Lexical Similarities and Differences in the Mathematics, Science and English Language Textbooks

Jayakaran Mukundan

Department of Language and Humanities Education, Faculty of Educational Studies, University of Putra Malaysia, Selangor, Malaysia
e-mail: jaya@educ.upm.edu.my

Sujatha Menon

English Language Department, Academy of Language Studies, UITM Melaka, Melaka, Malaysia
e-mail: sjathark@yahoo.com

Abstract: The teaching of Science and Math in English in Malaysia is an area of great concern to educators and students alike. This study looks, in particular, at the common word classes among keywords identified in the Science, Math and English language Form One textbooks used in Malaysia and the differences in language use identified in the Science and Math textbooks.

Key words: corpus, science, math, word class, keyword

With the sudden re-introduction in Malaysia of the teaching of Math and Science in English in January 2003, in standard One and forms One and lower Six, concerns regarding the effect of English language proficiency on the performance of students in these subjects were raised by various sectors of the community.

It cannot be denied that there is a need to study the language used and needed in the teaching of Math and Science in schools. In fact, in a study conducted on 88 Math and Science teachers in Perak regarding their attitudes and perceived readiness to teach Math and Science in English, it was found that most of these teachers were not clear about the linguistic features of their content subjects and had difficulty to communicate the linguistic elements of this form of discourse to their students (Pandian & Ramiah, 2004). According to Holme (2006), children who have been identified by their teachers as being intellectually able to learn Math and Science but are not proficient in English language, would take a longer

time to reach the level they would have achieved if the medium of instruction had been in a language they are familiar with, such as Bahasa Malaysia (national language). Therefore, there is a need for both the students and the teachers to recognize and understand Scientific and Mathematical concepts and also understand and have knowledge of the different types of lexical and grammatical constructions used in Science and Math.

Corpus Research

Insights from corpus research have revolutionized the way language is viewed, especially words and their relationship with each other in context (Schmitt, 2000). Corpus research allows researchers and learners to gain insights into the language, particularly the interconnection of lexical and grammatical patterns, collocations, colligations, the frequency of words and the use and functional behavior of these words (Tognini-Bonelli, 2001, Schmitt, 2000, & Sinclair, 1991).

One of the ways to improve the understanding of the Science and Math discourse and texts in the classroom and to learn about the specific sentence structures, lexis and grammar, is to get the students to engage with the data or texts. This would require not only the creation of a specific database or corpus, but also learning how to control it and incorporating it into the teaching and learning process (Tognini-Bonelli, 2001).

There is reasonable consensus that a corpus will not just provide insights into the contents but also that the results of the analyses will be claimed to be typical of the language from which the corpus was selected. Through corpora, teachers and learners would be able to check prescribed rules and generalizations against linguistic data and to make their own interpretations and generalizations of these patterns (Tognini- Bonelli, 2001). Corpora allow researchers, teachers and learners to use great amounts of real data in their study of language, instead of having to rely on intuition and made-up examples.

Language of Science

The languages of Science and Math are different from the languages that students use socially at home and with their peers, and in other subject areas at school (Laplante, 1997). Everyday words may mean something else in Math and Science, for example the words 'average' and 'divide'

may be everyday words but they acquire a more precise meaning in Math (Carlson, 2000, Khisty, 1995, & Thompson & Rubenstein, 2000). Trimble (1985) in his description of English for Science and Technology state that English for Science and Technology range from English for Occupational Purposes (EOP) to English for Academic Purposes (EAP) with a great deal of overlapping between the two. He states that there are two areas of problems concerning language in English for Science and Technology discourse for non-native learners. The first being the rhetorical-grammatical relationships and the second being the lexical elements of sub-technical vocabulary and the noun compounds.

Trimble (1985) restricts his discussion of lexis to three lexical areas, that is, the technical vocabulary, sub-technical vocabulary and the noun compounds. He believes that non-native learners do not usually have a problem with technical vocabulary as it is taught explicitly by content matter teachers. Sub-technical vocabulary is also considered not a very problematic one as they can be understood quickly with the use of specialist dictionaries. However, the most problematic area for students is compounds (Trimble, 1985).

Sub-technical vocabulary according to Trimble (1985) mean both context-independent words that occur with a high frequency across different disciplines of science, retaining the same meaning across these scientific disciplines and also words that have one or more 'general' English meanings and which in technical contexts take on extended meanings. The vocabulary of Science has been discussed and categorized by many other linguists. The most notable and one of the earliest categorization of the lexis of Science was by Cowan (1974), who was widely attributed with the introduction of the concept of sub-technical vocabulary. It was Cowan's definition of sub-technical vocabulary that Trimble extended in 1985.

Cowan (1974, p. 391) describes four categories of vocabulary ranging from highly technical words to sub-technical vocabulary which he defined as 'context independent' words which occur with high frequency across disciplines to semi-technical and finally, non-technical words, such as hospital, medicine, disease, which he grouped together making no clear distinction between them.

Nation (2001, p. 198) adds to these categories by declaring that there are degrees of 'technicalness' depending on how restricted a word is to a particular area. These degrees were categorized into four groups:

The first category is the most technical with the words appearing rarely outside its particular field such as 'morpheme' in the field of applied linguistics and 'pixel' in the computing field. This category can be described as being similar to Cowan's (1974) highly technical words.

The second category consists of words that are used both inside and outside this particular field but with different meanings, such as the sub-technical vocabulary of Trimble's (1985) and Cowan's (1974).

The third category consists of words that are used both inside and outside this particular field but the majority of its uses with a specific meaning are related to this field. The specialized meaning it has in this field is readily understood outside the field, such as the word 'accused' in the field of Law and 'memory' in the computing field (Nation, 2001, p. 199). This category is similar to that described by Cowan (1974) as semi-technical.

The final category consists of words that are more common in this field than elsewhere, similar to Cowan's (1974) non-technical words. There is little specialization of meaning, example 'judge' in the field of Law and 'print' in the field of computing.

The first step into looking at the type of language used and required of students for the study of Science and Math in English is to create a corpus of the language used in these subjects. A corpus would provide a convenient source from which to obtain evidence of the behavior of many different facets of language: lexical, grammatical and pragmatic (Schmitt, 2000). Once a corpus has been compiled then the language in it can be analyzed.

This work, therefore, aims to analyze the language used in a prescribed Form 1 Science textbook and Math textbook in Malaysian schools, as the first step to identifying the type of language students are required to understand and grasp in the process of learning Science and Math in Malaysian schools. This language will then be compared to the general English language used in the prescribed Form 1 English language textbook.

The distribution of prescribed textbooks in Malaysia is decided by the Textbook Bureau of the Malaysian Education Ministry. To standardize the type of textbooks used in schools and to allow more opportunities for different publishers to be involved in the production of textbooks for schools, the Bureau has divided all the schools in Malaysia into five textbook zones. These textbook zones are divided according to the states in

Malaysia. If schools from different states fall under one zone, then these schools would use the same textbooks. The five textbook zones are the northern, central, eastern, southern and east Malaysia zones. This study focuses on the prescribed Form 1 textbooks from the southern zone in Malaysia.

In order to find out what constitutes Scientific and Mathematical English in the textbooks used, and to see how it differs from the language used in the English language textbook, two research questions were formulated. The research questions posed were:

1   What are the most common word class among the keywords identified in the Science, Math and English language textbooks?
2   What are the differences in language use identified in the Science and Math textbooks?

METHODOLOGY

Research Design

The methodological base of a corpus research is diverse as it not only covers the fields of corpus linguistics but also involves looking into grammatical and lexical relationships and discourse analysis. The study is concerned with data of language used in textbooks and uses corpora to investigate the language of Science and Math.

The Use of Textbooks

The textbook is and has always been an important aspect of teaching in Malaysian schools. Students, while enjoying the benefits of the textbook as a teaching device that works alongside a teacher, would appreciate the role of the book as reference, for this enables the learner to revise and work on consolidation both inside and outside the classroom (Mukundan, 2004).

The criticisms leveled against textbooks are plenty and the fact remains that the textbook in most cases is indispensable and while teachers complain about them, they cannot do without them (Ansary & Babaii, 2003). As textbooks are used by students daily in their schools, it is important to analyze the language contained in the textbooks as this would be the language that would challenge the students the most, in the course of learning the English language used in Science and Math.

Population and Sampling in the Science, Math and English language Corpus

For the purpose of this study, the population for the Science, Math and English language corpus is defined as the prescribed Science, Math and English language textbooks used by Form 1 students in the southern zone of Malaysia.

The books used were:
1   Abdullah, K. I., Wee, D., & Teo, H, B. (2002). *English form 1*. Pustaka Alhas (M) Sdn. Bhd.: Kuala Lumpur.
2   Mat Saat, A. S., Salehin, M. S., & Wan Zaid, W. Z. (2002). *Science form 1*. Pustaka Sistem pelajaran Sdn. Bhd. Malaysia: Kuala Lumpur
3   Chua, M. K., Teh, E. K., & Ooi, S. H. (2002). *Mathematics form 1* (Vols. 1-2). Penerbitan Pelangi Sdn. Bhd. : Kuala Lumpur

Instrumentation: WordSmith Tools 4

WordSmith Tools was designed by Mike Scott (1996, 1997, & 1999) for students and researchers to be able to access and analyze corpora at their convenience on their PCs (Scott, 2001). The reliability of WordSmith Tools has been verified by numerous studies on various corpora which have used these tools to analyze texts (Flowerdew, 2003, Nelson, 2000, Mukundan, 2004, Scott, 2001, Henry & Roseberry, 2001, & Bondi, 2001). Other than these studies using WordSmith Tools, the reliability and wide capability of the software was verified by Mukundan, in his unpublished thesis (2004), in his exploration for suitable software to analyze prescribed textbooks.

The researchers decided to use the latest WordSmith Tools, version 4 for the purpose of text analysis in this study as this latest version has a larger capacity for concordancing and creation of word lists with more details, an improved word list cluster handling, enhanced tag handling and enhanced statistical functions for collocation, to name a few (http://www.lexically.net/wordsmith/version4.htm).

Data Collection

For this study, all the textbooks were first scanned, page by page, and then converted into text files. As there was distortion to text scanned, manual correction was carried out by entering (typed) words or phrases left out or undetected by the scanner. These text files were then analyzed using the WordSmith 4.0 software program.

FINDINGS AND DISCUSSION

Analysis of Key Words

A more accurate picture of any language can be gained by analysis of words that occur significantly more often in a particular linguistic area, in comparison to general language usage, rather than by looking at words that have high occurrence in terms of overall frequency. These words have been termed key words (Scott, 1997 &1999). Key words were arrived at in this study by using the key word function of WordSmith 4. A word will get into the list if it is unusually frequent (or unusually infrequent) in comparison to a larger word list.

Key Words by Chapters

This section of the discussion analyses the three largest word class categories found among the key words by chapters for each subject. Tables 1, 2 and 3 are a summary of the key words by chapters.

Table 1. Percentage of Word Class Distribution among Key Words: Science Text

Chapter	Nouns	Verbs	Adjectives
1	79	7	14
2	86	0	14
3	65	18	5
4	89	5	0
5	73	3	19
6	59	9	27
7	34	20	27
Total Average	**69**	**9**	**15**

Table 2. **Percentage of Word Class Distribution among Key Words: Math Text**

Chapter	Nouns	Verbs	Adjectives
1	17	0	33
2	41	6	29
3	43	14	29
4	75	25	0
5	64	18	5
6	71	0	14
7	43	21	29
8	80	7	0
9	35	4	48
10	70	20	5
11	81	10	0
12	84	5	11
Total Average	**57**	**11**	**17**

Table 3. **Percentage of Word Class Distribution among Key Words: English Language Text**

Chapter	Nouns	Verbs	Adjectives
1	40	0	10
2	50	10	10
3	33	33	33
4	56	22	22
5	100	0	0
6	67	11	11
7	50	50	0
8	75	25	0
9	60	0	20
10	100	0	0
11	50	17	33
12	25	0	25
13	50	17	17
14	100	0	0
15	43	0	43
16	100	0	0
17	80	20	0
18	50	25	25
Total Average	**63**	**12**	**14**

Overall, the three largest word class categories among the key words by chapters for all three subjects were similar, that is, nouns. This was followed by adjectives and then verbs. The Science text had the highest average percentage of nouns as keywords (69%), followed by the English language text (63%) and Math text (57%). The average percentage of adjectives as key words in all the subjects was quite similar, ranging from 14% (English language) to 17% (Math). There was not much difference in the percentage of verbs among the three subjects. The English language text had the highest average percentage of verbs as key words (12%), followed by Math (11%) and then Science (9%). What seems obvious is that in all the three texts, the percentage of key words which were nouns far outnumbered the percentage of key words from other word classes.

As the key words in this analysis were derived by using the entire text as a reference corpus, a complete picture of the lexis specific to Math and Science and which is different from general English language cannot be obtained. For this purpose, a key word analysis of the Science and Math word lists against the English language word list as reference corpus was carried out.

Key Words in Math and Science

This section of the discussion delves into the specialized vocabulary or lexis of Math and Science. The three largest word class categories among the key words found in the Math and Science texts are analyzed and presented as percentages of word class distribution and according to whether they were positive or negative key words, as seen in the table below.

Table 4. Percentage of Word Class Distribution among Key Words: Science and Math against Reference Corpus

SCIENCE	Positive	NOUNS	Negative	Positive	VERBS	Negative	Positive	ADJECTIVE	Negative
	55%	63%	8%	7%	13%	6%	9%	11%	2%
MATH	Positive	NOUNS	Negative	Positive	VERBS	Negative	Positive	ADJECTIVE	Negative
	37%	44%	7%	12%	24%	12%	8%	9%	1%

For the purpose of discussion, only the positive key words were looked into as these were the specialized words which occurred frequently in the Math and Science texts in comparison to the general English language text. Thus, these were words which could be considered specialized language used in Math and Science texts.

Major Word Classes

Similar to the findings in the previous keyword analysis, the main word class category in the Science and Math texts was nouns, 63% (Science) and 44% (Math). Positively keyed nouns (55%-Science, 37%-Math) occurred more than negatively keyed nouns. The next word class categories were verbs, 13% (Science) and 24% (Math), followed by adjectives (11%-Science, 9%-Math). Though there does not seem to be a great difference in the occurrence of verbs and adjectives in the Science text, there seems to be a large difference between the use of verbs and adjectives in the Math text.

There seems to be an equal or near equal percentage of positive and negative key verbs in both of the texts. Looking at positive key verbs, there is only 7% positively keyed verbs in the Science text in comparison to 12% positive key verbs in the Math text. The percentage of positively keyed adjectives (9%-Science, 8%-Math) nearly equals that of positively keyed verbs.

Nouns

The positive key nouns were further categorized into technical and sub-technical vocabulary as categorized by Cowan (1974) and Nation (2001). Tables 5 and 6 below list the nouns from the Math and Science lists according to the four categories: highly technical, sub-technical, semi-technical and non-technical.

Table 5. Categorization of Positively Keyed Nouns in the Science List

SCIENCE TEXT			
HIGHLY TECHNICAL	**SUB-TECHNICAL**	**SEMI-TECHNICAL**	**NON-TECHNICAL**
Bunsen burner	cell	apparatus	activity
cytoplasm	combustion	beaker	air
hydroxide	composition	bicarbonate	animal
litmus	fossil	carbon dioxide	ball
photosynthesis	nucleus	chemical	body
sodium		conical	candle
		copper	colour
		cylinder	conclusion

| | | SCIENCE TEXT | |
HIGHLY TECHNICAL	SUB-TECHNICAL	SEMI-TECHNICAL	NON-TECHNICAL
		density	earth
		experiment	elements
		filter	energy
		fuels	example
		gas	figure
		gauze	flame
		kerosene	flask
		laboratory	function
		lever	glass
		mass	heat
		matter	hot
		mercury	human
		metal	ice
		microorganisms	indicator
		microscope	input
		minerals	investigation
		nitrogen	jar
		organisms	length
		oxygen	level
		particles	life
		pendulum	light
		radian	lime
		respiration	liquid
		rod	material
		rubber	measurement
		science	method
		space	metre
		substance	mixture
		sulphur	movement
		thermometer	object
		vapour	observation
		variable	option
		volume	outcomes
			paper

SCIENCE TEXT			
HIGHLY TECHNICAL	SUB-TECHNICAL	SEMI-TECHNICAL	NON-TECHNICAL
			photo
			physical
			plant
			procedure
			process
			properties
			quantity
			red
			report
			resources
			salt
			soil
			solid
			solution
			sources
			splinter
			state
			structure
			sun
			surface
			system
			temperature
			test
			things
			tube
			unit
			vocabulary
			water
			weight

Table 6. Categorization of Positively Keyed Nouns in the Math List

MATH TEXT			
HIGHLY TECHNICAL	**SUB-TECHNICAL**	**SEMI-TECHNICAL**	**NON-TECHNICAL**
cuboid	angle	addition	activity
decimal	factor	cube	aim
denominator	fraction	digit	application
integers	prime	division	area
numerator		estimate	cost
parallelogram		mass	difference
perimeter		multiplication	example
protractor		number	faces
simulator		percentage	figure
trapezium		rectangle	finding
vertices		square	height
		subtraction	hour
		sum	length
		tonnes	line
		triangle	minutes
		volume	months
		zero	page
			plan
			problem
			procedure
			process
			question
			sides
			solution
			system
			temperature
			term
			terms
			unit
			unknown
			value

As can be seen from tables 5 and 6 above, the majority of the positively keyed nouns consisted of non-technical words or common general English language terms (Science-58%, Math-52%), such as 'water', 'air', 'solution', 'value', 'figure'. The next largest sub-category was the semi-technical words (Science-32%, Math-25%) or words which are usually used both inside and outside the Science and Math fields but the majority of its uses with a specific meaning are related to these two fields (Nation, 2001). This specialized meaning in its particular field is readily understood outside the field, such as words like 'triangle', 'mass', 'percentage', 'oxygen' and 'density'.

Only a small number of nouns were highly technical and sub-technical words. This finding indicates that the Math and Science texts contained more general English language and common nouns, which is appropriate for this level of students. However, semi-technical nouns may need to be given more attention to especially for L2 learners of low proficiency in English language.

Verbs

The positive verb key words in both texts were examined and categorized as lexicalized and delexicalised verbs as categorized by Sinclair (1991). Lexicalised verbs are verbs which carry specific meaning in relation to the field and delexicalised verbs are verbs which carry general meaning that is equally common to both the Science and Math and non-Science and non-Math language use. In short, lexicalized verbs are more technical and specialized, while delexicalised verbs are non-technical common English language verbs. Table 7 shows the categories according to text type.

The categories show that there is an equal number of lexicalized and delexicalised verbs in both the Science and Math texts. Even though, the verbs may be familiar to the students, the lexicalized verbs carry specific meanings when used in their respective fields and thus, need to be given more attention in class. A point to note is that there could be more delexicalised verbs used in these texts, but as they are equally common in general English language use, they are not found as key words because their usage is similar in both the English language text and the Science and Math texts.

Table 7. Delexicalised and Lexicalised Verbs by Text

SCIENCE		MATH	
Lexicalised	**Delexicalised**	**Lexicalised**	**Delexicalised**
balance	contains	angled	copy
change	happen	calculate	determine
effect	is	carry	draw
float	learn	convert	find
flow	need	devise	follow
measure	show	divided	increasing
observe	understand	evaluate	involving
record		measure	is
		multiply	mixed
		perform	represent
		shaded	try
		simplify	understand
		solve	

Adjectives

The adjectives were analyzed to find out if they were adjectives derived from changing nouns, changing verbs, adding suffixes to nouns and verbs, adding prefixes to verbs, or simple modifiers such as 'hot', 'red' (Quirk & Greenbaum, 1973, & Master, 1996). Table 8 lists the adjectives types.

The adjectives used both in the Science and Math texts mostly consist of modifiers such as 'red', 'hot', 'multiple', 'obtuse'. The Science text had more adjectives which were changed from nouns using suffixes such as 'science+ic','wood+en' and adjectives which were changed from verbs using suffixes such as 'renew+able', 'avail+able'.

The difference in adjective use between the Math and Science texts is that the adjectives used in the Science text is more complex ranging from simple modifiers to adjectives with different word class base forms. The Math text, on the other hand, is less complex involving basic modifiers related to this specific field. However, as the modifiers are specifically related to the Math field such as '*obtuse* angle', '*acute* angle', '*parallel*

line', these words can also be considered sub-technical and semi-technical words (Nation, 2001, & Cowan, 1974).

In summary, the analysis of key words has clearly identified that both Math and Science texts have more nouns as key words, with adjectives and verbs as the next two word classes more commonly used, and that the majority of the words used are semi-technical and non-technical words which are appropriate for this level of students. Table 9 shows the specific similarities and differences in language use between the Math, Science and English language texts.

Table 8. Adjective Types by Text

SCIENCE		MATH	
Word	**Adj Type**	**Word**	**Adj Type**
renewable	suffix to verbs	total	modifiers
available	suffix to verbs	straight	modifiers
soluble	suffix to nouns	multiple	modifiers
scientific	suffix to nouns	adjacent	modifiers
wooden	suffix to nouns	common	modifiers
natural	suffix to nouns	parallel	modifiers
basic	suffix to nouns	parallel	modifiers
potential	modifiers	acute	modifiers
kinetic	modifiers	lowest	modifiers
various	modifiers	obtuse	modifiers
convestion	modifiers	unknown	prefix+verbs
blue	modifiers	improper	prefix+modifier
red	modifiers	algebraic	suffix to nouns
dull	modifiers		
moist	modifiers		
hot	modifiers		
measuring	changing verb forms		
living	changing verb forms		

Table 9. Similarities and Differences in Language Use in Math, Science and English Language Texts

	ENGLISH	MATH	SCIENCE
Major key word , word class category, by chapters	Nouns	Nouns	Nouns
Major key word , word class category, against English language list as reference corpus		Nouns	Nouns
Major positively keyed nouns		Semi-Technical and Non-Technical	Semi-Technical and Non-Technical
Major positively keyed verbs		Equal number of lexicalized and delexicalised verbs	Equal number of lexicalized and delexicalised verbs
Major positively keyed adjectives		Modifiers	Equal number of modifiers and adjectives from noun and verb changes

CONCLUSION AND PEDAGOGICAL IMPLICATIONS

There is a need to determine the type of language the students are required to have, whether it is general English or technical language, to be able to cope with the everyday learning of Math and Science in English. This small corpus study on the language used in the Form 1 Math and Science textbooks used in one zone in Malaysia has provided insights into

the type of language students are required to know and need to be able to read and understand their textbooks. As textbooks are an integral part of teaching and learning in schools in Malaysia, it is important for these textbooks to be analyzed and assessed.

This study which focuses on key words has shown that there is a greater emphasis on the learning of nouns in the three texts, followed by verbs and adjectives. The nouns used in the Math and Science texts were mostly semi-technical and non-technical words, which are simple and considered to be appropriate for this level of students. This implies that there would not be much difficulty in learning new terms or words in the Math and Science texts as most of the words should be quite familiar to them.

However, semi-technical words, such as 'density', 'vapour', 'digit' and 'respiration', could pose a problem to L2 learners who have a low proficiency in English language; therefore, these words would have to be given more attention in class. There should also be caution in handling these words or nouns as individual words in Math and Science as many of them appear more as multi-word units in the texts and thus become more technical in appearance, for example, the words 'line' and 'angle' may be familiar but when in collocate form such as 'acute angle' and 'parallel line', these words then become sub-technical and semi-technical words which then need to be taught explicitly and learnt intentionally (Nation, 2001).

In the analysis of the verbs used in the Math and Science texts, it was found that there were an equal number of lexicalized and delexicalised verbs in both the Science and Math texts. Delexicalised verbs may not be a problem for students as these are common general English language verbs. However, special attention should be given to lexicalized verbs as these verbs carry specific meanings related to the Math and Science fields. It is only through small corpus studies such as this that differences in language can be observed and it is these differences which have to be brought into the classrooms and taught specifically.

The analysis of the adjective key words provided gainful insights into the language of Science and Math. The type of adjectives used in the Science text was more complex than the type used in the Math text. Overall, there were more simple modifiers used in both these texts but the Science text had many more derived adjectives. This implies that students would have to understand lexical derivatives, as this seems to be one of the

features of Scientific lexis (Thirumalai, 2003), to be able to produce them correctly.

In conclusion, this study reaffirms previous studies carried out on learning and teaching vocabulary for Math and Science which emphasizes the need for students to learn technical and sub-technical words and to recognize the differences between words in general English and their meanings in Math and Science (Khisty, 1995, Bernhardt, Hirsch, Teemant & Rodriguez-Munoz, 1996). There is no doubt that there is a need for the integration of language and Science and Math instruction for second language learners as the language demands facing Science and Math learners are very complex and different.

It is through corpus studies like this that teachers and material writers would be able to check and understand the differences in Scientific and Mathematical English from the general English language and be able to apply this knowledge to the teaching of Science and Math in English and the creation of better Science and Math textbooks. Corpora studies have allowed researchers, teachers and learners to use great amounts of real data in their study of language, instead of having to rely on intuition and made-up examples. This study proves that there is a need for small corpus studies to be carried out, especially on language for specific purposes, as these types of studies provide insights which would help in the production of better learning materials and in the teaching and learning process.

REFERENCES

Ansary, H., & Babaii, E. (2003). A 'Bell-Jar' consensus-reached set of universal characteristics of ELT textbooks. In J. Mukundan (Ed.), *Readings on ELT materials* (pp. 69-83). Serdang: Universiti Putra Malaysia Press.

Bernhardt, E., Hirsch, G., Teemant, A., & Rodriguez-Munoz, M. (1996). Language diversity and science: Science for limited English proficiency students. *Science Teacher*, *63*(2), 24-27.

Bondi, M. (2001). Small corpora and language variation: reflexivity across genres. In M. Ghadessy, A. Henry, & R. L. Roseberry (Eds.), *Small corpus studies and ELT* (pp. 135-174). Amsterdam/Philadelphia: John Benjamins Publishing Co.

Carlson, C. (2000). Scientific literacy for all: Helping English language learners make sense of academic language. *Science Teacher*, *62*(3), 48-52.

Cowan, J. R. (1974). Lexical and syntactic research for the design of EFL reading materials. *TESOL Quarterly*, *8*(4), 389-399.

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.

Henry, A., & Roseberry, R. L. (2001). Using a small corpus to obtain data for teaching a genre. In M. Ghadessy, A. Henry, R. L. Roseberry (Eds.), *Small corpus studies and ELT* (pp. 93-113). Amsterdam/Philadelphia: John Benjamins Publishing Co.

Holme, R. (2006). English for science or science for English: Designing content-based materials for language teaching. In J. Mukundan (Ed.), *Readings on ELT materials* (pp. 17-34). Serdang: Universiti Putra Malaysia Press.

Khisty, L. L. (1995). Making inequality: Issues of language and meanings in mathematic teaching with Hispanic students. In W.G. Secada, E. Fennema, & L.B. Adajian (Eds.), *New directions for equality in mathematics education* (pp. 279-297). New York, NY: Cambridge University Press.

Laplante, B. (1997). Teaching science to language minority students in elementary classrooms. *NYSABE Journal*, *12*, 62-83.

Master, P. (1996). *Systems in English grammar: An introduction for language teachers*. New Jersey: Prentice Hall.

Mukundan, J. (2004). *A composite framework for ESL textbook evaluation*. Unpublished doctoral dissertation. Universiti Putra Malaysia, Serdang.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nelson, M. (2000). *A corpus based study of business English and business English teaching materials*. Unpublished doctoral dissertation. University of Manchester, Manchester.

Pandian, A., & Ramiah, R. (2004). Mathematics and science in English: Teacher voice. *The English Teacher*, *33*, 50-61.

Quirk, R., & Greenbaum, S. (1973). *A university grammar of English.* London: Longman.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge: Cambridge 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.

Scott, M. (1996, 1997, 1999). *WordSmith tools.* Versions (1.0, 2.0, 3.0, 4.0). Oxford: Oxford University Press

Sinclair, J. M. (1991). *Corpus, concordance and collocation.* Oxford: Oxford University Press.

Thirumalai, M. S. (2003). Language in science. *Electronic Journal of Language in Science*, *3*(1). Retrieved July 6, 2006, from http://www. languageinin-dia.com/jan2003/languageinscience.html on 6 July 2006

Thompson, D. R., & Rubenstein, R.N. (2000). Learning mathematics vocabulary: Potential pitfalls and instructional strategies. *Mathematic Teacher*, *93*(7), 568-573

Tognini-Bonelli, E. (2001). *Corpus linguistics at work.* Philadelphia: John Benjamins

Trimble, L. (1985). *English for science and technology: A discourse approach.* New York, NY: Cambridge University Press.