

CHAPTER THREE: RESEARCH METHODOLOGY

3.0 INTRODUCTION

This chapter presents an overview of the methodology employed in carrying out the research. The selection of data, in relation to procedures and criteria for selection are explained. The details of the selected source for data, that is the movie, are also provided. Methods of data collection, presentation and analysis are demonstrated in the following sub-sections.

3.1 CRITERIA FOR DATA SELECTION

The selection of data is made based on the author's review of research previously done in the area of Malaysian English that is of her particular interest, as well as her own (qualitative) judgment and observation. Previous studies as reviewed in Chapter 2 have established characteristics of ME lexis and syntax, and in confirming these stable features, the data is collected accordingly. The data for the study is the language spoken (dialogues) by characters in a local English movie, "Ah Lok Café" (the movie particulars such as the producer, director, actors, setting, synopsis, and duration are presented in Appendix 1). The data is described in terms of its linguistic formation as well as comparison with the usage of SBE, thus contributing to the defining features of ME and the understanding of its nature as a variety of English. Justification for choosing a scripted data albeit its limitations is provided earlier in Chapter 1.

The main criterion for choosing this movie as the source for data is the presence of ample utterances of ME (in almost every scene), encompassing all the sub-varieties of ME continuum. Besides, the dialogues also reflect the language used daily as appropriate to the setting, thus intelligible and most of the time humorous to the general audience.

Another reason lies with the characters portrayed. Viewers can easily relate with the characters. Not only are they ethnically diverse, they also come from all walks of life: from a *char kuay teow* hawker to a doctor and lawyer, and from an educated penniless salesman to an uneducated wealthy contractor. It is important to distinguish the main characters' education and socio-economic background as these may influence their style of speaking and choice of words that can subsequently be associated with certain sub-division of the ME continuum. The major character is Ah Lok, the owner of the café, a single father whose life revolves around his two daughters, friends (who are mostly his regular customers), and neighbors (the husband and wife selling *char kway teow* and their son). With reference to the models of lectal continuum reviewed in Chapter 2, it is generally observed that Ah Lok loosely speaks various types of ME but his prominent feature is between basilectal to mesolectal. However, in certain scenes, such as when he is giving a speech in front of his customers in a serious mood (about his daughter's ailment), he mostly uses accurate structures that could turn him into an acrolectal speaker, at that point. A detailed description of the other characters is given in Appendix 2.

Although these observations were based on the author's initial review of the movie, they contribute to the reliability of the data to meet the purpose of study. The purpose of study, as

stated earlier, is to identify and describe features of ME utterances based on proven characteristics of ME, thus contributing to one's understanding of the nature of ME.

3.2 METHODS FOR DATA COLLECTION AND PRESENTATION

The effort in searching for a suitable movie for the study started by surfing the Internet, asking for colleagues' recommendations through SMS and e-mail, and purchasing some of the DVDs/VCDs suggested (i.e. *Sepet*, *Gubra*, *Mukhsin*, *Baik Punya Cilok*, *Goodbye Boys*). Upon reviewing these films, it was found that only a few scenes contained ME utterances that met all the sub-varieties of the lectal continuum, reducing their reliability as data in achieving the purpose of this study. Attempts subsequently continued by visiting various movie stores, and finally coming across the VCD of this particular movie. Indeed, this has been an exhaustive effort, yet of course for a good cause.

The data is gathered from dialogues in the movie that are transcribed and presented scene by scene. The movie is viewed on a laptop's Mpeg Windows Media Player. The dialogues are transcribed by listening through a multimedia headset, typing and saving the scripts using Microsoft Word in the same computer. A number of pauses and repetitions are required to ensure clarity of utterances, hence accuracy of transcription (the accuracy of this transcription is later verified by the supervisor). Each line of utterance is numbered for easy reference of the analysis. Within the transcribed dialogues, based on the established characteristics and framework presented in Chapter 2, ME lexical features are identified and labeled using specific fonts and colors (Appendix 3). Syntactic features are not specifically highlighted in

the transcription as they require more complex explanation beyond labelization, and this is presented within the analysis.

The transcription is done orthographically. The transcript notational conventions to represent the prosodic elements of the natural speech are not applied as they are not necessary for the purpose of the study which is mainly to examine the lexical and syntactical elements of the utterances. This is not to suggest that phonological aspects are not considered at all in the analysis. To a lesser extent, certain phonological factors such as intonation and pitch are taken into account in the interpretation of meaning and pragmatic functions of particles and exclamations.

3.3 METHODS FOR DATA ANALYSIS

The identified ME utterances are analyzed in Chapter 4 (Data Analysis) with reference to characteristic features of ME lexis and syntax as gathered in the literature. Data are presented and described under each sub-type of these lexical and/or syntactical features in comparison to the usage of SBE, where necessary. Although the study focuses on lexical and syntactic features of ME, some pragmatic and phonological associations are imperative in an attempt to describe the usage of certain items such as particles and exclamations. These are thus used as a method of analysis.

Another method of analysis is by using the model of ME continuum. This is because some lexical items can be significantly described in terms of their contexts of use and level of formality. They can be considered as either basilectal, mesolectal or acrolectal based on the

structural characteristics of each sub-variety as outlined by scholars in the reviewed literature. It is crucial to note that the model of ME continuum acts as one of the contributive methods of analysis, thus it is not the objective of the study to label each identified ME item in terms of these categorizations. Some items are described using this model as they can be closely associated with and applied on the defined characteristics of the model, based on the author's observation.

For the purpose of distinguishing the localization of certain items as compared to their usage (or absence of usage) in the SBE, corpus evidences are retrieved from Collin's Cobuild Concordancer⁴. In addition, Encarta World Dictionary⁵ (UK Edition) is also used as reference

⁴ A *concordancer* is a computer programme that shows examples of words in context drawn from a corpus of spoken and/or written texts. These examples allow one to see *collocations* --words that tend to co-occur in a language--in grammatical patterns, compounds or fixed phrases. The *Collins COBUILD* programme contains some 10,000 *node* words (a keyword that is the object of investigation); 140,000 node/collocate pairs; and approximately 2.6 million different examples. These examples are drawn from the *Bank of English* --a corpus of 200 million words from a variety of written and spoken sources (such as fiction, non-fiction, newspapers, magazines, transcriptions of radio broadcasts, everyday conversations and interviews) amassed by the University of Birmingham. The majority of the texts originated after 1990 are primarily British; approximately 25% are American English and 5% come from other native English varieties. A free online sampler of the concordancer software is available at <http://titania.cobuild.collins.co.uk/form.html> and a more detailed description of the *Bank of English* could be retrieved from http://www.cobuild.collins.co.uk/boe_info.htm

(Source:<http://www.collins.co.uk/Corpus/CorpusSearch.aspx>)

⁵ The *Encarta World English Dictionary*, spawn of the popular Microsoft CD-ROM and Internet reference products, covers the new development in the English language thoroughly and efficiently, creating a reference tool for anyone hooked into the new global culture. From the basics (American, British, Australian) to the fringes of English distribution in Africa and Asia, the *Encarta* staff track variant spellings, meanings, and pronunciations in over 100,000 entries comprising some 3.5 million words. If, for example, your Asian correspondent asks you for your "biodata," you can quickly and painlessly learn that she needs your résumé. The work of more than 250 lexicographers in 10 countries, the *Encarta World English Dictionary* is the first reference work that traces the global use of English in its written, spoken and electronic forms. With each word newly defined, the *Encarta World English Dictionary* is the most up-to-date dictionary on the market and will offer a unique perspective on English as the world's language. The *Encarta World English Dictionary* has: Over 400,000 entries, over 20,000 new words and definitions, over 3 million words of text, over 4,000 illustrations, over 1,000 boxed cultural and regional notes, and over 1,500 unique Word Stories on usage notes.

(Source: <http://www.amazon.com>)

in explaining the features identified. These two sources are chosen based on their reliability and reputation as recognized tools that have collected millions of entries, respectively focusing on the use of English lexis gathered from authentic corpora and on the variants in meanings, spellings and pronunciations of the New Englishes (see footnotes).

In order to provide a general statistical picture of the whole data, the information on frequency count for identified lexical items is provided using the 'Find' tool of the Microsoft Word. The overall total of ME utterances contributes to the reliability of the data in achieving the objectives of study. The individual frequency count, however, is not to suggest the items' different degrees of significance. The analysis of the word 'spend' (uttered twice) and the particle 'ah' (uttered 221 times), for instance, may carry the same weight of significance in confirming the presence of such items in ME and understanding their characteristics.

3.4 SUMMARY OF RESEARCH METHODOLOGY

The methodology employed in the study, concerning the sources for data collection, methods for data presentation and data analysis, as well as research for literature is summed up in Table 5. The review of literature consists of the framework/models and references as discussed in Chapter 2. Other references stated in the table that are not reviewed in Chapter 2 are used as additional sources in examining the features of ME within the analysis in Chapter 4. On the whole, the research methodology has accordingly been chosen and applied in carrying out the study in order to achieve its objectives.

METHODOLOGY OF RESEARCH			
RESEARCH SOURCES		RESEARCH REVIEW OF LITERATURE	
DATA COLLECTION & PRESENTATION	DATA ANALYSIS	FRAMEWORK/ MODELS	REFERENCES
<ol style="list-style-type: none"> 1. Data Selection: The Internet, e-mails, SMS, review of recommended VCDS. 2. Source for Data: Dialogues of the selected movie (“Ah Lok Café”) 3. Data Presentation: Transcription of the dialogues using Laptop (with multimedia headset, mpeg Windows Media Player, Microsoft Word 2007) 	<ol style="list-style-type: none"> 1. Collin’s Cobuild Concordancer 2. Encarta World Dictionary (UK Edition) 3. The ‘Find’ tool of the Microsoft word (for frequency count) 	<ol style="list-style-type: none"> 1. Models of Lectal Continuum: <ol style="list-style-type: none"> a. Baskaran’s (2005) Subdivision of ME. b. Pillah & Fauziah’s (2006) Sub-varieties of ME. 2. Morais’s (1997) Categorization of ME varieties based on socio-economic background. 3. Models of Concentric Circles: <ol style="list-style-type: none"> a. Kachru (1984,1985, 1989) b. Yano (2001) c. Ooi (2001) 4. Gupta’s (1998) Diglossia Model. 5. Firbas’s (1992) Framework of TP, MT, DT (Subject Omission) 	<ol style="list-style-type: none"> 1. Local Language Referents & Standard English Lexicalization: <ol style="list-style-type: none"> a. Baskaran’s (2005) Characterization of ME lexical features. b. Su’ad & Hajar (2008) c. Wikipedia Dictionary (2009) 2. Pragmatic/Discourse Particles: <ol style="list-style-type: none"> a. Tay (1977 & 1982) b. Kwan-Terry (1978) c. Smith (1985) d. Platt & Ho (1989) e. Gupta (1992) f. Wong (1994) g. Low & Brown(2005) h. Baskaran (2005) 3. Word Formation; <ol style="list-style-type: none"> a. Bauer (1983) b. Cannon (1985) c. Quirk (1985) d. Quirk & Greenbaum (1987) e. Heah (1989) f. Aitchison (1989) g. Bhat (1994) h. Asmah (1997) i. Quirk(1997) 4. Exclamations: <ol style="list-style-type: none"> a. Tsen (2004) 5. Syntactic Features: <ol style="list-style-type: none"> a. Platt (1980) b. Platt & Wong (1983) c. Firbas (1992) d. Alsagoff (2001) e. Leong, in Low & Deterding (2005) f. Baskaran (2005) g. Low & Brown (2005)

Table 5: Summary of Methodology of Research