

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

Dialectology can be approached in many ways. The tradition has been to look at dialectology in connection with history and development of the language in certain geographical locations. The newer approach is to emphasise on social variations in synchronic usage of a particular language especially in urban areas. This chapter draws attention to the relationship between these two approaches, globally and locally, on the Malay dialectology and specifically on the Malay language spoken in Sabah or Sabah Malay Dialect (SMD).

2.2 THE STUDY OF DIALECTS

Dialectology is the study of how sounds, words and grammatical forms vary from one dialect to the other within a single language. The layman's definition of a dialect is a non-standard, low status, often-rustic form of language, compared to the standard language. It is generally associated with the peasantry, the working class, or other groups lacking in prestige, and often it has no written form. It is regarded as some kind of deviation from the norm, as an aberration of the correct standard form of a language. On the other hand, the standard language is a form of communication, which is more prestigious than dialect (Chambers and Trudgill, 1980). Wardhaugh (1986) argues that it's the size of the community of speakers; dialect is smaller than language, whereas Hudson (1980) thinks it is actually prestige and size.

Dialects may differ from one another in many ways, usually those of a geographical or social nature. Dialect differences are viewed as continua. There are two types of dialect continua, namely the geographical dialect continuum and the social dialect continuum (Chambers and Trudgill, 1980). The degree of geographical differences varies from place to place, but is cumulative with distance. The further a language occurs from its origin, the larger the difference and the greater the difficulty of comprehension for speakers of the original version. At no point is there a complete break, so that geographically adjacent dialects are mutually unintelligible. This type of dialectology is referred to as dialect geography.

Traditionally, dialectology is closely related to the history of linguistics science and the development of comparative historical linguistics, which is diachronic in nature. This consists of three dimensions, namely geographical, structural, and historical. Dialectology later took a synchronic turn and added another two dimensions, namely social and stylistic (Chambers and Trudgill, 1980).

2.2.1 TRADITIONAL DIALECTOLOGY AND DIALECT GEOGRAPHY

The study of dialect variations according to geographical areas is called dialect geography (Bloomfield, 1933; Hockett, 1958; Hughes and Trudgill, 1979; Sankoff, 1973; Wakelin, 1972). It is sometimes simply called dialectology (Chambers and Trudgill, 1980), although this term is more generally used to refer to the study of language variety by any methodology. The geographical variety is also referred to as regional dialect (Hudson, 1980; Wardhaugh, 1986; Asmah Haji Omar, 1992).

Dialectology does not stop at comparing dialects in order to discover the similarities and differences, but a synchronic comparison also helps to answer

diachronic questions about language development (Asmah Haji Omar, 1988). Hence, dialectology comprises two dimensions, space and time, which form the foundation of the comparative study of languages. In other words, dialectology examines the relationship between languages or dialects in a geographical setting which also reflects their historical development.

The rationale for a discipline of dialect geography is simple and straightforward. It seeks to provide an empirical basis for conclusions about the linguistic variety that occurs in certain focal areas. The aims and methods of the dialect geography are as follows:

- i. **PURE FORM** – Dialect geography studies the relationship between language and geography; it identifies a local dialect particularly in the rural areas where “dialect is in its pure form, before it is further contaminated, weakened or totally lost” (Francis, 1983:76). It seeks the most ‘genuine’ and most ‘typical’ speech form spoken in a particular area.
- ii. **NORMS** – In order to get the purest form of a dialect, the purest informants are needed. Thus, the informants, in most cases, consist of Non-mobile Older Rural Males, a grouping often referred to by the acronym NORMs.
- iii. **RURAL AREA** – With the same aim of getting the purest form of dialect, studies have to be concentrated in rural areas, where language is less ‘contaminated’ by foreign elements.
- iv. **RAW DATA** – Data collected for the study are presented in its raw form. Each and every respondent is presented individually.
- v. **LINGUISTIC MAPPING** – Once the interviews are completed, data are collected, responses are tabulated and linguistic mappings are set in order to indicate dialectal variations. Linguistic mappings include literal maps onto

which tabulated responses for a particular study are transferred. This gives the tabulation a geographical perspective, producing interpretive maps. This is in order to derive a more general statement, by showing the distribution of predominant variants from region to region.

This method of studying dialects is also known as traditional dialectology (Francis, 1983). However, a new method emerged in the middle of the 20th century. This was due to the limitation of the way dialectological data were being collected and the criticism voiced by many linguists in the 1960s (Chambers and Trudgill, 1980). The rise of sociolinguistics, the mass-production of recording equipment, and the appreciation that linguistic factors may play a role in social problems in Western societies have led dialectologists to re-orientate their traditional approach to focus on social dialect.

2.2.2 URBAN DIALECTOLOGY AND SOCIAL DIALECT

The object of the study in urban dialectology is the social dialect. By adding social and stylistic dimensions to the traditional study of dialect or dialects, especially in the urban areas, the dialects are called social dialects. These social dialects or sociolects are known as non-regional differences (Hudson, 1980).

Because of the social factors, a speaker may show more similarity in his speech to people from his own social group in a different area than to people from a different social group in the same area. The differences may be termed sociolinguistic variations (Britian, 2005). Sociolinguistic dialectology focuses on the way language varies and changes in communities of speakers and concentrates particularly on the interaction of social factors and linguistic structures (Chambers and Trudgill, 1998).

Unlike dialect geography, which seeks the ‘pure’ form of a dialect from authentic speakers, sociolinguistic dialectology looks into variations to linguistic structures caused by social factors and focuses especially on cities or urban areas with high mobility. Hence, the term sociolinguistic dialectology is also referred to as urban dialectology (Chambers and Trudgill, 1980). As highlighted by Chambers and Trudgill (1980:56):

...linguistic studies suggest that there is probably no such thing as a ‘pure’ dialect, since most varieties of language appear to be variable and to show signs of influences from other varieties.

Urban dialectology does not exclude heterogeneous variations as corruptions of the pure local dialect; it attempts to measure them and fits them into an overall picture of the linguistic situation in the total community (Francis, 1983).

It is assumed that if there is a linguistic variability, there must be free variations. However, they are not so ‘free’ after all. This is because, if the variability has a social correlation, where the old people differ from the young people or the men from the women, then the analysis must be sociolinguistic. Therefore, the focus of the studies “shifts from the language itself to the people who use the language, their social orientations and contrasts” (Francis,1983:193). Two well-known and well-cited pioneering urban dialectologists are William Labov on New York City English in 1966 and Peter Trudgill on Norwich English in 1974.

Labov’s study of social stratification in the speech of lower East Side of New York City in 1966, is certainly one of the famous and detailed examples of urban sociolinguistic dialectology. In this New York department stores study, the

distribution of the variable (r) may be studied in terms of whether the postvocalic /r/ and word-final /r/ are realized as [r] or r-deletion [Ø], for example in 'fourth floor'. He illustrated that linguistic variation is not only geographically stratified, but it is socially stratified as well. According to his theory people having the same realizations of postvocalic /r/ and word-final /r/ would be in the same social group. The highest social group would realise [r] in most occurrences while the lowest class would show the least. The more careful the speaker, the more [r] is realised, while the less careful the speaker, the more r-deletion [Ø] is made. The results from the department store study highlight the main themes of the research. In summary, frequency of use of the prestige variable (r) in postvocalic or final position varied with level of formality and social class.

After the preliminary study of the department store in New York, Labov decided upon five phonological items: the absence of postvocalic /r/, tensing of 'short a', raising of /ɪ/, and substitution of stops or affricates for /t/ and /d/. These variables are clearly exhibited variations in New York, not only between individual informants but also in different social situations and differentiations through four different speech styles: (i) casual speech, (ii) careful speech, (iii) reading connected text, and (iv) word-list. Labov's questionnaire comprised traditional lexical questions, conversation of both folklore materials and free narrative of experiences (danger of death), and reading of formal lists and questions on subjective reactions to various types of speech. The questionnaire was completed by 122 informants, ranging from age 20 to 70. Each of the informants was assigned to one of the nine social ranks or stratifications ranging from lower class to upper middle class according to income, education, and occupation.

Trudgill's study on Norwich in 1974 modelled its procedure to resemble Labov's. Although his sample was quite close in size to Labov's, it represented a smaller percentage of the population. Unlike Labov who based his social criteria on the Mobilisation for Youth study, Trudgill formulated his own sociological analysis based on six criteria: occupation, income, education, housing, locality, and father's occupation. He subdivided his informants into 5 social classes ranging from lower working class to middle middle class. Trudgill recorded the performance of his informants on a set of phonological variables, which included 3 consonants and 13 vowels, in four different styles of speech similar to Labov's, which he terms: Word-list Style (WLS), Reading Passage Style (RPS), Formal Speech (FS), and Casual Speech (CS).

Trudgill's result shows that performance varies according to the variables of social class and type of speech. As social class becomes higher and register becomes more formal, the speech closely approaches the Received Pronunciation (RP) norm, for example the variation between variant [N] and [v] in verbal *-ing* ending. It can be seen that the Middle Middle-Class speakers used the [-Iv] form of the ending only in casual speech and only one-fourth of the time, while the lower working class always used it both in formal and casual speech, employing [-IN] only in reading style. The other three classes fell between these two. The overall pattern is quite similar to those obtained by Labov. Not all of Trudgill's variables, however, follow this pattern of social and stylistic variations. In some of them, the social variation is wide but there is little or no stylistic variation. For other variables, the reverse is true. The data shows little difference between the social classes, but considerable variations across the range from casual to formal language (Francis, 1983).

The sociolinguistic dialectology approach involves major shifts in basic assumptions and procedures from those of geographical dialectology. These affect the population sampling, the linguistic sampling, the technique of data collection, and the presentation of data (Francis, 1983; Chambers and Trudgill, 1980) as follows:

- i. **REPRESENTATIVE SAMPLING** –The concentration on the informant selection on NORMs group has also resulted in neglect of the other social groups. Consequently, the selection of informants is more representative of the population for the social and urban dialectology.
- ii. **SOCIAL DIFFERENTIATIONS** – Urban dialectology pre-selects the language sample, which includes a broad representation of a relatively small number of features, which are expected to correlate significantly with social differences such as age, gender, ethnicity, religion or social stratification.
- iii. **SPEECH DIFFERENCES** – Urban dialectology is not only concerned with social variation between individuals of various social differences, but also with stylistic variation within the speech of a single informant or socially equivalent group. Varying situations are set in order to put the informant into various speech differentiations such as formal, casual, careful, excited speech or reading.
- iv. **LINGUISTIC VARIABLE** – Based on foreknowledge or preliminary survey, the language sample is commonly selected for a broad representation of a relatively small number of features, which correlate significantly with social differences. Unlike the traditional dialectology that uses a broad questionnaire for establishing a full picture of local speech variations, social dialect investigates in detail fewer features for analytic value.
- v. **DATA IN TABLES** - The data of the social dialect are usually presented in tables. Results are very seldom presented in the raw form. Individual

informant's data often seem to disappear as his responses accumulate with those of others in his particular social grid.

- vi. **CHARTS AND GRAPHS** – The results are commonly presented statistically, summarized in charts and graphs. It is not the purpose of this kind of study to obtain the linguistic mappings as it is applicable in dialect geography.

The primary aim of sociolinguistic dialectology is to obtain a systematic description of linguistic variations and their significance in language structure and for language change. Hence, the addition of two more dimensions, i.e. the social and the stylistic to the traditional dialectology, which consist of geographical, structural and historical linguistics.

2.3 MALAY DIALECT STUDIES IN MALAYSIA

The Malaysian society has long realised the existence of different varieties of Malay dialects within and outside Malaysia. However, in Malaysia, linguistic research came very much later than the realisation of dialectal differences.

The setting up of the Malay Studies Department and teaching of sociolinguistics in Universiti Malaya in the early 1960s marked the beginning of the study of traditional dialectology. In 1977 the pioneering dialectal work of Asmah Haji Omar (1977), emerged with *Kepelbagaian Fonologi Dialek-dialek Melayu* and later the English version *The Phonological Diversity of the Malay Dialects* (Asmah Haji Omar, 1991b). This also resulted in many academic exercises, Master's theses, and PhD theses on various Malay dialects in Universiti Malaya, later to be followed by other local universities. Among other studies of the traditional dialectology are Ajib's

MA thesis (1977) on Pasir Mas Malay dialect which was published by Universiti Kebangsaan Malaysia in 1984.

Other MA and PhD theses of descriptive studies on dialects of Malay are Zaharani Ahmad's MA thesis (1988), which was done at the University of Hawaii on the Bota Malay dialect of Perak and was published by DBP in 1991; and this researcher's MA thesis on the Sabah Malay dialect, published in 2000 by Universiti Malaysia Sabah. At the undergraduate level, some of the academic exercises on descriptive studies on dialects have been published in books, for example, Madzhi Johari's academic exercise (1972) on the Kuching Malay dialect of Sarawak, published by Dewan Bahasa dan Pustaka in 1989. However, only one academic exercise (1984) of urban dialectology entitled *Persilihan Bahasa di Kampung Kerinchi* by Wong Khek Seng was published as a monograph by Universiti Kebangsaan Malaysia in 1987.

The contributions of the undergraduate academic exercises, Master's and PhD theses are significant and of high quality. However, more trained and experienced linguists should be the ones conducting detailed studies or accurate research, so that each dialect can be studied extensively and exhaustively. Among the renowned linguists that have contributed significantly to the studies of Malay dialectology in Malaysia are Asmah Haji Omar (1977, 1981, 1987a, 1988, 1991, 1992a, 1992b, 1997, 2004, 2005, 2006); Ajid Che Kob (1984, 1990a, 1990b); Nik Safiah Karim (1965, 1981b, 1985, 1992, 2004); and Collins (1983, 1986, 1988a, 1993a, 1993b, 1998b, 1999, 2006) just to name a few. Their research and writings have enriched the Malay heritage by demonstrating the importance and the value of the Malay language.

The standardisation of language, language planning and the teaching and learning of the national language and the rationalisation of Malay *Weltanschauung* are very much dependent upon the studies of Malay dialects. The studies of Malay dialectology are indeed an important addition to the studies of the standard variety. Each and every Malay dialect is equally important and should be preserved, as it is part of the heritage of the country. Malay dialectology studies can be divided into main studies: Traditional and Urban studies.

2.3.1 MALAY TRADITIONAL DIALECTOLOGY

For the past half a century, after the independence of Malaysia, dialect studies that use the traditional methods of dialectology have grown in number, from student theses to professional linguistic research and publications. The approaches to Malay traditional dialect studies in Malaysia can be divided into synchronic and diachronic, and qualitative and quantitative in nature.

Among the prominent, the synchronic Malay traditional dialect studies is Asmah Haji Omar's (1977) *Kepelbagaian Fonologi Dialek Melayu*. This book is comprehensive, it complies with the geographical dialect concepts and principles. It comprises a few mappings of language areas in the present time. Her classification of dialects does not go according to state boundaries, although state names are used. For example, the Kedah dialect does not confine to the state of Kedah but it goes from Perlis through Kedah to Penang and Northern Perak. In Asmah's (1991b) second edition in English entitled *The Phonological Diversity of the Malay Dialects*, she studies the dialects in comparison to STM in terms of their phonological correspondences, historical development and geographic distribution.

Among other synchronic studies in Peninsular Malaysia are by James Collins, such as the Hulu Terengganu dialect (1983a, 1983c), Pahang dialect (1983b, 1985b), Tioman dialect (1984, 1985a, 1992a, 1992c), Kedah dialect (1988c), Johor dialect (1992c), Hulu Tembeling dialect (1988), and Borneo (Collins 1987a, 1994, 1991a, 1991b, 1998a, 2000, 2001, 2004, 2005a, 2005b). More synchronic studies on Bornean Malay dialects are by Chong Shin, for example Kampung Kupang dialect (2002a), Sekadau dialect (2002b, 2003a, 2003c 2004, 2005, Chong Shin et al 2001), and Saribas dialect (2003b).

Besides synchronic studies, there are also great numbers of diachronic studies on Malay dialects. Asmah Haji Omar (1985) in her work, *Susur Galur Bahasa Melayu*, she quests for the history of the Malay language. This work also deals with the history of the national language development in the region and the people who speak it. Thus, based on thorough research, she describes the Malay dialects of Peninsular Malaysia, Sabah and Sarawak by using the internal construction method. Each of these dialects is systematically studied in terms of phonology and is subdivided into sub-dialects based on the similarity of phonemes. In addition, Asmah also describes various Malay dialects in an innovative and perceptive way and includes some dialects and subdialects such as Patani, Muar and Kedayan, which have not been included in earlier publications on dialectal studies.

SMD as said earlier is a result of migration of ethnic groups within a geolinguistic area. In this belief, Asmah Haji Omar later intensifies her studies into geolinguistics of the Malay world looking at the migration of Malays. As further emphasized by Asmah Haji Omar (2005:83):

...geolinguistik, iaitu cabang linguistik yang mengkaji bahasa daripada segi penyebaran atau difusinya dalam ruang geografi... pertimbangan geografi tidak dapat menisihkan pertimbangan masa (... geolinguistics, a branch of linguistics that studies language in terms of the spread or diffusion of geographical space ... consideration of geography cannot exclude consideration of time).

According to Asmah Haji Omar, the focus of geolinguistics is geographical space, while other parameters such as population movement, direction of movement, and time are supporting factors. It combines both synchronic and diachronic studies, as synchronic curiosity could be answered with diachronic background information and vice versa.

Most recently, Asmah Haji Omar focuses her studies more on the Malay world and language diffusion. From the expedition programme of *Jejak Bahasa*, which she participated in, a book emerged entitled *Alam dan Penyebaran Bahasa Melayu* (Asmah Haji Omar, 2005). In discussing the diffusion of Malay, Asmah introduces the concepts of '*kawasan teras*' (core area), '*kawasan bukan teras*' (non-core area) '*kawasan tradisional*' (traditional area) and '*kawasan bukan tradisional*' (non-traditional area). According to her, an area can be dynamic and it can change its status accordingly. She also uses the concept of '*tadahan penyebaran bahasa*' (receptacle area) as a situation that provides a conducive environment for a language to grow such as an organization or social system, worldview, social-economic activities, education, religion and semiotics.

From another expedition programme, *Semarak Bahasa*, comes another book entitled *Bahasa Melayu di Samudera* (Asmah Haji Omar, 2006). Here, she discusses the diffusion of Malay in Cocos Islands (Keeling), Christmas Island, Western Australia, and Eastern Australia, which she categorises as "*kawasan bukan tradisional*

penyebaran bahasa Melayu”, meaning non-traditional spread area of Malay (Asmah Haji Omar, 2003, 2006). Asmah highlights the language maintenance, language shift and language reversing-shift in the islands and mainland Australia and emphasises that “*benih yang baik, jatuh ke laut menjadi pulau jua*”, meaning that Malay will continue to diffuse and grow even in the non-core or non-traditional areas. The latest book of Asmah Haji Omar (2008) entitled *The Malays in Australia: Language, Culture, Religion*, highlighted three parameters of the typology of an area of language spread: culture area, time-depth and density of first-language speakers. These parameters are used to define a core area or non-core area and traditional or non-traditional area of Malay language spread. In explaining the Malay spread from Brunei to Sabah, Sabah is defined as a non-core but traditional area of Malay language spread.

2.3.2 MALAY URBAN DIALECTOLOGY

Realising the importance of social factors in determining a language choice, use, and variation, Asmah Haji Omar (1981:1) has opened the window to sociolinguistics, by stating:

Sekarang ini dengan bertambah majunya ilmu linguistik dan bertambahnya perkembangan kajian-kajian bahasa dalam hubungan dengan kehidupan masyarakat, tegasnya dengan meningkatnya bidang sosiolinguistik, kajian-kajian mengenai dialek tidak terhad kepada kajian dalam ruang geografis tetapi meliputi juga kajian dalam ruang sosial, yakni kajian yang memberi perhatian pada penggunaan dialek-dialek dalam pelbagai konteks sosial. Dengan itu, istilah dialektologi itu mengalami pertambahan pada bidang cakupannya dengan memasukkan kajian dari segi ruang social (Now with more advanced knowledge and increasing development of linguistic studies in relation to society, and with rapid expansion of sociolinguistics field, research on the dialects are not limited to the study of the geographical space but also includes research in the social space, the study provides attention on the use of dialects in various social contexts. Thus, the term dialectology in the field with additional coverage to include research from the social space).

It was only after the realization of the importance of sociolinguistics that curiosity and interest in the history of languages as well as the relations between language and social factors in the ‘almost neglected’ urban area developed. Linguists then began to pay attention to urban dialectology. As Asmah Hj Omar (1978:109) reminds us:

...dialek bukanlah hanya satu kelainan bahasa yang aneh yang digunakan oleh orang yang tidak berpelajaran, yang hanya ditemukan dalam masyarakat-masyarakat yang terpencil, tetapi adalah bentuk pertuturan yang mana pun jua, yang dituturkan baik oleh orang-orang bandar atau desa, orang-orang tinggi atau rendah dalam tatatingkat penggolongan sosial (...dialect is not just a different variety of language used by people who are not educated, who only knew in remote communities, but is a form of speech of any kind, well spoken by the city or village, the people of higher or lower hierarchy in social classification).

Linguists began to study and discuss urban dialectology by looking at dialectal differences in relation to the social factors that cause the structural uniqueness of social dialectology, especially in urban and city environments. The two typical examples of urban dialectology, specifically the urban dialectology, are the pioneering works of Wong (1982) followed by Idris Aman (1995). These studies were done in the urban areas of Kerinchi and Shah Alam respectively.

Wong’s (1987) study was originally an undergraduate academic exercise done in Universiti Kebangsaan Malaysia in 1982. It was later published by Penerbit UKM in 1987. As a pioneering work (Wong, 1987: 24), his task was to:

...menimbulkan kesedaran sedikit banyaknya tentang kajian sosiolinguistik di bandar kerana penulis berpendapat bahawa kajian dialek bandar daripada segi sosiolinguistik bukan saja menarik bahkan merupakan suatu tugas yang perlu daripada segi pengumpulan data linguistik (...to raise a little awareness of the sociolinguistic study of the city as the writer’s opinion of the sociolinguistic urban dialect is not

only interesting but is a necessary task in terms of linguistic data collection).

Although urban dialectology is a new field in Malaysia and it can be interesting, there are limitations to it as Wong (1987:25) points out:

... pada hal bidang inilah yang menarik kerana masyarakat bandar merupakan masyarakat yang telah banyak mengalami perubahan daripada pertembungan dialek dan bahasa serta interaksi social (... in fact, this is an interesting field for urban society which is a society that has greatly changed from the contact of languages and dialects and social interactions).

Wong (1987) chose to study Kampung Kerinci in Kuala Lumpur Federal Territory, as most of its inhabitants are migrants who have moved there from rural areas. This is also because of urbanisation and industrialisation as well as community relocation of *Dasar Ekonomi Baru* according to the Second Malaysia Plan. Using the number list of Jendall and Smith (1954), he randomly selected 100 informants, aged between 22 and 55 out of 590 families who lived in the focal area. Only data from 72 out of 100 informants were utilised in his study. A total of 62 recordings were ‘*sedar*’ (conscious) while 10 were ‘*tidak sedar*’ (non-conscious) recordings. Informants were categorised into five different categories i.e. male labour, female labour, male non-labour, female non-labourer and housewife. Here, labour refers to blue-collar workers, and non-labour refers to white-collar workers, while housewife refers to homemakers who do not have jobs outside the home.

The main objectives of Wong’s (1987) study were to look at the correlation between linguistic variation and social variation, and to observe code-switching patterns. For the linguistic variatioparameters such as social stratification, age, gender and ethnic membership was also carefully determined in the study.

Informant sampling involves two major decisions. First, the number of informants, which is the size of the sample, and second, the procedure of selecting the informants, which is the method of sampling. For this study, the researcher examined informant sampling methods of Labov on the Lower East Side of New York City (1966), Trudgill on the city of Norwich (1974), the local pioneering work of Wong (1987) on Kampung Kerinchi and Idris bin Aman (1995) on Shah Alam. Each one of them has been adapted where applicable.

3.3.1.1 SIZE OF SAMPLING

Kota Kinabalu covers an area of 350,702 square km with a population of 452,940 (*Yearbook of Statistics Sabah 2010*). By adapting Labov's study in the Lower East Side, where he eliminated "those who had moved away, and those who lived less than two years in the area" (1966: 157), the researcher chose only those who had been residing in Kota Kinabalu for more than two years to be the informants.

The numbers of informants used in previous urban dialectology studies, were ranged from 0.03- 0.14 % of the total population in the selected area. Labov's sample comprised of 195 informants which was about 1:690 or 0.14% of the population of Lower East Side; Trudgill's consisted of 50 which was about 1:3,200 or 0.03% of the population of Norwich, while Wong's was 72, about 1:42 or 0.14% of the Kerinchi's population. Idris, on the other hand, had 32 informants which was about 1:4,062 or 0.03% of Shah Alam's population. For this study, 120 informants were involved, i.e

n, he investigated the phonetic aspect, which is only the variable (a) in word final with three variants [↔], [α] and [α?]. As for the morphology aspect, he only

investigated the affixes or more precisely the bound morphemes on verbs. The social variations comprise gender of informant, informant's occupation, and speech style.

On the other hand, Idris Aman's (1995) study is a Master's thesis done in Universiti Kebangsaan Malaysia. This is Malay dialectology done in an urban area, Shah Alam, using the social approach. This study investigates the correlation between linguistic and non-linguistic factors. These comprise of the stylistic variation and social stratification variation. The linguistic variations consist of seven variables: (ρ), (υ), (ϵ), (ι), ($\alpha 1$), ($\alpha 2$) and ($\alpha 3$). The stylistic differentiations consist of Word-List Style (WLS), Reading-Passage Style (RPS), Careful or Formal Style (FS) and Casual Style (CS). The social stratification variations are made up of four social indices; highest education, occupation, income and type of housing. Idris categorises his informants into four social stratifications: Upper Middle Class (UMC), Middle Middle Class (MMC), Lower Middle Class (LMC) and Upper Working Class (UWC). He chose 32 informants semi-randomly based on house numbers. Of these informants, 28 were male and only four were female. Five informants were of UMC, eight informants were of MMC, 13 informants were of LMC and six informants were of UWC.

In terms of findings, Wong (1987) found out that linguistic variables are subject to gender variable. Males are more likely to use the non-standard form of speech, which is the variant [α]. On the other hand, females, especially those with higher education, use more of the standard form of speech which includes the frequent use of the variant [\leftrightarrow] and more affixes. In terms of occupation, labourers use more of the variant [\leftrightarrow] than those in other occupations. However, the differences between occupations are not great. Although occupation is often determined by one's education, it is interesting to note that linguistic variations are not particularly susceptible to the

effect of education in this case. The ways of recording also do not have any influence on the speech of informants. However, this is not conclusive as only data from one social category i.e. housewife was recorded in ‘conscious’ and ‘non-conscious’ manner while others were all recorded consciously. Females engaged in code-switching more than males. However, other factors such as the topic of conversation, gender, education, psychology, and ethnicity of the researcher also played important roles. As a whole, other than the gender factor, the findings are not very significant. Perhaps this is due to the urbanisation of the focal area where everyone is equally exposed to the mass media and other social factors.

The findings of Idris (1995) show that only the variable (ρ) exhibits stylistic and social strata differentiations. The more formal the situation is, the more the variable (ρ) is realized as $[\rho]$. The lower the social stratification, the more the variable (ρ) is realized as $[\rho]$. The variable (t) and variable (α_3) also exhibit stylistic variation, while the variable (α_1) is significant in social strata differentiation. However, the variable (v), variable (α_2) and variable (ϵ) are not subject to stylistic and social strata differentiations.

In the 1970s, there was a shift in research approach, moving from the qualitative to the quantitative specifically in research of a sociolinguistic nature when a choice of norm has to be made. In quantitative approach, Asmah Haji Omar (2001:34) explains:

... penghitungan dibuat untuk menunjukkan gejala atau ciri yang mana lebih banyak atau lebih kerap berlaku berbanding dengan ciri atau gejala lain. Dalam membuat kesimpulan biasanya yang lebih banyak atau kerap itulah yang diberi perhatian khas (...a calculation is made to show symptoms or characteristics which are more or common than

the other characteristics or symptoms. In concluding usually with more or often it is given special attention).

Besides Zulkifli Yahya (1982), Wong (1987) and Idris Aman (1995), whose findings are presented in figures and statistics, not many quantitative studies have been done since.

Zulkifli tries to employ the quantitative approach in his research in Pasir Mas, Kelantan. With the recording of conversations and distribution of questionnaires to 80 Siamese informants in Pasir Mas, Zulkifli studies the frequency of language choice and use in different differentiations. His findings conclude that: (i) a high percentage of the Kelantanese dialects are used as compared to other dialects in all situations except at home (ii) a high percentage of the Kelantanese dialects are used between Siamese and Kelantanese speakers, and surprisingly, this is also the case among Siamese speakers, and (iii) there is a decrease in the use of the Kelantanese dialect among the Siamese speakers of age 35 and above as compared to those in the age group of 15-35 year olds.

On the other hand, Wong (1987) and Idris Aman (1995) investigated the percentage of certain linguistic features used by social class differentiation, in different speech styles, in urban areas. As further emphasised by Wong (1987:22).

... penyelidikan sociolinguistik yang bersifat kuantitatif dan saintifik telah banyak dilakukan di Amerika Syarikat dan Britian tetapi tidak sedemikian di Malaysia. Oleh yang demikian, penulis merasakan wajar kita sama-sama meneliti hubungan di antara bahasa dengan masyarakat secara saintifik demi kepentingan perkembangan bahasa Malaysia, iaitu bahasa kebangsaan kita (...sociolinguistic research which is quantitative and scientific in nature has done much in the United States and Britain but not of such in Malaysia. Therefore, the author feels it is rational that we examine the relationship between the language and society scientifically in the interest of Malaysia's language development, which is our national language).

2.4 SABAH MALAY DIALECT STUDIES

During the colonial period, most colonial writers, linguists, scholars and missionaries were more interested in the non-Malay communities than in the Malays mainly for the purpose of Christian evangelism. Thus, studies on Malay communities have been neglected in East Malaysia (Wong, 2000). This is evident as Nathesan's study (1995), out of the 115 entries in *Kajian Bahasa-bahasa Bumiputera di Sabah: Satu Renungan*, there were only three entries on *Bahasa Melayu dialek Sabah* (Sabah Dialect of the Malay language) and another three entries on studies of *Bahasa Melayu di Sabah* (Malay language in Sabah). The other 109 entries were on non-Malay indigenous languages and dialects. All the six entries mentioned above were in the form of articles in various local journals. This shows how little attention has been given to the Malay language spoken in Sabah or particularly the Sabah Malay Dialect. This is also indicated by the fact that there was no entry for SMD in any bibliography of Malay dialectology, such as *Bibliografi Dialek Melayu* (Collins, 1988a) or *Bibliografi Dialek Melayu di Pulau Borneo* (Collins, 1990).

Since then, not much has been researched or written on the Malay language spoken in Sabah or the Sabah Malay dialect (SMD). Most of the studies are of minor importance and the methodologies used were not clear. These studies are mainly discussed on distinctive features, formation, influences, interference, and deviation.

The linguistic features of the Malay language spoken in Sabah or SMD are distinctive, especially in contrast with Standard Malay. Nathesan (1992b and 1993a) points out four phonological changes from STM to SMD, there are namely /↔/ > /α/, /τΣ/ > /σ/, /ʔ/ > /κ/ in all word-final positions, and /O/ > /ρ/ in all word-final position such as [απ↔] > [απα], [τΣυτι] > [συτι], [καταʔ] >

[κατακ], ανδ [s↔da:] > [σαδαρ] in SMD. Here, Nathesan talks about sound contrasts as ‘changes’, by supposing that SMD has changed from STM. He also discusses the wide use of *kita*, *mama*, *jam*, *bilang* and *sudah* in SMD. For example, the use of *sudah* in SMD is practically taking over the functions of the word *akan* and *telah* of the STM. Nathesan also examines the semantic differences of a set of vocabulary and differences in certain grammatical structures of SMD.

Zainal Abidin (1993) discusses the distinctiveness of SMD of having only four vowel phonemes /α/, /ι/, /υ/ and /ο/, where /ο/ is often realized as [□], and the absence of vowel phoneme /ε/ and /↔/; the extensive use of the insertion of [?] in open syllable word-final position and deletion of /η/ and some consonant change. Like Nathesan, Zainal Abidin talked about them as ‘changes’, by generalising more changes from STM to SMD such as: /ε/ > [ι], /↔/ > [α], /↔/ > [υ], /ο/ > [υ], /ο/ > [□], /τΣ/ > [σ], /φ/ > [π], /η/ > [Ο], /Ο/ > [?], /?/ > [κ], /ν/ > [N], /ϖ/ > [β] and /ζ/ > [δZ]. As for grammar, he also discusses the different uses of *sudah*, the use of the clitic *bah* as an intensifying particle (*kata penguat*), emphatic particle (*kata penegas*) and affirmative particle (*kata membenar*) by the speakers of SMD. Zainal also highlights some words that are uniquely and preferably used by Sabahans such as *limpas*, *buyuk*, *bubut*, *sigup* and others. Many established forms of address and reference especially *kau* and *aku*, considered as sociolinguistically neutral forms in Sabah, are also briefly discussed by Zainal.

Aishah Saugan (1996) highlighted 21 phonemes in SMD which include 4 vowels: /α/, /↔/, /ι/ and /υ/, and 17 consonants: /π/, /β/, /τ/, /δ/, /κ/, /γ/, /?/, /δZ/, /μ/, /ν/, /N/, /j/, /σ/, /ρ/, /λ/, /ω/ and /ψ/. There is an absence of vowel phonemes /ε/ and

/o/, and consonant phonemes /ç/, /φ/ and /τΣ/¹. These consonant phonemes are regarded loan phonemes. Aishah explains that most the phonological distinctiveness of SMD as going through a ‘replacement’ process from STD such as: [e] > [i]; [o] > [u]; [tΣ] > [s]; [f] > [p]; [z] > [dZ]; and [h] > [ʔ]. She also highlighted the h-deletion and ʔ-insertion as part of the SMD distinctiveness.

In Wong (2000), the researcher highlights in more detail the distinctiveness of SMD phonology by comparing it to STM by using the traditional dialectology method. She formulates a SMD sound inventory of vowels: /α/, /ι/ and /υ/, and consonants: /π/, /β/, /τ/, /δ/, /κ/, /γ/, /μ/, /ν/, /N/, /l/, /h/, /σ/, /tΣ/, /dZ/, /l/, /ρ/, /φ/ and /ω/. She uses the concept of ‘cognate’ instead of ‘replacement’ or ‘change’. She explains the vowel differences as: STM [α], [↔]: SMD [α]; STM [↔]: SMD [α], [ι], [υ]; STM [ε]: SMD [ι]; STM [ι], [I]: SMD [ι]; STM [o]: SMD [υ]; and STM [υ], [Y]: SMD [υ]. While, the consonant differences are as: STM[ʔ], [κ]: SMD[κ]; STM[⊗], [↔], [⌐]: SMD [ρ]; α ν δ STM[η]: SMD [η], [O], [ʔ]. In the effort to regard SMD as a dialect of Malay, she has also emphasised on the social functions of SMD, where its initial function as a lingua franca is slowly extending to become the first language of many Sabahans.

What causes the distinctiveness of SMD? The heterogeneous people of Sabah, many of whom are non-native speakers of Malay and immigrants, are in one way or another responsible for the forming of SMD (Nathesan, 1992b; Wong, 2000). Yusof Mohamad et al (1988) argue that indigenous languages play a great role in influencing the development of SMD. It is further agreed by Abdul Hamid Ahmad (1992) and Nathesan (1993a), that Brunei Malay and Indonesian Malay can be added onto the list

¹ In fact, the /τΣ/ is not a loan sound but a native phoneme in Malay.

of languages which have influenced the formation and development of SMD. Realising that language influence works both ways, Zainal (1993) emphasises that “*pengaruh antara bahasa Melayu dengan bahasa/dialek peribumi ini [Sabah] sebenarnya pengaruh dua hala*”, meaning the influence between Malay and native language/dialect in Sabah is indeed a two-way.

Yusof Mohamad et al, Abdul Hamid Ahmad, Nathesan and Zainal Abidin, have claimed that SMD has been influenced by Brunei Malay. They might have misunderstood the concept of ‘influence’. SMD is not only influenced by Brunei Malay as claimed, but Brunei Malay is actually the base of SMD. Historically, Malay has long been established in Sabah since the 14th century, as it was the lingua franca and the administrative language of the kingdom of Brunei, which was a centre of trade in the region. Sabah is contiguous to the kingdom of Brunei, consequently the Malay in Brunei has spread to Sabah, and making Sabah a large extension of the existing area of Malay spread (Asmah Haji Omar, 2008). Hence, it is not surprising that Brunei Malay has become the base of SMD. This Brunei-based Sabah Malay dialect is then influenced by the speakers’ mother tongue, which includes the non-Austronesian indigenous languages of Sabah and the languages of various immigrant groups (Wong, 2000).

This is where the influence of Indonesian Malay comes in. This is not surprising as there are Indonesian immigrants in most parts of Sabah, especially in Sandakan, Tawau and Kota Kinabalu (Ajijah Alidin, 1993). Perhaps it is for this reason that Abdul Hamid Ahmad, Nathesan and Zainal Abidin conclude that SMD is influenced by Indonesian Malay, especially in pronunciation and lexical choice. This is also supported by Nik Safiah Karim (1988), who states “*Bahasa Melayu di Sabah kelihatan lebih mendekati sifat-sifat bahasa Indonesia daripada bahasa Melayu*

Malaysia”, meaning the languages in Sabah are closer to the characteristics of the Indonesian Malay rather than Malaysian Malay.

The similarity of SMD to Brunei Malay and Indonesian Malay could be due to the choice of the pronunciation of certain sounds such as /α/ and /ρ/ in word-final position. As highlighted by Asmah Haji Omar (1992b:159):

In Brunei and Indonesia “a” is pronounced as the low front vowel [α], whereas in Malaysia it is realised as the schwa ...the final “r” is clearly rolled in Brunei and Indonesia, whereas in Malaysia it ranges from the flapped “r” to a total silence.

However, this is not the case in SMD. Although SMD is spoken in certain parts of East Malaysia, particularly in Sabah, the /α/ is often realised as [α], and the /ρ/ is trill clearly as [ρ] in word-final position (Yusof Mohamad et al, 1988; Zainal Abidin, 1993; Nathesan, 1992b and 1993a, Abdul Hamid Mahmood and Noor Aina Dani, 1994; Aishah Saugan, 1996; Wong 2000). Perhaps for this reason, it is assumed that SMD is more akin to Brunei Malay and Indonesian Malay than to other varieties of Malay in Malaysia.

The Malay spoken in Sabah has been considered as deviation in that speakers are regarded as lacking fluency in and deviating from the standard variety. As such, it has always been associated with language learning interference. Yusof Mohamad et al (1988) in their *Pengaruh Bunyi Bahasa Ibunda dalam Kegunaan Bahasa Malaysia* conclude that the extensive use of [α] as in [πανιN] and [ταλυρ]; the insertion of [ʔ] in final open syllable as in [μωʔ] and [λακʔ]; and insertion of [η] at word final [α] as in [δZιωαη] and [δZαωαη] are deviations from the standard Malay. Similarly,

Abdul Hamid Mahmood and Noor Aina Dani (1994) in their article entitled *Pengaruh Bahasa Ibunda dalam Proses Pemelajaran Bahasa Melayu di Kalangan Pelajar Dusun*, highlight that the failure of the Dusun students to pronounce [ɛ], [↔], [ζ], [ϕ] and [τΣ] in learning Malay ɪσ due to the interference of L1. Noor Aina Dani (1996) further extends her study to the PhD level on a similar topic.

Suzani Sabran (2004) discusses the influences of SMD on learning STM. She emphasises the influences of the morphological aspects of SMD in acquiring Standard Malay by Year Five students of Sungai ethnicity. She studied the Malay writing or composition of 60 Sungai students in two primary schools. She concluded that there are influences of SMD in the Malay writing or composition by these students. The influences of SMD morphology are then categorised according to form and group of words.

The above studies show that SMD is indeed distinctive, especially in comparison with the standard Malay. However, very often the distinctiveness of the Malay language spoken in Sabah or SMD is mistaken for deviation of the standard variety. For this very reason, SMD is considered as the cause of the interference in the standard Malay learning. However, the issue of deviation and interference can be dealt with by looking at SMD as a variety or a dialect of Malay. Subsequently, all the previous studies show different findings. There are indeed variations within the variety itself. Hence, attention should also be drawn to the factors contributed to the language variations.

2.5 CONCLUSION

Although other Malay dialects have been widely studied, not many studies have been done on the Malay language spoken in Sabah or specifically the Sabah Malay dialect. Being a language of multi-ethnic communication, SMD has distinctive phonological features, which to some extent display phonological variations. Hence, it is not an easy task to define and prescribe what the sound system of SMD is. By using the methodology of urban dialectology, it is hoped that through this study we can observe the factors influencing the variations existing in the Malay language spoken in Sabah or more particularly the Sabah Malay Dialect or SMD.