UNIVERSITI PUTRA MALAYSIA

CORPUS-BASED ANALYSIS OF LEXICAL PATTERNS IN MALAYSIAN SECONDARY SCHOOL SCIENCE AND ENGLISH FOR SCIENCE AND TECHNOLOGY TEXTBOOKS

SUJATHA MENON A/P N.S. MENON

FPP 2006 16

CORPUS-BASED ANALYSIS OF LEXICAL PATTERNS IN MALAYSIAN SECONDARY SCHOOL SCIENCE AND ENGLISH FOR SCIENCE AND TECHNOLOGY TEXTBOOKS

By

SUJATHA MENON A/P N. S. MENON

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

June 2009

DEDICATION

This thesis is dedicated to my parents for their continuous support, love and encouragement throughout my post-graduate studies. They have been and I know will always be my pillars of support.

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

CORPUS-BASED ANALYSIS OF LEXICAL PATTERNS IN MALAYSIAN SECONDARY SCHOOL SCIENCE AND ENGLISH FOR SCIENCE AND TECHNOLOGY TEXTBOOKS

By

SUJATHA MENON A/P N.S. MENON

June 2009

Chairman: **Associate Professor Jayakaran Mukundan, PhD**

Faculty: **Faculty of Educational Studies**

The teaching of Science in English in Malaysia has become an issue yet both teachers and material writers are operating in the dark, as they are teaching Science in English, when it is not quite known what scientific English is actually needed in schools. The first step into looking at the type of language used in schools and required of students for the study of Science in English language and of English for Science and Technology (EST) is, to create a corpus of the language used in these subjects.

As there is no existing corpus of the language used in the teaching and learning of Science subjects, nor for the English for Science and Technology subject, this study aims to develop two corpora: one for the Science subjects and one for the English for Science and Technology subject. These corpora will be based on the language used in both the upper secondary Science textbooks and the English for Science and

Technology textbooks from two textbook zones in Malaysia. These corpora will create a reference point for scientific and English for Science and Technology lexical patterns used in existing prescribed textbooks and would also aid materials writers and curriculum builders in the process of re-designing teaching materials in future.

This thesis analyses the similarities and differences between the lexico-grammatical patterns of scientific English and 'everyday' general English language and between scientific English and the language used in the English for Science and Technology textbooks. As this study intends to look at word relationships which involve collocations and multi-word clusters, the concordance software, WordSmith Tools version 4.0 is used for the purpose of text analysis in this study. Through analyses of wordlists, concordance lines and keywords, the lexico-grammatical patterns of both the corpora are identified.

The findings indicated that the English for Science and Technology textbooks from both zones were quite dissimilar, not only in vocabulary loading and distribution but also in content focus. This work has also uncovered the inadequacy of the English for Science and Technology textbook to cope with the language needs of the upper secondary Science students as even though there were many patterns shared between the English for Science and Technology and Science textbooks, the language used in the English for Science and Technology textbooks lacked variety.

The Science corpus had a greater variety of phrases, prepositions and phrasal verbs than those formed in the English for Science and Technology corpus. The study also found that even though some of the collocations in the Science corpus had predictable flexible combinations, many other collocations could not be easily predicted as they were arbitrarily blocked by usage, thus creating genre specific collocations. The study also found that many of the words when in collocation or compound form often acquire extended meanings which are content specific. General English language grammar rules could not be used to infer the meanings of compound nouns as many of the elements in a compound do not retain their literal meaning and in fact acquire extended meanings.

The creation of the Science and English for Science and Technology corpora together with the work done in this thesis on the lexis and phraseology of Scientific English used in prescribed textbooks, should be used as a platform for materials writers to design and develop more relevant and accurate English for Science and Technology textbooks and materials.

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

KAJIAN DATA KORPUS BERASASKAN CORAK LEKSIS DALAM BUKU TEKS SAINS DAN INGGERIS UNTUK SAINS DAN TEKNOLOGI

Oleh

SUJATHA MENON A/P N.S. MENON

June 2009

Pengerusi: **Profesor Madya Jayakaran Mukundan, PhD**

Fakulti: **Pengajian Pendidikan**

Pengajaran dan pembelajaran Sains dalam bahasa Inggeris di Malaysia menghadapi pelbagai masalah terutamanya berkenaan kesukaran para pendidik dan penulis-penulis buku teks untuk memberi kefahaman Sains dalam bahasa Inggeris. Para pengajar masih kabur tentang jenis bahasa Inggeris saintifik yang diperlukan oleh pelajar-pelajar dan yang patut diberi focus di sekolah. Langkah pertama untuk mengenali jenis bahasa saintifik yang digunakan di sekolah dan diperlukan oleh pelajar-pelajar untuk memahami pembelajaran Sains dalam bahasa Inggeris dan juga pembelajaran subjek Inggeris untuk Sains dan Teknologi (EST), adalah dengan membentuk satu sistem pangkalan data korpus berasakan bahasa yang digunakan dalam subjek-subjek Sains dan EST.

Oleh kerana tidak ada pangkalan data korpus yang sedia ada berasaskan bahasa Inggeris yang digunakan dalam pengajaran dan pembelajaran Sains dan EST di sekolah, kajian ini bertujuan untuk membentuk pangkalan-pangkalan data korpus Sains dan Inggeris untuk Sains dan Teknologi (EST) berasaskan bahasa yang digunakan dalam buku-buku teks Sains dan EST yang dicadangkan untuk pelajar-pelajar tingkatan 4 dan 5 di dua zon buku teks di Malaysia. Pangkalan-pangkalan korpus tersebut akan menjadi pangkalan rujukan leksis dan tatabahasa yang terkandung dalam bahasa saintifik dan EST yang digunakan di sekolah. Data korpus ini juga akan membantu penulis bahan pengajaran-pembelajaran dan pembentuk kurikulum dalam proses mengkaji semula bahan-bahan pengajaran-pembelajaran Sains dan EST.

Tesis ini menganalisa perbezaan dan kesamaan corak-corak 'lexico-grammatical' di antara bahasa Inggeris saintifik dan bahasa Inggeris 'harian' dan di antara bahasa Inggeris saintifik dan bahasa Inggeris yang digunakan dalam buku teks EST. Kajian ini telah menggunakan perisian konkordans, WordSmith Tools versi 4.0, untuk menganalisa teks dan mencerap hubungan leksis yang melibatkan kolakasi dan 'multi-word clusters'. Corak 'lexico-grammatical' kedua-dua korpus dikenalpasti melalui penganalisaan senarai kata, baris-baris konkordans dan kata kunci.

Hasil kajian menunjukkan bahawa kedua-dua buku teks EST mempunyai kemampatan dan pengagihan kosakata yang jauh berbeza. Kajian ini juga menunjukkan bahawa bahasa yang sedia ada dalam buku teks EST tidak memadai untuk menolong pelajar-pelajar dalam proses pembelajaran Sains. Walaupun terdapat beberapa corak

penggunaan bahasa yang sama di antara korpus Sains dan korpus EST, korpus Sains mempunyai pelbagai jenis corak leksis berbanding dengan korpus EST.

Korpus Sains mempunyai lebih jenis frasa, kata preposisi dan 'phrasal verbs' berbanding dengan korpus EST. Kajian ini juga mendedahkan beberapa kombinasi kolakasi yang dapat diramalkan dengan senang (flexible) berbanding dengan beberapa kombinasi kolakasi yang tidak boleh diramalkan (fixed) disebabkan oleh penggunaan yang spesifik. Kolakasi-kolakasi ini yang tidak fleksibel merupakan kolakasi khusus bagi genre Sains. Kajian ini juga mendapati bahawa perkataan-perkataan dalam bentuk kata majmuk (compound noun) biasanya akan memperolehi makna yang lebih luas (extended meaning) atau makna yang lain. Didapati bahawa peraturan tatabahasa bahasa Inggeris tidak dapat digunakan untuk meramalkan makna kata majmuk tersebut oleh kerana beberapa elemen dalam kombinasi-kombinasi kata majmuk tersebut tidak mengekalkan maknanya yang asal.

Pembangunan korpus Sains dan korpus EST bersama-sama dengan hasil kajian ini ke atas leksis dan frasa bahasa Saintifik harus digunakan oleh para penulis bahan pengajaran-pembelajaran untuk mengkaji semula dan membina bahan pengajaran-pembelajaran EST yang lebih relevan dan lebih berkesan.

ACKNOWLEDGEMENTS

First and foremost I wish to thank the Chairman of my supervisory committee, Associate Professor Dr. Jayakaran Mukundan. This work would not have been possible without the support and encouragement of Dr. Jayakaran. He has been abundantly helpful and has assisted me in numerous ways but more than the help given to me to complete this thesis, he has given me the encouragement and opportunity to challenge myself and realize my potential.

I would also like to thank my thesis committee members, Dr. Ghazali bin Hj. Mustapha and Dr. Ismi Arif bin Ismail who have guided me through all these years and for their substantial contributions. Thank you to Wan Hafizi bin Wan Umar, who painstakingly scanned all my texts and manually corrected them. I could not have asked for a better research assistant. Many thanks to Anealka Aziz Hussin who tutored me on the intricacies of the WordSmith concordance software. I owe a debt of gratitude to many people who have helped me along the way to completion of this project. Foremost, to my employers Universiti Teknologi MARA for giving me time off to complete my thesis and for their continuous support. I cannot end without thanking my family, on whose constant encouragement and love I have relied on. I would like to thank my husband, sons and my parents for their unconditional support all these years; having given up many things for me in the process of obtaining my PhD. They have cherished with me every great moment and supported me whenever I needed it.

I CERTIFY THAT AN Examination Committee has met on **19 June 2009** to conduct the final examination of **Sujatha Menon A/P N. S. Menon** on her **Doctor of Philosophy** thesis entitled **Corpus-Based Analysis of Lexical Patterns in Malaysian Secondary School Science and English for Science and Technology(EST) Textbooks** in accordance with Universiti Pertanian Malaysia (Higher Degree) Act 1980 and Universiti Pertanian Malaysia (Higher Degree) Regulations 1981. The Committee recommends that the student be awarded the Doctor of Philosophy.

Members of the Examination Committee were as follows:

Arshad Abd. Samad, PhD
Associate Professor
Faculty of Educational Studies
University Putra Malaysia
(Chairman)

Malachi Edwin Vethamani, PhD
Associate Professor
Faculty of Educational Studies
University Putra Malaysia
(Internal Examiner)

Roselan Baki, PhD
Senior Lecturer
Faculty of Educational Studies
Universiti Putra Malaysia
(Internal Examiner)

Alan Maley
Professor
Department
University
United Kingdom
(External Examiner)

_____

BUJANG KIM HUAT, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

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 were as follows:

Jayakaran A/L A.P. Mukundan, PhD
Associate Professor
Faculty of Educational Studies
Universiti Putra Malaysia
(Chairman)

Ghazali Bin Hj. Mustapha, PhD
Faculty of Educational Studies
Universiti Putra Malaysia
(Member)

Ismi Arif Bin Ismail, PhD
Faculty of Educational Studies
Universiti Putra Malaysia
(Member)

HASANAH MOHD. GHAZALI, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 17 July 2009

DECLARATION

I declare that the thesis is my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously, and is not concurrently, submitted for any other degree at Universiti Putra Malaysia or at any other institution.

_____

SUJATHA MENON A/P N. S. MENON

Date: 22 June 2009

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES