

## Emergent Grammatical Structures of Bahasa Pasar: Based on Hakka-Malay and Cantonese-Malay Glossaries

Wang Xiaomei

[wangxm@um.edu.my](mailto:wangxm@um.edu.my)

Faculty of Languages and Linguistics  
University of Malaya

### Abstract

The Malay language belongs to the Malayo-Polynesian branch of the Austronesian language family. It is the mother tongue of the Malay ethnic group and the second language of the Chinese and Tamil ethnic groups in Malaysia (Asmah 1976). Since the mass migration of the Chinese from China to Malaysia in the 19<sup>th</sup> century, a colloquial form of Malay, *Bahasa Pasar*, has come into being with the contact of Malay and Chinese language. Various text books on colloquial Malay were published to meet the needs of communication between the local residents and migrants. However, the research on *Bahasa Pasar* from the perspective of language contact is still not adequate. This paper aims to describe the grammatical features of *Bahasa Pasar* based on two glossaries published one century ago, *Ma La Yu Yue Yin Yi Yi* (1890/1912) and *Zheng Ke Yin Yi Yi Mu Lai You Hua* (1909/1926). Four grammatical structures emergent from the two glossaries will be analyzed, the nominal structure, verb-object structure, yes-no question, and passive structure. The analysis shows that *Bahasa Pasar* adopts many grammatical features of Chinese dialects. This result indicates that language contact has induced a new variety of Malay, which has syntactic features transferred from Chinese dialects.

**Keywords:** bahasa pasar; language contact; Hakka; Cantonese; *Zheng Ke Yin Yi Yi Mu Lai You Hua*; *Ma La Yu Yue Yin Yi Yi*

### Introduction

The Malay language belongs to the Malayo-Polynesian branch of the Austronesian language family. Currently, Malay is spoken as the mother tongue of the Malay ethnic group and the second language of the Chinese and Tamil ethnic groups in Malaysia (Asmah, 1976). According to the Indian and Chinese records, Malay was the lingua franca in the Malay Archipelago region as early as in the seventh century (Nik, 1995). Since the mass migration of the Chinese from China to Malaysia in the 19<sup>th</sup> century, a colloquial form of Malay, *Bahasa Pasar*<sup>1</sup>, has come into being with Malay and Chinese language as its source languages. Malay as the prestige language supplies most of the vocabulary and is known as the lexifier. Chinese dialect, most probably Hokkien as many scholars have claimed, is the substrate which influences the grammatical structure (Holmes, 2001). Hassan (1999) contends that the term *Bahasa Pasar* probably was first coined by the Europeans. Even earlier, Baba Malay (Shellabear, 1913; Tan, 1980; Pakir, 1986; Thurgood, 1998), which mixed both Malay and Southern Min dialect<sup>2</sup>, was spoken among the Babas<sup>3</sup> in Malacca, Penang and Singapore. With the influx of Chinese

immigrants in the 19<sup>th</sup> Century, various text books on colloquial Malay were published to meet the needs of communication between the local residents and migrants, such as *A handbook of Malay colloquial as spoken in Singapore, being a series of introductory lessons for domestic and business purposes* (Dennys, 1878) and *A manual of the Malay colloquial: such as spoken by all nationalities in the colonies of the Straits Settlements, and designed for domestic and business purposes* (Lim, 1887). These two text books were written in English and served those who were educated in English. Chinese language also has a rich semantic reservoir (Chan, 2003), and hence, to cater to the needs of the migrants from southern China, a series of bilingual dictionaries or glossaries were compiled. Among them, *Ma La Yu Yue Yin Yi Yi* (1890/1912) is a glossary on Cantonese and Malay, while *Zheng Ke Yin Yi Yi Mu Lai You Hua* (Mu 1916/1926) is a glossary on Hakka and Malay, which will be introduced in detail in section 3.

Language is a human communication system. Its basic function is to facilitate the communication between people. This function is even more important in a multilingual and multiethnic society especially during the initial stage of language contact. Malaysia as a multilingual country is an ideal context to carry out research on the communication across ethnic groups either in spoken or written form. This study is such an endeavor to explore the communication between Malays and Chinese a century ago.

This paper aims to analyze the grammatical structures of *Bahasa Pasar* emerging in *Ma La Yu Yue Yin Yi Yi* and *Zheng Ke Yin Yi Yi Mu Lai You Hua* and further look into its relation with Malay, Baba Malay, and Chinese dialects. These two glossaries are written in dialects which were spoken by the two major Chinese dialect groups in the 19<sup>th</sup> Century. Since no glossary in Hokkien has been found yet, these two glossaries remain the only written materials for dialect speakers to learn Malay. Being the first study on *Bahasa Pasar* utilizing written Chinese resources, this article particularly discusses the impact of the first language (Cantonese and Hakka) on the target language (Malay) in a language contact situation. The paper starts with a brief review of the relevant studies in section 2, followed by the introduction to *Ma La Yu Yue Yin Yi Yi* and *Zheng Ke Yin Yi Yi Mu Lai You Hua* in section 3. Section 4 introduces the methodology of current study. Subsequently in section 5, four grammatical structures are analyzed with detailed examples from the glossaries respectively. Finally the study is summarized in the last section.

## Literature Review

The previous studies on the contact languages in Malaysia focus on Baba Malay which is reported to be a mixture of Southern Min dialect and Malay (Tan, 1980; Pakir, 1986; Thurgood, 1998). This variety of Malay has gained native speakers, Babas in Malacca, Penang, and Singapore since 15<sup>th</sup> Century (Shellabear, 1913). Therefore, it could be regarded as a variety of creole language (Pakir, 1986). Tan (1980), Pakir (1986) and Thurgood (1998) describe the phonological and grammatical structures of Baba Malay and claim that Southern Min dialect has influenced the linguistic structure of Baba Malay to a great extent. However, with the decline of the Baba group and its culture, this variety of Malay has very few speakers nowadays.

Shellabear's (1913) work was the first study on Baba Malay. He gave a list of the differences between Baba Malay and the colloquial language of the Malays: (1) a number of words of Chinese origin; (2) not acquainted with a large number of Malay words; (3) mispronunciation of many Malay words; (4) the usage of Chinese idioms rather than the Malay. He also noticed that there was a frequent use of the possessive marker *punya* in Baba Malay which he claimed to be the influence from Southern Min dialect (Shellabear, 1913). The other grammatical feature of Baba Malay mentioned by Shellabear was the word order in the nominal structure. The pronouns *itu* (that) and *ini* (this) always precede the nouns (Shellabear, 1913). Shellabear introduced Baba Malay as a new research field and provided the solid background information for further studies. Although the description on the grammatical features of Baba Malay is not systematic yet, it did shed light on the subsequent studies.

Tan (1980) compared Baba Malay with standard Malay in the following five aspects with a focus on the loanwords and phonological differences: (1) loanwords and loan translations; (2) phonemic differences; (3) lexical differences; (4) semantic differences; and (5) syntactic differences. Grammatical characteristics of Baba Malay were not described in details in Tan's study. Some features were mentioned occasionally, such as the usage of personal pronouns *gua* (first person pronoun, I) and *lu* (second person pronoun, you).

Pakir (1986) and Thurgood (1998) made use of transcription of interviews of Babas, newspapers published by the Babas in late 19<sup>th</sup> century and *A manual of the Malay colloquial* (Lim 1887) respectively to look into the syntactic features of Baba Malay.

Nowadays, *Bahasa Pasar* (hereafter BP) or Bazaar Malay<sup>4</sup>, a pidginized variety of Malay, is commonly used among Malaysians. In fact, since the mass migration in the 19<sup>th</sup> century, Chinese migrants adopted this pidginized version of Malay in order to communicate with the local residents. Za'ba (cf. Hassan 1999: 10) describes this variety as follows, "*This Bahasa Pasar, the words are also Malay words but not sufficient; they are understood and their use is resorted to out of sheer necessity in all simple communication as regards matters of buying and selling and restricted interaction.*" Pakir (1986) and Tan (1980) contend that *Bazaar Malay* may be the root of Baba Malay. The current study aims to investigate such a variety of Malay by looking at its grammatical structures with the reference to modern Malay and the Hakka dialect.

Although there are a few studies on the syntactic structures of Baba Malay (Shellabear, 1913; Pakir, 1986; Thurgood, 1998), systematic studies on the grammatical features of BP are not adequate (Aye, 2006; Sasi, 2008). It may be because of the traditional negative attitude toward this variety by the society (Collins, 1987). Both Aye and Sasi's studies were carried out in Singapore. Aye (2006) provided a detailed description of BP in Singapore and explained it using the substratists' approach, which highlighted the influence of Hokkien on BP. Sasi (2008) described the linguistic features of Singapore Indian Malay, which is spoken by ethnic Indians and is claimed to be a pidgin of Malay. A very recent study on BP was conducted by Bao and Aye (2010) in which they discussed three topic constructions in BP based on the analysis of spoken BP corpus recorded in Singapore and addressed this topic in the light of the convergence of Chinese

and BP. Despite the inadequate literature, the studies in the relevant fields such as Baba Malay could shed light on the current study.

### **Zheng Ke Yin Yi Yi Mu Lai You Hua and Ma La Yu Yue Yin Yi Yi**

*Zheng Ke Yin Yi Yi Mu Lai You Hua* (hereafter the Glossary, cf. Figure 1 for its title page) in this study was published by *Di Qi Fu Yi Wen Shu Ju* in 1926 in Guangzhou. However, the author *Mu Tao* wrote the preface in 1909. This shows that this dictionary was published earlier than 1926. Among various versions, an earlier version (1916) is kept in Stanford University, which was published by *Di Shi Ba Fu Shi Jing Tang* in Guangzhou. No matter which version is chosen for research, it could reflect the usage of the Hakka dialect and Malay language in early twentieth century.

As stated in the preface of the dictionary (cf. Figure 1 right page), it aims to help Hakka people to learn Malay. Therefore, it adopts a style of direct annotation, using Chinese characters with similar pronunciation to imitate the Malay word or phrase or sentence. The reader may learn the corresponding Malay word through its Hakka pronunciation. For instance, 一 (*one*) is annotated as 沙都. The Hakka reader may pronounce it as *satu* according to their mother tongue. For phrases or sentences, the reader needs to read faster to catch the meaning. For instance, 尔想打我 (*You want to beat me*) is annotated as 鲁帽布骨三马沙也 which may be transcribed as *lu mau pukul sama saya* in Malay.

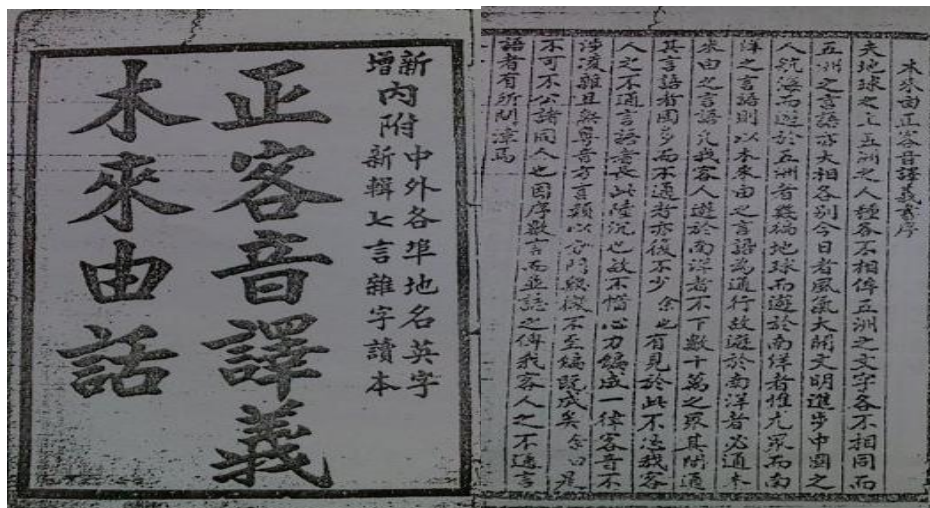


Figure 1: The title page and the preface of the Glossary

This glossary is composed of 27 categories [一], such as numbers (795[i]), astronomy (13), geography and direction (41), time and season (67), goods (40), food (32), instruments (65), clothes (23), instruments in ships (11), craftsman (12), colors (8), instruments for construction (17), houses (23), body parts (44), disease (15), occupation (56), metals (14), mining (12), fruits and vegetables (31), animals (30), one-word (246), two-words (190), three-words (78), long-short sentence (89), asking-the-way (55), dialogue-with-the-chef (26), and other category (203). In total, there are 1,520 entries. The content of the ISSN: 1675-8021

Glossary focuses on the daily communication and needs. For this study, we choose four categories<sup>6</sup>, which are three-word category, long-short sentence category, asking-the-way category, and dialogue-with-the-chef category since the purpose of the study is to investigate the grammatical features of *Bahasa Pasar*.

Within three-word category, there are 78 items which are all composed of three characters. For instance, 我唔信(No. 61, *I don't believe*, cf. Figure 2 right page) is annotated as 沙也拚不植也(*Saya tak percaya*). Long-short sentence category includes 89 sentences which are long or short ones. For instance, the equivalent of 尔能定夺否 (No. 62, *Can you decide?*) is 鲁窝利丹都(*Lu boleh tentu*). Asking-the-way category is about how to ask a way in Malay which includes 55 items. For instance, 此条是正路否 (No. 22, *Is this the right way?*) is presented as 以尔夜烂勿犀帝拚(*Ini Jalan betul tidak*). Dialogue-with-the-chef category is the dialogue between the chef and the master, which is composed of 26 sentences, for example 我七点钟要食饭, 尔到此时能备办此饭餐否(No. 24, *I would like to have dinner at 7. Could you get it ready by then?*) is annotated as 沙也布骨都州淹帽马间拿四三卑如都淹鲁窝利未敬锡(*Saya pukul tujuh mahu makan nasi, sampai itu jam, lu boleh bikin siap*).

*Zheng Ke Yin Yi Yi Mu Lai You Hua* is not the only glossary for the purpose of learning Malay found in Malaysia. *Ma La Yu Yue Yin Yi Yi* 马拉语粤音译义, which was compiled in late Qing Dynasty (1890), serves the same purpose, i.e. helping Cantonese speakers from China learn Malay language. Both glossaries adopt the same compiling style, i.e. using Chinese characters which have the similar pronunciation as its Malay equivalent to represent the corresponding items. Besides, glossaries on learning English through Chinese are also found elsewhere such as *The Redhaired Glossary* 红毛通用番话 in mid 19<sup>th</sup> Century and *The Chinese-English Instructor* 英语集全 around 1862 (cf. Ansaldo, Matthews & Smith 2010), both of which provide valuable resources for Chinese pidgin studies.

## Methodology

In this study, *Zheng Ke Yin Yi Yi Mu Lai You Hua* is the main reference. Examples from *Ma La Yu Yue Yin Yi Yi* will be drawn when necessary. The glossary was transcribed into a corpus in Excel format. In the corpus, we have five different types of information: (1) original text in Hakka, (2) its Chinese annotation, (3) original annotation in Malay, (4) its corresponding Roman writing in pidgin Malay, and (5) its corresponding spoken form in modern Malay. For instance,

(1) (2) (3) (4) (5)

15 乜谁叫尔 是谁叫你 西亚坝邦英三马鲁 siapa panggil sama lu siapa panggil kamu?

This sentence is from Long-short sentence category, No. 15. The original text in Hakka means *who called you*. According to the Glossary, the sentence was explained as 西 亞坝邦英三马鲁(3) in Malay. In the corpus, it was transcribed into Roman spelling system, *Siapa panggil sama lu(4)* based on the pronunciation of Meixian Hakka7[i]. The modern spoken Malay (5) was added to the corpus. In the corpus, it was transcribed into Roman spelling system, *Siapa panggil sama lu(4)* based on the pronunciation of Meixian Hakka<sup>8</sup>. The modern spoken Malay (5) was added to the corpus.

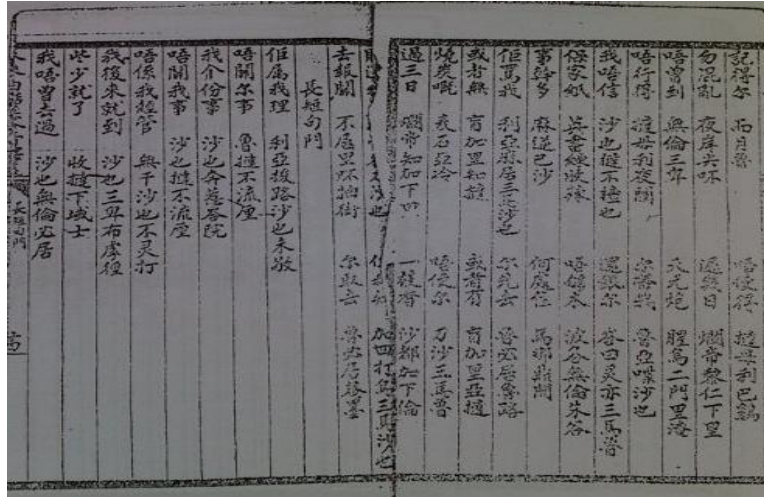


Figure 2: Long-short sentence category

The current study focuses on the grammatical structures of the pidginized Malay language in early 20<sup>th</sup> century. Both qualitative and quantitative approaches were taken in this study. For qualitative approach, to be specific, we will analyze the nominal structure, verb-object structure, yes-no question, and passive structure in the corpus. Then we will compare these structures with modern colloquial Malay and identify the difference. Comparison will also be made between this variety of Malay and the Hakka dialect in order to find any influence from it due to language contact. Wherever necessary, Baba Malay (Lim, 1887; Pakir, 1986; Thurgood, 1998) will also be drawn for comparison. In addition, *Ma La Yu Yue Yin Yi Yi*, a Cantonese-Malay glossary will also be a comparison source. For quantitative approach, the number of each structure in the glossary is counted. Moreover, different forms of each structure are also calculated.

In the following sections, the nominal structure, passive structure, yes-no question and verb-object structure will be analyzed respectively.

### Grammatical Structures in the Glossaries

#### The Nominal Structure

The nominal structure in Malay is constructed as (M)H(Q)(D) (Lewis, 1969). Among the four elements of a nominal structure, Head is obligatory; while Measurer, Qualifier, and Determiner is not. For instance,

- (1) Semua rumah besar itu  
 All houses large those  
 ‘All those large houses’

*Semua* is the Measurer; *rumah* is the Head; *besar* is the Qualifier; and *itu* is the Determiner. We notice that the word order of the nominal structure in Malay is different from the Hakka dialect and standard Chinese. In Chinese, the Head comes after all the modifiers. This difference may lead to mistakes when Chinese learn Malay (Chen 1965: 126). In *Zheng Ke Yin Yi Yi Mu Lai You Hua*, we found some examples of nominal structures.

	Hakka	Mandarin	Malay annotation	Roman writing	Modern Malay
63	保家纸	挂号信	英书练收辣	insurans surat	surat insurans

This is no.63 from three-word category, which means *insurance letter* literally<sup>9</sup>. We notice that its annotation in Malay is *insurans surat* which adopts the same word order as Chinese with the modifier preceding the head. Obviously this breaks the grammatical rule of Malay. The explanation may go to the influence from the Chinese structure. With the contact between Chinese and Malay in the 19<sup>th</sup> century, Malay was influenced by Chinese to a great extent especially at the lexical level (Jones, 2009). With the deep contact of the two languages, a different variety of Malay has come into being gradually. When Pakir (1986) analyzed Baba Malay language, she also found that the word order in its nominal constructions is similar to Southern Min dialect. The above example indicates that Chinese nominal structure rule was applied to Malay. More examples are shown below.

	Hakka	Mandarin	Malay annotation	Roman writing	Modern Malay
22	此条是正路否	这条路正确吗	以尔夜烂勿犀帝挞	ini jalan betul tidak	jalan ini betulkah?

This example is from Asking-the-way category, no.22. We notice that the nominal structure, 此条 was annotated as *ini Jalan (this road)* with the demonstrative preceding the head. Other combinations are also found in the corpus, such as *ini mata* (this policeman), *ini malam* (this evening), *ini orang* (this person), *ini hari* (this day), *ini barang* (this thing), *ini pasal* (this thing). We checked all the items with demonstratives

(14 instances) and found no exceptions<sup>10</sup>. Shellabear (1913, p. 59) claimed that “the Babas always make the adjectival pronouns *itu* and *ini* precede the noun which they qualify”. Pakir (1986) also recorded such finding in her study on Baba Malay. However, if the modifier is not a demonstrative, the word order of the nominal structure tends to follow the grammatical rule of Malay. For instance,

	Hakka	Mandarin	Malay annotation	Roman writing	Modern Malay
31	头家住在高山	老板住在高山	端鼎间无咭鼎如	tuan tinggal bukit tinggi	tuan tinggal di bukit tinggi

In the above example, the nominal structure 高山(*high hill*) is explained as 无咭鼎如(*bukit tinggi*). *Bukit* (hill) is the head, followed by the modifier *tinggi* (high). One more example from two-word category may also illustrate this structure: 别人(*other people*, No.27) is annotated as 阿冷黎仁(*orang lain*) in which *orang* is the head followed by the modifier *lain*.

To sum up, we found both types of nominal structures in *Zheng Ke Yin Yi Yi Mu Lai You Hua*: (1) Head + Modifier (Adjective), (2) Modifier + Head (Demonstrative/Noun + Head). Table 1 summarizes the distribution of these two structures in the glossary. Head-initial structure is the majority, which is the default of Malay. Non head initial structure may be influenced by Chinese. However, the proportion is not very high.

Table 1: Occurrence of nominal structures in the glossary

Nominal structures	Number of examples	Percentage
Head initial (Head + Modifier)	285	90.2%
Non head initial (Modifier + Head)	31	9.8%

This shows that the Hakka speakers applied the rule of nominal structure both in Chinese and Malay in early 20<sup>th</sup> century. In the following sections, more grammatical structures are to be analyzed in order to identify the syntactic characteristics of *Bahasa Pasar*.

### The Passive Structure

In Malay, the passive structure is constructed in three different ways (Nik, 1995, pp. 217-221): (1) with the passive prefixes, *di-* or *ter-*; (2) with personal pronouns<sup>11</sup>; and (3) with *kena* (to suffer)<sup>12</sup>. The typical passive structure is (1), “Object + *di-*verb + *oleh*(by) + Agent”. For instance,

- (2) Buku itu dibaca oleh Ali.  
 Book DEM PASSread by Ali  
 ‘That book was read by Ali.’

*Baca* is the basic form of the verb *read*; *di-* is the prefix showing the passive case; *oleh* introduces the agent Ali. Nik (1995, p. 218) mentions that “the word *oleh* may be deleted in this structure”. Therefore, sentence (2) can also be written as below.

- (3) Buku itu dibaca Ali.  
 Book DEM PASS read Ali  
 ‘That book was read by Ali.’

In this form, the passive reading could only be understood by the prefix *di*.

In *Zheng Ke Yin Yi Yi Mu Lai You Hua*, there are four instances of passive structure.



	Hakka	Mandarin	Malay annotation	Roman writing	Modern Malay
118	被打	被打	阿冷布骨	orang pukul	dipukul
15	被人捉	被人捉	阿冷丹甲	orang tangkap	ditangkap
16	被人抢	被人抢	阿冷三望	orang samun	dirompak
81	佢被捉去	他被捉去	利亚收挞阿冷单甲	dia sudah orang tangkap	dia ditangkap

Number 118 is from Two-word category; number 15 and 16 are from Three-word category; number 81 is from Long-short sentence category. With the transcription in Roman writing system, it is noticed that the passive structure is constructed as follows.

(4) Subject + agent + verb (base form)

For instance, 佢被捉去 (No. 81, S/he was *caught by someone*.) is annotated as 利亚收挞阿冷单甲 (dia sudah orang tangkap). The internal structure is shown below.

dia sudah orang tangkap  
3SG already people catch  
'S/he was caught by someone already.'

The verb *tangkap* is maintained in its base form, no prefix attached. To indicate the agent of the action, *orang* (someone) was added before the verb. However, it is noticed that even there is no agent, the word *orang* (someone) still occurs before the main verb (see no. 118). Therefore, the passive structure in *Zheng Ke Yin Yi Yi Mu Lai You Hua* may be revised into the following.

(5) Subject + *orang* + verb (base form)

In (5), *Orang* may be used as a passive marker. However, more examples are needed to confirm this hypothesis, especially those sentences with an agent other than *orang*. It is also possible that topicalization strategies are employed here. The topic is the patient followed by the agent and verb phrase. Lim (1887) gave some examples of passive expression in his book on the colloquial Malay (Baba Malay) which was spoken by the Babas and Straits-born Chinese<sup>13</sup>.

(6) Saya kasi dia tunang skali. (Saya kasih dia tendang sekali.)  
1SG give 3SG kick once  
'I was kicked by him once.'

(7) Kummaren dia kunna tigga orang pukol. (Kemarin dia kena tiga orang pukul.)  
Yesterday 3SG touch three people assault  
'Yesterday he was assaulted by three people.'

These two examples show that the passive structure in Baba Malay is constructed in a different way. *Kasi*<sup>14</sup> (*kasih*) and *kunna* (*kena*) are used as two agent markers (Pakir 1986).

(8) Kasih / kena + agent + verb (base form)

This finding is also confirmed by Pakir's (1986) data on Baba Malay in Singapore. Besides, the passive structures in *Zheng Ke Yin Yi Yi Mu Lai You Hua* (A Cantonese-Malay Glossary) were checked. There are four instances of passive structures, which are exactly the same as those in *Zheng Ke Yin Yi Yi Mu Lai You Hua*.

Entries	Annotation	Transcription	English translation
被打	斤𪗇布决	Kena pukul	Beaten
被人捉	根𪗇阿冷 登甲	Kena orang tangkap	Caught by someone.
被人抢	根𪗇阿冷 心门	Kena orang samun	Robbed by someone.
佢被捉去	厘亚根𪗇阿冷 登甲	Dia kena orang tangkap	He was caught by someone.

Different from *Zheng Ke Yin Yi Yi Mu Lai You Hua*, the passive structures in *Ma La Yu Yue Yin Yi Yi* manifest the same pattern as Baba Malay, i.e. "kena + agent + verb (in base form)". To compare this structure and (5) with the Chinese counterpart, we find some similarity. In Chinese, the passive structure may be expressed in the following construction.

(9) Bei 被 + agent + verb

Although the agent marker in the Hakka dialect<sup>15</sup> is different from standard Chinese, the basic construction of passive structure is the same. The passive structure in *Ma La Yu Yue Yin Yi Yi* seems to resemble the Chinese pattern. However, in *Zheng Ke Yin Yi Yi Mu Lai You Hua*, this structure is constructed without the agent marker. This indicates that there was variation in *Bahasa Pasar* in terms of the passive expression in early 20<sup>th</sup> century. Nevertheless, both types of structures reveal the influence from the substratum language, which is the so called *substratum transfer* (cf. Bao, 2005), which is also exemplified in next section.

### The Yes-no Question

There are two types of interrogative sentences in Malay: (1) with question words, (2) without question words (Nik, 1995). In this section, type (2) or Yes-no question is the focus. According to Nik (1995, p. 201), "type (2) is actually a declarative sentence with a rising intonation". An interrogative suffix *-kah* may be attached to the end of the sentence in spoken Malay.

For instance,

(10) Pegawai itu di pejabatkah?  
 officer DEM in office Q  
 'Is that officer in the office?'

In *Zheng Ke Yin Yi Yi Mu Lai You Hua*, there are a number of Yes-no questions which are shown below.

	Hakka	Mandarin	Malay annotation	Roman writing	Modern Malay
77	佢有证人否	他有证人吗	利亚亚挞石 四帝 挞	dia ada saksi tidak	dia ada saksi kah?
22	此条是正路否	这条路正确吗	以尔夜烂勿 犀帝 挞	ini jalan betul tidak	jalan ini betul kah?
20	尔欢喜否	你喜欢吗	鲁收 隔帝 挞	lu suka tidak	adakah kamu suka?

No.77 is from Long-short sentence category; no. 22 from Asking-the-way category; no 20 from Dialogue-with-the-chef category. All these three interrogative sentences are constructed in the same way.

(11) Declarative sentence + *tidak* (no/not)

This structure is also available in colloquial Malay. However, a phonologically reduced negator *tak* is used to form the question. For instance, *Dia suka tak?* (Does s/he like (it)?) In Bahasa Pasar, a full form *tidak* is used. This may resemble the structure in Chinese as exemplified in (12).

(12) 佢有证人否

3SG has witness NEG  
dia ada saksi tidak  
'Does he have witness?'

否 as a negative marker is put in the end of the sentence to construct an interrogative structure in Hakka<sup>16</sup>. It is noticed that the word order of the above sentence in Hakka and *Bahasa Pasar* is exactly the same. Therefore, the expression of interrogative structure in this colloquial variety of Malay may result from the influence of the Hakka dialect. It could also be explained in another way: the Hakka speakers at that time applied the grammatical rules of the Hakka dialect when they learned or spoke Malay. This phenomenon is commonly observed in second or foreign language learning which is termed *transfer* (cf. Li Wei 2010, p. 5). As a result, the grammatical structures of the Hakka dialect were transferred into *Bahasa Pasar*. In *Zheng Ke Yin Yi Yi Mu Lai You Hua*, there are 9 such examples as the above structure<sup>17</sup>. *Ma La Yu Yue Yin Yi Yi* was also checked and the same structure was found.

(13) 佢有证人否 厘亚亚打杀矢地打 (Dia ada saksi tidak)

3SG has witness NEG  
'Does he have witness?'

(13) indicates that Cantonese speakers underwent the same process as Hakka speakers when they learned Malay. They also transferred the syntactic rule of Cantonese to the target language and that resulted in the grammar of *Bahasa Pasar*.

In both glossaries, yes-no question is also constructed in a simpler way, attaching a rising intonation to a declarative sentence. For instance, 你能定夺否 (No. 62 from Long-short sentence, *Can you decide?*) is annotated as *Lu boleh tentu* (you can decide) without the

use of *tidak* in the end of the sentence. In this case, the speaker has to use a rising intonation. This particular example also appears in *Ma La Yu Yue Yin Yi Yi* with the use of *tidak* in the end of the sentence (the transcription is *Lu boleh tentu tidak*). This may imply that there is a free choice between the two patterns of yes-no question, i.e. (1) a declarative sentence with a rising intonation, or (2) a ‘declarative sentence + *tidak*’.

In addition to yes-no question, “V-*tak*-V” (V-not-V) is also found in the glossary. It is similar to the “V-唔-V” structure in Hakka.

	Hakka	Mandarin	Malay annotation	Roman writing	Modern Malay
1	肯唔肯	肯不肯	冒挞冒	mau tak mau	mahu atau tidak

This example is from Three-word category, no.1. In Hakka, 唔[m] is one frequently used negative marker. This structure constructs an affirmative-negative question by combining an affirmative and a negative verb phrase. It is annotated as *mau tak mau* (want not want). The same entry in *Ma La Yu Yue Yin Yi Yi* uses a different structure to annotate, *boleh tidak*, i.e. sentence final negator structure. Whether V-*tak*-V is influenced by the Chinese language is not clear because this structure is also found in colloquial Malay. For example, *Suka tak suka dia kena belajar Bahasa inggeris juga* (*Whether he likes it or not, he’s got to learn English.*) (King, 1960, p. 133). It may result from the contact between the Chinese language and Malay in earlier days, which needs more investigation.

### Verb-object Structure

Malay follows SVO word order. The object comes after the verb. In *Zheng Ke Yin Yi Yi Mu Lai You Hua*, a special marker in Verb-object structure is found as exemplified below.

	Hakka	Chinese	Malay annotation	Roman writing	Modern Malay
47	尔想打我	你要打我	鲁帽布骨三马沙也	lu mau pukul sama saya	kamu mahu pukul saya
15	乜谁叫你	是谁叫你	西亚坝邦英三马鲁	siapa panggil sama lu	siapa panggil kamu

Both No. 47 (*You want to beat me*) and No. 15 (*Who called you?*) are from Long-short sentence category. In both sentences, the inflections of the transitive verbs, *pukul* (to beat) and *panggil* (to call), are lost; both objects are pronouns, *saya* (me) for No. 47 and *lu*<sup>18</sup> (you) for No. 15. Between the verb and the object, *sama* is inserted. In Baba Malay (Lim 1887), similar structure is also found.

- (14) Tanya sama dia.  
Ask sama 3SG  
‘Ask him.’

*Sama* means ‘together with’ or ‘same’ in Malay. Tan (1980) claimed that *sama* in Baba Malay is used to replace *dengan* (with) and *kepada* (to). However, the above examples do not require any prepositions between the transitive verb and its object. Therefore, Tan’s explanation needs more clarification. In *Zheng Ke Yin Yi Yi Mu Lai You Hua*, *sama* tends

to appear between the transitive verb (in its base form) and the object. Most of the time, the object is a person. Otherwise, it is not necessary to use *sama* before the object. For instance,

	Hakka	Mandarin	Malay annotation	Roman writing	Modern Malay
18	点蜡烛	点蜡烛	巴声厘灵	pasang lilin	nyalakan lilin

No.18 (from Three-word category, *to light a candle*) is a Verb-object structure with an object *lilin* (candle). The object is connected with the verb *pasang* (to light) directly without the use of *sama*. This seems to imply that in a Verb-object structure, when the object does not associate with a person, there is no need to use *sama*. Therefore, the use of *sama* can be summarized as follows.

(15) Verb + *sama* + object (animate)

This structure also appears in *Ma La Yu Yue Yin Yi Yi* as exemplified in (16). The same condition for the use of *sama* is observed in (16). It always precedes a personal pronoun.

(16) 求你 明德三孖老  
Minta sama lu  
Beg *sama* 2SG  
'Beg you.'

Among the 35 such structures in *Zheng Ke Yin Yi Yi Mu Lai You Hua*, 12 entries are observed not to follow the above rule. Out of the 12 entries, 5 of them adopt *orang*<sup>19</sup> or a phrase containing *orang* (No. 12 from the Dialogue-with-the-chef category) as its object such as No. 25 in the following.

	Hakka	Annotation	Transcription	Translation
25	寻人	楂里阿冷	cari orang	Look for someone.
53	记得尔	而月鲁	ingat lu	Remember you.
12	无要紧的使个帮手去	帝拈亚巴鲁窝利邦英沙都阿冷完都仁必居巴虱	tidak apa, lu boleh panggil satu orang bantu pergi pasar	It doesn't matter. You may ask a person to go to the market.

No. 25 is from the Two-word category, which does not insert *sama* between the main verb and its object *orang*. This may indicate that *sama* is not necessary when the object is not a pronoun. Only two exceptions are found<sup>20</sup>. For comparison, the usage of *sama* in *Ma La Yu Yue Yin Yi Yi* was checked. The result<sup>21</sup> shows that all the Verb-object structures with *orang* and other nouns as its object do not use *sama*. Therefore, (16) is to be revised as follows.

(17) Verb + *sama* + personal pronoun

No. 53 from the Three-word category is also an exception although its object is a personal pronoun. There are 7 similar entries in the Glossary. To check if they require different conditions for the use of *sama*, the corresponding entries in *Ma La Yu Yue Yin Yi Yi* were compared. It shows that *sama* is used for all the same entries as *Zheng Ke Yin Yi Yi Mu Lai You Hua*.

- (18) 记得你 意𠵼乞三孖老 ingat sama lu  
remember *sama* 2SG  
'Remember you'

In *Ma La Yu Yue Yin Yi Yi*, *sama* is consistently used in dative context. The indirect object must be a personal pronoun as explained above. Here are some examples from the glossary.

- (19) 你去话佢知 Lu bergi kata sama dia tahu  
2SG go tell *sama* 3SG know  
'You go and tell her/him.'

- (20) 还银你 Bayar ringgit sama lu.  
pay money *sama* 2SG  
'Pay back the money to you.'

The use of *sama* may be summarized as follows: (1) used as a patient marker if the patient is a personal pronoun, (2) tend to be used before a dative personal pronoun. The inconsistent use of *sama* in *Zheng Ke Yin Yi Yi Mu Lai You Hua* may result from the variation of the grammatical structures of *Bahasa Pasar*. This type of variation reflects the instability of this pidgin variety.

## Conclusion

### *Bahasa Pasar* as a Pidgin Malay

As described above, *Bahasa Pasar* (Bazaar Malay) manifests a different grammar with the reference to colloquial Malay. Based on the analysis of its nominal structure, passive structure, yes-no question, and verb-object structure<sup>22</sup>, the contact between the Hakka dialect, Cantonese and Malay in late nineteenth century and early twentieth century can be traced to some extent.

- (1) BP tends to use both "Modifier + Head" and "Head + Modifier" word order in nominal structures. However the latter is the majority (90.2%).
- (2) The passive structure is constructed as "Orang + verb" or "kena + agent + verb".

- (3) Attaching *tidak* (no) to a declarative sentence is one of the ways to construct yes-no question. V-*tak*-V is used in both Chinese dialects and bahasa pasar.
- (4) It inserts *sama* between a transitive verb and its personal pronoun object especially in a dative context.

With above syntactic characteristics, *Bahasa Pasar* does exhibit the nature of a pidgin language<sup>23</sup>. (1) First of all, it was developed out of the needs of communication between two groups who do not share a common language (Holmes, 2001). As stated by the author of *Zheng Ke Yin Yi Yi Mu Lai You Hua* in the preface of the glossary, it aims to help those Hakka immigrants to learn Malay in order to make a better living in Nanyang (Southeast Asia). (2) It has no native speakers. This variety of Malay was only spoken between Chinese (including Babas and other Strait-born Chinese) and Malays or Indians. As for the intra-communication within the Chinese community, they spoke different varieties of Chinese dialects at that time. The prestige dialect in the region was used as the lingua franca between dialect groups, such as Hokkien in Penang. (3) Its functions are limited. Most of the content of *Zheng Ke Yin Yi Yi Mu Lai You Hua* is regarding the usage of the language in the settings where the reader needs to communicate with the local residents, such as Asking-the-way setting and Dialogue-with-the-chef setting. (4) Its grammar is relatively simpler than its source language (Sebba, 1997). For BP, almost all the inflections in verbs are dropped. Only stem verbs are used in all the examples from both glossaries.

In addition to the above four grammatical features reviewed in this paper, the morphological, phonological, semantic, and lexical features should also be analyzed as a full analysis of pidgin language. To understand BP as reflected in the Chinese-Malay glossaries, this article is only a first step. Sebba (1997) identified four *design features* of pidgins: (1) lack of surface grammatical complexity, (2) lack of morphological complexity, (3) semantic transparency, and (4) vocabulary reduction. Whether these features are found in BP, more analysis needs to be done in the future. *Bahasa Pasar*<sup>24</sup> is still in use today in Malaysia as stated by Asmah (1987)<sup>25</sup>, although in a slightly different<sup>1</sup> form from the variety in *Zheng Ke Yin Yi Yi Mu Lai You Hua*. For instance, the passive structure and yes-no question are constructed differently from the current variety of colloquial Malay. According to Bakker's (1994) typology of pidgin language, *Bahasa Pasar* may be grouped into *interethnic contact language* since it is mainly used across ethnic groups, i.e. Chinese, Indians and Malays. However, if its initial stage is taken into account, *Bahasa Pasar* will fall into the group *work force pidgin*<sup>26</sup> (Bakker 1994) or *immigrants' pidgin*<sup>27</sup> (Sebba 1997). However, Collins (1987) held a different opinion towards the status of *Bahasa Pasar*. He contends that it complements the standard Malay and even resisted Pidgin English in the history. No matter how the status of BP is defined, its communicational value in the Malaysian society cannot be denied. It has truly served as a communication tool across ethnic groups in Malaysia for the past century.

## Acknowledgement

This project was supported by UMRG057/10SBS, University of Malaya. This paper was presented at the Fifth ASEAN Linguistics Conference (21-22 December 2011, Dewan Bahasa dan Pustaka).

## Endnotes

1. *Bahasa (Malayu) Pasar* literally means the language of the market. Alternatively, *Bazaar Malay* is also used in the literature. Hassan (1999: 9-11) also mentioned other equivalent terms *Bahasa Dagang* (trade language) and *Bahasa Kacukan* (mixed language). In this paper, *Bahasa Pasar* is used.
2. Southern Min dialect is known as Hokkien in Malaysia.
3. Babas refer to those descendants of late 15<sup>th</sup> and 16<sup>th</sup> Chinese immigrants who adopted partial Malay customs due to the mix-marriage with local women. Females are known as Nyonyas; males are Babas.
4. The two terms will be used interchangeably in this paper. However, the term in its original language is preferred.
5. The numbers in the brackets are the number of entries in each category. For instance, there are 79 entries in the category of numbers.
6. We also checked Two-word category for the analysis of nominal structure, Verb-object category and Passive structure.
7. The research assistant of this project is a native Meixian Hakka speaker from Malaysia. He also speaks fluent Malay and Mandarin.
8. The research assistant of this project is a native Meixian Hakka speaker from Malaysia. He also speaks fluent Malay and Mandarin.
9. In modern Hakka, 保家纸 may be understood as registered mail.
10. There are two exceptions in Two-word category, No. 10 今晚(this evening) is annotated as *malam ini* and 本年(this year) as *tahun ini*. However, when *malam ini* is used in the utterance (Dialogue-with-the-chef No. 21), it becomes *ini malam*. This shows the instability of *Bahasa Pasar* on the one hand and the variation at the pragmatic level. Generally speaking, 'demonstrative + noun' is still a preferred word order for *Bahasa Pasar*.
11. Structure (1) is applicable with the third person. As for the first and second person, there is no need to add prefix to the main verb. The pronoun is followed by the base word (Nik 1995: 219). For example, *Hutang itu telah saya jelaskan minggu lalu* (The debt was paid by me last week.).
12. In this structure, the word *kena* (to suffer) is used before the base verb (Nik 1995: 220). For example, *Pencuri itu kena tangkap oleh polis* (The thief was caught by the police).
13. Straits-born Chinese refer to those who were born in Malay Archipelago region before the mass migration in 19<sup>th</sup> century; their counterpart is China-born Chinese who were born in China.
14. *Kasi* and *kunna* follow the old spelling system; *kasih* and *kena* are spelt according to the new spelling system of Malay. The basic meaning of *Kasih* is to give; *kena* means to touch or suffer. For the standardization of spelling system of Malay, please refer to Jacobson (1992).
15. The passive marker in Hakka is [pun] and [pei] in Cantonese.
16. In the Hakka dialect, the negative marker in yes-no question is [mau], while in Cantonese the negator is [mou].
17. In total, there are 13 entries which contain the negative marker 否 in the Glossary. 9 out of them are annotated as 'declarative sentence + *tidak*'. The rest 4 entries omit the negative marker *tidak*.
18. *Lu* is a loan word from Southern Min dialect (Lim 1887:133). It is found in both glossaries and extensively used in Bahasa Pasar.
19. *Orang* means person in Malay, which is a noun.
20. There are altogether 7 entries with *orang* as their objects. One instance is No. 48 from Long-short sentence 尔想杀人( You want to kill someone) which is annotated as *Lu mau potong sama orang*. The other one is No. 56 from the same category 尔莫话别人知(You don't tell others) which is transcribed as *Lu jangan cakap kasih tahu sama orang*.



- <sup>21</sup>. There are 4 such instances in Ma La Yu Yue Yin Yi Yi.
- <sup>22</sup>. In addition to these four structures, the possessive marker *punya* is frequently mentioned in the literature of Baba Malay or Bazaar Malay. In *Zheng Ke Yin Yi Yi Mu Lai You Hua*, *punya* is also used as a genitive marker. For instance, 这是我物 (This is my thing) (No 73 from Long-short sentence category) is annotated as 以尔峇冷沙也奔惹 (*ini barang saya punya*). Many scholars claimed that the use of *punya* (own or possess) is modeled on the Hokkien particle *e* (Lim 1887, Shellabear 1913, Tan 1980, Bao and Aye 2010). In Hakka dialect, the genitive marker is *ke* which constructs a similar possessive structure (modifier + *ke* + head) as other dialects or Mandarin. Therefore, *punya* is also commonly used by Hakka speakers when they communicate with Malays or Indians. Many scholars claimed that the use of *punya* (own or possess) is modeled on the Hokkien particle *e* (Lim 1887, Shellabear 1913, Tan 1980, Bao and Aye 2010). In Hakka dialect, the genitive marker is *ke* which constructs a similar possessive structure (modifier + *ke* + head) as other dialects or Mandarin. Therefore, *punya* is also commonly used by Hakka speakers when they communicate with Malays or Indians.
- <sup>23</sup>. For a full definition of pidgin, please refer to Holm (2000:5).
- <sup>24</sup>. This variety of pidgin is listed in the annotated list of creoles, pidgins, and mixed languages by Smith (1994: 358). It also appears in Hancock's (1971) survey in which it is believed to be a pidginized variety of High Malay in widespread use in Malaysia and Indonesia.
- <sup>25</sup>. According to Asmah (1987), there were varieties of pidgins in Malaysia, such as pidgin English, pidgin Cantonese, and pidgin Hokkien. The most common pidgin is pidgin Malay or the Bazaar Malay.
- <sup>26</sup>. It refers to the pidgin which came into being in work situations (Bakker 1994: 28).
- <sup>27</sup>. It refers to the pidgin spoken by new immigrants, such as Guest-worker German (Sebba 1997: 31).

## References

- Ansald U., Matthews S. & Smith G. (2010). China coast pidgin. *Journal of Pidgin and Creole languages*, 25(1), 63-94.
- Asmah Haji Omar (1976). *The teaching of Bahasa Malaysia in the context of national language planning*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Asmah Haji Omar (1987). Patterns of language communication in Malaysia. In Asmah H. O. (Ed.), *National language and communication in multilingual societies* (pp. 13-25). Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Aye Khin Khin (2006) *Bazaar Malay: History, grammar and contact*. Unpublished PhD Thesis. National University of Singapore.
- Bakker P. (1994). Pidgins. In Arends J., Muysken P. & Smith N. (Eds.), *Pidgins and Creoles: An introduction*, (pp. 25-40). Amsterdam: John Benjamins Publishing Company.
- Bao Zhiming (2005). The aspectual system of Singapore English and the systemic substratist explanation. *Journal of Linguistics*, 41(2), 237-267.
- Bao Zhi Ming and Aye Khin Khin (2010). Bazaar Malay topics. *Journal of Pidgin and Creole languages*, 25(1), 155-171.
- Chan, Sheung-Wai. (2003). Some crucial issues on the translation of poetic discourse from Chinese to English. *GEMA Online Journal of Language Studies*, 3(2), 1-25.

- Chiang, Min-hua (2006) Grammatical characteristics of tung and bun in Dongshi Hakka and the relatedness of the two markers. *Languages and Linguistics*, 7(2), 339-364.
- Chen, Meng He (1965). *Malaiyu yufa duben*. Singapore: Shanghai Shuju.
- Collins J. (1987). Malaysian and Bazaar Malay: Polarity, continuity and communication. In Asmah H. O. (Ed.), *National language and communication in multilingual societies* (pp. 151-174), Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Dennys N. B. (1878). *A handbook of Malay colloquial as spoken in Singapore, being a series of introductory lessons for domestic and business purposes*. Singapore: London : Trubner.
- Feng Zhao Nian (1912). *Ma la yu yue yin yi yi* (A Malay-Cantonese Glossary). Shanghai: Huaying Shu Ju.
- Hancock I. (1971). A survey of the pidgins and creoles of the world. In Hymes D. (Ed.), *Pidginization and creolization of the languages* (pp. 509-523). Cambridge: University Press.
- Hassan Ahmad (1999). *The language policy of Malaysia*. Kuala Lumpur: Art Printing Works.
- Holm J. (2000). *An introduction to pidgins and creoles*. Cambridge: Cambridge University Press.
- Holmes J. (2001). *An introduction to sociolinguistics*. Harlow: Pearson Education Limited.
- Jacobson R. (1992). In search of status: Bahasa Malaysia for national unification. In Ammon W. and Hellinger M. (Eds.), *Status change of languages* (pp. 200-226). Berlin: Walter de Gruyter.
- Jones R. (2009). *Chinese loan-words in Malay and Indonesian: A background study*. Kuala Lumpur: University of Malaya Press
- King E. S. (1960). *Speak Malay*. London: University of London.
- Lai Huei-ling (2001). On Hakka Bun: A case of polygrammaticalization. *Languages and Linguistics*, 2(2), 137-153.
- Lewis M. B. (1969). *Sentence analysis in modern Malay*. Cambridge: Cambridge University Press.
- Li Wei. (2010). BAMFLA: Issues, methods and directions. *International Journal of Bilingualism*, 14(1), 3-9.
- Lim Hiong Seng (1887). *A manual of the Malay colloquial*. Singapore: Koh Yew Hean Press.

- Mu Tao (1926). *Zheng Ke Yin Yi Yi Mu Lai You Hua*. Guangzhou: Di Qi Fu Yi Wen Shu Ju.
- Nik Safiah Karim (1995). *Malay grammar for academics and professionals*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Pakir A. (1986). *A linguistic investigation of Baba Malay*. Unpublished Ph.D thesis. University of Hawaii.
- Sasi Rekha D/O Muthiah (2008). *A grammar of Singapore Indian Malay*. Unpublished MA thesis, National University of Singapore.
- Sebba M. (1997). *Contact language: Pidgins and creoles*. New York: Palgrave.
- Shellabear W. G. (1913). Baba Malay: An introduction to the language of the Straits-born Chinese. *Journal of Straits Branch of the Royal Asiatic Society*, 65, 49-63.
- Smith N. (1994). An annotated list of creoles, pidgins, and mixed languages. In Arends J., Muysken P. & Smith N. (Eds.), *Pidgins and Creoles: An introduction* (pp. 331-374). Amsterdam: John Benjamins Publishing Company.
- Tan Chee Beng. (1980). Baba Malay Dialect. *Journal of the Malaysian Branch of the Royal Asiatic Society*, 53(1), 150-166.
- Thurgood E. A. (1998). *A description on nineteenth century Baba Malay: A Malay variety influenced by language shift*. Unpublished Ph.D thesis. University of Hawaii.

### **About the author**

Wang Xiaomei is a senior lecturer at University of Malaya. Her research interest includes language spread, language maintenance and language shift, multilingualism and teaching Chinese as a second language. She teaches courses in Chinese linguistics and culture.<sup>28</sup>

