



UNIVERSITI PUTRA MALAYSIA

***A CORPUS-BASED STUDY OF MALAYSIAN ESL STUDENTS' USE OF
DISCOURSE CONNECTORS IN UPPER AND POST-SECONDARY
ARGUMENTATIVE WRITING***

SEYED ALI REZVANI KALAJAHI

FPP 2014 12



**A CORPUS-BASED STUDY OF MALAYSIAN ESL STUDENTS' USE OF
DISCOURSE CONNECTORS IN UPPER AND POST-SECONDARY
ARGUMENTATIVE WRITING**

By

SEYED ALI REZVANI KALAJAHI

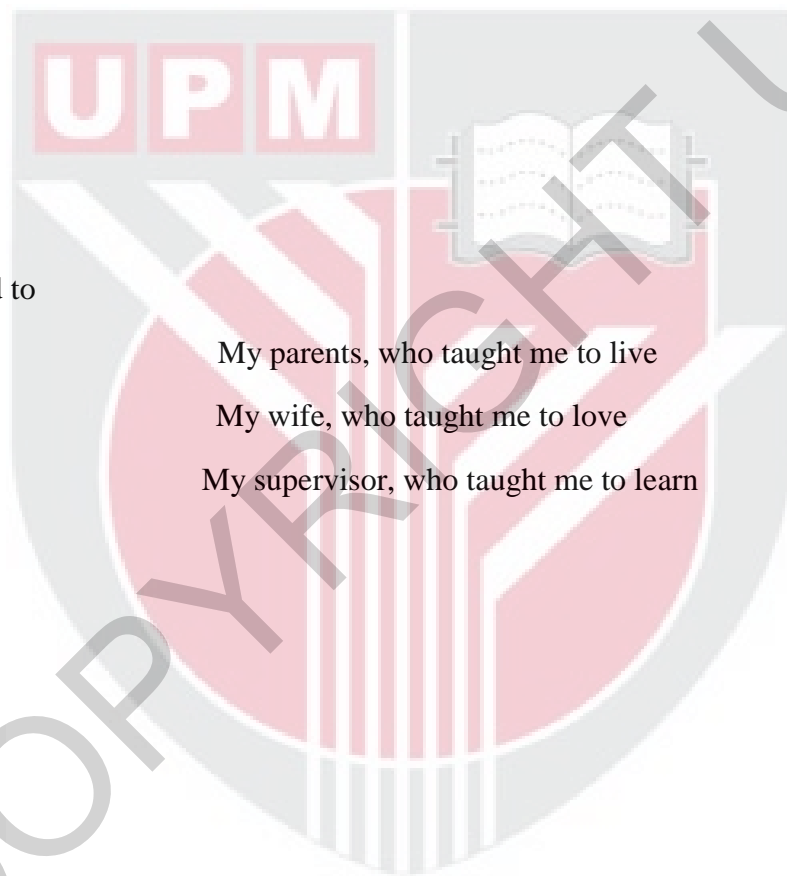
**Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia,
in Fulfilment of the Requirement for the Degree of Doctor of Philosophy**

July 2014

All material contained within the thesis, including without limitation text, logos, icons, photographs and all other artwork, is copyright material of Universiti Putra Malaysia unless otherwise stated. Use may be made of any material contained within the thesis for non-commercial purposes from the copyright holder. Commercial use of material may only be made with the express, prior, written permission of Universiti Putra Malaysia.

Copyright © Universiti Putra Malaysia





Dedicated to

My parents, who taught me to live

My wife, who taught me to love

My supervisor, who taught me to learn



Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment
of the requirement for the degree of Doctor of Philosophy

**A CORPUS-BASED STUDY OF MALAYSIAN ESL STUDENTS' USE OF
DISCOURSE CONNECTORS IN UPPER AND POST-SECONDARY
ARGUMENTATIVE WRITING**

By

SEYED ALI REZVANI KALAJAHI

July 2014

Chairman: Associate Professor Ain Nadzimah Abdullah, PhD
Faculty: Educational Studies

Discourse connectors (DCs) are one of the elements of cohesive devices that bring about cohesion to a piece of writing or speech. They are potentially useful means for writers, particularly in ESL and EFL writing pedagogic settings. DCs usefulness is two-pronged. First, they help and guide readers through the text, and then they are tools for writers to engage with their readers. It has been well-documented that appropriate and efficient use of DCs will create a coherent flow of the text. However, second/foreign language learners have some difficulties to use them efficiently and systematically in their writing. Literature review shows that the Malaysian ESL students are also suffering from improper and efficient use of DCs which leads them in failing to produce a cohesive text. Surprisingly, no single study was found in the context of Malaysia to investigate Malaysian ESL students' understanding and use of DCs.

Hence, this study attempted to investigate and understand the nature and the use of the DCs in the Malaysian student writing compared with Native speakers writing. The study also was set to examine the correlation between the frequency use of the DCs and the quality of writing. The final goal of this research study was to find out to what extent Malaysian ESL students are committing errors while using DCs.

A corpus-based approach was adopted to meet the objectives of the study. To this end, an argumentative topic was assigned to the Form 4, Form 5 (upper-secondary) and the first year college students (post-secondary) and they were asked to write about the given topic in the classroom and submit their works to the instructors. They were required to write 250 words within 45 minutes. Upon compilation of the essays, the Malaysian Corpus of Students' Argumentative Writing (MCSAW) was built with \approx 600,000 tokens. To compare and find out a vivid picture of Malaysian ESL students use of DCs with Native English Speakers, the Louvain Corpus of Native Essay Writing [LOCNESS] corpus was used.

Oxford Wordsmith Tools (5) was employed to extract data from corpus for analysis, by using frequency count and concordance functions. Aiming to identify what type of DCs is used by Malaysian ESL students, Discourse Connector List developed by Rezvani

Kalajahi and Neufeld (2014) was used. To be able to examine the relationship between the quality of writing and the frequency of the use of the DCs, ESL composition profile offered by Jacobs et al. (1981) was utilized. Finally, a framework of identification of DCs error type was developed by the researcher to explore the errors that students commit while using DCs.

Findings of this study entail three phases. First, it was observed that Malaysian students tend to use DCs more frequently than native students. The overall frequency of the use of the DCs between Malaysian and native students was statistically significant at $p < .05$. However, the native students used more variety of DCs types than Malaysian students (398vs.328). It was also found that Malaysian students use DCs in some categories frequently and infrequently. Based on the findings in the native students writing (LOCNESS Corpus), the most frequent DCs in written English were offered. Second, there was a very weak negative but insignificant correlation between writing quality and the frequency of the use of the DCs in the writing of Malaysian ESL students. Finally, the qualitative analysis revealed that the erroneous use of DCs made by Malaysian ESL student writers mainly manifested in eight different categories. They had problems with the use of these devices which involved semantic, syntactic, stylistic, positional and mechanical errors. They also appeared to have tendency for unnecessary addition, omission, and redundant repetition of the DCs.

In conclusion, this study demonstrated that Malaysian ESL students' use of DCs was still at an evolving level. It is vitally important that the accurate use of DCs in writing among Malaysian students be further highlighted in the classrooms through using concordance lines and adopting explicit instruction technique. Besides, material developers may take the outcome of the research into consideration and could find out possible ways to distribute and introduce DCs systematically across the educational levels.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Doktor Falsafah.

**KAJIAN BERASASKAN KORPUS TENTANG PENGGUNAAN
PENGHUBUNG WACANA DALAM KALANGAN PELAJAR ESL
MENENGAH ATAS DAN PASCA MENENGAH MALAYSIA DALAM
PENULISAN PERBINCANGAN BAGI**

Oleh

SEYED ALI REZVANI KALAJAHI

Julai 2014

Pengerusi: Prof. Madya Ain Nadzimah Abdullah, PhD
Fakulti: Pengajian Pendidikan

Penghubung wacana - Discourse Connectors (selepas ini DC) merupakan salah satu daripada unsur-unsur peranti kohesif membawa kepada penulisan atau ucapan yang padu. Perkara ini sangat berguna bagi penulis, terutamanya dalam pedagogi penulisan ESL dan EFL. DC dianggap sebagai serampang dua mata. Pertama, mereka membantu dan membimbing pembaca dalam pembacaan teks, kemudiannya menjadi alat untuk penulis melibatkan diri dengan pembaca mereka. Hal ini telah didokumenkan dengan baik bahawa kesesuaian dan kecekapan penggunaan DC akan mewujudkan aliran koheren teks. Walaubagaimanapun, pelajar bahasa kedua/pelajar bahasa asing menghadapi kesukaran menggunakannya dengan cekap dan sistematik dalam penulisan mereka. Satu sorotan literatur menunjukkan bahawa pelajar ESL Malaysia juga mengalami penggunaan DC yang kurang cekap dan kurang berkesan yang kemudiannya mengakibatkan kegagalan untuk menghasilkan teks yang padu. Yang menghairankan, tiada pun satu kajian dijalankan dalam konteks Malaysia untuk mengkaji kefahaman dan penggunaan DC oleh pelajar ESL Malaysia.

Oleh itu, kajian ini bertujuan untuk mengkaji dan memahami sifat dan penggunaan DC bagi pelajar Malaysia berbanding dengan penutur asli. Kajian ini juga telah direka untuk mengkaji korelasi antara kekerapan penggunaan DC dan kualiti penulisan. Matlamat akhir kajian ini adalah untuk mengetahui sejauh mana pelajar ESL Malaysia melakukan kesilapan semasa menggunakan DC.

Pendekatan berasaskan korpus telah digunakan bagi mencapai objektif kajian. Untuk tujuan ini, satu topik penulisan perbincangan ditugaskan kepada pelajar tingkatan 4, Tingkatan 5, dan pelajar kolej tahun pertama. Pelajar-pelajar diminta untuk menulis tentang topik yang diberikan di dalam kelas dan menghantar karya mereka kepada pengajar. Mereka dikehendaki menulis sebanyak 250 patah perkataan dalam masa 45

minit. Setelah selesai penulisan itu disusun, korpus penulisan perbincangan bagi pelajar-pelajar Malaysia (MCSAW) telah dibina dengan $\approx 600,000$ token. Untuk perbandingan dan memperolehi gambaran yang jelas antara pelajar-pelajar ESL Malaysia dengan penutur asli bahasa Inggeris dalam menggunakan DC, pendekatan Louvain Corpus of Native Essay Writing [LOCNESS] telah digunakan.

Oxford WordSmith Tools (5) telah digunakan untuk meperoleh data daripada korpus untuk dianalisis, dengan menggunakan kekerapan dan fungsi konkordans. Bertujuan untuk mengenal pasti jenis DC yang digunakan oleh pelajar ESL Malaysia, senarai penghubung wacana (DC) yang dibuat oleh Rezvani Kalajahi dan Neufeld (2014) telah digunakan. Untuk mengkaji hubungan di antara kualiti penulisan dan kekerapan penggunaan DC, profil penulisan ESL yang dihasilkan oleh Jacobs et al. (1981) telah digunakan. Akhir sekali, satu rangka kerja pengenalan jenis kesalahan DC telah dibangunkan oleh penyelidik untuk meneroka kesilapan yang pelajar-pelajar lakukan semasa menggunakan DC.

Hasil kajian ini melibatkan tiga fasa. Pertama, ia telah memperlihatkan bahawa pelajar Malaysia menggunakan DC lebih banyak berbanding dengan pelajar-pelajar penutur asli. Kekerapan keseluruhan penggunaan DC antara pelajar Malaysia dan pelajar penutur asli didapati mempunyai statistik yang ketara pada $p < .05$. Walau bagaimanapun, pelajar penutur asli menggunakan pelbagai jenis DC berbanding dengan pelajar Malaysia (398vs.328). Kajian ini juga mendapati bahawa terdapat pelajar-pelajar Malaysia yang menggunakan beberapa kategori DC secara berlebihan dan berkurangan. Berdasarkan penemuan kajian ini penulisan pelajar penutur asli (LOCNESS Corpus), yang paling kerap menggunakan DC telah ditawarkan. Kedua, terdapat korelasi yang menunjukkan hubungan yang negatif dan lemah tetapi tidak ketara antara kualiti penulisan dan kekerapan penggunaan DC dalam penulisan pelajar ESL Malaysia. Akhir sekali, analisis data kualitatif pula menunjukkan bahawa penggunaan DC yang salah telah dilakukan oleh penulis pelajar ESL Malaysia terutamanya dimanifestasikan dalam lapan kategori yang berbeza. Mereka mempunyai masalah tentang penggunaan alat-alat ini yang melibatkan semantik, sintaksis, gaya, kedudukan dan kesilapan mekanikal. Mereka juga mempunyai kecenderungan untuk penambahan yang tidak perlu, tidak menggunakan, dan pengulangan penggunaan DC.

Kesimpulannya, kajian ini menunjukkan bahawa penggunaan DC oleh pelajar ESL Malaysia masih pada tahap perkembangan. Ia amat penting untuk dititikbertkan penggunaan DC yang tepat dalam bilik darjah melalui penggunaan talian konkordans dan menerima pakai teknik arahan-arahan terperinci. Selain itu, pemaju bahan pengajaran boleh mempertimbangkan hasil kajian ini dan mengenalpasti pelbagai pelaksanaan untuk mengedar dan memperkenalkan DC secara sistematik di seluruh peringkat pendidikan.

ACKNOWLEDGEMENTS

I would like to express my gratitude to a number of people for their help, support, and cooperation in conducting this dissertation.

First of all, I would like to thank my supervisor, Associate Professor Dr. Ain Nadzimah Abdullah who took her precious time and rendered me valuable advice and constructive comments in writing this thesis. She was a great help in shaping my study. A profound gratitude is also owed to the feedback of the committee members, Professor Dr. Jayakaran Mukundan and Associate Professor Dr. Arshad Abd. Samad for their valuable recommendations, worthwhile suggestions and encouragement.

Special thanks go to my wife and my family for their constant support and encouragement all through my life when I needed it most.

My thanks are also expressed to Dr. Vahed Zarifi and Dr. Vahid Nimehchisalem, Bakhtiar Naghdipour, and Steve Neufeld for their feedback and rating the students' essays. I am also grateful to the teachers and students who patiently helped me in building the MCSAW corpus.

Last but not least, I am truly grateful to Allah for giving me patience, strength, and wisdom to complete this thesis.

I certify that a Thesis Examination Committee has met on 18 July 2014 to conduct the final examination of Seyed Ali Rezvani Kalajahi on his thesis entitled "A Corpus-Based Study of Malaysian ESL Students' use of Discourse Connectors in Upper and Post-Secondary Argumentative Writing" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Thesis Examination Committee were as follows:

Ismi Arif bin Ismail, PhD

Associate Professor
Faculty of Educational Studies
Universiti Putra Malaysia
(Chairman)

Chan Swee Heng, PhD

Professor
Faculty of Modern Language and Communication
Universiti Putra Malaysia
(Internal Examiner)

Nooreen binti Noordin, PhD

Senior Lecturer
Faculty of Educational Studies
Universiti Putra Malaysia
(Internal Examiner)

Randi Reppen, PhD

Professor
Northern Arizona University
United States
(External Examiner)



NORITAH OMAR, PhD

Associate Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 18 August 2014

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfillment of the requirement for the degree of Doctor of Philosophy. The members of the Supervisory Committee are as follows:

Ain Nadzimah Abdullah, PhD

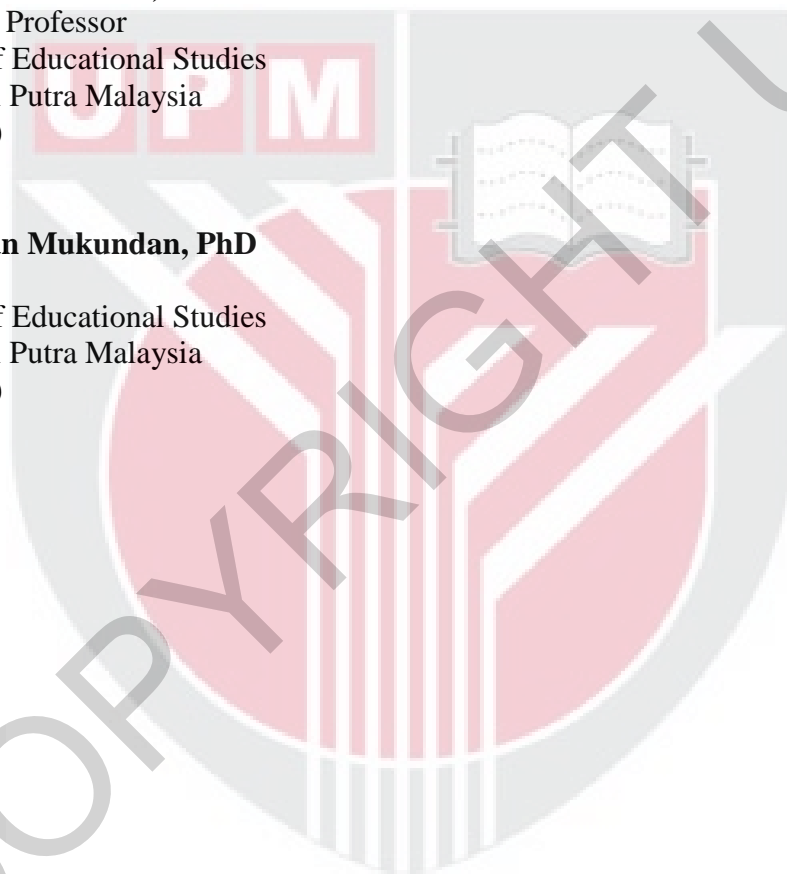
Associate Professor
Faculty of Modern Languages and Communication
Universiti Putra Malaysia
(Chairman)

Arshad Abd. Samad, PhD

Associate Professor
Faculty of Educational Studies
Universiti Putra Malaysia
(Member)

Jayakaran Mukundan, PhD

Professor
Faculty of Educational Studies
Universiti Putra Malaysia
(Member)



BUJANG BIN KIM HUAT, PhD

Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

DECLARATION

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Piitra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____

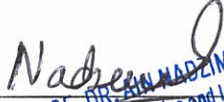
Date: _____

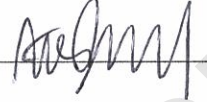
Name and Matric No: Seyed Ali Rezvani Kalajahi GS29486

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: 
Name of **ASSOC. PROF. DR. AIN NADZIMAH ABDULLAH**
Chairman (Deputy Dean (Graduate Studies and International Affairs))
Supervisory Faculty of Modern Languages and Communication
Committee: Associate Professor Dr. Ain Nadzimah Abdullah
Universiti Putra Malaysia
43400 UPM Serdang, Selangor

Signature: 
Name of **ASSOC. PROF. DR. ARSHAD ABD. SAMAD**
Member of Dept. of Language & Humanities Education
Supervisory Faculty of Educational Studies
Committee: Associate Professor Dr. Arshad Abd. Samad
Universiti Putra Malaysia
43400 Serdang, Selangor D. E.
Tel No: 03-89467278 Fax No: 03-89480119

Signature: 
Name of **PROF. DR. JAYAKARAN MUKUNDAN**
Member of Professor
Supervisory Jabatan Pendidikan Bahasa Dan Kemanusiaan
Committee: Prof. Dr. Jayakaran Mukundan
Fakulti Pengajian Pendidikan
Universiti Putra Malaysia
43400 Serdang, Selangor.

TABLE OF CONTENTS

	Page
ABSTRACT	I
ABSTRAK	III
ACKNOWLEDGEMENTS	V
APPROVAL	VI
DECLARATION	VIII
LIST OF TABLES	XIII
LIST OF FIGURES	XIV
CHAPTER	
1 INTRODUCTION	1
1.1 Background of the Study	1
1.2 Discourse Connectors	3
1.3 Corpus and its Connection to Language Teaching	5
1.4 Nature of the Present Study	6
1.5 Rationale of the Present Study	7
1.6 Statement of the Problem	8
1.7 Objectives	11
1.8 Research Questions	11
1.9 Significance of the Study	12
1.10 Limitations	12
1.11 Definition of Terms	14
2 LITERATURE REVIEW	17
2.1 Theoretical Underpinnings	17
2.1.1 Theories of Writing	17
2.1.1.1 Social Constructivism Theory	17
2.1.1.2 Social Cognitive Theory	18
2.1.2 Two Approaches to DCs	21
2.1.2.1 Coherence-based Approach	21
2.1.2.2 Relevance-theoretic Account Approach	23
2.1.2.2.1 Relevance Theory and DCs	24
2.1.3 Conclusion of Two Approaches	26
2.1.4 Theoretical Framework of the Study	27
2.2 Term, Definition and Taxonomy of DCs in the Literature	28
2.2.1 What are DCs?	34
2.2.2 DCs List	36
2.2.2.1 Description of DCs List	39
2.3 Text Coherence and DCs	41
2.3.1 The Importance of DCs in Written Discourse	41
2.3.2 The Sources of Difficulty in Learning DCs	43
2.4 Error Analysis	47
2.4.1 Relevance of Error Analysis in English Language Teaching	49
2.4.2 Model for Error Analysis	50
2.4.2.1 Sources of Errors	50
2.4.3 Advantages of Error Analysis	51
2.5 Corpus Linguistics	52
2.5.1 What is a Corpus?	52

2.5.2	Advantages of Using Computers	52
2.5.3	Nature of Corpus Linguistics	53
2.5.4	Corpus Processing	53
2.5.4.1	Input of Information to a Corpus	53
2.5.4.2	Output of Information from a Corpus	54
2.5.5	Approaches to Corpus Linguistics	54
2.5.5.1	Corpus-based Approach	54
2.5.5.2	Corpus-driven Approach	55
2.5.6	Characteristics of a Modern Corpus	55
2.5.7	Types of Corpora	56
2.5.7.1	Learner Corpus	56
2.5.7.1.1	Need for a Learner Corpus	58
2.5.7.1.1.1	Corpus Work in Malaysia	60
2.5.7.2	Reference Corpora	61
2.5.8	Learner Corpora and English Language Teaching	61
2.6	Previous Studies on DCs in Different Settings and Genres	62
2.7	Conclusion	76
3	METHODOLOGY	77
3.1	Corpus Design	77
3.1.1	Purpose of Building the MCSAW Corpus	77
3.1.2	Register of Text	78
3.1.3	Task Setting	78
3.1.4	Genre and Topic Choice	78
3.1.5	Essay Length	79
3.1.6	Corpus Size and Subjects	80
3.1.7	Corpus Sampling: Representativeness and Generalization	81
3.1.8	Corpus Collection Procedure	82
3.2	Reference Corpus	82
3.3	Research Design	82
3.4	Identification of DCs	84
3.5	Method of Data Analysis	85
3.5.1	Instrumentation	86
3.5.1.1	Wordsmith	86
3.5.1.2	ESL Composition Profile	87
3.6	Data Analysis	88
3.6.1	Quantitative Phase	88
3.6.1.1	Frequency Count and Proportion Analysis	88
3.6.1.2	Discrepancy Analysis	88
3.6.1.3	Correlational Analysis	89
3.6.2	Qualitative Phase	90
3.6.2.1	Identification and Classification of Error Use of DCs	90
3.7	Research Framework	92
3.8	Conclusion	94
4	RESULTS AND DISCUSSION	95
4.1	Frequency of Occurrence	95
4.1.1	Overall Frequencies	95
4.1.2	Individual DCs	100
4.1.3	Position of DCs	105

4.1.4 Frequently and Infrequently Used DCs	107
4.2 Frequency of Semantic Types of DCs	110
4.3 The Most Frequent DCs of Written English	118
4.4 Relationship between Quality of Writing and Frequency of DCs	121
4.4.1 Summary and Discussion	123
4.5 Error Analysis of DCs in Malaysian Students' Writing	124
4.5.1 Error Types	125
4.5.1.1 Semantic Misuse of DCs	125
4.5.1.2 Syntactic Misuse of DCs	126
4.5.1.3 Stylistic Misuse of DCs	130
4.5.1.4 Positional Misuse of DCs	135
4.5.1.5 Unnecessary Addition of DCs	137
4.5.1.6 Omission of DCs	137
4.5.1.7 Redundancy in the use of DC	140
4.5.1.8 Mechanical Misuse of DCs	142
4.5.2 Summary and Discussion	143
4.6 Concluding Remarks	147
5 CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS	148
5.1 Summary of Findings, Discussions and Conclusions	148
5.2 Implications of the Study	151
5.3 Recommendations for Future studies	154
5.4 Contributions of the Study	155
5.4.1 Assigning DC to Primary and Secondary Levels of Education	155
5.5 Concluding Remarks	156
REFERENCES	157
APPENDICES	177
Appendix A : Discourse Connectors List	177
Appendix B: LOCNESS Description	201
Appendix C: ESL Composition Profile	208
Appendix D: List of the Sorted and the Ranked DCs	210
Appendix E: List of the Common DCs between MCSAW and LOCNESS with an Approximate Frequency of Occurrence Per Million	234
Appendix F: Pie Chart for Distribution of the DCs Categories within Sub-corpus	236
Appendix G: Distribution of the DCs Sub-categories within Sub-corpus	239
Appendix H: Two Raters Scores and DCs Frequencies in the Selected Writing	240
Appendix I: The Discourse Connector List mapped into the primary and secondary levels in Malaysia (Standard 1-6 & Form 1- 5)	243
BIODATA OF STUDENT	251
LIST OF PUBLICATIONS	252

LIST OF TABLES

Table	Page	
2. 1	Lists of the classification of linking adverbials (Biber et al., 1999)	32
2. 2	Quirk et al.'s (1985) framework of conjunctions	33
3. 1	MCSAW Information	77
3. 2	Number of essays and word size according to level in MCSAW	80
3. 3	Frequency and percent of students according to gender	81
3. 4	Frequency and percent of students according to ethnicity	81
3. 5	Cohn's Kappa Value Scales and Interpretations	89
3. 6	Reliability Statistics	89
4.1	Overall comparative frequencies of DCs in the MCSAW and LOCNESS and their sub-corpora	95
4.2	Total Types of DCs used in each corpus and their sub-corpora	96
4.3	Chi-square test of goodness-of-fit for discrepancy analysis	97
4.4	Top ten DCs in the MCSAW	101
4.5	Top ten DCs in the LOCNESS	101
4.6	Top 7 DCs in the MCSAW Corpus	103
4.7	Top 7 DCs in the LOCNESS Corpus	104
4.8	The positions of top 7 DCs in the MCSAW and LOCNESS Corpus	105
4.9	Significantly Overused DCs in the MCSAW	107
4.10	Significantly Underused DCs in the MCSAW	108
4.11	Distribution of the Semantic Types of DCs within 8 categories in the MCSAW and its sub-corpora	110
4.12	Distribution of the Semantic Types of DCs within 8 categories in the LOCNESS and its sub-corpora	110
4.13	Distribution of the Semantic Types of DCs sub-categories in the MCSAW and the sub-corpus	112
4.14	Distribution of the Semantic Types of DCs sub-categories in the LOCNESS and the sub-corpus	112
4. 15	Some additive common DCs in Malay (Source, Ahmad, 2012)	115
4.16	Frequency and Coverage of Top 100 DCs in LOCNESS	118
4.17	Overall Correlation between writing quality and the frequency of DCs use	122
4. 18	Correlations between writing quality and the frequency use of DCs in three educational levels	122

LIST OF FIGURES

Figure		Page
2.1	A Model of Discourse Construction (Flower, 1994, p.53)	20
2.2	Theoretical Framework of the Study	27
3.1	A framework of identification of DCs error type	91
3.2	Research Framework	93
4.1	The concordance lines of the position of the “IF” DC in the MCSAW corpus	106
4.2	The concordance lines of the position of the “IF” DC in the LOCNESS corpus	106
4.3	Distribution of the Semantic Types of DCs within 8 categories in the LOCNESS and the MCSAW	111
4.4	The Comparison of the Semantic Types of DCs Sub-categories Distribution in the LOCNESS and the MCSAW	114
4.5	The frequency of “in my point of view” DC in the COCA and BNC	129
4.6	The concordance lines for accurate use of ‘for example, from my point of view, beside’	130
4.7	Frequency of “Actually” as a DC in COCA and BNC	131
4.8	Concordance lines of “other than that” DC in the MCSAW Corpus	133
4.9	The frequency of “Other Than That” DC in COCA and BNC Corpus	134
4.10	The concordance lines of “especially” DC in the LOCNESS Corpus	135

CHAPTER 1

INTRODUCTION

“Use of DCs appears to be the last part of grammar. To make an analogy, the relationship between grammar and DCs is like the relationship between train and rail. Coal is morphology, skeleton of train is syntax, chains that join the compartments together are semantics, waiters are vocabulary items, passengers are readers, the person who is leading the train is writer, and the rail is DMs. If all parts of train work well, then the train is able to move on the rail, and if the person who is leading the train knows the interpretation of signs along the rails, then he is able to keep the train on the right track (Jalilifar, 2008).”

This chapter begins with a review of the background of the study. Next, a brief account of corpus and its link to language teaching is presented. Later, the nature of the present study is explained and then it is continued with the review of the rationale of the present study. The chapter continues with the statement of the research problem, objective, and research questions. Next comes the significance of the study followed by some limitations. The chapter concludes with the definitions of some of key words used throughout the study.

1.1 Background of the Study

English is seen as an important language that is taught in schools for 11 years in Malaysia. English Language has also become an important and a compulsory subject matter for primary and secondary educational levels (Malaysia Education Blueprint 2013 – 2025). In fact, majority of the Malaysian children find an opportunity to be exposed to English language even at the age of four or five at the kindergarten. Then they continue learning the language until they complete SPM or Malaysian Education certificate at the age of 17.

Although the majority of the Asian students appear to be capable of using English communicatively, their performance may not be satisfactory in relation to a number of language skills. Some studies have indicated that Malaysian ESL students have problems mainly in writing tasks, at both school level and tertiary level many learners tend to be incapable of choosing appropriate conjunctions, or what is called as Discourse Connectors (henceforth DCs) in this study. Therefore, inexplicable absence or inappropriate use of the DCs tends to result in lack of cohesion in the text (Ghabool, Mariadass, & Kashef, 2012; Maros, Hua, & Khazriyati, 2007; Musa, Lie, & Azman, 2012; Ahour & Mukundan, 2012; Ellis & Barkhuizen, 2005).

Norrish (1983) claims that writing has been proven to be the most difficult language skill even for native speakers. By the same token, it comes as no surprise that ESL or EFL students may suffer from inappropriate writing. Ostensibly, writing is more difficult than speaking since, in written communication, there are no additional nonverbal cues (e.g. facial expression, gesture) to ensure that the message comes across. Hence, it is very important for the student writers to put their ideas into words in a way that are clear, succinct, and easily comprehensible for the readers.

A written text requires the writer not only to be able to construct sentences accurately but also to be able to use DCs appropriately to build a coherent text. It has been well documented that DCs are one of the important devices for signaling the connections between clauses or beyond in textual communication. Biber, Johansson, Leech, Conrad, & Finegan (1999) observe that the primary function of DCs is to state the speaker/writer perception of the relationship between two and more units of discourse and thus help to create textual cohesion. Appropriately used DCs are considered to be contributable to the clarity and comprehensibility of discourse.

With the status of English as an international language and the increasing expansion of the use of the English language throughout the world, second and foreign language learners find themselves in different contexts that require them to write in English. It is, therefore, the concern of this study to investigate the Malay students' writing to see how they deal with the DCs in compositions and how their use of DCs compares with native students.

L2 writers must learn that the readers are able to follow the ideas expressed in the text more easily if they properly signal the relations of their utterances to those which precede and follow. Therefore, DCs constitute an important component of communicative competence, which L2 learners must acquire if they want to communicate effectively. This implies that the nonnative speakers competent in using the DCs of the L2 will be more successful in interaction than those who are not (Rahimi, 2011).

DCs function as cohesive devices in discourse that help guide the reader or listener through the message by signaling how successive units are related (Leech, 1998). Some examples of DCs are: *but* (indicating contrast), *because* (reason) and *therefore* (result). Halliday & Hasan (1976) describe these expressions as connectors of logical or semantic relations between units of discourse. Results of the studies have attempted to show that DCs actually contribute to a better understanding of discourse (Flowerdew, 1998; Hartnett, 1986; Mauranen, 1993; Mosenthal & Tierney, 1984). In addition, some research results report that DCs may be very significant in terms of how a text is perceived. For instance, Mauranen (1993) found that a sample of

academic writing with DCs to be more logical, convincing and authoritative than the same sample with all the DCs removed.

It, therefore, seems reasonable to assume that inappropriate use of DCs could, to a certain degree, hinder successful communication and lead to a misunderstanding. Therefore, as part of communicative competence, L2 learners should acquire the appropriate use of DCs of their target language (TL). Warsi (2000) remarks that it is plausible to suppose that L2 learners who are competent in the use of DCs of the target language will be more successful in both verbal and non-verbal interaction than those who are not. For these reasons, the use of DCs needs to be studied and should receive sufficient attention in language research and instruction.

As highlighted earlier, writing skill turns out to be a major concern for non-natives as they more often than not find it difficult to produce a piece of coherent text. Numerous researchers in the realm of writing and discourse claim that DCs may lead to cohesion (Halliday and Hassan 1976; Tannen, 1982; Quirk et al., 1985; Cook, 1989; McCarthy, 1991; Biber et al., 1999; Celce-Murcia & Larsen-Freeman, 1999; Brown, 2001; Ozono and Ito's, 2003). However, students often experience difficulty applying them in their writing. Considering this, it can be assumed that Malaysian students are not an exception. Although extensive researches have been carried out on the use of DCs both in EFL and ESL contexts, no single study was found to address the use of DCs by Malaysian ESL students.

1.2 Discourse Connectors

There is a plethora of research studies done on linguistic devices writers use to connect their ideas; yet, researchers have to agree on common terms for such devices, let alone a defining framework of reference for language users. Rezvani Kalajahi, Nadzimah Abdullah, Mukundan, and Tannacito (2012) state that current debate over the terms 'linkers, DCs or discourse markers' illustrates the ambiguity linguists face due to the lack of universal definitions and terms to label these linguistic units. A trend to use the term 'discourse connector' (henceforth DC) has emerged in the literature, being applied and fully justified by Prommas and Sinwongsuwat (2011), Cowan (2008), and Biber (2006). Cowan additionally features the ability of DCs to connect a sentence to a greater piece of discourse with less restriction in terms of where they may occur in a sentence. It is worthwhile to clarify that there are three major types of DCs, namely discourse markers, adverbial connectors or linking adverbials, and coordinating conjunctions. In fact, DCs is an umbrella term and in this study it entails the instances of all three types.

Generally, without considering the three aforementioned types DCs have been studied under different terms including, but not limited to connectives (Huddleston & Pullum, 2002), discourse markers (Fraser, 1999), connectors (Biber et al., 1999), conjunctive adverbials (Celce-Murcia, Larsen-Freeman, & Williams, 1983), pragmatic particles (Östman, 1981), pragmatic operators (Ariel, 1994), cue phrases (Knott & Dale, 1994), pragmatic expressions (Erman, 1987), discourse operators

(Redeker, 1990 & 1991), phatic connectives (Bazanella, 1990), pragmatic markers (Fraser, 1988) and (Schiffrin, 1987), discourse marker (Schiffrin, 1987), pragmatic formatives (Fraser, 1987), discourse connectives (Blakemore, 1987; Blakemore, 1992), discourse particles (Schourup, 1985), semantic conjuncts (Quirk, Greenbaum, Leech & Svartvik, 1985), discourse signaling devices (Polanyi & Scha, 1983), pragmatic connectives (Stubbs, 1983), and sentence connectives (Halliday & Hasan, 1976).

This study chose to use the term ‘discourse connector’ and apply it to cover DCs of both spoken and written English according to the definition put forward by Rezvani Kalajahi and Neufeld (2014), that “[discourse connectors] are words and expressions that can be accommodated within the text [or speech] in order to join one sentence to another sentence or one paragraph to another paragraph.” Obviously, the appropriate use of DCs will help the reader to easily follow the direction of the text or the speech and understand his points of view, arguments, examples, etc., and yield cohesion both in spoken and written language.

As Biber (2006) puts forward:

DCs are devices used to bridge between turns and sentences, indicating the logical relations among the parts of a discourse, and providing an interpretive framework for the listener/reader. There are two major classes of DCs: discourse markers and linking adverbials. DMs — forms like *ok, Well, and now* — are restricted primarily to spoken discourse. These forms have distinct discourse functions, but it is difficult to identify the specific meaning of the word itself. In contrast, linking adverbials — forms like *however, thus, therefore, for example (e.g.), and that is (i.e.)* — are found in both spoken and written registers, and they have greater inherent meaning than discourse markers (p.133).

By using appropriate DCs, a reader is able to follow easily the direction of the text and understand a writer’s points of view, arguments, examples, etc. Cohesion is a text feature upon which good writing heavily relies, but it is not limited to DCs because they are one of the means of cohesive devices proposed by Halliday and Hasan (1976) that help the text or piece of writing to be coherent.

A number of taxonomies have been offered which take into account the various functions and meanings that these devices may assume (Fraser 1999; Celce-Murcia and Larsen-Freeman, 1999; Biber et al., 1999; Quirk et al, (1985). As these accounts appeared to fail to provide a comprehensive description of these cohesive elements, Rezvani Kalajahi and Neufeld (2014) proposed a new taxonomy of DCs validated against the two corpora of British National Corpus (BNC) and Corpus of Contemporary American English (COCA). The researcher has decided that this new taxonomy serves as the major tool in this study for the analysis of the treatment of the DCs as used in Malaysian ESL students’ writing. They have built on this pioneering effort with the intention that it is comprehensive enough for material

developers, teachers, learners and non-native writers and speakers of English to depend on. As a result, the taxonomy offered by Rezvani Kalajahi and Neufeld (2014) has been adopted and expanded to encompass eight broad classes, subdividing six of the broad classes into a total of seventeen categories.

It is worth mentioning here that there exist plenty of research studies carried out on DCs. For example, DCs have been investigated in classroom oral discourse (Hays, 1992), informal settings (Lee, 1999; Müller, 2004; Trillo, 2002), reading (Abdollah Zadeh, 2006; Jalilifar, 2009; Jalilifar & Alipour, 2007), lectures (Dailey-O'Cain, 2002; Pérez & Macià, 2002), academic genres (Abdi, 2002; Blagojevic, 2004; Bunton, 1999; Longo, 1994; Mauranen, 1993; Ventola & Mauranen, 1991), and student writings (Connor, 1984; Field & Oi, 1992; Intaraprawat & Steffensen, 1995; Johns, 1984; Johnson, 1992; Karasi, 1994; Norment, 1994; Steffensen & Cheng, 1996). These studies have targeted the frequency counts and patterns of use of these devices.

It is perhaps of pedagogical interest to point out that the great bulk of recent research on DCs have focused on the use of these linking devices in learner written discourse (Granger and Tyson, 1996; Altenberg and Tapper, 1998; Tankó, 2004; Narita, Sato, and Sugiura; 2004; Tang and Ng, 1995; Yaochen, 2006; Milton, 2001; Yoon, 2006; Chen, 2006; Leńko-Szymańska, 2007; etc.). Findings of these studies are, nevertheless, inconsistent. While Chen (2006), Tankó (2004), Yoon (2006) and Milton (2001) reported on a general overuse of connectors by non-native speakers of English, Altenberg and Tapper (1998) found a general tendency for underuse of the connectors among the Swedish. Granger and Tyson (1996), on the other hand, noticed neither general overuse nor underuse of connectors in the essays composed by the French students. Despite the observed differences in the general use, over/underuse and misuse of individual connectors were observed by all the authors.

1.3 Corpus and its Connection to Language Teaching

Corpus or 'corpora' are large and principled collections of natural texts (Reppen, 2010). For most linguists, a corpus cannot only be equated with just a large collection of texts but needs to be justified in linguistic terms. Therefore, a corpus can be defined as a collection of texts assumed to be representative of a given language or subset of a language to be used for linguistic analysis (Tognini-Bonelli, 2001).

The aim of corpus linguistics is to analyze and describe the language use as realized in the selected texts (Tognini-Bonelli, 2001). Schmitt (2000) pointed out "insights from corpus research have revolutionized the way language is viewed, especially words and their relationship with each other in context".

Corpus linguistics does not begin by accepting certain rules as given; it rather defines its own sets of rules before being applied. It provides new rules and parameters for linguistic description (Tognini-Bonelli, 2001). Corpus research permits researchers and learners to gain insights into different aspects of the language. It, for example, enables the researchers to notice how DCs are used and to observe how ESL/EFL learner patterns as well as frequency of use of the DCs compare with the native writer norm. Through the use of powerful concordance software, corpus analysis has enabled the linguists to gain new insights into the language and its use (Ghadessy, Henry & Roseberry, 2001).

Corpus use has also contributed to language teaching, especially in the compilation of pedagogical grammars and dictionaries (Hunston & Francis, 2000; Kennedy, 1990;) (Gabrielatos, 2005). Flowerdew (2003) adds that corpus linguistics contributes to language teaching by selecting the features of language which seem worth teaching in a given pedagogical content. Research on corpus, particularly learner corpus, consists of almost all the instances of language that a learner has been exposed to in a classroom, from language in teacher instructions to language in textbooks (Hunston, 2002). Such enquiries have also contributed a lot to the understanding of language learning processes (Granger, Hung, & Petch-Tyson, 2002) and to the construction and evaluation of language tests (Alderson, 1996). Most importantly, research on corpus from language learners has enabled the comparison of the language to which learners are used with reference corpora or real-language corpora, resulting in the development of more effective pedagogical materials (Gabrielatos, 2005).

In fact, an instructional advantage of a corpus is that when an item is met in a text, examples of the item from similar text can be used as evidence for the learner to draw conclusions about the syntagmatic and paradigmatic use of that item in that particular genre. In other words, it helps learners to recognize patterns or phraseology particular to that discourse (Hunston, 2002; Sinclair, 1991). Other than benefiting learners and teachers, a learner corpus would be useful in re-designing teaching materials in the future. For instance, identification of the phraseological patterns used by the learners and their comparison with real world language use could aware material writers, teachers, and even language learners of how similar to or different from the native speakers the learners are using the language.

1.4 Nature of the Present Study

The present research focused primarily on the so far understudied DCs in written discourse of Malaysian ESL students through a corpus-based approach. To this end, it adopted both qualitative and quantitative methods. First and foremost, the main goal of this current research was building a written corpus of Malaysian ESL students. Next in importance for the study was to carry out a thorough corpus-based study of DCs use in the writing of Malaysian ESL students. The study sees it as its third objective to determine if employing an adequate number of DCs will result in

higher quality of the written text. It finally aimed to seek whether or not Malaysian students' could appropriately utilize DCs in, syntactic, semantic and stylistic terms.

1.5 Rationale of the Present Study

Several reasons motivated the present study. First of all, as Granger (2004) states that research into the use of corpus for language teaching is almost entirely done by linguists; the contribution of SLA researchers to – and the participation of ESL teachers in – what happens in corpus linguistics is still relatively low. It is, however, crucial to involve SLA researchers and the ELT community, including teachers and learners, to a much larger extent in the actual work on corpus and in the systematic evaluation and improvement of corpus-based activities in the classroom.

When it comes to the Malaysian context, internet search revealed that there was no empirical research on DCs by using corpus-based approach. Besides corpus-based study of DCs, the researchers was interested to find out to what extent writing quality in the writing of Malaysian students depend on DCs use and to what extent students misuse DCs in their writing since there was no single research conducted in the context. Aforementioned shortcomings made the researcher think about modification of method and carrying out a comprehensive study that may make new contribution to the current context of English language teaching in Malaysia.

Previous researchers studying DCs attempted only to extract the frequency of DCs manually without comparing with any reference corpus, to explore errors of DCs use randomly without adopting or following a scientific framework, and also to correlate writing quality with the quantity of DCs in the texts by using Halliday and Hassan's taxonomy (1976), Fraser's classification (1999), Quirk et al's (1985) framework, Biber et al. (1999) taxonomy, Celce-Murcia and Larsen-Freeman (1999) framework. However, there are two major criticisms of these taxonomies. First, the classification and taxonomies are not thorough enough to take into account all the possible DCs in the real language use. For instance the above mentioned taxonomies failed to include comprehensive list of DCs and they were only introduced some instance that were mainly common among those taxonomies. The previous studies, therefore, managed to develop and use a list of DCs incorporating elements from different taxonomies. Second, learners' erroneous use of DCs which deteriorated the quality of text were not taken into account. It is, however, often argued that it is not only the quantity of DCs but also the proper use of them that makes a text hang together or improves its readability. Despite this, review of the existing literature revealed that majority of the studies including ESL and EFL context (Johns 1984; Crewe, 1990; Field, 1994; Field & Yip, 1992; Tickoo, 1998; Green, Christopher, Lam, 2000; Lake, 2004) investigated DCs mainly quantitatively and failed to address exploration of the use of DCs qualitatively.

To date various methods have been developed and introduced to investigate DCs, but there were two major reasons why corpus-based approach was adopted in the current research to investigate DCs with a different list. First, computer text-processing capabilities have provided an opportunity to investigate large samples of learner writing through corpus-based research. Text processors have also enabled the researchers to compare learner-created texts with those written by native speakers in an attempt to corroborate their intuitions about different aspects of language use (Chen, 2006; Granger & Tyson, 1996). Moreover, the main strengths of using corpus in language research have been identified by researchers interested in corpus studies (Biber, Conrad & Reppen, 1996). They highlight that computerized corpora “provide large databases of naturally occurring discourse, enabling empirical analyses of the actual patterns of use in a language; and, when coupled with (semi-) automatic computational tools, the corpus-based approach enables analyses of a scope not otherwise feasible”.

1.6 Statement of the Problem

The English language had played a dominant role as a compulsory subject and medium of instruction in English medium schools which formed the backbone of the education system of pre-independence Malaysia (Pandian, 2006, cited in Voon Foo, 2007, p. 1). “Writing is such an important learning tool because it helps students to understand ideas and concepts better” (Voon Foo, 2007, p. 4).

From theoretical perspective, social constructivist theory and social cognitive theory provide deep insight into how writing is happening. It can be noticed that writing is not simply a cognitive activity but a social act as well. ESL/EFL writers have to take into account the social-cognitive aspects in composing, organizing and analyzing the written discourse. A challenge that has received relatively little attention in studies dealing with written discourse is that ESL/EFL students generally face difficulties in using higher cognitive skills which include the use of DCs in their writing (Tan, 2010).

As students manage to put their ideas into words, they are expected to make their message clear to the reader. The writer has to narrow the topic and explain it in as much detail as possible. Nothing should remain vague. In doing so, the ultimate goal is to form a well-organized, unified, and coherent composition in which DCs can play a pivotal role in yielding cohesion to the text (Demirci & Kleiner, 1997). Despite this, empirical studies have confirmed that Malaysian ESL students have major problems in building cohesion in their writings. There is, therefore, a considerable need for the future generations of Malaysian students to master in writing skill, knowing how to write coherently (Ghabool, Mariadass & Kashef, 2012).

The need for studying DCs was due to their importance in writing. ESL and EFL students need to use DCs appropriately. Proper use of DCs allows students to differentiate their multiple functions and various grammatical constructions. Failing

to utilize them appropriately, students will face problems composing a coherent text. If this issue is not addressed at the right time during school years, it may eventually affect learner writing at the tertiary level.

Wu (2006), in an empirical study, found a main problem with nonnative speakers' use of DCs in writing tasks. Asian students, in particular, tended to overuse or under use some types of DCs. It was also found that students placed DCs in sentence-initial position more often than native speakers do. Furthermore, the use of DCs has been reported to be problematic both for native and non-native speakers (Heino, 2010; Crew 1990, Altenberg & Tapper, 1998). Heino (2010) argues that DCs usage may cause difficulties for ESL/EFL students due to several reasons. First of all, appropriate use of DCs calls for right recognition of discourse type and register, which is, to a great extent, complicated for the EFL/ESL learner to gain. In order to know which DCs to use in a certain situation, the student writer needs to know about different registers and text types. For instance, certain connectors like *therefore*, *thus*, *hence*, etc. belong to the formal registers whereas the resultive connector *so* and the contrastive connector *anyhow* are generally used in informal registers (Altenberg & Tapper, 1998; Quirk et al., 1985). Another problem is that mere use of DCs does not contribute to the quality and richness of the text and languages differ from one another in their preference for the use of these devices. Yet there is no piece of evidence to show whether or not the quality of writing depends on the quantity of DCs in the context of Malaysia. Finally, correct use of DCs tends to be influenced by factors like the development of the learners' communicative competence and how language is taught (Altenberg and Tapper, 1998).

In a cross-cultural study, Johnson (1992) examined cohesion in expository Malay and English essays composed by the native speakers. The research findings showed that "Good" Malay compositions had more intersentence semantic ties (e.g., reiteration and collocation) than "weak" compositions. On the other hand, "good" English compositions turned out to have more intersentence syntactic ties (e.g., reference and conjunction [connectors]) than "weak" compositions. In other words, languages seem to differ from one another in their use of techniques to create coherent texts. Implied here is that Malay learners' use of DCs to create coherent texts in English might, under the influence of their cultural-linguistic background, be radically different that of the English native speakers.

Another issue of pedagogical concern in ESL/EFL use of DCs has to do with misuse of DCs. Errors in language learning have always been the center of research attention and dealing with errors have become one of the most actively discussed issues in language pedagogy (Stapa & Izahar, 2010). Johnson (1992) indicates that forming coherent texts presents a problem to Malaysian ESL learners. This is supported by Ahour and Mukundan (2012) when they found that majority of the students in Malaysia experienced problems with the appropriate use of DCs in their writing. Therefore, analyzing learners' errors provides useful insights into the features of the language that Malaysian ESL students are using.

According to Leki (1992), ESL students are also disappointed with difficulties in understanding the implications and various meanings of DCs. It is obvious that there is a divergence between writing in the first and the second language. Musa, Lie, and Azman, (2012) believes that ESL students may encounter with a variety of complicated problems, which may be either culturally or linguistically motivate. These problems may create difficulty for both teachers and ESL student writers, therefore paying attention to different aspects of variation of use can aid ESL students adapt themselves to the writing standards.

As DCs make a text coherent, it is logical to suppose that the lack of use or inappropriate use DCs in an L2 could, to a certain degree, hinder understanding or even lead to misunderstanding (Lahuerta Martínez, 2004). Malaysian students often cannot construct an organized and coherent text in English. The researcher's observations also confirmed that it is difficult for students to make use of various DCs in their writings. Certainly, the fault cannot be put on the shoulders of the students mainly because there are research evidence that developers of materials and dictionaries in particular as well as teachers are not fully aware of them. As a result, they are often introduced as a list requiring the learners to use them interchangeably (Bikelienè, 2008; Rezvani Kalajahi, Abdullah, Baki, 2012).

An important observation that inspired the present study is the lack of literature on how Malaysian ESL students use DCs. Despite some studies in written discourse by L2 learners with different L1 backgrounds in other contexts, the DCs use by Malaysian ESL students of English has not been the focus of research. Unfortunately, there is no comprehensive study and no single published evidence of corpus-based study of DCs by Malaysian students, although a bulk of studies have discussed the roles and importance of DCs in relation to the comprehension of a text or written discourse or listening comprehension around the world. Romero Trillo (2002) and Muller (2005) opined that in spite of the highly important role of DCs in achieving coherence, it is vitally important to note that they have been rather neglected in the ESL curriculum.

More importantly, the current blueprint and English language curriculum of Malaysia under language content category clearly state that DCs have to be taught and assessed systematically as they bring cohesion to a piece of writing. However, school English textbooks, which are main source of teaching and learning in the context of Malaysia, offer only a limited number of DCs. Word processing enquiries revealed that less than 90 DCs have been introduced to the Malaysian pupils in both Primary and Secondary levels (Rezvani Kalajahi & Abdullah, 2012). Enquiries also showed that they are not systematically distributed across different stages even though their role is emphasized in both levels in the Curriculum Specification of each respective level. The textbooks also seem not to have stressed on the pedagogical aspects of DCs either explicitly or implicitly. As a result, the learners may not realize at what level of education they should learn and use them efficiently and appropriately.

In order to cultivate this unstudied field, this is perhaps the first study to use a corpus-based approach to investigate the use of DCs by Malaysian ESL students. This study contributes to practical implications of raising students' understanding of what DCs are, teachers' understanding of how they can be taught in classroom and material developers' understanding of how suitable and appropriate teaching material should be developed in order to help students to learn and use DCs effectively.

1.7 Objectives

Generally, writing sentence, paragraphs or larger units of discourse and connecting them together requires aspects more than the patterns within the sentence level (Dülger, 2007). Therefore, this study mainly concerns writing beyond the sentence level, because efficient use of the DCs proves to have a great contribution to stick the sentences together, creating a cohesive, coherent and unified text. Appropriate use of DCs enables the writer to make strong connections among ideas and deliver them clearly to the readers (Cavender & Weiss, 1987). That being said, the current study seeks:

1. To build Malaysian students' written English corpus
2. To investigate the frequency of the use of DCs in the students' writing and to compare it with that in the Native students corpora (Louvain; the Louvain Corpus of Native Essay Writing [LOCNESS]).
3. To find out semantic distribution of DCs in Malaysian school students' writing and native students
4. To examine the relationship between the use of DCs frequency and the quality of writing
5. To explore the erroneous use of DCs in the writing of Malaysian ESL students

1.8 Research Questions

The ultimate goal of this study is to explore this unknown area and attempts to provide a continuum of the importance of a corpus-based analysis to make a principled decision on using DCs by students or input for instructional purposes that have not been previously studied. To achieve the purpose of this study, the following research questions are addressed.

1. What is the overall frequency of DCs use in argumentative essays of Malaysian school students' writer and native students' writer?
 - 1.1 What is the frequency of DCs use in three different educational levels, namely Form 4, Form 5, and College?
 - 1.2 Is there any significant difference in the frequency of use of DCs between Malaysian school students and the native speakers of English?
 - 1.3 What are the most frequent DCs used in Malaysian ESL students' writing and native writing?

2. What is the sematic distribution of DCs in Malaysian school students' writing and native students?
3. Is there any relationship between the use of DCs frequency and the quality of the essays in different educational levels?
4. What kind of errors do Malaysian ESL student commit in using DCs?

1.9 Significance of the Study

The significance of this study can be viewed in a number of ways. This study would provide insightful evidence on the knowledge base of DCs among Malaysian ESL students in their writing, particularly in argumentative compositions. Writing instructors may gain insights from the results and get to know how to help the students to improve their writing skills in different ways in terms of DCs use. Studying learner use of DCs is useful for writing instructors mainly because it empowers them to assist students to enhance their knowledge on how to “engage with their topic and their readers” and the proper means of conveying their “attitude, marking structures and engaging with their readers” in their writing (Hyland, 2004; p.146). In addition, comparison between native speakers and ESL learner would provide helpful insight into the extent to which L2 learner use of DCs is similar to or different from that of the native speaker. Recognition of the use of DCs displayed in the concordance lines would make such comparison plausible.

It is hoped that through the assistance of informative insights provided by the results of the present study, the problems encountered by Malaysian students of different levels of education in the use of DCs could be effectively addressed. Being aware of the different problems with the learner use of DCs, writing instructors both at school and tertiary level may find it useful to design specific writing materials which would satisfy learners' individual needs as to the acquisition of DCs in order to improve their writing proficiency remarkably. The correct use of DCs in writing is vitally important as they act as a signpost in writing that help readers to the intended message of the writer. A failure in the appropriate use of DCs will definitely cause confusion and the text might appear incomprehensible to the readers.

Finally, it is anticipated that assigning a right DC to a right educational level would help both English teachers and learners to realize the appropriate time for learning a certain number of DCs. Furthermore, material developers may take the list even more serious to be able to incorporate them efficiently into the right level so that both teachers and learners may know which DCs are required to be taught and learnt by the end of the course.

1.10 Limitations

In spite of the research and pedagogical significance of the current study, it is threatened by several limitations. First, the findings would be more reliable if there

were a possibility of creating a writing corpus of native students with the same age, education level, in the same genre type with the same task setting, etc.

Second, This research could not use BNC, COCA or Brown corpus as a reference corpus since the coverage of the topics in these general corpora are totally different and they cover broad themes and topics from newspapers, plays, TV shows and so forth. Hence, they could not be assumed as suitable reference corpora against which the results of Malaysian school students' findings might be compared.

Third, the researcher was unable to develop a spoken corpus of Malaysian students mainly because it turned out to be costly, time consuming and labour intensive. The fourth limitation was the issue of sampling in building the corpus. Just because the researcher could not access all the Malaysian school students with the same equal number from three ethnic groups (Malay, Chinese, and Indian) and request them to write about a topic for the corpus purposes, opportunistic sampling was used to build the learner corpus and the level of proficiency was confined to Form 4, Form 5, and College students who are known as Upper Secondary level.

Fourth, the data were collected only from the students attending the schools in Selangor, Sembilan and Malacca. Unfortunately, other states could not be included in the corpus compilation because creating a larger corpus required a team with certain amount of budget to be allocated by pertinent organizations, a great deal of time and energy.

Fifth, as is the case for all research methods, there are potential limitations of this corpus-based approach. One of the main disadvantages of such approach lies in the ways in which linguistic information can be retrieved (Leech, 1998). When investigating large corpora, the researcher is for all practical purposes confined to investigating linguistic features which are feasible to search for by computer software. There are in principle two factors that control the searchability of a corpus. The first deals with the restrictions set by the available search and retrieving software. The second has to do with the corpora that have not been annotated in some way, e.g. tagged or parsed. Such corpora primarily leave the researcher to search for those linguistic features that are visible in the electronic record of the text.

Furthermore, the entire instances of discourse markers, adverbial connectors and coordinating conjunctions which are three major types of DCs were not considered in detail and all instances were studied under the term of DCs.

Finally, the qualitative analysis of this research was primarily carried out by the researcher and one non-native instructor. The researcher, though with the help of various reliable references such as grammar reference books and authentic language corpus, is not an English native speaker and may not be able to identify every problem learners encounter. The learners' problems are re-confirmed by the instructors who is already an experienced instructor in ELT, but this may reflect one-sided opinions.

1.11 Definition of Terms

Prior to reviewing the related literature in the next chapter, the key words related to the study are defined both conceptually and operationally.

1. Discourse connectors

Discourse markers, adverbial connectors, and coordinating conjunctions form DCs. They are words and expressions that can be accommodated within the text in order to join one sentence to another sentence or one paragraph to another paragraph. By using appropriate DCs, a reader/listen is able to follow easily the direction of the text/speech and understand a writer's/speakers points of view, arguments, examples. Cohesion is a text/speech feature upon which a good writing heavily relies on, but it is not limited to DCs because they are only one of the cohesive devices proposed by Halliday and Hasan (1976) that help the text or writing to be coherent. In this study, DCs includes all instances of 'discourse markers, adverbial connectors, and coordinating conjunction' and it is referred to words and expression such as "but, yet, moreover, also, nevertheless, in addition to, and for example" that Malaysian ESL students used in their writing to connect one sentence to another through giving examples, refuting an earlier idea or adding some more information.

2. Corpus

Corpus is "a large, principled collection of naturally occurring texts (written or spoken) stored electronically" (Reppen, 2010). Corpus is perhaps best described in simple terms as "the study of language based on examples of real-life language use" (McEnery & Wilson, 2001: 1). Linguists have always utilized the word corpus to elucidate and describe a collection of naturally occurring examples of language, containing of anything from a few sentences to a set of written texts or recordings which have been composed for linguistic study. Corpus in this study involves written composition of Malaysian ESL students who were in Form 4, Form 5, and College level.

3. Corpus-based approach

Corpus-based studies use a corpus as a source of examples to check researcher intuition or to examine the frequency and/or plausibility of the language contained within a smaller data set. The researcher following a corpus-based approach does not question pre-existing traditional descriptive units and categories (Tognini-Bonelli, 2001).

4. Learner corpus

The really authentic texts for foreign language learning are not those produced by native speakers for native speakers, but those produced by learners learning the target language. James (1998) holds that a learner corpus consists of language output produced by learners of a language. Most learner corpora consist of written essays using pre-set topics produced in language-teaching classrooms. Learner corpus, in this study, refers to the words, phrases and language that Malaysian ESL students produced from three educational levels (Form 4, Form 5, and the first year students in College) via asking them to write about an argumentative during a class time.

5. Reference Corpus

A reference corpus is one that is designed to provide comprehensive information about a language (Leech, 2002; Sinclair, 1996). It aims to be large enough to represent all the relevant varieties of the language, and the characteristic vocabulary, so that it can be used as a basis for reliable grammars, dictionaries, thesauri and other language reference materials. In this study, reference corpus refers to the Louvain Corpus of Native English Essays (LOCNESS), composed of samples of argumentative written English texts by British and American student writers consisting of a total of 325,000 tokens.

6. Persuasive Versus Argumentative writing

Persuasion and argument are often used interchangeably. The former one is a broad term, which embraces many tactics designed to move people to a position, a belief, or a course of action, whereas the latter one is a specific kind of persuasion based on the principles of logic and reasoning and present an argument with the PROS (supporting ideas) and CONS (opposing ideas) or comparison and contrast of an argumentative issue (Lai, 2008). For the corpus collection in this study, the Malaysian students were asked to write about an argumentative topic in which they were assigned by the researcher.

7. ESL writing

English as a second language (ESL) is the use or study of English by speakers with different native languages. A piece of writing can be considered as ESL writing when a text is written by ESL writers. English in Malaysia is considered as a second language and therefore, in this study, ESL students are referred Malaysian students who contributed to this study.

8. Concordance

Concordance or Key Word in Context (KWIC) is a list of all of the occurrences of a particular search term in a corpus with a few words to either side (Baker, Hardie, & McEnery, 2006). In this study, it includes an electronic list of entries from the corpus containing the DCs along with their context. This study used concordance function in the WordSmith version 5.0.

9. Type and token

The number of tokens in a corpus refers to the total number of words while the number of types refers to the total number of unique words. For example, the word 'however' may occur 177 times in a corpus, but it only counts as one type of word.

10. Normalized or Standardized Frequency

A frequency expressed relative to some other value as a proportion of the whole -- for example, a frequency of a word relative to the total number of words in the corpus. As there were two corpus in this study with different size, normalized frequency was set to per 1.000.000 tokens.

11. Correlational Analysis

Samah (2013) states that correlation analysis is to examine, evaluate and assess strength or the weakness of the relations between variables. A positive correlation can be found where the high values of one variable are linked with the high values of the other variable(s). In contrast, a negative correlation implies association of high values of one with the low values of the other(s). Correlation can vary from +1 to -1. Values which are close to +1 represent a high-degree of positive correlation, and values close to -1 show a high degree of negative correlation. Values close to zero indicate poor correlation of either kind, and 0 indicates no correlation at all. While correlation is useful in discovering possible connections between variables, it does not prove or disprove any cause-and-effect (causal) relationships between them. In this study correlational analysis is performed to determine if there is any relationship between quality of writing and the frequency of DCs use.

12. Error Analyses

Making errors for the second and foreign language learners could be unavoidable part of their language production. From the perspective of norm or well-formedness of a sentence, error is observed as “an infringement or deviation of the code of the formal system of communication through which the message is conveyed” (Catalan, p. 66). Errors are alterations of the rules of the accepted norm and are termed as surface errors which may be further classified as semantically, syntactically, positionally, etc. In this study, error refers to the unintentional misuse of the DCs that are observable and clearly identifiable in the Malaysian ESL students’ writing.

REFERENCES

- Abdi, R. (2002). Interpersonal meta-discourse: An indicator of interaction and identity. *Discourse Studies*, 4(2), 139-145.
- Abdollahzadeh, E. (2006). The role of textual signals in L2 text comprehension. *ESP Malaysia*, 12, 1-18.
- Abdollahzadeh, E. (2010). Poring over the findings: Interpersonal authorial engagement in applied linguistics papers. *Journal of Pragmatics*, 43, 875-890.
- Ahmad, J. (2012). Fungsi wacana dalam bahasa Melayu. Retrieved May 30, 2013, from <http://jaafar2026.blogspot.com/p/fungsi-wacana-dalam-bahasa-melayu.html>.
- Ahour, T., & Mukundan, J. (2012). Errors and Variations of TESL Students' Written Description. *Pertanika J. Soc. Sci. & Hum*, 20 (1), 55 – 64.
- Aidinlou, N. A. (2012). The Effect of Discourse Markers Instruction on EFL Learners' Writing. *World Journal of Education*, 2(2), 10-16.
- Akbari, O. (2009). A Corpus Based Study on Malaysian ESL Learners' Use of Phrasal Verbs in Narrative Compositions. (Unpublished doctoral thesis). Serdang: Universiti Putra Malaysia.
- Alarcon, J. B., & Morales, K. N. S. (2011). Grammatical cohesion in students' argumentative essay. *Journal of English and Literature*, 2(5), 114-127.
- Alderson, J.C. (1996). Do corpora have a role in language assessment?. In T. M. H. Short (Ed.), *Using corpora for language research* (pp. 248-259). London: Longman.
- Altenberg, B. (1997). Exploring the Swedish component of the International Corpus of Learner English. In *Proceedings of PALC'97 Practical Applications in Language Corpora (Lódz, 10-14 April 1997)*, eds. B. Lewandowska-Tomaszyk & P.J. Melia, 119-132. Lódz: Lódz University Press.
- Altenberg, B., & Tapper, M. (1998). The use of adverbial connectors in advanced Swedish learners' written English. In S. Granger (Ed.), *Learner English on computer*. (pp. 80-93). London and New York: Longman.
- Ancker, W. (2000). Errors and corrective feedback: Updated theory and classroom practice. *English Teaching Forum*. 38(4), 20-24.
- Anscombe, J.C., & Ducrot, A. (1983). *The Arguments in the language*. Brussels: Madarga.
- Ariel, M. (1994). *Pragmatic operators*. The Encyclopedia of Language and Linguistics, 3250-3253. Oxford: Pergamon Press.
- Ary, D., Jacobs, L. C., Razavieh, A., & Sorensen, C. (2006). *Introduction to Research in Education*: Thomson Wadsworth. Belmont, CA.
- Arshad, A.S., Mukundan, J., Ghazali, K. Sharifah, Z. Juidah, R., & Edwin, M.V. (2002). *The English of Malaysian School Students (EMAS) corpus*. Serdang: Universiti Putra Malaysia Press.

- Blakemore, D. (1987). *Semantic constraints on relevance* (Vol. 57): Blackwell Oxford.
- Blakemore, D. (1988). So as a constraint on relevance. In R. Kempson, ed., *Mental representations: The Interface between Language and Reality*. Cambridge: Cambridge University Press, 183-195.
- Blakemore, D. (1989a). Denial and contrast: A relevance theoretic analysis of but. *Linguistics and Philosophy*, 12, 28-51.
- Blakemore, D. (1989b). Linguistic form and pragmatic interpretation: the explicit and the implicit. In Leo Hickey, ed., *The Pragmatics of Style*. London: Routledge.
- Blakemore, D. (1992). *Understanding utterances: An introduction to pragmatics*: Oxford: Blackwell.
- Blakemore, D. (1993). The relevance of rdoTmteúón. *Language and Literature*, 2(2), 101-220.
- Blakemore, D. (2000). Indicators and procedures: nevertheless and but. *Journal of Linguistics*, 36(03), 463-486.
- Blass, R. (1990). *Relevance Relations in Discourse: A Study with Special Reference to Sissala*. Cambridge: Cambridge University Press.
- Bolton, K., Nelson, G., & Hung, J. (2002). A corpusbased study of connectors in student writing: Research from the International Corpus of English in Hong Kong ICEHK. *International Journal of Corpus Linguistics*, 7(2), 165-182.
- Bondi, M. (2001). Small corpora and language variation: Reflexivity across genres. In M. Ghadessy, A. Henry, & A.L. Roseberry, (Eds), *Small corpus studies and ELT*, (pp. 135-174). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Botley, D. A., Metom, D., & Izza, A. (2005). CALES: a corpus based archive of learner English in Sarawak. A Report. Universiti Teknologi MARA.
- Brown, H. D. (2001). *Teaching by principles: An introduction approach to language pedagogy* (2nd ed.). New York: Longman.
- Brown, D. B. (1994). *Principles of language learning and teaching*. Third edition. New Jersey: Prentice Hall.
- Bunton, D. (1999). The use of higher level metatext in Ph. D theses. *English for Specific Purposes*, 18, 41-56.
- Candling, R. B. (2001). *Vocabulary and language teaching*. New York: Longman Inc.
- Carlsen, C. (2010). Discourse connectives across CEFR-levels: A corpus based study. *Monographs of the European Second Language Association*, 191(1), 191-209.
- Carter, R., & McCarthy, M. (2006). *Cambridge grammar of English*. Cambridge: Cambridge University Press.
- Carrier Pastor, L. M. (2006). The use of connectors in scientific articles by native and non-native writers. *ESP-world*, 5(2).

- Cavender, N. & Weiss, L. (1987). *Thinking/Writing*. U.S.A.: Wadsworth Co.
- Celce-Murcia, M., Larsen-Freeman, D., & Williams, H. A. (1983). *The grammar book: An ESL/EFL teacher's course*: Newbury House Rowley, Massachusetts.
- Celce-Murcia, M. & Larsen-Freeman, D. (1999). *The Grammar Book*. Boston: Heinle & Heinle.
- Chen, C.W. (2006). The use of conjunctive adverbials in the academic papers of advanced Taiwanese EFL learners. *International Journal of Corpus Linguistics*, 11(1), 113-130.
- Chen, W. (2011). *An Investigation of Topic Sentences in Chinese Students' Argumentative Essays: A Multidimensional Probe* (Doctoral dissertation, Kristianstad University).
- Cobb, P. (1996). Where is the mind? A coordination of sociocultural and cognitive constructivist perspectives. In C. T. Fosnot, *Constructivism: Theory, perspectives, and practice* (pp. 34-52). New York: Plenum
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1), 37-46.
- Connor, U. (1984). A study of cohesion and coherence in English as a second language students' writing. *Research on Language & Social Interaction*, 17(3), 301-316.
- Conrad, S. R., & Biber, D. (2004). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Conrad, S. (2005). Corpus linguistics and L2 teaching. In E. Hinkel (Ed.), *Handbook of Research in Second Language Teaching and Learning*, 1. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cook, G. (1989). *Discourse*. Oxford: Oxford University Press.
- Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.
- Corder, S. P. (1974). Idiosyncratic dialects and error analysis. in Richards, J. C. (Ed.), *Error Analysis* (pp.158-171). London: Longman.
- Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics*, 5(4), 161-169.
- Cowan, R. (2008). *The Teacher's Grammar of English: A Course Book and Reference Guide, with answers*: UK: Cambridge University Press.
- Crewe, W. J. (1990). The illogic of logical connectives. *ELT Journal*, 44(4), 316-325.
- Cumming, A., & Riazi, A. (2000). Building models of adult second-language writing instruction. *Learning and Instruction*, 10(1), 55-71.
- Dagneaux, E., Denness, S., & Granger, S. (1998). Computer-aided error analysis. *System*, 26(2), 163-174.
- Dailey-O'Cain, J. (2002). The sociolinguistic distribution of and attitudes toward focuser like and quotative like. *Journal of Sociolinguistics*, 4(1), 60-80.

- Darus, S., & Subramaniam, K. (2009). Error analysis of the written English essays of secondary school students in Malaysia: A case study. *European Journal of Social Sciences*, 8(3), 483-495.
- Davies, M. (2004). *BYU-BNC*. (Based on the British National Corpus from Oxford University Press). Available online at <http://corpus.byu.edu/bnc/>.
- Davies, M. (2008). The Corpus of Contemporary American English: 450 million words, 1990-present. Available online at <http://corpus.byu.edu/coca/>.
- Demirci, M., & Kleiner, B. (1997). Discourse Markers in Second Language Research. *Journal of Intensive English Studies*, 11, 131-142.
- De klerk, V. (2004). The use of 'actually' in spoken Xhosa English: A corpus study. *World Englishes*, 24(3), 275-288.
- De klerk, V. (2005). Procedural meanings of 'well' in a corpus of Xhosa English. *Journal of Pragmatics*, 37(8), 1183-1205.
- De Monnik, I. (2004). Using Corpora and Experimental Data: A Multimethod Approach. In Lejung, Magnus (ed.) *Corpus-based Studies in English. Papers from the 17th International Conference on English Language Research in Computerized Corpora*. Amsterdam: Rodopi, 227-244.
- Deng, F. (2006). The effect of the use of adverbial connectors on Chinese EFL learners' English writing quality. *CELEA Journal (Bimonthly)*, 29(1), 105-111.
- Deng, Y. (2006). Use of Adverbial Connectors in Chinese EFL Learners Argumentative Writing. *CELEA Journal*, 29(6), 32.
- Ducrot, O. (1972). *Diré et nepasDiré*. Paris: Hermann.
- Ducrot, O. (1973). *La Freuve et le Diré*. Paris: Mame.
- Ducrot, O. (1984). *Le Diré et le Dit*. Paris: Minuit.
- Dülger, O. (2007). Discourse markers in writing. Retrieved 20 December 2013, from http://www.sosyalbil.selcuk.edu.tr/sos_mak/makaleler/Osman%20D%3%9CLGER/D%3%9CLGER,%20OSMAN.pdf
- Education in Malaysia. (n.d.). In Wikipedia. Retrieved 13 March 2013, from http://en.wikipedia.org/wiki/Education_in_Malaysia.
- Eldridge, J., Neufeld, S., & Hancioğlu, N. (2011). Towards a Lexical Framework for CLIL. *International CLIL Research Journal* 1, 80-95.
- Ellis, R. (1997). *SLA Research and Language Teaching*. Oxford: Oxford University Press.
- Ellis, R. (1995). *Understanding Second Language Acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press.
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford University Press.
- Erman, B. (1987). *Pragmatic expressions in English: A study of you know, you see and I mean in face-to-face conversation*. Stockholm, Almqvist & Wiksell.

- Erman, B. (1992). Female and male usage of pragmatic expressions in same-sex and mixed-sex interaction. *Language Variation and Change*, 4(02), 217-234.
- Farghal, M. (1992). Naturalness and the Notion of Cohesion in EFL Writing Classes. *IRAL*, 30(1), 45-50.
- Ferris, D. R., Hedgcock, J., Ferris, D., & Hedgcock, J. S. (2004). *Teaching ESL composition: Purpose, process, and practice*. Psychology Press.
- Field, Y. (1994). Cohesive conjunctions in English writing of Cantonese speaking students from Hong Kong. *Australian Review of Applied Linguistics*, 17(1), 125-139.
- Field, Y. & Yip, L. M. O. (1992). A comparison of internal conjunctive cohesion in the English essay writing of Cantonese speakers and native speakers of English. *RELC Journal*, 23(1), 15-28.
- Flower, L. (1985). *Problem -solving strategies for writing (2nd)*. San Diego: Harcourt Brace Jovanovich.
- Flower, L. S. (1994). *The construction of negotiated meaning: A social cognitive theory of writing*. Carbondale: Southern Illinois University Press.
- Flower, L. S., & Hayes, J. R. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 31-50). Mahwah, NJ: Erlbaum.
- Flowerdew, L. (1998). Corpus linguistic techniques applied to textlinguistics. *System*, 26(4), 541-552.
- Flowerdew, L. (2003). A Combined Corpus and Systemic-Functional Analysis of the Problem-Solution Pattern in a Student and Professional Corpus of Technical Writing. *Tesol Quarterly*, 37(3), 489-511.
- Flowerdew, L. (2005). An integration of corpus-based and genre-based approaches to text analysis in EAP/ESP: countering criticisms against corpus-based methodologies. *English for Specific Purposes*, 24(3), 321-332.
- Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International journal of corpus linguistics*, 14(3), 393-417.
- Fosnot, C. T. (1996). Constructivist: A psychological theory of learning. In C. T. Fosnot (Eds.), *Constructivism: Theory, Perspective and Practice* (pp. 8-13). New York: Teacher College Press.
- Fox, G. (1998). Using corpus data in the classroom. In B. Tomlinson (Ed.), *Materials development in language teaching* (pp. 25-43). Cambridge: Cambridge University Press.
- Fraser, B. (1987). Pragmatic formatives. In V. a. M. Bertuccelli-Papi (Ed.), *The pragmatic perspective* (pp. 179-194). Amsterdam: Benjamins.
- Fraser, B. (1988). Types of English discourse markers. *Acta Linguistica Hungarica*, 38(1-4), 19-33.
- Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics*, 14, 383- 395.
- Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, 31(7), 931-952.

- Fung, L. P. Y. (2003). *The use and teaching of discourse markers in Hong Kong: students' production and teachers' perspectives* (Doctoral dissertation, University of Nottingham).
- Gabrielatos, C. (2005). Corpora and language teaching: Just a fling, or wedding bells? *TESL-EJ*, 8(4), 1-37.
- Gardner, D., & Davies, M. (2007). Pointing Out Frequent Phrasal Verbs: A Corpus-Based Analysis. *Tesol Quarterly*, 41(2), 339-359.
- Gavioli, L. (1997). Exploring texts through the concordancer: Guiding the learner In S. F. A. Wichmann, T. McEnery & G. Knowles (Ed.), *Teaching and language corpora* (pp. 83-89). New York: Addison Wesley Longman.
- Geva, E. (1992). The role of conjunctions in L2 text comprehension. *TESOL*, 26(4), 731-747.
- Ghahool, N., Mariadass, M.E., & Kashef, S. H. (2012). Investigating Malaysian ESL Students' Writing Problems on Conventions, Punctuation, and Language Use at Secondary School Level. *Journal of Studies in Education*, 2(3), 130-143.
- Ghadessy, M., Henry, A., & Roseberry, R. L. (2001). *Small Corpus Studies and ELT: theory and practice*. Amsterdam: John Benjamins Publishing Company.
- Gillespie, M. K. (2001). Research in writing: Implications for adult literacy education. *The Annual Review of Adult Learning and Literacy*, 2, 63-110.
- Giora, R. (1997). Understanding figurative and literal language: The graded salience hypothesis. *Cognitive Linguistics*, 7, 183-206.
- Giora, R. (1999). On the priority of salient meanings: Studies of literal and figurative language. *Journal of Pragmatics*, 31, 919-929.
- Granger, S. (2002). A bird's-eye view of computer learner corpus research. In S. Granger, J. Hung & S. Petch-Tyson (eds.) *Computer Learner Corpora, Second Language Acquisition and Foreign Language Teaching*, 3-33. John Benjamins, Amsterdam.
- Granger, S. (2004). Computer learner corpus research: current state and future prospects. In C. Upton (Ed.), *Applied corpus: A multidimensional perspective* (pp. 123-145). Amsterdam: Rodopi.
- Granger, S., Hung, J., & Petch-Tyson, S. (2002). *Computer learner corpora, second language acquisition and foreign language teaching* (Vol. 6). Amsterdam: John Benjamins Publishing Company.
- Granger, S., & Rayson, P. (1998). Automatic profiling of learner texts. In *Learner English in computer*, ed. S. Granger, 119-131. London: Longman.
- Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17-27. doi: 10.1111/j.1467-971X.1996.tb00089.x.
- Gray, J. H., & Densten, I. L. (1998). Integrating quantitative and qualitative analysis using latent and manifest variables. *Quality & Quantity*, 32, 419-431.

- Green, C. F., Christopher, E. R., & Lam, J. K. M. (2000). The incidence and effects on coherence of marked themes in interlanguage texts: a corpus-based enquiry. *English for Specific Purposes*, 19(2): 99-113.
- Grice, P. H. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Speech acts: Syntax and semantics* (Vol. 3, pp. 41–58). New York: Academic.
- Gries, S. T. (2009). *Quantitative corpus linguistics with R: a practical introduction*. Routledge.
- Gustilo, L., & Magno, C. (2012). *Learners' errors and their evaluation: The case of Filipino ESL writers*. Philippine ESL Writers, 8, 96-112.
- Haiyang, A., & Gong, P. (2006). *A Corpus-based Study of Connectors: Research from the CAS Learner Corpus of English Essays*. Paper presented at the Symposium conducted at the meeting of the first International symposium of computer assisted learning, Beijing, China.
- Halliday, M., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Handoyo, P.W. (2009). *The use of adverbial conjunctions*. Unpublished manuscript.
- Hansen, M. B. M. (1997). Alors and done in spoken Frenen: A reanalysis. *Journal of Pragmatics*, 28, 153-187.
- Hardy, C., Harley, B., & Phillips, N. (2004). Discourse analysis and content analysis: Two solitudes?. *Qualitative Methods*, 2(1), 19-22.
- Hartnett, C. G. . (1986). Static and Dynamic Cohesion: Signals of Thinking in Writing. In B. Couture (Ed.), *unctional Approaches to Writing: Research Perspectives*. London: Frances Pinter.
- Hartnett, C. (1980). Cohesion and mental processes in writing competence. *ERIC Ed 202-028*.
- Hays, P.R. (1992). Discourse markers and L2 acquisition. *Papers in Applied Linguistics-Michigan*, 7, 24-34.
- Heino, P. (2010). *Adverbial Connectors in Advanced EFL Learners' and Native Speakers' Student Writing*. (Bachelor Degree Project in English Linguistics), Stockholms Universitet, Stockholm.
- Henry, A. & Roseberry, R.L. (2001). Using a small corpus to obtain data for teaching a genre. In M. Ghadessy, A. Henry, & A.L. Roseberry,(Eds), *Small Corpus Studies and ELT* (pp. 93-113). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Hinds, J. (1987). Reader versus writer responsibility: A new typology. In U. Connor & R.B. Kaplan (eds.) *Writing across Languages: Analysis of L2 Text*, 141-152. Addison-Wesley, Reading, MA.
- Hinkel, E. (2002). *Second language writers' text: Linguistic and rhetorical features*: Lawrence Erlbaum Associates Mahwah.
- Hofland, K. and Johansson S. (1982) *Word Frequencies in British and American English* (Bergen: Norwegian Computing Centre for the Humanities).

- Hovy, E. H., & Maier, E. (1995). Parsimonious or profligate: how many and which discourse structure relations? [Unpublished manuscript; Retrieved 20 May 2012 from <http://www.isi.edu/natural-language/people/hovy/papers/93discproc.pdf>].
- Howarth, P. (1998). *Phraseology in English academic writing. Some implications for language learning and dictionary making*. Tübingen: Niemeyer.
- Hubbard, P., Jones, H., Thornton, B. & Wheeler, R. (1996). *A Training Course for TEFL*. Oxford: Oxford University Press.
- Huddleston, R., & Pullum, G. K. (2002). *The Cambridge Grammar of English*: Cambridge: Cambridge University Press.
- Hultman, T.G., & Westman. M. (1977). *Gymnasistsvenska*. Lund: Liber Läromedel.
- Hung, M. (2000). A Study on the Use of Conjunctions in Compositions by Taiwanese Senior-high-school Students. [Unpublished manuscript; Retrieved 09 December 2013 from <http://dspace2.lib.nccu.edu.tw/>].
- Hunston, S. (2002). *Corpora in applied linguistics*: Cambridge University Press.
- Hunston, S., & Francis, G. . (2000). *Pattern grammar. A corpus-driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.
- Hyland, K. (Ed.). (2005). *Metadiscourse: Exploring interaction in writing*. Continuum International Publishing Group.
- Hyland, K. (2004). Disciplinary interactions: Metadiscourse in L2 postgraduate writing. *Journal of Second Language Writing*, 13(2), 133-151.
- IBM Corp. Released. (2010). IBM SPSS Statistics for Windows, Version 19.0. Armonk, NY: IBM Corp.
- Intaraprawat, P., & Steffensen, M.S. (1995). The use of metadiscourse in good and poor ESL essays. *Journal of Second Language Writing*, 4(3), 253-272.
- Iten, C. (1998). The meaning of although: a relevance theoretic account. *UCL Working Papers in Linguistics*, 10, 81-108.
- Jacobs, R. M. (2003). Educational Research: Correlational Studies. Retrieved June 05 2012 from www83.homepage.villanova.edu/richard.../correlation.ppt.
- Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F. and Hughey, F. B. (1981). *Testing ESL composition: a practical approach*. Massachusetts: Newbury House Publishers.
- Jafarpur, A. (1991). Cohesiveness as a basis for evaluating compositions. *System*, 19: 459-65.
- Jalilifar, A. (2009). Discourse Markers in Composition Writings: The case of Iranian learners of English as a foreign language. *English Language Teaching*, 1(2), 114-122.
- Jalilifar, A., & Alipour, M. (2007). How explicit instruction makes a difference: Metadiscourse markers and EFL learners' reading comprehension skill. *Journal of College Reading and Learning*, 38(1), 35-52.
- James, C. (1988). *Errors in language learning use: Exploring error analysis*. Harlow, Essex: Addison Wesley Longman Limited.

- Johns, A. M. (1984). Textual cohesion and the Chinese speakers of English. *Language Learning and Communication*, 3(1), 69-73.
- Johnson, P. (1992). Cohesion and coherence in compositions in Malay and English. *RELC journal*, 23(2), 1-17.
- Joseph, B. (2004). On change in Language and change in language. *Language*, 80, 381-3.
- Karasi, M. (1994). *Cohesive features in the expository essays of secondary four (Express) and secondary five (Normal) students in Singapore*. Unpublished M.A. thesis. Singapore: Nanyang Technological University.
- Kaszubski, P. (1998). Enhancing a written textbook: A national perspective. In S. Granger (Ed.), *Learner English on computer* (pp. 172-185). London: Longman.
- Kehler, A. (2004). Discourse coherence. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (pp. 241-265). Malden: Blackwell.
- Kementerian Pendidikan Malaysia (2001). *Sukatan Pelajaran Kurikulum Bersepadu Sekolah Rendah*. Pusat Perkembangan Kurikulum.
- Kennedy, G. D. (1990). Collocations: Where Grammar and Vocabulary Teaching Meet. In S. Anivan (Ed.), *Language Teaching Methodology for the Nineties*. (pp. 215-229). Singapore: SEAMEO Regional Language Centre.: RELC Anthology Series.
- Kennedy, G. D. (1998). *An Introduction to Corpus Linguistics*. London and New York: Longman.
- Khojasteh, L. (2011). *A corpus-based study of modal auxiliary verbs used in the Malaysian English language textbooks*. (Unpublished doctoral thesis). Serdang: Universiti Putra Malaysia.
- Knott, A., & Dale, R. (1994). Using linguistic phenomena to motivate a set of coherence relations. *Discourse processes*, 18(1), 35-62.
- Knowles, G., & Zuraidah, M. D. (2004). Introducing MACLE: The Malaysian Corpus of learner English. 1st National Symposium of corpus linguistics and foreign language education. South China Normal University, Guanzhou, China.
- Knowles, G., & Don, Z. M. (2004). The notion of a lemma: Headwords, roots and lexical sets. *International journal of corpus linguistics*, 9(1), 69-81.
- Krippendorff, K. (2004). Reliability in content analysis. *Human Communication Research*, 30(3), 411-433.
- Lahuerta Martínez, A. C. (2002). The use of Discourse Markers in EFL learners' Writing. *Revista alicantina de estudios ingleses*, 15, 123-132.
- Lai, Y. Y. (2008). *A corpus-based investigation of conjunctive use in the Taiwanese students' writing* (MA Thesis. Ming Chuan University).
- Lake, J. (2004). Using on the contrary': the conceptual problem for EAP students. *ELT Journal*, 58(2), 137-144.
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

- Lee, H. (1999). The Acquisition of Colloquial English Features by Korean Americans. *Journal of Pan-Pacific Association of Applied Linguistics*, 3, 71-87.
- Lee, I. (2002). Teaching coherence to ESL students: A classroom inquiry. *Journal of Second Language Writing*, 11(2), 135–139.
- Leech, G. (1998). Learner English on Computer. In S. Granger (Ed.), *Computational Linguistics* (Vol. 25, pp. 302-303). London: Longman.
- Leech, G. (1991). The state-of-the-art in corpus linguistics. In B. A. K. Aijmer (Ed.), *English corpus linguistics, Studies in honour of Jan Svartvik* (pp. 8-29). London/New York: Longman.
- Leech, G. (2002). The importance of reference corpora. *Hizkuntza-corporak. Oraina eta geroa*, 10(24/25), 1-11.
- Leech, G. (2004). Adding linguistic annotation. In M. Wynne (Ed.), *Developing linguistic corpora: A guide to good practice* (pp. 17-29). Oxford: Oxbow Books. Retrieved from <<http://ahds.ac.uk/linguistic-corpora>> [03/05/12]
- Leki, I. (1992). ESL composing: Strategies and perceptions. In B. Leeds (Eds.), *Writing in a Second Language: Insights from First and Second Language Teaching and Research*. Longman: Addison-Wesley.
- Leńko-Szymańska, A. (2008). Non-native or non-expert? The use of connectors in native and foreign language learners' texts. *Acquisition et interaction en langue étrangère*(27), 91-108.
- Lennon, P. (1991). Error: some problems of definition, identification, and distinction. *Applied linguistics*, 12(2), 180-196.
- Li, F. (2010). Discourse Markers in English Writing. *Journal of International Social Research*, 3, 299-305.
- Lim Ho Peng. (1976). Errors and error analysis in TESL: The Malaysian experience. *RELC Journal*. 9(2), 335-351.
- Linnarud, M. (1979). Lexikal täthet och betydelsesamband. In *Svenska i invandarperspektiv*, ed. K. Hyltenstam, 219-230. Malmö: Gleerups.
- Liu, S. X. (2008). Teachers' discourse markers and classroom context. *Journal of Hunan Institute of Humanities (Science and Technology)*, 4: 148-150.
- Longman dictionary of contemporary English* (4th ed.). (2003). Harlow, England: Longman.
- Longo, B. (1994). Current research in technical communication: The role of meta-discourse in persuasion. *Technical Communication*, 4(2), 348-352.
- Louvain, UCL Université Catholique de. (2012). Learner corpora around the world. 2012, from <https://www.uclouvain.be/en-cecl-lcworld.html>.
- Mair, C. (1991). "Quantitative or qualitative corpus analysis? Infinitival complement clauses in the Survey of English Usage corpus." In: Johansson, Stig & Anna-Brita Stenström (eds.). *English computer corpora. Selected papers and research guide*. Berlin & New York: Mouton de Gruyter. 67-80.
- Mann, W. C., & Thompson, S. A. (1986). Relational propositions in discourse. *Discourse Processes*, 9, 57-90.

- Mariko, A. (2001). *A Corpus-based Study of the Use of Conjunctions. Pac3 at Jalt symposium.* (541-547). Kitakyushu, Japan.
- Maros, M., Tan, K. H., & Salehuddin, K. (2007). Interference in learning English: Grammatical errors in English essay writing among rural Malay secondary school students in Malaysia. *e-BANGI: Jurnal Sains Sosial dan Kemanusiaan*, 2(2), 15.
- Martínez, A.C.L. (2004). Discourse markers in the expository writing of Spanish university students. *Ibérica*, 8, 63-80.
- Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied linguistics*, 33(3), 299-320.
- Mauranen, A. (1993). Contrastive ESP rhetoric: Meta-text in Finnish-English economic texts. *English for Specific Purposes*, 12, 89-103.
- Meisuo, Z. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC Journal*, 31(1), 61-95.
- McCarthy, M. (1991). *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.
- McEnery, T., Wilson, A. (2001). *Corpus Linguistics*. An Introduction. Second Edition. Edinburgh. Edinburgh University Press.
- McEnery, T., Wilson, A., & Barnbrook, G. (2003). Corpus linguistics. *Computational Linguistics*, 24(2).
- Meisuo, Z. (2000). Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC journal*, 31(1), 61-95.
- Menon, S. (2009). Corpus-based analysis of lexical patterns in Malaysian secondary school Science & English for Science & Technology textbooks. Unpublished doctoral dissertation, Serdang: Universiti Putra Malaysia, Malaysia.
- Meyer, C. (2002). *English Corpus Linguistics*. Cambridge: Cambridge University Press.
- Michael, S. (1996). *Text and corpus analysis*. Oxford: Blackwell.
- Mihaljević Djigunović, J., & Vickov, G. (2011). Acquisition of Discourse Markers- Evidence from EFL Writing. *Studia Romanica et Anglica Zagrabienis*, 55, 255-278.
- Milton, J. (2001). Elements of a Written Interlanguage. from The Hong Kong University of Science and Technology: Language Centre.
- Milton, J.C.P., & Tsang, E.S. (Eds.). (1993). *A corpus-based study of logical connectors in EFL students' writing: directions for future research*. In Richard Pemberton and Elza S. C. Tsang (Eds.), *Studies in Lexis*. Hong Kong University of Science and Technology. Language Center.
- Ministry of Education. (2012). Malaysian education blueprint 2013-2025. <http://www.moe.gov.my/userfiles/file/PPP/Preliminary-Blueprint-Eng.pdf>
- Mitchell, R. and Myles, M. (2004). *Second language learning theories*. New York: Hodder Arnold.

- Miri, H. (2008). *Two accounts of discourse markers in English*. Retrieved November 11, 2011, from <http://semanticsarchive.net/Archive/TljODdhM/DMs%20in%20English.pdf>.
- Mizuno, H. A. (1985). Psycholinguistic approach to the article system in English. *JACET Bulletin*, 16, 1-29.
- Mizuno, H. (1986). *Interlanguage analysis of article errors in English among Japanese adult learners in an acquisition-poor environment*. (Doctoral Dissertation, Columbia University). University Microfilms International.
- Modhish, A. S. (2012). Use of Discourse Markers in the Composition Writings of Arab EFL Learners. *English Language Teaching*, 5(5), 56-61.
- Mosenthal, J. H., & Tierney, R. J. . (1984). Commentary: Cohesion: Problems with Talking about Text. *Reading Research Quarterly*, 19(2), 240-244.
- Mukundan, J. (2004). *A composite framework for ESL textbook evaluation*. Unpublished Doctoral Thesis. Serdang: Universiti Putra Malaysia.
- Mukundan, J. & Menon, S. (2006). *The language of form one Mathematics, Science and English language textbooks. A comparative corpus-based study*. Unpublished project work (PRPi). Serdang: University Putra Malaysia.
- Müller, S. (2004). Wellyouknowthattypeofperson’: functions of well in the speech of American and German students. *Journal of Pragmatics*, 36(6), 1157-1182.
- Müller, S. (2005). *Discourse markers in native and non-native English discourse*. Amsterdam/Philadelphia: John Benjamins.
- Murray, J. D. (1995). Logical connectives and local coherence. In R. F. Lorch, Jr. & E. J. O’Brien (Eds.), *Sources of coherence in reading* (pp. 107–125). NJ: Lawrence Erlbaum.
- Musa, C.N., Lie, Y.K., & Azman, H. (2012). Exploring English Language Learning and Teaching in Malaysia, *Journal of Language Studies*. 12(1), 35-51.
- Narita, M., Sato, C., & Sugiura, M. (2004). *Connector usage in the English essay writing of Japanese EFL learners*. Paper presented at the Proceedings of 4th International Conference on Language Resources and Evaluation.
- Nemser, W. (1974). Approximate systems of foreign language learners. In Richards, J. (Ed.). *Error analysis: Perspectives on second language acquisition*. pp: 55-63. Essex: Longman.
- Nesi, H., Gardner, S., Forsyth, R., Hindle, D., Wickens, P., Ebeling, S., Leedham, M., Thompson, P. & Heuboeck, A. (2005). ‘Towards The Compilation of a Corpus of Assessed Student Writing: An Account of Work in Progress’. Paper presented at Corpus Linguistics Conference, University of Birmingham, UK. Published in the Proceedings from the Corpus Linguistics Conference Series 1(1): www.corpus.bham.ac.uk/PCLC.
- Neuendorf, K. (2004). “Content Analysis: A Contrast and Complement to Discourse Analysis,” *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods* 2:1 (Spring 2004).

- Norment, N. (1994). Contrastive analysis of cohesive devices in Chinese and Chinese ESL in narrative and expository written texts. *Chinese Language Teaching Association Journal*, 29(1), 49-81.
- Norrish, J. (1983). *Language learners and their errors*. London: The Macmillan Press Ltd.
- Nyikos, M., & Hashimoto, R. (1997). Constructivist theory applied to collaborative learning in teacher education: In search of ZPD. *The Modern Language Journal*, 81(4), 506-517.
- Olasehinde, M. O. (2002). Error analysis and remedial pedagogy. In Babatunde S. T. and D. S. Adeyanju (eds.). *Language, meaning and society*. Ilorin: Itaytee Press and Publishing Co., Nigeria.
- Ong, J. (2011). Investigating the Use of Cohesive Devices by Chinese EFL Learners. *The Asian EFL Journal Quarterly*, 13(3), 42-65.
- Östman, J. O. (1981). *You know: A discourse functional approach* (Vol. 2): John Benjamins Publishing Company.
- Ostman, J.O. (1995). Pragmatic particles twenty years after. In: B. Warvik et al., Organization in discourse: Proceedings from the Turku Conference, 1995. *Anglicana Turkuensia*, 14, 95-108.
- Owen, C. (1993). Corpus-Based Grammar and the Heineken Effect: Lexico-grammatical Description for Language Learners1. *Applied Linguistics*, 14(2), 167-187.
- Owen, C. (1996). Do concordances require to be consulted. *ELT Journal*, 50(3), 219-224.
- Ozono, S., & Ito, H. (2003). Logical connectives as catalysts for interactive L2 reading. *System*, 31(2), 283-297.
- Ouyang, R. (2012). Basic Concepts of Quantitative Research. Retrieved October 20 2012 from <http://ksuemail.kennesaw.edu/~rouyang/ED-research/methods.htm>.
- Pandian, A. (2006). English language teaching in Malaysia today. *Asia Pacific Journal of Education*, 22(2), 35-52.
- Patanasorn, A.T. (2010). The Use of Linking Adverbials in the Argumentative Essays of Thai EFL Learners. *KKU Res J*, 15 (78), 751-767.
- Penanda wacana. (n.d.). In Wikipedia. Retrieved 18 October 2013, from http://ms.wikipedia.org/wiki/Penanda_wacana.
- Pérez, M. A., & Macià, E. A. (2002). Metadiscourse in Lecture Comprehension: Does it Really Help Foreign Language Learners? *Atlantis*, 7-21.
- Perkins, K. (1983). On the use of composition scoring techniques, objective measures, and objective tests to evaluate ESL writing ability. *TESOL quarterly*, 17(4), 651-671.
- Petch-Tyson, S. (1998). Writer/reader visibility in EFL written discourse. In *Learner English in computer*, ed. S. Granger, 107-118. London: Longman.
- Parrot, M. (2002). *Grammar for English Language Teachers*. Cambridge University Press.

- Philip, A. , Mukundan, J. , Nimehchisalem, V. (2012). Conjunctions in Malaysian Secondary School English Language Textbooks. *International Journal of Applied Linguistics & English Literature*, 1(1), 1-11. doi:10.7575/ijalel.v.1n.1p.1
- Polanyi, L., & Scha, R. (1983). The syntax of discourse. *Text*, 3, 261-270.
- Pouriyeh, A., & Allami, H. (2011). *Discourse Markers in Types of Compositions Written by Iranian EFL Learners*. Paper presented at the ICELT, Malaysia.
- Pravec, N.A. (2002). Survey of learner corpora. *ICAME journal*, 26(81), 114.
- Prommas, P., & Sinwongsuwat, K. (2011). *A comparative study of discourse connectors used in argumentative compositions produced by Thai EFL learners and English-native speakers*. Paper presented at the The 3rd International Conference on Humanities and Social Sciences, Faculty of Liberal Arts, Prince of Songkla University.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.
- Rahimi, M. (2011). Discourse Markers in Argumentative and Expository Writing of Iranian EFL Learners. *World Journal of English Language*, 1(2), p68.
- Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of Pragmatics*, 14(3), 367-381.
- Redeker, G. (1991). Linguistic markers of discourse structure. *Linguistics*, 29(6), 1139-1172.
- Reid, M. J. (1993). *Teaching ESL writing*. New Jersey: Prentice Hall.
- Reppen, R. (2001). Review of MonoConc Pro and WordSmith Tools. *Language Learning and Technology*, 5(3), May, 32-36. Retrieved November 12 2012 from <http://llt.msu.edu/vol5num3/review4/>
- Reppen, R. (2010). *Using corpora in the language classroom*. Cambridge: Cambridge University Press.
- Rezvani Kalajahi, S. A., Neufeld, S. (2014). The Discourse Connector List: A multi-genre cross-cultural corpus model. *International journal of corpus linguistics*. In review
- Rezvani Kalajahi, S. A., Abdullah, A. N., Mukundan, J., & Tannacito, D. J. (2012). Discourse Connectors: An Overview of the History, Definition and Classification of the Term. *World Applied Sciences Journal*, 19(11), 1659-1673.
- Rezvani Kalajahi, S. A. R., & Abdullah, A. N. (2012). Perceptions of Iranian English Language Teachers towards the Use of Discourse Markers in the EFL Classroom. *Theory and Practice in Language Studies*, 2(10), 2002-2010.
- Rezvani Kalajahi, S. A. R., & Abdullah, A. N. (2012). *A Corpus-based Study of Discourse Connectors in Iranian and Malaysian English Language Textbooks*. Paper presented at MICELT2012. Ipoh, Perak , Malaysia

- Rezvani Kalajahi, S. A., Abdullah, A. N., & Baki, R. (2012). Constructing an Organized and Coherent Text: How Discourse Markers Are Viewed by Iranian Post-Graduate Students?. *International Journal of Humanities and Social Science*, 2(9), 196-202.
- Rezvani Kalajahi, S. A., Abdullah, A. N., Arshad, A.S, Mukundan, J. K. (2012, October). A Corpus-based Study of discourse connectors in Iranian and Malaysian English language textbooks. Paper presented at the MICELT 2012, Malaysia.
- Richards, J.C. (ed.) (1974). *Error Analysis. Perspectives on second language acquisition*. London: Longman.
- Richards, J.C., & Schmidt, J. (2002). *Longman dictionary of language teaching and applied linguistics*. Longman, UK
- Richardson, V. (Ed.). (1997). *Constructivist teacher education: Building a world of new understandings*. Uk. Routledge.
- Ringbom, H. (1998). Vocabulary frequencies in advanced learner English: A cross-linguistic approach. In *Learner English on Computer*, ed. S. Granger, 41-52. London & New York Longman.
- Romero-Trillo, J. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics*, 34, 769-784.
- Roslim, N. (2013). A corpus-based study on an English preposition. Unpublished PhD dissertation. Universiti Putra Malaysia.
- Sáez, F. T. (2003). Culture in writing: discourse markers in English and Spanish student writing. *Departamento de Didáctica de la Lengua y la Literatura*, 345-364.
- Samah, B.A. (2013). Basic/intermediate statistics- pearson correlation. Retrieved 20 December, 2013 from <http://psm.upm.edu.my/TrainingMaterials.htm>.
- Sanders, T. J. M., Spooren, W. P. M. S., & Noordman, L. G. M. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, 15(1), 1-35.
- Sanders, T., & Spooren, W. (1999). Communicative intentions and coherence relations. In W. Bublitz, M. Lenz, & E. Ventola (Eds.), *Coherence in spoken and written discourse* (pp. 235–350). Amsterdam: John Benjamins.
- Sasaki, M., & Hirose, K. (1999). Development of an analytic rating scale for Japanese L1 writing. *Language Testing*, 16(4), 457-478.
- Schiffrin, D. (1987). *Discourse Markers. (Studies in International Sociolinguistics 5)*. Cambridge: Cambridge University Press.
- Schmied, J. (1993). Qualitative and quantitative research approaches to English relative constructions. In Souter, C. and Atwell, E. (eds.) *Corpus Based Computational Linguistics*. Amsterdam: Rodopi.
- Schmit, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge University Press.
- Schneider, E. W. (2007). *Postcolonial English: varieties around the world*. Cambridge University Press.

- Schourup, L. C. (1985). *Common discourse particles in English conversation*. New York: Garland.
- Schourup, L. (1999). Discourse markers: Tutorial overview. *Lingua*, 107, 227-265.
- Scollon, R. (1993). Cumulative ambiguity: Conjunctions in Chinese-English intercultural communication. *Perspectives*, 5(1), 55-73.
- Scott, M. (1996,1997, 1999, 2008). *Versions (1.0, 2.0, 3.0, 4.0, 5.0) WordSmith Tools*. Oxford: Oxford University Press.
- Scott, M. (2001). Comparing corpora and identifying key words, collocations and frequency distributions through the WordSmith Tools suite of computer programs. In M. Ghadessy, A. Henry, & R.L. Roseberry, (Eds), *Small Corpus Studies and ELT* (pp. 47-67). Amsterdam/Philadelphia: John Benjamins Publishing Co.
- Sercombe, P. G. (2000). Learner language and the consideration of idiosyncracies by students of English as a second or foreign language in the context of Brunei Darulsalam. In A.M. Noor et al. (eds.) *Strategising teaching and learning in the 21st century*. Proceedings of the International Conference on Teaching and Learning. Faculty of Education: Universiti Kebangsaan Malaysia, Bangi.
- Sharpling, G. (2010). *Discourse markers*. Retrieved May 11, 2012, from Warwick university, Learning English Online at Warwick Web site: http://www2.warwick.ac.uk/fac/soc/al/learning_english/leap/grammar/discourse.
- Shea, M. (2009). A corpus-based study of adverbial connectors in learner text. Retrieved November 28, 2011, from Second language studies department of Michigan State University website slls.msu.edu/soslap/journal/index.php/sls/article/download/4/2.
- Shen, Y. (2006). *An analysis of discourse makers used by non-native English learners: Its implication for teaching English as a foreign language*. Paper presented at the Paper presented at The 19th National Conference of the Japanese Association for Asian Englishes. Retrieved from <http://www.jafae.org/Englishes/19thNationalconference.html>.
- Shin, D., & Nation, I.S.P. (2008). Beyond single words: the most frequent collocations in spoken English. *ELT Journal* 62(4), 339-348.
- Sinclair, J. (1991). *Corpus, concordance and collocation*. Oxford: Oxford University Press.
- Sinclair, J. (1996). Preliminary recommendations on corpus typology. *EAGLES Document TCWG-CTYP/P* (available from <http://www.ilc.pi.cnr.it/EAGLES/corpusyp/corpusyp.html>).
- Sperber, D. & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell. (Second edition 1995).
- Sperber, D., & Wilson, D. (1995). *Relevance: Communication and cognition*. Oxford, England: Blackwell. (Original work published 1986).
- Spivey, N.N. (1995). Written discourse: A constructive perspective. In L. P, Setffe & J. Gale (Eds.) *Consructivisim in education*. (pp.313-330). Mahwah NJ: Lawrence Erlbaum Associates.

- Spooren, W., & Sanders, T. (2008). The acquisition order of coherence relations: On cognitive complexity in discourse. *Journal of Pragmatics*, 40, 2003–2026.
- Stapa, S. H., & Izahar, M. M. (2010). Analysis of errors in subject-verb agreement among Malaysian ESL learners. *3L: Language, Linguistics and Literature, The Southeast Asian Journal of English Language Studies.*, 16(1), 56-73.
- Stark, L. (2001). *Analyzing the interlanguage of ASL natives*. Newark: University of Delaware.
- Steffani, S. A., & Nippold, M. A. (1997). Japanese speakers of American English: Competence with connectives in written language. *Journal of Speech, Language, and Hearing Research*, 40, 1048–1055.
- Steffensen, M., & Cheng, X. (1996). *Meta-discourse and text pragmatics: How students write after learning about meta-discourse*. ERIC Document Reproduction Service, No. ED400709.
- Stern, H. H. (1992). *Issues and options in language teaching*. Oxford: OUP.
- Stubbs, M. (1983). *Discourse analysis*. Chicago: The University of Chicago Press.
- Talif, R., & Edwin, M. (1989). Error analysis of Form Four English compositions. *The English Teacher*, XVIII.
- Tan, H. (2010). Metadiscourse features in the persuasive essays of undergraduate writers. Unpublished Doctoral Dissertation. UPM Serdang, Malaysia.
- Tannen, D. (1982). Oral and literate strategies in spoken and written narratives. *Language*, 58(1), 1–21.
- Tang, E., & Ng, C. (1995). A Study on the Use of Connectives in ESL Student's Writing. *Perspectives Working Papers in Applied Linguistics-Michigan*, 5(1), 38-45.
- Tankó, G. (2004). The use of adverbial connectors in Hungarian university students' argumentative essays. In J. M. Sinclair (Ed.), *How to use corpora in language teaching* (Vol. 12, pp. 157-181). Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Tapper, M. (2005). Connectives in advanced Swedish EFL learners' written English—preliminary results. *The Department of English: Working Papers in English Linguistics*, 5, 116-144.
- Tickoo, M. L. (1998). ESP: State of the Art. SEAMEP Regional Language Centre.
- Ting, F. (2003). An investigation of cohesive errors in the writing of PRC tertiary EFL students. Retrieved December, 15, 2012. from http://etds.lib.ncku.edu.tw/etdservice/view_metadata?etdun=U0026-2405201018572400.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work*. Philadelphia: John Benjamins.
- Tognini-Bonelli, Elena. (2001). *Corpus linguistics at work* (Vol. 6): Amsterdam: John Benjamins Publishing Company.
- Tomlinson, B. (1998). Access-self materials. In Tomlinson (Ed.), *Materials Development in Language Teaching* (pp. 320-337). Cambridge: Cambridge University Press.

- Trillo, J.R. (2002). The pragmatic fossilization of discourse markers in non-native speakers of English. *Journal of Pragmatics*, 34(6), 769-784.
- Umi Kalthom, M. (2007). The use of modals in Malaysian ESL learners' writing. Unpublished Doctoral Dissertation. UPM Serdang, Malaysia.
- Urbaniak, G. C., & Plous, S. (2013). Research Randomizer (Version 4.0) [Computer software]. Retrieved on June 22, 2013, from <http://www.randomizer.org/>
- Vahdatinejad, S. (2008). Students' error analysis and attitude towards teacher feedback using a selected software: a case study. Unpublished Masters thesis. Universiti Kebangsaan Malaysia, Bangi.
- Ventola, E., & Mauranen, A. (1991). Non-native writing and native revising of scientific articles. *Functional and systematic linguistics*, 457-492.
- Virtanen, T. (1998). Direct questions in argumentative student writing. In *Learner English in computer*, ed. S. Granger, 94-106. London: Longman.
- Vygotsky, L. S. (1964). Thought and language. Cambridge, MA: M.I.T. Press
- Warsi, J. (2000). *The acquisition of English contrastive discourse markers by advanced Russian ESL students*. (Ph.D. Thesis), Boston University, USA.
- Waters, J. (2013). Correlational Research Guidelines, Conducting Correlational Research. Retrieved from <http://www.capilanou.ca/psychology/student-resources/research-guidelines/Correlational-Research-Guidelines/>.
- Weber, R. P. (1990). Basic Content Analysis, 2nd ed. Newbury Park, CA.
- Werth, P. (1984). *Focus, coherence and emphasis*. London/Sydney: Croom Helm.
- Widdowson, H.G. (1990). *Aspects of language teaching*. Oxford: Oxford University Press.
- Widodo, H.P. (2009). The use of adverbial conjunctions. Retrieved 20 November 2013. Available online: <http://www.asiatefl.org/journal/board/content.php?no=32>.
- Witte, S., & Faigley, L. (1981) Coherence, cohesion and writing quality. *College Composition and Communication*, 22,189-204. <http://dx.doi.org/10.2307/356693>
- Voon Foo, C. T. (2007). The Effects of the Process-Genre Approach to Writing Instruction on the Expository Essays of ESL Students in a Malaysian Secondary School. Unpublished PhD thesis, Universiti Sains Malaysia. Penang, Malaysia.
- Weireesh, S. (1991). How to analyze interlanguage. *Journal of Psychology & Education*. 9: 113-22.
- William, F. (1996). Language form and communicative practices. In: John J. Gumperz and Stephen C. Levinson (eds.). *Rethinking Linguistic Relativity*. Cambridge: Cambridge University Press, 232-70.
- Willis, J. (1998). Concordances in the classroom without a computer: Assembling and exploiting concordances of common words. In B. Tomlinson (Ed.), *Materials Development in Language Teaching* (pp. 44-67). Cambridge: Cambridge University Press.

- Wilson, D. & Sperber, D. (1993). Linguistic form and relevance. *Lingua*, 90, 1-25.
- Wood, T., Cobb, P., & Yackel, E. (1995). Reflections on learning and teaching mathematics in elementary school. In L. Steffe & J. Gale (Eds.), *Constructivism in education* (pp. 401–422). Hillsdale, NJ: Erlbaum.
- Wong, R. (1992). *Strategies for the construction of meaning: Chinese students in Singapore writing in English and Chinese*. Ph.D. thesis. University of Toronto.
- Wu, S. R. (2006). *Connectives and topic-fronting devices in academic writing: Taiwanese EFL student writers vs. international writers*. Paper presented at the 2006 International Conference and Workshop on TEFL and Applied Linguistics.
- Xiao, R., McEnery, T., & Qian, Y. (2006). Passive constructions in English and Chinese: A corpus-based contrastive study. *Languages in contrast*, 6(1), 109-149.
- Yaochen, D. (2006). Use of Adverbial Connectors in Chinese EFL Learners' Argumentative Writing. Retrieved 20 July 2012. Available online: <http://www.celea.org.cn/teic/70/70-32.pdf>.
- Yip, G. S. (1992). *Total Global Strategy – Managing for Worldwide Competitive Advantage*. Prentice Hall, New Jersey.
- Yoon, H. (2006). A Corpus-based Analysis of Connectors in Korean Students' Essay Writing. *dbpia*, 22(2), 159-178.
- Yum, J. O. (2000). The impact of Confucianism on international relationships and communication pattern in East Asia. In L. A. Samovar & R. E. Porter (Eds.), *Intercultural communication: A reader* (pp. 63–73). Belmont, CA: Wadsworth.
- Zamel, V. (1983). Teaching those missing links in writing. *ELT Journal*, 37(1), 22–29.
- Zarifi, A. (2013). A corpus-based study of phrasal verbs in Malaysian ESL textbooks. Unpublished PhD dissertation. Universiti Putra Malaysia. Selangor.
- Zhuang-lin, H. (1982). Some linguistic Differences in the written English of Chinese and Australian students. *Language learning and communication*, 1(1), 39-49.
- Zuraidah, M.D. (2006). Investigating an emerging Asian variety of English: the Malaysian corpus of learner English. In Zuraidah (ED.) *English in globalised environment: investigating an emerging variety of English* (21-31). KL: University Malaya Press.
- Zwicky, A. M. (1985). Clitics and particles. *Language*, 61, 283-305.