. (2008). Assessment of the educational environment at the College of Medicine of King Saud University, Riyadh. *Eastern Mediterranean Health Journal*, 14 (4): 953-959.
- Al-Enazi, M. (2002). *The Syntactic form and social function of Saudi Arabic- English Code Switching among Bilingual Saudis in the United States*. Unpublished PhD dissertation, Indiana University of Pennsylvania.
- Al-Gindan Y., Al-Sulaiman A., Al-Faraidy A. (2000). Undergraduate curriculum reform in Saudi medical schools: which direction to go?. *Saudi Medical Journal*, 21 (4): 324-6.
- Al-Hazimi A., Zaini R, Al-Hyiani A., Hassan N., Gunaid A., et al. (2004). Educational environment in traditional and innovative medical schools: a study in four undergraduate medical schools. *Education Health*, 71: 192-203.
- Al-Jarf, Reima (2008). The impact of English as an International Language (EIL) upon Arabic in Saudi Arabia. *Asian EFL Journal*, 10 (4): 193-210
- Al-Khatib, H. (2003). Language alternation among Arabic and English youth bilinguals: Reflecting or constructing social realities?. *International Journal of Bilingual Education and Bilingualism*, VI (6): 409-422.
- Al-Nofaie, H. (2010). Attitudes toward code-switching as facilitating tool in English classes. *Journal of Novitiate-Royal*, 4 (1): 64-95.
- Al-Seghayer, K., (2012). *Status and functions of English in Saudi Arabia*. Saudi Gazette, published December 11, 2012. Available at <http://goo.gl/70EVY>. (Accessed on October 18, 2013).
- Al-Shehri, M. (2001). Medical curriculum in Saudi medical colleges: current and future perspectives. *Annals of Saudi Medicine* ; 21: 320-3.

- Al-Shehri, M. and Campbell, S. (2013) Development of Medical Education in Saudi Arabia. In Smith, L. and Abouammoh, A. (Eds.), *Higher Education in Saudi Arabia: Achievements, challenges and opportunities*, 40: 137-149. Springer, New York.
- Al-Sulaiman, A. (2000). Saudi medical education: Challenges in the new millennium. *Journal of Family Community Medicine*, 7 (1): 15-20.
- Albar, A. and Assuhaimi, S. (1996). Attitude of medical students and postgraduate residents at King Faisal University towards teaching medicine in Arabic. *Saudi Medical Journal*, 17 (2): 230-234.
- AlEnizi, A. (2010). Students' language attitude towards using code-switching as a medium of instruction in the college of health sciences: An exploratory study. *The Annual Review of Education, Communication, and Language Sciences (ARECLS)*, 7: 1-22
- Allott, N. (2013). Relevance Theory. In Capone, A., Lo Piparo, F., and Carapezza, M. (Eds.), *Perspectives on Linguistic Pragmatics (Perspectives in Pragmatics, Philosophy & Psychology, Vol. 2)*: 57-98. Springer International Publishing.
- Allwood, J. (Ed.) (1996), *Talspråksfrekvenser (Frequency dictionary of spoken language). Gothenburg papers in theoretical linguistics S 20*. University of Göteborg: Department of Linguistics
- Almoallim, H., Aldahlawi, S., Alqahtani, E., Alqurashi, S., and Munshi, A. (2010). Difficulties facing first-year medical students at Umm Alqura University in Saudi Arabia. *Eastern Mediterranean Health Journal*, 16 (12): 1272-1277.
- Alnassar, S. and Dow, K. (2013) Delivering High-Quality Teaching and Learning for University Students in Saudi Arabia, Smith, L. and Abouammoh, A. (Eds.) *Higher Education in Saudi Arabia: Achievements, challenges and opportunities*, 40: 49-60. Springer, New York.
- Amr, A. (1996). Arabic Language and Translation: Challenges and Solutions. Proceedings of the 13th conference on *language, linguistics and translation*. Yarmouk University: Jordan.
- Anderson, A. and Lynch, T. (1988). *Listening*. Oxford University Press.
- Andersson, L. and Trudgill, P. (1990). *Bad Language*. New York: Penguin Books
- Anthony, L. (2012). AntConc (Version 3.3.5) [Macintosh]. Tokyo, Japan: Waseda University. Available from <http://www.antlab.sci.waseda.ac.jp/>
- Apperson, A. (2014) How to drink from a fire-hose without drowning – Successful study strategies in medical school (Online). Available at http://meded.ucsd.edu/index.cfm/ugme/oess/study_skills_and_exam_strategies/how_to_study_actively/ (Accessed on May 10, 2014).
- Aşık, A. and Cephe, P. (2013). Discourse Markers and Spoken English: Nonnative Use in the Turkish EFL Setting. *English Language Teaching*, 6 (12): 144-155
- Aston, G. and Burnard, L. (1998). *The BNC Handbook: Exploring the British National Corpus with SARA*. Edinburgh: Edinburgh University Press.

- Athanasidou, A. (1991). The discourse function of questions. *Pragmatics*, 1 (1): 107-122
- Atkinson, D. (1987). The mother tongue in the classroom: a neglected resource?. *ELT Journal*, 41 (4): 241-247.
- Atkinson, D. (1993). *Teaching Monolingual classes*. London: Longman.
- Auer, P. (1984). On the meaning of conversational code-switching. In Auer P and Luizo A (Eds.), *Interpretive sociolinguistics: Migrants, childre, migrant children*: 87-112. Tübingen, Germany: Gunter Narr Verlag.
- Auer, P. (1991). Bilingual in/as social action: a sequential approach to code-switching. In Papers for the symposium on *Code-switching in bilingual studies: theory, significance and perspectives*. Barcelona: European Science Foundation.
- Auer, P. (1992). Introduction: John Gumperz's approach to contextualization. In P. Auer and A. di Luzio (Eds.), *The contextualisation of language*: 1-37, Amsterdam: John Benjamins.
- Auer, P. (1995). The pragmatics of code-switching: A sequential approach. In L. Milroy and P. Muysken (Eds.), *One speaker, two languages: Cross-disciplinary perspectives on code-switching*: 115-135. Cambridge, UK: Cambridge University Press.
- Auer, P. (1998). *Code switching in conversation: Language, interaction and identity*. New York, NY: Routledge.
- Auer, P. (1999). From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism*, 3: 309-332.
- Auer, P. (2002). *Code-Switching in Conversation: Language, Interaction, and Identity*. London: Routledge
- Baker, P. (2010). *Sociolinguistics and Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Barbieri, F. (2013). Involvement in University Classroom Discourse: Register Variation and Interactivity. *Applied Linguistics*: 1-24. Oxford University Press.
- BASE. The recordings and transcriptions used in this study come from the British Academic Spoken English (BASE) corpus. The corpus was developed at the Universities of Warwick and Reading under the directorship of Hilary Nesi and Paul Thompson. Corpus development was assisted by funding from BALEAP, EURALEX, the British Academy and the Arts and Humanities Research Council.
- Bellés-Fortuño, B. (2006). *Discourse markers within the university lecture genre: A contrastive study between Spanish and North-American lectures*. Unpublished PhD dissertation. Universitat Jaume I.
- Bennett, G. (2010). *Using Corpora in the Language Learning Classroom: Corpus Linguistics for Teachers*. Michigan: University of Michigan Press.

- Benson, M. (1994). Lecture listening in an ethnographic perspective. In Flowerdew, J. (Ed.), *Academic Listening: Research Perspectives*: 181-198. Cambridge University Press: Cambridge.
- Bhatia, T. and Ritchie, W. (2004). Social and Psychological Factors in Language Mixing. In Ritchie, W. and Bhatia, T. (Eds.), *Handbook of Bilingualism*: 336–352. Wiley-Blackwell Publishers
- Bhatia, V. (1993). *Analysing genre: language use in professional settings*. London, Longman.
- Bhatia, V. (2002). A generic view of academic discourse. In Flowerdew, J. (Ed.), *Academic discourse*: 1-39. London: Longman Pearson Education.
- Bhatia, V. (2004). *Worlds of Written Discourse: A Genre-Based View*. London: Continuum.
- Biber, D. (1988). *Variation across speech and writing*. Cambridge: Cambridge University Press.
- Biber, D. (1990). Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and Linguistic Computing*, 5: 257-269.
- Biber, D. (1993). The multi-dimensional approach to linguistic analyses of genre variation: An overview of methodology and findings. *Computers and the Humanities*, 26: 331-345.
- Biber, D. (1995). On the role of computational, statistical, and interpretive techniques in multi-dimensional analyses of register variation: A reply to Watson (1994). *Text*, 15: 341-370.
- Biber, D. and Conrad, S. (2004). Corpus-Based Comparisons of Registers. In C. Coffin, A. Hewings, and K. O'Halloran (Eds), *Applying English Grammar: Functional and Corpus Approaches*: 40-56. London: Arnold
- Biber, D., Conrad, S. and Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25 (3): 371–405.
- Biber, D., Conrad, S., Reppen, R. (1994). Corpus-based approaches to issues in applied linguistics. *Applied Linguistics*, 15: 169-189.
- Biber, D., Johanson, S., Leech, G., Conrad, S., and Finegan, E. (Eds.) (1999). *Longman grammar of spoken and written English*. London: Longman.
- Blakemore, D. (1987). *Semantic Constraints on Relevance*. Oxford: Blackwell.
- Blakemore, D. (1992). *Understanding Utterances: An Introduction to Pragmatics*. Oxford: Blackwell.
- Blakemore, D. (1995). On so-called 'discourse connectives'. Unpublished lecture.
- Blom, J. and Gumperz, J. (1972). Social meaning in structure: code switching in Norway. In Gumperz, J. and Hymes, D. (Eds.) *Directions in sociolinguistics: the ethnography of communication*, 407–434. New York: Holt, Rinehart, Winston.
- Boas, H. (2010). On the equivalence and multifunctionality of discourse markers in language contact situations. In Harden, T. and Hentschel, E. (Eds.), *40 Jahre Partikelforschung*, 301-315. Tuebingen: Stauffenburg Verlag.

- Bolden, G. (2006). Little words that matter: Discourse markers 'so' and 'oh' and the doing of other-attentiveness in social interaction. *Journal of Communication*, 661-668.
- Breck, E. (2004) Soundscriber Software (Online). Available at <http://http://www-personal.umich.edu/~ebreck/code/sscriber/> (Accessed on September 10, 2010).
- Brinton, L. (1996). *Pragmatic markers in English: Grammaticalization and discourse function*. Berlin/New York: Mouton de Gruyter.
- Brock, C. (1986). The effects of referential questions on ESL classroom discourse. *TESOL Quarterly*, 20: 47-59.
- Brown, P. and Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
- Bussman, H. (2000). *Routledge Dictionary of Language and Linguistics*. London: Routledge.
- Buysse, L. (2010). Discourse markers in the English of Flemish university students. In I. Witczak-Plisiecka (Ed.), *Pragmatic Perspectives on Language and Linguistics*, (Vol. 1: Speech Actions in Theory and Applied Linguistics, 461-484). Newcastle-upon-Tyne: Cambridge Scholars Publishing.
- Camiciottoli, B. (2004). Interactive discourse structuring in L2 guest lectures: Some insights from a comparative corpus-based study. *English for Specific Purposes*, 3: 39-54.
- Canagarajah, S. (1995). Functions of Codeswitching in ESL Classrooms: Socialising Bilingualism in Jaffna. *Journal of Multilingual and Multicultural Development*, 16 (3): 173-195.
- Carter-Thomas, S. and Rowley-Jolivet, E. (2008). *If*-conditionals in medical discourse: From theory to disciplinary practice. *Journal of English for Academic Purposes*, 7: 191-205.
- Chambers, F. (1991). Promoting use of the target language in the classroom. *Language Learning Journal*, 4: 27-31.
- Chang, Y. (2010). English-medium instruction for subject courses in tertiary education: Reactions from Taiwanese undergraduate students. *Taiwan International ESP Journal*, 2 (1): 55-84.
- Chang, Y. (2012). The use of questions by professors in lectures given in English: Influences of disciplinary cultures. *English for Specific Purposes*, 31: 103-116.
- Chaudron, C. (1984). Evaluating writing: Effects of feedback on revision. *RELC Journal*, 15: 1-14.
- Chaudron, C. and Richards, J. (1986). The effect of discourse markers on the comprehension of lectures. *Journal of Applied Linguistic*, 7 (2): 113-127.
- CL, Corpus Linguistics (2012). Corpus linguistics (Online). Available at www.cl2011.org.uk (Accessed on February 10, 2013)
- Clerehan, S. (1995). Taking it down: Note-taking practices of L1 and L2 students. *English for Specific Purposes*, 14: 137-155.

- Clyne, M. (1972). *Perspectives on language contact*. Melbourne: Hawthorne Press.
- Cole, S. (1998). The use of L1 in communicative English classroom. *The Language Teacher*, 22: 11-13
- Conrad, S. (1996). A comparison of register patterns in biology and history academic writing. In paper presented in the Colloquium on *Corpus Linguistics and Language Teaching: Concordancing and Beyond*, International TESOL Convention, Chicago, IL.
- Conrad, S. (2004). Corpus linguistics, language variation, and language teaching. In J. Sinclair (Ed.), *How to use corpora in language teaching*: 67-85. Amsterdam: John Benjamins.
- Cook, V. (1975). Strategies in the comprehension of relative clauses. *Language and Speech*, 18 (3): 204-12.
- Cook, V. (2001). *Second language learning and language teaching*. Third Edition. London: Arnold
- Coulthard, M. and Montgomery, M. (1981). The Structure of Monologue. In Coulthard, M. and Montgomery, M. (Eds.), *Studies in Discourse Analysis*: 30-39. London: Routledge Kegan Paul.
- Creese, A. (2005). Is this content-based language teaching?. *Linguistics and Education*, 16 (2): 188-204.
- Crystal, D. (1988). Another look at, well, you know ... *English Today*, 13: 47-9.
- Crystal, D. (1999). World English: Past, Present, Future (Online). Paper to the ASKO Europa-Stiftung Symposium, 11-13 June. University of Wales, Bangor, UK.
Available at <http://goo.gl/64F5pv> (Accessed on October 12, 2012).
- Crystal, D. (2003). *English as a global language* (Online). Available at <http://NCL.ebib.com/patron/FullRecord.aspx?p=221205> (Accessed on September 27, 2013).
- Csomay, E. (2002). *Episodes in university classrooms: a corpus-linguistic investigation*. Unpublished PhD dissertation, Northern Arizona University.
- Csomay, E. (2005). Linguistic variation within university classroom talk: A corpus-based perspective. *Linguistics and Education*, 15: 243-274.
- Dafouz, E., Núñez, B., and Sancho, C. (2007). Analysing stance in a CLIL university context: Non-native speaker use of personal pronouns and modal verbs. *International Journal of Bilingualism and Bilingual Education*, 10 (5): 647-662.
- De Rooij, V. (1996). *Cohesion through contrast: Discourse structure in Shaba Swahili/French conversations*. Amsterdam: IFOTT.
- Del Saz-Rubio, M. (2007). *English discourse markers of reformulation: An analysis and description*. Peter Lang, Bern, Switzerland.
- Dudley-Evans, A. (1994) Genre analysis: an approach for text analysis for ESP. In M. Coulthard (Ed.), *Advances in Written Text Analysis*: 219-228. London: Routledge.

- Dudley-Evans, A. and Johns, T. (1981). A team teaching approach to lecture comprehension for overseas students. In the *Teaching of Listening Comprehension*, 30-46 (ELT Documents Special). London: The British Council.
- Dunkel, P. and Davis, J. (1994). The effects of rhetorical signalling cues on the recall of English lecture information by speakers of English as a native or second language. In Flowerdew J. (Ed.), *Academic listening: Research perspectives*, 55-74. Cambridge: Cambridge University Press.
- Eid, M. (1992). Directionality in Arabic-English code-switching. In A. Rouchdy (Ed.), *The Arabic Language in America*, 50–71. Detroit: Wayne State University Press.
- Eldridge, J. (1996). Code-switching in a Turkish secondary school. *ELT Journal*, 50 (4): 303-311
- Elfaki, E. (2004). Undergraduate curriculum reform in Saudi medical schools. *Saudi Medical Journal*, 21: 324–326.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1997). *SLA research and language teaching*. Oxford: Oxford University Press.
- Ferguson, C. (1959) Diglossia. *Word*, 15: 325-340.
- Ferguson, G. (2001). If you pop over there: A corpus-based study of conditionals in medical discourse. *English for Specific Purposes*, 20 (1): 61-82.
- Ferguson, G. (2003). Classroom codeswitching in post-colonial contexts: Functions, attitudes and policies. *AILA Review*, 16 (1): 38-51.
- Ferris, D. and Tagg, T. (1996). Academic listening/speaking tasks for ESL students: Problems, suggestions, and implications. *TESOL Quarterly*, 30 (2): 297-320.
- Flowerdew, J. (1994). Research of relevance to second language comprehension: an overview. In Flowerdew, J. (Ed.), *Academic Listening: Research Perspectives: 7-29*. New York: Cambridge University Press.
- Flowerdew, J. and Tauroza, S. (1995). The effect of discourse markers on second language lecture comprehension. *Studies in Second Language Acquisition*, 17: 435–458.
- Fortanet-Gómez, I. and Bellés-Fortuño, B. (2005). Spoken academic discourse: An approach to research on lectures. *Volumen Monográfico*, 25: 161-178.
- Fox, J. and Schrock, E. (2002). Basic meanings of *you know* and *I mean*. *Journal of Pragmatics*, 34: 727-747.
- Fraser, B. (1990). An Approach to Discourse Markers. *Journal of Pragmatics*, 14: 383- 395.
- Fraser, B. (1998). Contrastive discourse markers in English. In Jucker and Ziv (Eds.), *Discourse Markers: Descriptions and Theory*. Pragmatics and Beyond, John Benjamins.
- Fraser, B. (1999). What are discourse markers?. *Journal of Pragmatics*, 31: 931-952.

- Fraser, B. (2005). Towards a theory of discourse markers. In K. Fischer (Ed.), *Approaches to discourse particles*, 189–204. Amsterdam: Elsevier.
- Fraser, B. (2006). On the conceptual-procedural distinction. *Style*, 40 (1&2): 24-35.
- Freelon, D. (2013). ReCal OIR: Ordinal, interval, and ratio intercoder reliability as a web service. *International Journal of Internet Science*, 8 (1): 10-16.
- Fuller, J. (2001). The principle of pragmatic detachability in borrowing: English-origin discourse markers in Pennsylvania German. *Linguistics*, 39: 351-369.
- Fuller, J. (2003). Discourse marker use across speech contexts: a comparison of native and non-native speaker performance. *Multilingua*, 22: 185-208.
- Fung, L. (2003). *The use and teaching of discourse markers in Hong Kong: students' production and teachers' perspectives*. Unpublished PhD dissertation, University of Nottingham.
- Fung, L. and Carter, R. (2006). Cantonese e-discourse: A new hybrid variety of English. *Multilingua* 26 (1): 35–66.
- Fung, L. and Carter, R. (2007). Discourse markers and spoken English: Native and learner use in pedagogical settings. *Applied Linguistics*, 28 (3): 410-439.
- Gal, S. (1988). The Political Economy of Code Choice. In Monica Heller (Ed.), *Codeswitching: Anthropological and Sociolinguistic Perspectives*, 243-261. Berlin: Mouton de Gruyter.
- Gallagher, E. (1995). Medical education in Saudi Arabia: A sociological perspective in modernization and language. *Journal of Asian and African Studies*, XX (1-2): 1-12.
- Geertzen, J. (2012). Inter-Rater Agreement with multiple raters and variables (Online). Available at <https://mlnl.net/jg/software/ira/> (Accessed on February 4, 2014).
- Goffman, E. (1981). *Forms of Talk*. Philadelphia: University of Pennsylvania Press.
- Goldberg, J. (1980). *Discourse particles: An analysis of the role of y'know, I mean, well, and actually in conversation*. Unpublished PhD dissertation, Cambridge University.
- Goss, E. and Salmons, J. (2000). The evolution of a bilingual discourse marking system: Modal particles and English markers in German-American dialects. *International Journal of Bilingualism*, 4: 469-484.
- Grabe, W. (1987). Contrastive rhetoric and text-type research. In Connor, U. and Kaplan, R. (Eds.), *Writing across languages: Analysis of L2 texts*. Reading, MA: Addison-Wesley Publishing Co.
- Graddol, D. (1997). *The Future of English? A guide to forecasting the popularity of the English language in the 21st century*. The British Council.
- Graddol, D. and Meinhof, U.H. (Eds.) (1999). Introduction. In Graddol, D. and Meinhof, U. (Eds.). *English in a Changing World. AILA*, Oxford: The English Company, UK Ltd.
- Greenbaum, S. (1996). Introducing ICE. In Greenbaum (Ed.), *Comparing English Worldwide: The International Corpus of English*, 3-12. Oxford: Clarendon Press.

- Grice, H. P. (1975). Logic and conversation. In Cole, P. and Morgan, J. (Eds.), *Syntax and Semantics, Speech Acts*, 3: 41-58. New York City: Academic Press.
- Griffiths, R. (1990). Speech rate and NNS comprehension: a preliminary study in time-benefit analysis. *Language Learning*, 40: 311-336.
- Grosjean, F. (1982). *Life with two languages. An introduction to bilingualism*. Cambridge, Mass: Harvard University.
- Gumperz, J. (1977). Sociocultural knowledge in conversational inference. In M. Saville-Troike (Ed.), *Georgetown Round Table on Languages and Linguistics*: 191-212. Washington, D.C.: Georgetown University Press.
- Gumperz, J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.
- Halliday, M. A. (1994). *Functional grammar*. London: Edward Arnold.
- Halliday, M. and Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Halliday, M. and Hassan, R. (1985). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Victoria: Deakin University
- Hamdy, H., Telmesani, A., Al Wardy, N., Abdel-Khalek, N., Carruthers, G., et al. (2010). Undergraduate medical education in the Gulf Cooperation Council: a multi-countries study (Part 1). *Medical Teacher*, 32: 219-224.
- Heller, M. (1992). The politics of codeswitching and language choice. *Journal of Multilingual and Multicultural Development*, 13: 123-142.
- Heller, M. (1995b). Code-switching and the politics of language. In Milroy, L. and Muysken, P. (Eds.), *One speaker two languages: Cross-disciplinary perspectives on code-switching*: 158-174. New York: Cambridge University Press.
- Heller, M. (Ed.) (1988). *Codeswitching: anthropological and sociolinguistic perspectives*. Berlin: Mouton de Gruyter.
- Hellerman, J. and Vergun, A. (2006). Language which is not taught: The discourse marker use of beginning adult learners of English. *Journal of Pragmatics*, 39 (1): 157-179.
- Henry, A. and Roseberry, R. (2001). Using a small corpus to obtain data for teaching a genre. In Ghadessy, M., Henry, A. and Roseberry, R. (Eds), *Small Corpus Studies and ELT: Theory and practice*, 93-133.
- Holmes, J. (1997). Women, language and identity. *Journal of Sociolinguistics*, 1: 195-224.
- Holmes, J. (2000). *An Introduction to Sociolinguistics* (2nd Ed.). Wellington: Longman.
- House, J. (2010). The Pragmatics of English as a lingua franca. In Trosborg, A. (Ed.), *Pragmatics across languages and cultures*: 363–387.
- Howe, M. (1991). *Topic Change in Conversation*. Unpublished PhD Dissertation, University of Kansas.
- Hudson, R. (1996). *Sociolinguistics*. Cambridge: Cambridge University Press.

- Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.
- Hussein, R. (1999). Code-alteration among Arab college students. *World Englishes*, 18 (2): 281-289.
- Hussein, R. and Shorrab, G. (1993). Syntactic constraints on the code switching of Arabic English bilinguals. *International Review of Applied Linguistics Teaching*, 31 (3): 236-41.
- Hussein, R. and Zughoul, M. (1993). Lexical Interference in Journalistic Arabic in Jordan. *Language Sciences*, 5 (3): 239-54.
- Hyland, K. (1997). Is EAP necessary? A Survey of Hong Kong Undergraduates. *Asian Journal of English Language Teaching*, 7: 77-99.
- Hymes, D. (1967). Models of the interaction of language and social settings. *Journal of Social Issues*, 23 (2): 8-28.
- Hymes, D. (1972). Models of the interaction of language and social life. In Gumperz, J. and Hymes, D. (Eds.), *Directions in sociolinguistics: The ethnography of communication*, 35-71. New York: Holt, Rinehart and Winston.
- Jabeen, F., Rai, M., and Arif, S. (2011). A corpus based study of discourse markers in British and Pakistani speech. *International Journal of Language Studies*: 69-86.
- Jacobsson, C. (2010). *A Study of Code Switching in Four Swedish EFL Classroom* (Online). Available at <http://dspace.mah.se/bitstream/handle/2043/9576/Astudyofcodeswitching.pdf?sequence=1> (Accessed on December 17, 2013).
- Jake, J. (1994). Intrasentential code switching and pronouns: On the categorial status of functional elements. *Linguistics*, 32: 271-98.
- Jenkins, S. (2010). Monolingualism: An uncongenial policy for Saudi Arabia's low-level learners. *ELT Journal*, 64 (4): 459-461.
- Johns, A. (1997). *Text, role, and context: Developing academic literacies*. Cambridge University Press.
- Johnson, A. (2002). So . . . ? : Pragmatic Implications of So-Prefaced Questions in Formal Police Interviews. In Cotterill, J. (Ed.), *Language in the legal process*: 91-110. London: Palgrave.
- Jordan, R. (1989). English for academic purposes (EAP). *Language Teaching*, 22 (3): 151-164.
- Jucker, A. and Smith, S. (1998). And people just you know like “wow”: Discourse markers as negotiating strategies. In Jucker, A. and Ziv, Y. (Eds), *Discourse markers: Descriptions and theory*. Amsterdam. John Benjamins B.V.
- Jucker, A. and Ziv, Y. (1998). *Discourse Markers: Descriptions and Theory*. Amsterdam and Philadelphia: John Benjamins.
- Jung, E. (2003). The role of discourse signaling cues in second language listening comprehension. *The Modern Language Journal*, 87: 562-575

- Khalid B. (2008). The current status of medical education in the Gulf Cooperation Council countries. *Annals of Saudi Medicine*, 28: 83–88.
- Khuwaileh, A. (1999). The role of chunks, phrases and body language in understanding coordinated academic lectures. *System*, 27: 249-260.
- Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford, UK: Pergamon.
- Krashen, S. (1985). *The input hypothesis*. London: Longman
- Krashen, S. (1987). *Principles and practice in second language acquisition*. Prentice-Hall International.
- Krashen, S. (1988). *Second Language Acquisition and Second Language Learning*. Prentice-Hall International.
- Krieger, D. (2003). Corpus Linguistics: What is it and how it can be applied to teaching. *The ITESL Journal*, 9 (3). (Online). Available at <http://iteslj.org/Articles/Krieger-Corpus.html> (Accessed on November 10, 2012).
- Kwo, O. (1987). Bilingual secondary education in Hong Kong: What are the options?. *Working Papers in Linguistics and Language Teaching*, 10: 1-10.
- Labov, W. (1972). *Language in the inner city: Studies in the Black English vernacular*. Philadelphia University of Pennsylvania Press.
- Lawson, S. and Sachdev, I. (2000). Codeswitching in Tunisia: Attitudinal and behavioural dimensions. *Journal of Pragmatics*, 32: 1343-1361.
- Lebauer, R. (1983). Using Lecture Transcripts in EAP Lecture Comprehension Courses. Paper presented at the *Annual Convention of TESOL*, 1-19 (Online). Available at <http://files.eric.ed.gov/fulltext/ED230040.pdf> (Accessed on April 20, 2014).
- Leech, G. (1992). Corpora and theories of linguistic performance. In Startvik, J. (Ed.), *Directions in Corpus Linguistics*, 105-122. Berlin: Mouton de Gruyter.
- Lenk, U. (1998). *Marking discourse coherence: Functions of discourse markers in spoken English*. Gunter Narr Verlag: Tübingen.
- Levin, H. and Gray, D. (1983). The lecturer's OK. *American Speech*, 58: 195-200.
- Levinson, S. (1979). Pragmatics and social deixis: Reclaiming the notion of conventional implicature. In Proceedings of the Fifth Annual Meeting of the *Berkeley Linguistics Society*, 206-223. Berkeley, CA.
- Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Lightbown, P. (2001). L2 Instruction: time to teach. *TESOL Quarterly*, 35: 598-99.
- LINDSEI. Louvain International Database of Spoken English Interlanguage (Online). Available at <http://www.uclouvain.be/en-cecl-lindsei.html> (Accessed on May 10, 2014).
- Long, M. (1981). Questions in foreigner talk discourse. *Language Learning*, 31 (1): 135–158.

- Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics*, 4 (2): 126–141.
- Lutzky, U. (2006). Discourse markers? Well,... Delimiting the basic features of discourse markers. Vienna English Working Papers, 15 (1): 3-24 (Online). Available at http://anglistik.univie.ac.at/fileadmin/user_upload/dep_anglist/weitere_Uploads/Views/views0601.pdf
- Lynch, T. (1994). Training lecturers for international audiences. In Flowerdew, J. (Ed.), *Academic Listening: Research Perspectives*, 269-289. Cambridge. Cambridge University press.
- MacDonald, C. (1993). *Using the target language*. Cheltenham, UK: Mary Glasgow Publications.
- Martin-Jones, M. (1995). Codeswitching in the classroom: two decades of research. In Milroy, L. and Muysken, P. (Eds.), *One Speaker; Two Languages: Cross-Disciplinary Perspectives on Codeswitching*, 90-111. Cambridge University Press, Cambridge.
- Maschler, Y. (1998). Rotse lishmoa keta? Wanna hear something weird/funny? Segmenting Israeli Hebrew talk-in-interaction. In Jucker, A. and Ziv, Y. (Eds.), *Discourse Markers: Description and Theory*, 13-60. Amsterdam: John Benjamins.
- Maschler, Y. (2000). What can bilingual conversation tell us about discourse markers? Introduction. *International Journal of Bilingualism*, 4 (4): 437-445.
- Matras, Y. (2000). Fusion and the cognitive basis for bilingual discourse markers. *International Journal of Bilingualism*, 4 (4): 505–528.
- Mauranen, A. (2003). The corpus of English as lingua franca in academic settings. *TESOL Quarterly*, 37 (3): 513–527.
- Mazi, A. and Altbach, P. (2013). Dreams and Realities: The World-Class Idea and Saudi Arabian Higher Education. In Smith, L. and Abouammoh, A. (Eds.), *Higher Education in Saudi Arabia: Achievements, challenges and opportunities*, 40: 13-26. Springer, New York.
- McCarthy, M. (1991). *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.
- McEnery, T., Xiao, R., and Tono, Y. (2006). *Corpus-based language studies*. London and New York: Routledge.
- Mendelsohn, D. (1984). There ARE strategies for listening. *TEAL Occasional Papers*, 8: 63-76.
- Merritt, M., Cleghorn, A., Abagi, J. and Bunyi, G. (1992). Socialising Multilingualism: Determinants of Codeswitching in Kenyan Primary Classrooms. *Journal of Multilingual and Multicultural Development*, 13 (1-2): 103-121.
- MOH, Ministry of Health – Saudi Arabia. (2009). *Statistical Book for 2009* (Online). Riyadh: Ministry of Health. Available at <http://www.moh.gov.sa/en/Ministry/Statistics/Indicator/Pages/default.aspx>. (Accessed on July 22, 2013)

- Morell, T. (2000). EFL content lectures: a discourse analysis of an interactive and a non-interactive style. Spain: Departamento de Filología Inglesa, Universidad de Alicante (Working papers, 7).
- Morell, T. (2004). Interactive lecture discourse for university EFL students. *English for Specific Purposes*, 23: 325-338.
- Morrison, J.W. (1974). *An investigation of problems in listening comprehension encountered by overseas students in the first year of postgraduate studies in the University of Newcastle-upon-Tyne*. Unpublished MA thesis, University of Newcastle-upon-Tyne, England.
- Mujiono, M., Poedjosoedarmo, S., Subroto, E., and Wiratno, T. (2013). Code Switching in English as foreign language instruction practiced by the English lecturers at universities. *International Journal of Linguistics*, 5 (2): 46-65
- Müller, S. (2005). *Discourse markers in native and non-native English discourse*. Amsterdam: Benjamins.
- Murphy, D., and Candlin, C. (1979). Engineering lecture discourse and listening comprehension. *Practical papers in Language Education*, 2: 1-79. Lancaster: University of Lancaster.
- Myers-Scotton, C. (1983). The negotiation of identities in conversation: A theory of markedness and code choice. *International Journal of the Sociology of Language*, 44: 116-136.
- Myers-Scotton, C. (1993a [1997]). *Duelling languages: Grammatical structure in codeswitching* (1997 edition with a new Afterword). New York: Oxford University Press.
- Myers-Scotton, C. (1993b). *Social motivations for codeswitching: Evidence from Africa*. Oxford: Oxford University Press.
- Myers-Scotton, C. (1998). A theoretical introduction to the markedness model. In Myers-Scotton, C. (Ed.), *Codes and consequences*. New York: Oxford University Press, 18–38.
- Myers-Scotton, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Myers-Scotton, C. (2006). *Multiple voices: An introduction to bilingualism*. Malden, MA: Blackwell Publishing.
- Myers-Scotton, C. and Jake, J. (1995). Matching lemmas in a bilingual language competence and production model: evidence from Intrasentential Codeswitching. *Linguistics*, 33: 981-1024
- Naffee, I. (2013). *KSA has 9.2 million expatriates, says study*. Arab News article, published October 11, 2013. Available at <http://www.arabnews.com/news/467423>. (Accessed on October 18, 2013)
- Nattinger, J. and DeCarrico, J. (1988). Lexical phrases for the comprehension of academic lectures. *English for Specific Purposes*, 7: 91-102.

- Nattinger, J. and DeCarrico, J. (1992). *Lexical Phrases and Language Teaching*. Oxford University Press.
- Nerghes, A. (2011). *The impact of code-switching on persuasion: An elaboration likelihood perspective*. Wageningen University.
- Nesi, H. (2001). A corpus-based analysis of academic lectures across disciplines. In Cotterill, J. and Ife, A. (Eds), *Languages Across Boundaries*, London: BAAL in association with Continuum Press: 201-218.
- Nesi, H. (2012). Laughter in university lectures. *Journal of English for Academic Purposes*, 11 (2): 79-89.
- Nesi, H. and Basturkmen, H. (2006). Lexical bundles and discourse signaling in academic lectures. *International Journal of Corpus Linguistic*, 11 (3): 169-187.
- Nesi, H., and Thompson, P. (2003). The British Academic Spoken English (BASE) corpus (Online). Available at <http://www.rdg.ac.uk/slals/base/> (Accessed on September 1, 2010).
- Nilep, C. (2006). “Code Switching” in sociocultural linguistics. *Colorado Research in Linguistics*, 19: 1-22. Boulder: University of Colorado.
- Northcott, J. (2001). Towards an Ethnography of the MBA classroom a consideration of the role of interactive lecturing styles within the context of one MBA Programme. *English for Specific Purposes*, 20: 15-37
- Olsen, L., and Huckin, T. (1990). Point-driven understanding in engineering lecture comprehension. *English for Specific Purposes*, 9: 33–47.
- Othman, Z. (2010). The use of *okay*, *right* and *yeah* in academic lectures by native speaker lecturers: Their ‘anticipated’ and ‘real’ meanings. *Discourse Studies*, 12 (5): 665-681
- Partington, A. (1998). *Patterns and meanings: Using corpora for English language research and teaching*. Amsterdam: John Benjamins.
- Passot, F. (2007). ‘A because B so A’. Circularity and discourse progression in conversational English. In Celle, A. and Huart, R. (Eds.), *Connectives as Discourse Landmarks*, 117-134. Amsterdam/Philadelphia: John Benjamins.
- Pearson, J. (1998). *Terms in Context*. John Benjamins Publishing.
- Pfaff, C. (1976). Functional and structural constraints on syntactic variation in code-switching. In Steever, B., et al. (Eds.), *Papers from the Para session on diachronic syntax*: 248-59. Chicago: Chicago Linguistic Society
- Pfaff, C. (1979). Constraints on language-mixing: Intrasentential code-switching and borrowing in Spanish/English. *Language*, 55: 291-318.
- Pica, T. (1994). Review article, research on negotiation what does it reveal about second-language learning conditions, processes, and outcomes?. *Language Learning*, 44 (3): 493-527.

- Pica, T., Young, R., and Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly*, 21: 737-758.
- Poplack, S. (1980). Sometimes I start a sentence in English y termino en espanol: toward a typology of code-switching. *Linguistics*, 18: 581–616.
- Quirk, R. (1960). Towards a description of English usage. *Transactions of the Philological Society*, 59: 40–61.
- Redeker, C. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, 14: 387-381.
- Redeker, G. (1991). Review article: linguistic markers of discourse structure. *Linguistics*, 29 (6): 1139-1172.
- Rendle-Short, J. (2000). When 'okay' is okay in computer science seminar talk. *Australian Review of Applied Linguistics*, 22: 19-33.
- Rendle-Short, J. (2003). So what does that show us?: Analysis of the discourse marker “so” in the academic monologue. *Australian Review of Applied Linguistics*, 26 (3): 46–62.
- Riehl, C. (2005). Code-switching in Bilinguals: Impacts of Mental Processes and Language Awareness. In Cohen, J., McAlister, K., Rolstad, K., and MacSwan, J. (Eds), *ISB4: Proceedings of the 4th International Symposium on Bilingualism, 1945-1959*. Cascadilla Press, Somerville, MA.
- Romaine, S. (1995). *Bilingualism*. Second edition. Oxford: Blackwell.
- Romero-Trillo, J. (1997). Your attention please: Pragmatic mechanisms to obtain the addressee’s attention in English and Spanish conversations. *Journal of Pragmatics*, 28: 205-221.
- Romero-Trillo, J. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics*, 34: 769-784.
- Rounds, P. (1987a). Multifunctional personal pronoun use in an educational setting. *English for Specific Purposes*, 6 (1): 13–29.
- Rounds, P. (1987b). Characterizing successful classroom discourse for NNS teaching assistant training. *TESOL Quarterly*, 21 (4): 643–669.
- Sabbour, S., Dewedar, S. and Kandil, S. (2010). Language barriers in medical education and attitudes towards Arabization of medicine: student and staff perspectives. *Eastern Mediterranean Health Journal (EMHJ)*, 16 (12): 1263-1271
- Sacks, H., Schegloff, E. and Jefferson, G. (1974). A Simplest Systematics for the Organisation of Turn-Taking for Conversation. *Language*, 50 (4i): 696-735.
- Safi, S. (1992). Functions of code-switching: Saudi Arabic in the United States. In A. Rouchdy (Ed.), *The Arabic Language in America*: 72–80. Detroit: Wayne State University Press.
- Salmons, J. (1990). Bilingual discourse markers: Code switching, borrowing, and convergence in some German-American dialects. *Linguistics*, 28: 453-480.

- Sanchez, R. (1994). *Chicano Discourse: Socio-Historic Perspectives*. Houston, TX: Arte Publico Press.
- Sankoff, D., and Poplack, S. (1981). A formal grammar for code-switching. *Papers in Linguistics, International Journal of Human Communication*, 14 (1): 3-45
- Schegloff, E. (2006). Interaction: The infrastructure for social institutions, the natural ecological niche for language and the arena in which culture is enacted. In Enfield, N. and Levinson, S. (Eds), *The roots of human sociality: Culture, cognition and interaction*. New York: Berg.
- Schiffrin, D. (1986). Functions of and in discourse. *Journal of Pragmatics*, 10: 41-66
- Schiffrin, D. (1987). *Discourse markers*. Cambridge: Cambridge University Press.
- Schiffrin, D. (1994). *Approaches to Discourse*. Oxford: Blackwell.
- Schiffrin, D. (2001). Discourse markers: language, meaning and context. In Schiffrin, D., Tannen, D, Hamilton, H. (Eds), *The Handbook of Discourse Analysis: 54-75*. Oxford: Basil Blackwell.
- Schleef, E. (2005). Gender, power, discipline, and context: On the sociolinguistic variation of *okay, right, like, and you know* in English academic discourse. In Sunakawa, C., Ikeda, T., Finch, S. and Shetty, M. (Eds.), *Texas Linguistic Forum*, 48: 177-186. Austin, Texas: Department of Linguistics.
- Schleef, E. (2008). The 'lecturer's ok' revisited: Changing discourse conventions and the influence of academic division. *American Speech*, 83 (1): 62-84.
- Schleppegrell, M. (1991). Paratactic because. *Journal of Pragmatics*, 16: 121–135.
- Schmitt, N. (2000). *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Schourup, L. (1985). *Common discourse particles in English conversation: like, well, y'know Fundamentals*. Garland, New York.
- Schourup, L. (1999). Discourse markers. *Lingua*, 3 (4): 227-265.
- Seargeant, P. (2009) *The Idea of English in Japan: Ideology and the evolution of a global language* (Online). Available at <http://NCL.ebib.com/patron/FullRecord.aspx?p=449895> (Accessed on October 13, 2013).
- Sebai, Z. (1982). Why Do Not We Teach Medicine in Arabic? (Arabic). *Proceedings of VII Saudi Medical Meeting: 74–75*. Dammam: King Faisal University.
- Sert, O. (2005). The Functions of code switching in ELT classrooms. Hacettepe University. *The Internet TESL Journal* (Online). Available at <http://iteslj.org/Articles/Sert-CodeSwitching.html> (Accessed on July 25, 2012).
- Simpson, R., Briggs, S., Ovens, J. and Swales, J. (2002). *Michigan Corpus of Academic Spoken English* (Online). Ann Arbor, MI: The Regents of the University of Michigan Available at <http://micase.umdl.umich.edu/m/micase/> (Accessed on July 25, 2010).

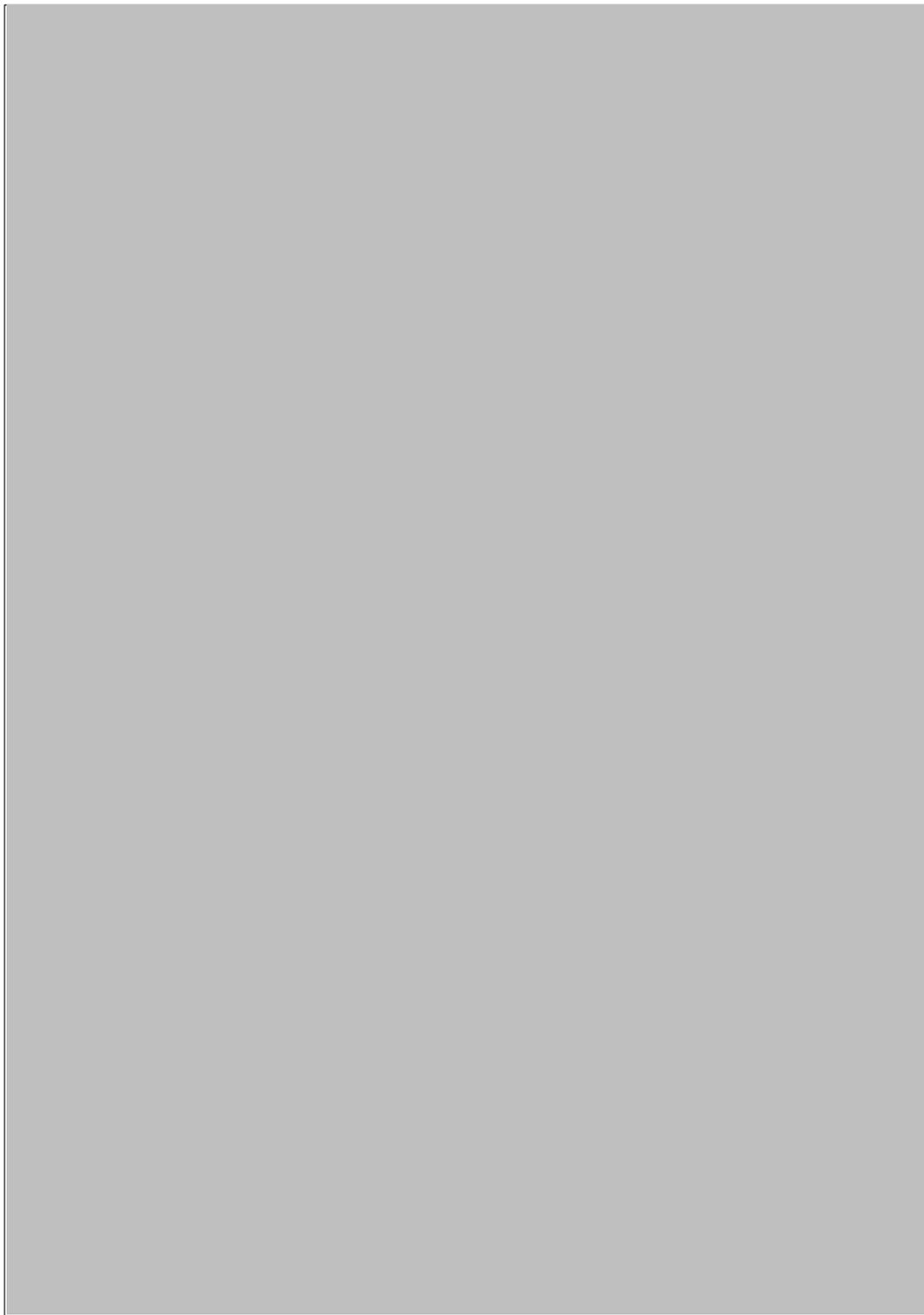
- Sinclair, J. and Coulthard, R. (1975). *Towards an analysis of discourse: the English used by teachers and pupils*. Oxford: Oxford University Press
- Smit, T. (2009). The Academic Lecture: Learn to Listen or Listen to Learn?. *Nawa Journal of Language and Communication*, 3 (1): 1-18
- Smith, L. and Abouammoh, A. (2013). Challenges and opportunities for Higher Education in Saudi Arabia: An exploratory focus group, Smith, L. and Abouammoh, A. (Eds.) *Higher Education in Saudi Arabia: Achievements, challenges and opportunities*, Springer, New York, 40: 167-179.
- Smith, L. and Abouammoh, A. (2013). Higher Education in Saudi Arabia: Reforms, challenges and priorities, Smith, L. and Abouammoh, A. (Eds.) *Higher Education in Saudi Arabia: Achievements, challenges and opportunities*, Springer, New York, 40: 1–12.
- Steinert and Snell (1999). Interactive lecturing: strategies for increasing participation in large group presentations. *Medical Teacher*, 21 (1): 39-42
- Stroud, C. (1998). Perspectives on Cultural Variability of Discourse and some Implications for Code-switching. In Auer, P. (Ed.), *Code-Switching in Conversation: Language, Interaction and Identity*, 321-348. London: Routledge.
- Suleiman, Y. (1999). *Language and society in the Middle East and North Africa: Studies in variation and identity*. UK: Curzon Press
- Suleiman, Y. (2003). *The Arabic language and national identity: A study in ideology*. Edinburgh University Press.
- Suliman, W. and Tadros, A. (2010). Nursing students coping with English as a foreign language medium of instruction. *Nurse Education Today*, doi:10.1016/j.nedt.2010.07.014
- Suviniitty, J. (2012). *Lectures in English as a Lingua Franca: Interactional Features*. Unpublished PhD dissertation, University of Helsinki.
- Swain, M. and Lapkin, S. (2000). Task-based second language learning: the uses of the first language. *Language Teaching Research*, 4 (3): 251-274.
- Swales, J. (1981). Aspects of article introductions. *ESP Monograph No. 1 Language Studies Unit*: Aston University
- Swales, J. (1990). *Genre Analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Swales, J. (2001). Metatalk in American academic talk. The cases of point and thing. *Journal of English Linguistics*, 29 (1): 34–54.
- Swales, J. (2004). *Research genres: Explorations and applications*. Cambridge: Cambridge University Press.
- Swales, J. and Malczewski, B. (2001). Discourse management and new-episode flags. In R.C. Simpson and J.M. Swales (Eds), *Corpus Linguistics in North America*. The University of Michigan Press.

- Swales, J., and Malczewski, B. (1995). Attention-getters and new episode flags in MICASE. NASCLLT abstracts. Available at <http://www.lsa.umich.edu/eli/micase/abstracthtml>
- Taboada, M. (2006). Discourse markers as signals (or not) of rhetorical relations. *Journal of Pragmatics*, 38: 567–592
- Tardy, C. (2004). The role of English in scientific communication: lingua franca or Tyrannosaurus Rex? *Journal of English for Academic Purposes*, 3: 247–269.
- Tauroza, S. and Allison, D. (1990). Speech rates in British English. *Applied Linguistics*, 11 (1): 90-105.
- Taweel, A. and Btoosh, M. (2010). Burke and Austin: Comparing Insights on Studying Meaning. *European Journal of Scientific Research*, 40 (3): 478-491.
- Taweel, A. and Btoosh, M. (2012). Syntactic aspects of Jordanian Arabic-English intrasentential code-switching. *Linguistica Online Journal*, 4 (13): 1-21.
- Tehrani, A. and Dastjerdi, H. (2012). The pedagogical impact of discourse markers in the lecture genre: EFL learners' writings in focus. *Journal of Language Teaching and Research*, 3 (3): 423–429.
- Telmesani, A., Zaini, R., and Ghazi, H. (2011). Medical education in Saudi Arabia: a review of recent developments and future challenges. *Eastern Mediterranean Health Journal*, 17 (8): 703-707.
- Thompson, G. (1996). Voices in the text. Discourse perspectives on language reports. *Applied Linguistics*, 17 (4): 501-530.
- Thompson, S. (1994). Frameworks and contexts a genre-based approach to analysing lecture introductions. *English for Specific Purposes*, 13 (2): 171–187.
- Thompson, S. (2003). Text-structuring metadiscourse, intonation and the signaling of organisation in academic lectures. *Journal of English for Academic Purposes*, 2 (1): 5-20.
- Thornbury, S. and Slade, D. (2006). *Conversation: From description to pedagogy*. Cambridge: Cambridge University Press.
- Tien, C and Liu, K. (2006). Code-switching in two EFL classes in Taiwan. In Hashim, A. and Hassan, N. (Eds), *English in Southeast Asia: prospects, perspectives and possibilities*. Kuala Lumpur: University Malaysia Press.
- Timm, A. (1975). Spanish-English code-switching: El Porque Y How-Not To. *Romance Philology*, 18 (4): 437–82.
- Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam and Philadelphia: John Benjamins.
- Toribio, A. (2004). Spanish/English speech practices: Bringing chaos to order. *Bilingual Education and Bilingualism*, 7 (2&3): 133-154.
- Trudgill, P. (1992). *Introducing language and society*. London: Penguin.

- Trudgill, P. (2000). *Sociolinguistics: an introduction to language and society*. Fourth Edition. Harmondsworth: Penguin
- Trudgill, P. (2006). *New-dialect formation: The inevitability of colonial Englishes*. Edinburgh, Edinburgh.
- Turnbull, M. and Dailey-O'Cain, J. (2009). Introduction. In Turnbull, M. and Dailey-O'Cain, J. (Eds), *First language use in second and foreign language learning*, 1-14. Clevedon: Channel View Publications.
- UoL. University of Leicester Medicine (Online). Available at <http://www2.le.ac.uk/departments/msce/undergraduate/medicine/international-entry-requirements> (Accessed on January 20, 2014).
- Üstünel, E. and Seedhouse, P. (2005). Why that, in that language, right now? Code-switching and pedagogical focus. *International Journal of Applied Linguistics*, 15 (3): 302-325.
- Waer, H. (2012). *Why that language, in that context, right now?: The use of the L1 in L2 classroom interaction in an Egyptian setting*. Unpublished PhD dissertation. Newcastle University.
- Webber, P. (2005). Interactive features in medical conference monologue. *English for Specific Purposes*, 24: 157–181.
- Williams, G.C. (2003). From Meaning to Words and Back: Corpus Linguistics and Specialised Lexicography. *Asp [En ligne]*: 91-106.
- Yang, S. (2011). Investigating discourse markers in pedagogical settings: A literature review. *ARECLS*, 8: 95-108.
- Yang, X. (2007). The Influence of Discourse Organizational Patterns on Chinese EFL Learners' Listening Comprehension. *US-China Foreign Language*, 5 (3): 22-31.
- Young, L. (1994). University lectures: macro-structure and micro-features. In Flowerdew, J (Ed.), *Academic Listening: Research Perspective*, 159-176. Cambridge: Cambridge University Press.
- Yu, L. (2008). Interpersonal meaning of discourse markers by foreign language teachers in classroom. *Journal of the Graduates of Sun Yat-Sen University (Social Sciences)*, 29 (3): 129-135.
- Zabrodskaja, A. (2007). Russian-Estonian code-switching in the university. *Arizona Working papers in SLA and Teaching*, 14, 123-139.

APPENDICES

APPENDIX A: INFORMED CONSENT LETTER





APPENDIX B: EXPLORATORY STUDY

Appendix B1. Exploratory Study NNS English SAMEX Corpus Summary

Lecture Title	Academic Dep.	Lec. Gender	Native Speaker Status	Nationality	Audience	Tokens
1. <i>Growth & Development</i>	Nursing	F	NNS	Egyptian	UG4	3857
2. <i>Medical Terminology</i>	Basic Sciences	F	NNS	Lebanese	UG2	2654
3. <i>Bioenergetics</i>	Basic Sciences	F	NNS	Lebanese	UG2	4427
4. <i>Family Health & Child Protection</i>	Family Medicine	M	NNS	Saudi	UG4	3603
5. <i>Fungi</i>	Microbiology	M	NNS	Saudi	UG2	2405
6. <i>Antibiotics</i>	Pharmacology	M	NNS	Saudi	UG2	2455
7. <i>Tuberculosis</i>	Pathology	M	NNS	Sudanese	UG2	3142
8. <i>Cell Mediated Immunity</i>	Pathology	M	NNS	Sudanese	UG2	2055
9. <i>Amino Acids</i>	Biochemistry	F	NNS	Syrian	UG2	1577
10. <i>Complementary System</i>	Medicine	M	NNS	Sudanese	UG2	2377
11. <i>Pathology of Microbes</i>	Microbiology	M	NNS	Sudanese	UG2	2838
12. <i>Pathology</i>	Pathology	F	NNS	Saudi	UG2	3406
						34796

Appendix B2. Exploratory Study NS English BASEX Corpus Summary

Lecture Title	Academic Department	Gender	Native Speaker Status	Nationality	Audience	Tokens
---------------	---------------------	--------	-----------------------	-------------	----------	--------

1. <i>Immunology</i>	Animal & Microbial Sciences	M	NS	UK	UG	6380
2. <i>HIV & AIDS</i>	Biological Sciences	M	NS	UK	UG	6278
3. <i>Intro to the Laboratory</i>	Biological Sciences	M	NS	UK	UG1	5060
4. <i>Sources of Variation I</i>	Statistics	F	NS	UK	UG/PG	5671
5. <i>Sources of Variation I</i>	Statistics	F	NS	UK	UG/PG	3710
6. <i>Abdominal Aortic Aneurysms Trial & Error</i>	Medicine	F	NS	UK	UG/PG	6053
7. <i>The Kidney & Calcium Homeostasis</i>	Medicine	F	NS	UK	UG2/PG	5090
8. <i>Kidney Stones</i>	Medicine	M	NS	UK	UG2/PG	6570
9. <i>HIV Infection</i>	Biological Sciences	M	NS	UK	UG3	5829
10. <i>Introduction to DISC (Developing Interview Skills in the Consultation)</i>	Medicine	F	NS	UK	UG1/PG	6428
						57069

Appendix B3. Description of *Okay1* with Exploratory Corpus Excerpts

Okay1 confirmation	The first and most frequent function of the discourse marker ‘ <i>okay</i> ’ was designed to elicit confirmation of understanding, meaning it acted as a method of checking student comprehension.
---------------------------	--

<p>SAMEX <i>Medical Terminology</i></p>	<p><i>...we started talking about cytology which is the study of what and then we studied trends related to tissue what is the study of the tissue histology okay so we went from cell to tissue and we said that we have four types of tissue what are the four types of tissue who can remind us who can create actually because it is written in front of you...</i></p>
<p>BASEX <i>HIV & Aids</i></p>	<p><i>...infection that is not due to a drop in C-D-four cells it s due actually to an increase in C-D-eight cells okay you may see some atypical lymphocytes in blood smears but you may not okay so the picture I m painting then of stage one is somebody who s infected probably has no idea they re infected has no symptomology okay and...</i></p>

Appendix B4. Description of *Okay2* with Exploratory Corpus Excerpts

<p>Okay2 <i>opening/ closing frame</i></p>	<p>The second ‘<i>okay</i>’ discourse marker functioned as a signal which indicated the beginning or end of a new topic or idea by framing the proposition.</p>
<p>SAMEX <i>Fungi</i></p>	<p><i>...nuclear material is dispersed into the cell and it cannot be distinguished under the microscope or you cannot see it and identify it also you can not see the other organisms like the mitochondrion and so on and so forth okay additionally bacteria are small bacteria are much smaller than fungi they are one micron wide or less right the fungi are bigger they are about maybe 2 μ microns or even more right so these are the major differences between the bacteria and the fungi...</i></p>
<p>BASEX <i>Kidney Stones</i></p>	<p><i>...this is a hard one isn t it tell me the normal calcium there you go that s an easier question gentleman at the front you’re going to get poor chaps [laughter] going to get picked on chap behind him normal calcium level [sm0393:] two-point-five in plasma [nm0370:] okay not bad okay two-point-two to two-point-six so you should know the l-, the normal levels of all basic biochemical parameters by now and if you don t I d like you to learn them because you re going to need them by the time you hit the wards...</i></p>

Appendix B5. Description of *Okay3* with Exploratory Corpus Excerpts

<p>Okay3 <i>acceptance</i></p>	<p>The third ‘<i>okay</i>’ functioned as an indicator of agreement or encouragement with a statement or action, in other words, lecturers utilized this discourse marker to indicate that they were listening and to prompt students to action.</p>
---	---

SAMEX <i>Growth & Development</i>	<i>[student begins reading, then hesitates]</i> <i>[lecturer] please read your part</i> <i>[student continues reading]</i> <i>[lecturer] okay is there a relationship between intelligence and in this theory we will reach in what</i> <i>[student answers]</i> <i>[lecturer] okay</i> <i>[student continues explaining answer]</i> <i>[lecturer] this is the process okay thank you namex yes who else can give me another definition</i>
BASEX <i>HIV & Aids</i>	<i>[lecturer] cause they re completely non-thinking people who couldn't be physicians <laughter> is that what you said</i> <i>[student] yeah</i> <i>[lecturer] <laughter> okay okay sorry I don t I don t don t let me put words in your mouth though <laughter> er oh that s on camera as well <laughter> er no surgeons are good because they know how to get rid of the ...</i>

Appendix B6. Functional Distribution of DM 'okay' in Exploratory Corpora

	<i>Okay1</i>	<i>Okay2</i>	<i>Okay3</i>
DM Function	<i>confirmation</i>	<i>open/closer frame</i>	<i>acceptance</i>
SAMEX raw freq.	187	96	36
Freq. per 1,000	6.2	3.2	1.2
%	59%	30%	11%
BASEX raw freq.	163	36	8
Freq. per 1,000	3.1	0.7	0.2
%	79%	17%	4%

Appendix B7. Distribution of DM 'okay' Clusters in Exploratory Corpora

<i>okay clusters</i>	<i>okay so</i>	<i>okay and</i>	<i>okay now</i>
SAMEX raw frequency	96	25	12
Per 1,000 tokens %	2.8	0.7	0.3

BASEX raw frequency	50	12	12
Per 1,000 tokens ‰	0.9	0.2	0.2

Appendix B8. Arabic DMs used in the SAMEX Exploratory Corpus

Arabic DMs	يعني	تمام	لكن	مثل	الآن
	{ya3ni}	{tamam}	{lakin}	{mithl}	{alan}
English DM equivalent	<i>I mean/ meaning</i>	<i>okay</i>	<i>however</i>	<i>for example</i>	<i>now</i>
ADM Frequency	41	10	12	13	6
Per 1,000 tokens ‰	1.2	0.3	0.3	0.4	0.2
SAMEX DM Raw Frequency	5	187	6	25	68
Per 1,000 tokens ‰	0.1	5.4	0.2	0.7	2
English DMs	<i>I mean/ meaning</i>	<i>okay</i>	<i>however</i>	<i>for example</i>	<i>now</i>
BASEX DM Raw Frequency	24	163	6	20	-
Per 1,000 tokens ‰	0.5	3.2	0.1	0.4	-

APPENDIX C: METHODOLOGY

Appendix C1. Tagging method used to mark up occurrences in SASME

Tag Category	Function	Tag Category Members	Example
<intro>	Introduction	Introduction to start of lecture	<intro> Lecturer: <i>The evening session is starting</i>
<topcon>	Topic Content	Start of Lecture topic content	<topcon> <i>So now we will go to the names of the particulars.</i>
<aside>	Aside	Utterances not related to lecture - off topic	<aside> <i>Can you wake him up</i> </aside>
<overview>	Overview	Course discussion, reminders related to lecture, summary of lecture, etc	<overview> <i>I think you have gone through the signaling through physiology. This I assume...</i>
<conclusion>	Conclusion	Indicating end or wrapping up of lecture topic	<concl> <i>So today what I'll stop here, today what we talk is introduction...</i>
<q>	Confirmation Checks	<okq> <rq> <cq>	<i>forms, is kind of more common, <okq>okay</okq>?</i>
<fs>	Closer	<okfs>	<i>in the west it's really common. <okfs>Okay</okfs>.</i>
<c>	Topic Initiator	<okc> <sc> <nc>	<i><okc>Okay</okc>, obviously you have the idea as to what the term is...</i>
<m>	Summarizer	<sm> <tfm>	<i><sm>So</sm> what I'm saying here is causes of dementia can have...</i>
<td>	Topic Developer	<std> <btd> <ctd> <atd>	<i><atd>Also</atd> some gene duplication or triplications.</i>
<e>	Elicitor	<se> <aqe>	<i><se>so</se> what is the difference between staging and grading?</i>
<rp>	Rephraser	<imrp> <owrp>	<i>The patient may become [unclear] and may go, again, <imrp>I mean</imrp> , in seizure.</i>

APPENDIX D: RESULTS

Appendix D1. SASME Structural DM Functions across NNS English Lectures %

#	SASME Lecture	<i>Topic Initiator</i> Freq. per 1000 ‰	<i>Topic Developer</i> Freq. per 1000 ‰	<i>Summarizer</i> Freq. per 1000 ‰	<i>Closer</i> Freq. per 1000 ‰
1	<i>Anatomy & Physiology of Sympathetic Nervous System (APSNS)</i>	10	3.5	0.7	3
2	<i>Cerebral TB (CTB)</i>	5.3	3.8	0.8	0
3	<i>Vitamin A (VitA)</i>	1.8	4.7	3	1.5
4	<i>Intro to Degenerative Brain Diseases (IDBD)</i>	3.2	8	5.2	2.6
5	<i>Drugs used in Epilepsy (DE)</i>	2.5	13	3.1	0.3
6	<i>Congenital Malformations and Hydrocephalus (CMH)</i>	5.5	7.3	5.3	2.3
7	<i>Anatomy of the Ear (AE)</i>	2	0.7	0	0.3
8	<i>Drugs used in Anxiety Disorders (DAD)</i>	2.9	8.9	6	0
9	<i>Physiology of Synapses and Receptors (PSR)</i>	1	4.4	15.6	9.3
10	<i>Pathology of Brain Tumors (PBT)</i>	6.1	4.6	2.1	3.2
11	<i>Drugs used in Balance System (DBS)</i>	9.5	18	10.9	0
	Total Freq. per 1000 ‰	4.7	7.5	4.2	1.6

Appendix D2. BASME Structural DM Functions across NS English Lectures %

#	BASME Transcripts	Topic Initiator Freq. per 1000 ‰	Topic Developer Freq. per 1000 ‰	Summarizer Freq. per 1000 ‰	Closer Freq. per 1000 ‰
1	<i>Abdominal Aortic Aneurysms. Trial and Error</i> (lslct017)	5	5.7	1.7	0.2
2	<i>Diabetic Nephropathy</i> (lslct032)	2.7	9.6	2.2	1.4
3	<i>Glomerulonephritis</i> (lslct034)	1.9	3.9	2.9	0.4
4	<i>HIV and AIDS</i> (lslct008)	3.4	5.4	2.2	0.5
5	<i>HIV Infection</i> (lslct037)	2.9	8.8	3.3	0.7
6	<i>Introduction to Renal Function. The Concept of Clearance</i> (lslct026)	4.1	12.3	7.8	0.1
7	<i>Kidney Stones</i> (lslct028)	1.8	10.4	4.9	1.5
8	<i>Hepatitis B Endemicity</i> (lslct018)	4.5	12.4	8.5	1.1
9	<i>Concentration and Dilution of Urine</i> (lslct029)	3.6	19.3	12.3	0.8
	Total Freq. per 1000 ‰	3.3	10.1	5.3	0.7

Appendix D3. SASME Interactional DM Functions across NNS English Lectures ‰

#	SASME Lecture	Confirmation Checks Freq. per 1000 ‰	Rephrasers Freq. per 1000 ‰	Elicitors Freq. per 1000 ‰
1	<i>Anatomy & Physiology of Sympathetic Nervous System (APSNS)</i>	4	0.2	0
2	<i>Cerebral TB (CTB)</i>	1.5	0	0
3	<i>Vitamin A (VitA)</i>	22.6	0	0.8
4	<i>Intro to Degenerative Brain Diseases (IDBD)</i>	4.3	0	1.4
5	<i>Drugs used in Epilepsy (DE)</i>	1.4	0.7	2.2
6	<i>Congenital Malformations and Hydrocephalus (CMH)</i>	1.3	0	0.3
7	<i>Anatomy of the Ear (AE)</i>	0.3	0	0
8	<i>Drugs used in Anxiety Disorders (DAD)</i>	1.1	0	1.6
9	<i>Physiology of Synapses and Receptors (PSR)</i>	0	1.5	0.5
10	<i>Pathology of Brain Tumors (PBT)</i>	0.7	0.4	0.7
11	<i>Drugs used in Balance System (DBS)</i>	1.2	0	1.4
	Total Freq. per 1000 ‰	4.4	0.3	1.2

Appendix D4. BASME Interactional DM Functions across NS English Lectures ‰

#	BASME Transcripts	Confirmation Checks Freq. per 1000 ‰	Rephrasers Freq. per 1000 ‰	Elicitors Freq. per 1000 ‰
1	<i>Abdominal Aortic Aneurysms. Trial and Error</i> (IsIct017)	0	0.2	0
2	<i>Diabetic Nephropathy</i> (IsIct032)	1.1	2	0
3	<i>Glomerulonephritis</i> (IsIct034)	1.7	1.8	0.1
4	<i>HIV and AIDS</i> (IsIct008)	2.9	2.5	0.2
5	<i>HIV Infection</i> (IsIct037)	7.2	0.6	0.7
6	<i>Introduction to Renal Function. The Concept of Clearance</i> (IsIct026)	1.1	0.8	0
7	<i>Kidney Stones</i> (IsIct028)	3.7	1.1	0.2
8	<i>Hepatitis B Endemicity</i> (IsIct018)	0	1.1	0.2
9	<i>Concentration and Dilution of Urine</i> (IsIct029)	0.8	0.8	1
	Total Freq. per 1000 ‰	1.9	1.2	0.4