

Some introductory notes on the development and characteristics of Sabah Malay

TOM G. HOOGERVORST

ABSTRACT

This is a preliminary description of the Malay variety used as a lingua franca in the Malaysian state of Sabah at the northernmost top of Borneo. The paper discusses a number of common linguistic features that distinguish Sabah Malay from other Malay varieties and analyses these features from a historical linguistic perspective. While it is argued that Sabah Malay has a close historical relation with other Malay dialects spoken in Borneo, especially Brunei Malay, the vernacular is also influenced phonologically and lexically by Sabah's indigenous and immigrant speech communities. Words and sentences recorded or elicited during fieldwork in various parts of Sabah illustrate these points.

KEYWORDS

Sabah Malay, colloquial Malay varieties, Borneo, Brunei Malay, contact linguistics, phonology, lexical borrowing, morphosyntax, periphrastic causatives, head-final alternatives, discourse particles.

1 INTRODUCTION¹

As a part of a volume on charting borders, the current paper aims to offer new perspectives on the linguistic boundaries and continuities of Borneo. The island is home to various indigenous Malayic varieties and three Malay-

¹ I would like to express thanks to Junki Suaka, Kartini Abd. Wahab, Mark Miller, Helen Brunt, Azam Othman, Hein Steinhauer, and Sander Adelaar for their valuable suggestions and to the European Research Council for financial support.

TOM G. HOOGERVORST (1984) is a junior scholar in the field of Southeast Asian Studies. He studied Languages and Cultures of Southeast Asia and Oceania (BA) and Asian Studies (MPhil) at the University of Leiden. He is currently completing his DPhil dissertation on linguistic evidence for the early presence of insular Southeast Asian traders in the Indian Ocean (Sealinks Project, University of Oxford). His interests include the Malay language and dialects, Austronesian linguistics, Southeast Asian archaeology, maritime trade and archaeobotany. Tom Hoogervorst can be reached at: tomhoogervorst@hotmail.com.

derived national languages, which correspond to its political division between Malaysia, Indonesia and Brunei Darussalam. However, the formation of Borneo's national and provincial borders is antedated by the presence of several Malay-speaking communities, many of which are still poorly described. This paper aims to address some common features of the Malay variety spoken in Sabah, Malaysia's easternmost state. Sabah Malay is used among speakers from various linguistic backgrounds as a means of interethnic communication. I will not here attempt to estimate how many people use this variety as their mother tongue, second or third language, but it suffices to say that the vast majority of Sabah's more than 2.8 million inhabitants are able to use Sabah Malay in everyday unmarked speech situations. The variety is increasingly becoming the first language of people of mixed heritage and those who have lost their ancestral language.

Sabah probably acquired its name from Brunei Malay *saba* 'downriver', alluding to its location down the Brunei River (Maxwell 1981-1982). The Malay language has been in use in the northernmost regions of Borneo since at least the fourteenth century as a result of the expansion of the Brunei Sultanate. At present, several villages in south-western Sabah are inhabited by first language speakers of Brunei Malay varieties. Sabah is also home to "Cocos-Malay" communities, consisting of people from various Malay-speaking backgrounds sent to the Cocos Islands under the British colonial administration. After several generations, many Cocos-Malay families moved to Sabah, bringing with them a distinct and not closely related Malay variety (Adelaar 1996). The Malay variety discussed in this study, on the other hand, is a second language for most of its speakers, facilitating interethnic contact between Sabah's over 50 different ethno-linguistic groups.² As a result, Sabah Malay is used in markedly different ways from one ethnicity to another. However, the shared features in phonology, lexicon and morphosyntax discussed in this paper delimit Sabah Malay from other Malay varieties and establish its status as a distinct dialect. In the cities and peripheries alike, Sabah Malay (henceforth SM) is preferred over standard Malay or Bahasa Malaysia (henceforth BM) as a means of in-group communication (Wong 2000). As a general rule of thumb, standard Malay is only used with outsiders, including people from the Malay Peninsula and Sarawak. Having said that, there is a complicated situation of diglossia between standard Malay and Sabah Malay, which is not comprehensively treated in this study. Although this paper focuses on the variety at the bottom of this continuum, extensive code-mixing and code-switching between "high" and "low" language is very much a part of daily life in Sabah.

Sabah does not possess an ancient cultural or literary tradition comparable to other regions of Borneo, such as the historical sultanates of Brunei and Banjarmasin. Perhaps for this reason, Sabah's Malay variety is rarely acknowledged as a distinct dialect. Instead, it has been called a pidgin (Moody 1984) or an offshoot from Brunei Malay (Collins 1991: 689 footnote 2). In

² See J.K. King and J.W. King (1984) for a detailed analysis of Sabah's linguistic situation.

fact, the Malay varieties spoken in the area between Limbang (Sarawak) and Labuan (Sabah) are part of a continuum, for the sake of convenience usually labelled “Brunei Malay”. Bearing this in mind, the arbitrary division into Sabah Malay, Sarawak Malay and Brunei Malay has been rejected by some scholars, claiming that it does not do justice to Borneo’s linguistic complexity (Collins 1989: 240). The Malay varieties spoken in Brunei have been relatively well documented (Collins 1990). Those spoken in Lembah Baram, North Sarawak, are analysed by Chong Sin (2009) in a short paper. Not much descriptive work has hitherto been done on the Malay varieties of Sabah proper. Chuchu (2003, 2005) provides lexical data of varieties spoken around the Gulf of Brunei, but looks exclusively at mother tongue speech communities. D.J. Prentice (1969) provides a “colloquial Malay” wordlist intended for linguistic fieldwork in Borneo, but the words listed can be identified exclusively as Sabah Malay and would not necessarily be understood in Sarawak or Kalimantan. Finally, we can consult two short articles on Sabah Malay (Nathesan 1993; Zainal Abidin Bakar 1993) and a very useful dissertation on its phonology and social functions (Wong 2000).

In contrast to the high density of primary Malayic branches in West Borneo, there is a relative paucity of dialects in other parts of the island. The known varieties share several common features and, as will be argued in this paper, might be part of an historical dialect continuum. Until recently most Malay varieties spoken outside of West Borneo, such as Brunei Malay, Berau Malay, Kutai Malay, Banjar Malay, Bukit Malay, Kotawaringin Malay, Pontianak Malay, Sambas Malay and Sarawak Malay, were exclusively spoken around river deltas (Collins 1991: 690), suggesting relatively recent settlement by Malay-speaking communities from overseas. These Malay-speaking communities presumably settled down at their present locations with the intend to conduct trade with people from the inlands, famed for their precious stones, tortoise shells, gold, pepper and valuable forest products (Broek 1962: 145-146). Collins (1991: 693-694) notices two important linguistic features that set Borneo’s alleged dialect continuum apart from other Malay varieties and Proto-Malayic. First, the Proto-Malayic /ə/ developed into /a/ in Brunei Malay, Berau Malay, Kutai Malay and Banjar Malay. It has been argued in another context that this innovation might have taken place in relatively recent times (Blust 2006: 23 footnote 13). Second, the split of Proto-Malayic /u/ into /u, o(, ɔ)/ and /i/ into /i, e(, ε)/ attested in extra-Bornean varieties did generally not take place in Borneo.³ In addition to these phonological similarities, Collins (1991: 691-692) provides examples of lexical innovations and retentions shared by several Malay dialects of Borneo, some of which will be examined in this paper. According to Chuchu (2008: 155), the Malay variety of Berau is so similar to his native Brunei dialect that people would mistake

³ To complicate this generalization, it is attested in the Borneo varieties of Kutai and Sarawak, but not in several peripheral Malay varieties in the west, including Besemah (McDonnell 2008) and Ulu Terengganu (Collins 1983a: 24-30). In Old Malay inscriptions, the *akṣaras* for /o/ and /e/ were only used for Indic loanwords (Mahdi 2005: 188).

him for a local during his fieldwork in the Berau region. Be that as it may, we are still facing large gaps in our knowledge of Borneo's Malay dialects, obscuring a comprehensive survey of their interrelatedness.

This paper is divided in several sections. The first section examines the phonology of Sabah Malay compared to other Malay varieties. The second section addresses its lexicon, focusing on lexical similarities with other Bornean varieties and lexical borrowing from Sabah's indigenous and immigrant languages. The third section elaborates on several morphosyntactic features of Sabah Malay, including periphrastic causatives, head-final alternatives and discourse particles. Some points for discussion and suggestions for further research conclude this study. The linguistic data in this study are collected during a one-month fieldwork trip to Sabah in November 2009. Such a timeframe is, of course, insufficient to acquire an excellent working knowledge of the local lingua franca, but it allowed me to notice some of its more salient characteristics. I conducted my fieldwork in Labuan, Kota Kinabalu, Tuaran, Ranau, Telupid, Sandakan, Lahad Datu, and Semporna. At first, I used standard Malay (Bahasa Malaysia) for communication. Some consultants responded in the standard language, but others felt more at ease switching to Sabah Malay while explaining unfamiliar words. Illustratively, the latter used the local form *inda?ada* 'there is not', whereas the former used the standard Malay equivalent *tiada*.⁴ I made an effort to quickly familiarise myself with the pronunciation and some frequently used lexical items of Sabah Malay, often by overhearing everyday conversations. This eventually allowed me to pursue communication in a more localized type of Malay.

This paper provides comparative data from several Malay varieties of Borneo. The speech locations of these varieties are indicated in the inset of Figure 1. Unless indicated otherwise, the following works have been used:

- 1) Banjar Kuala (Hudson 1967; Abdul Djebar Hapip 1977)
- 2) Berau Malay (Kurniawati et al. 2002)
- 3) Brunei Malay (including Kedayan; Haynes 1900; Chuchu 2003, 2005, *Kamus* 2006)
- 4) Delang (Hudson 1967)
- 5) Iban (V. Sutlive and J. Sutlive 1994)
- 6) Kedayatn (Salako variety; Adelaar 2005b)
- 7) Kutai Malay (Suryadikara et al. 1979)
- 8) Mualang (Tjia 2007)
- 9) Sarawak Malay (Collins 1987)
- 10) Tamuan (Hudson 1967)
- 11) *Proto-Malayic (Adelaar 1985)

⁴ This form is slightly formal and rarely occurs in spoken language; colloquial Peninsular Malay has *tak ada* or *takda*.



Figure 1. Map of Sabah; The locations of the discussed Malay varieties are indicated in the inset (Figure by the author).

2 PHONOLOGY

This section examines some of the phonological features of Sabah Malay. The original vowel inventory consists of /a/, /i/ and /u/, whereas /e/ and /o/ predominantly occur in borrowings or instances of code-switching or code-mixing with standard Malay. In addition, there are three isolated examples of Sabah Malay words with /o/ given in (13). Compare the following examples of loanwords with /e/ and /o/ and their original source:

- | | | | |
|-----|----------|---------------|--|
| (1) | bos | 'bos, father' | English |
| | mem | 'madam' | English (< ma'am) |
| | pemes | 'famous' | English |
| | tontolou | 'penis' | Kadazandusun |
| | amoi | 'waitress' | Hakka (< 阿妹 <i>a¹ moi³</i> 'younger sister') |

In most other cases, Sabah Malay shows no signs of having developed the distinctions /u, o(, ɔ)/ and /i, e(, ɛ)/. Even in borrowings, only the high vowels are used:

- | | | | |
|-----|---------|------------|-------------------------------|
| (2) | jandila | 'window' | Portuguese (< <i>janela</i>) |
| | inum | 'handsome' | English |
| | numbur | 'number' | English |

As mentioned before, Sabah Malay and various other Bornean varieties are

characterized by the absence of the schwa /ə/, although this sound is often heard during code-switching or code-mixing with standard Malay. In Sabah Malay, the reconstructed Proto-Malayoic /ə/ is generally replaced by /a/ and in some cases by /i/ or /u/.⁵ As Sabah Malay has not developed syllable stress or consonant gemination, we find homonyms such as *gali* 'to dig' (< **gali*) versus *gali* 'itch' (< **gəli*) and *tapuk* 'to hide' (< *tapuk*; compare (33)) versus *tapuk* 'to clap' (< **təpuk*). There are certain rules that decide whether /ə/ becomes /a/, /i/ or /u/, to which the vowel or consonant in the following syllable is influential. The historical schwa becomes /a/ when the subsequent vowel is /a/, as illustrated in (3):

(3)	*-ə(C)Ca-	>	-a(C)Ca-	
	pənat	>	panat	'tired'
	bəsar	>	basar	'big'
	təgap	>	tagap	'robust'
	kətawa	>	katawa	'to laugh'
	dəngar	>	dangar	'to hear'

The historical schwa generally becomes /a/ when the vowel in the following syllable is /i/, as illustrated in (4):

(4)	*-ə(C)Ci-	>	-a(C)Ci-	
	bəli	>	bali	'to buy'
	kəring	>	karing	'dry'
	pədih	>	padi	'smarting'
	səmpit	>	sampit	'narrow'
	gəli	>	gali	'itch'

In addition to four isolated examples given in (12), I have found three words in which the historical schwa has become /i/. In all three examples, the subsequent vowel is /i/, the intervocalic consonant or consonant cluster is anterior and the final syllable is open:

⁵ Where there are no Proto-Malayoic reconstructions, I am compelled to compare my Sabah Malay data with their equivalents in standard Malay. These cognate standard Malay examples regularly reflect the historical /ə/.

- (5) * $\text{-}\text{ə}(\text{C}_1)\text{C}_1\text{i}$ > $\text{-i}(\text{C}_1)\text{C}_1\text{i}$ (in which C_1 represent an alveolar consonant)
- | | | | |
|-------|---|-------|---------------|
| pəti | > | piti | 'case, chest' |
| məsti | > | misti | 'have to' |
| bəsi | > | bisi | 'iron' |

The historical schwa generally becomes /a/ before /u/, as the examples in (6) demonstrate:

- (6) * $\text{-}\text{ə}(\text{C})\text{Cu-}$ > $\text{-a}(\text{C})\text{Cu-}$
- | | | | |
|--------|---|--------|--------------|
| jəmur | > | jamur | 'to dry' |
| kəntut | > | kantut | 'fart' |
| təbu | > | tabu | 'sugar cane' |
| təpung | > | tapung | 'flour' |
| ləsung | > | lasung | 'mortar' |

If the intervocalic consonant is a liquid (/r/ or /l/), the historical schwa before /u/ becomes /a/ or /u/. Compare the examples provided in (7), in which the pronunciation of the historical schwa fluctuates between /a/ and /u/:

- (7) * $\text{-}\text{ə}\text{C}_1\text{u-}$ > $\text{-a}\sim\text{uC}_1\text{u-}$ (in which C_1 represent a liquid consonant)
- | | | | |
|-------|---|---------------|--------------------|
| təlur | > | talur ~ tular | 'egg' |
| pərut | > | parut ~ purut | 'belly' |
| pəluk | > | paluk ~ puluk | 'embrace' |
| tərus | > | tarus ~ turus | 'straight, direct' |
| bəlum | > | balum ~ bulum | 'not yet' |

There are two seemingly isolated examples in which the historical schwa has become /u/. These are given in (8):

- (8)
- | | | | |
|-------|---|-------|----------|
| gəruk | > | gumuk | 'fat' |
| bətul | > | butul | 'really' |

The historical schwa becomes /i/ before a postalveolar affricate (/c/ or /j/) regardless of the subsequent vowel. This innovation also takes place when the historical /c/ is deaffricated to /s/, an innovation addressed under (18).

Compare the following examples:

- (9) * $-\text{ə}(\text{C})\text{C}_1\text{V}-$ > $-\text{i}(\text{C})\text{C}_1\text{V}-$ (in which C_1 represent a palatal affricate)
- | | | | |
|---------|---|---------------------|--------------|
| kəncing | > | kincing ~ kinsing | 'to urinate' |
| kəcil | > | kici (1) ~ kisi (1) | 'small' |
| bənci | > | binci ~ binsi | 'to hate' |
| səjuk | > | sijuk | 'cool' |
| səkəjap | > | sakijap ~ skijap | 'a moment' |

As is the case in several other Malay varieties, the antepenultimate schwa in Sabah Malay is neutralized in fast speech, provided that the resulting consonant cluster is phonotactically permitted. Compare the examples given in (10):

- (10)
- | | | | |
|-----------|---|----------|--------------------|
| bərani | > | brani | 'brave' |
| pərəmpuan | > | prampuan | 'woman' |
| kəluar | > | kluar | 'go outside' |
| bəlas | > | blas | 'the number 11–19' |
| sətəngah | > | stengah | 'half' |

The cluster $/\text{CərC}/$ becomes $/\text{CaraC}/$ in Sabah Malay. Similar innovations can be noticed in other Malayic varieties, such as Minangkabau. In general, clusters of the $/\text{CərC}/$ type are atypical in Malay phonology and are probably a relatively recent development. None of the words that display this cluster can unambiguously be reconstructed to Proto-Malayic, instead being either morphologically complex or borrowed. Compare the following examples in Sabah Malay:

- (11) * $-\text{CərC}-$ > $-\text{CaraC}-$ ⁶
- | | | | |
|---------|---|----------|-----------------|
| kərbau | > | karabau | 'water buffalo' |
| kərja | > | karaja | 'work' |
| tərjun | > | tarajun | 'waterfall' |
| tərbang | > | tarabang | 'to fly' |
| mərcon | > | maracun | 'fireworks' |

⁶ This innovation does not take place if the post-schwa consonant is a liquid, as demonstrated by the examples *parlu* 'need' and *camarlang* ~ *samarlang* 'brilliant'.

A small set of attestations given in (12) do not follow any of the phonological rules listed above and appear to be more general substandard forms, as they are also widespread in other Malayic varieties. Compare the following examples:

(12)	SM		BM
	piluru	'bullet'	pəluru
	cilaka ~ silaka	'accident'	cəlaka
	pigang	'to hold'	pəgang
	pi(gi)	'to go, towards'	pərgi

The historical diphthongs /au/ and /ai/ remain largely intact. I have found only three isolated (yet frequently occurring) examples in which /au/ has become /o/:

(13)	kalau	>	kalo	'if'
	ma(h)u	>	mo	'want'
	kau	>	ko	'you'

Vowel lengthening in Sabah Malay conveys excessiveness, for example, *dulu* 'earlier' > *du:lu* 'way back', *jauh* 'far' > *ja:uh* 'very far', *sana* 'there' > *sa:na* 'way over there'. The lengthened vowel is pronounced in a high pitch. The occurrence of vowel lengthening in Sabah Malay could be the result of regional influence, as I have also noticed it in West Coast Bajau. On the other hand, it is also attested in Malay varieties outside or Borneo. Example (14) illustrates the usage of vowel lengthening in Sabah Malay:

(14)	Urang	sampai	di	sabla	sa<:>na	suda	paam.
	people	until	LOC	side	there<EXC>	already	understand.
	'People way over there already understand it.'						

Whereas the historical word-final /k/ has become a glottal stop /ʔ/ in several Malay varieties, in Sabah Malay it remains /k/. The glottal stop in Sabah Malay is used in another context. It is generally added to an intonational segment ending in a vowel or an elided /h/, an innovation addressed under (17). As a result, all words ending in an open syllable are given a word-final glottal stop in word-for-word elicitation, but are not necessarily pronounced so in a natural sentence. I have noticed a similar tendency in Kadazandusun and suspect it to be a regional feature. The negative marker *ndaʔ* is the only word pronounced with a word-final glottal stop in all positions. Examples (15) and (16) illustrate the use of a word-final glottal stop (underlined) at the end of an intonational segment:

(15) Jangan ko pigi sana? ada pangait!
 VET 2s go there EXIST headhunter
 'Don't go there, there are head-hunters.'

(16) Maka susa la? bagi parau mo limpas di sana?
 So difficult DP for boats want go.past LOC there.
 'So it's difficult for boats to go past there.'

The /h/ elision in Sabah Malay offers interesting linguistic comparisons across Borneo. Collins (1991: 695-696) comments on the realisation of the historical /h/ in several Borneo varieties, claiming that it is retained in Banjar and Brunei Malay but lost in Berau Malay.⁷ In Sabah Malay, it is dropped in all positions. The only exception I have found is the discourse particle *bah*, addressed in examples (73) to (94). This particle is mostly (but not always) pronounced with an audible word-final /h/. The realization of the reconstructed /h/ in several Bornean varieties is illustrated in (17):

(17)	Proto-Malayic	Banjar, Kedayan	Brunei, SM		BM
	*tuha	tuha	tua	'old'	tua
	*tihadang	tihadang	tiang	'pole, mast'	tiang
	*jahit	jahit	jait	'to sew'	jahit
	*hidung	hidung	idung	'nose'	hidung
	*buah	buah	bua	'fruit'	buah
	(from Sanskrit)	sudah	suda	'already'	sudah

Many of Sabah's regional languages, including Kadazandusun and West Coast Bajau, lack a voiceless postalveolar affricate /c/ in their consonant inventory. As a result, speakers at the lower end of Sabah's diglossic continuum are inclined to pronounce /c/ as /s/. On the other hand, speakers with a Chinese background and those who regularly speak standard Malay rarely pronounce /c/ as /s/. Compare the following examples of /c/ ~ /s/ fluctuation in Sabah Malay:

(18)	SM		BM
	cantik ~ santik	'pretty'	cantik
	macam ~ masam	'kind of, like'	macam
	binci ~ binsi	'to hate'	bənci

As is the case in many other Malay varieties, the consonants /f/, /z/ and

⁷ Nowadays, the historical /h/ is only realized in Kedayan, whereas it is omitted in other Brunei Malay varieties (Azam Othman, personal communication).

/v/ are absent in Sabah Malay. In loanwords they are realized respectively as /p/, /j/ and /b/, as demonstrated in (19):

(19)	SM		BM
	upis	'office'	ofis
	paam	'understand'	faham
	bija	'difference'	beza
	gibap	'to give up'	-(from English)

To conclude this section on Sabah Malay phonology, many speakers impose their native phonology on their Malay speech. For instance, speakers of Kadazandusun tend to realize /j/ as /z/, speakers of Ubian and Chinese dialects pronounce /r/ as /l/ and speakers of Philippine languages fluctuate between /p/ and /f/. Specific patterns in phonology, intonation and word choice can give away someone's ethno-linguistic background to a local expert. However, it is beyond the scope of this paper to examine this issue in more detail.

3 LEXICON

This section elaborates on vocabulary specific to Sabah Malay. The words shared with Brunei Malay and other Bornean varieties will be examined first, before moving on to language contact and lexical borrowing. Although Sabah Malay is believed to be derived from Brunei Malay (Collins 1991: 689 footnote 2), a lot of specific Bruneian vocabulary is not used or understood in Sabah. I have confronted Sabah Malay speakers with a Brunei Malay wordlist compiled by Haynes (1900) and most of the words were unfamiliar to them. Having said that, I do not know which of these words are still understood by modern Brunei Malay speakers. Counterexamples of lexical items shared by Sabah Malay and Brunei Malay are listed in (20):

(20)	SM, Brunei Malay		BM
	akun	'to confess, to agree'	məngaku
	ampus	'asthma'	səsak nafas
	andang-andang	'naturally'	memang
	bida	'not pretty, ugly'	buruk, hodoh
	bubut	'to chase'	kəjar
	buyuk	'to cheat'	tipu
	giuk ⁸	'worm, maggot'	ulat
	kalatiak	'armpit'	kətiak
	kapayas	'papaya'	bətik
	lanjang	'cooking pot'	bəlanga
	sapak	'mixture of ingredients'	campuran
	sigup	'cigarettes'	rokok
	sikui	'water melon'	təmbikai

⁸ Probably borrowed from a local source, compare for 'worm, maggot' in Mukah *giyuʔ*, Murut *giyuk* (Nothofer 1987 Sub Appendix), and Kadazandusun *gizuk*.

tamaha	'greedy'	tamak
tamu	'local market'	pasar
umban	'to throw'	lempar (umban = 'to sling')

In addition to the above examples, some lexical items are also shared with other Malay varieties spoken in Borneo. To the best of my knowledge, Collins (1991: 691-692) was the first scholar to use lexical data from Bornean dialects to explore their interrelatedness. The lexical correspondences between Sabah Malay and other varieties, listed in (21) to (34), shed new light on this issue:

- (21) Sabah Malay *bangas* 'stale'; Brunei Malay *bangas* 'sour, stale', Banjar *bangas* 'stale', Kedayatn *bangas* 'any animal which is harmful to rice plants (insects, birds, mice); term of abuse for people who cause damage', but standard Malay *basi* 'stale'.
- (22) Sabah Malay *bantut* 'gay, transsexual'; Kedayatn *bantut* 'castrated', but Banjar *bantut* 'unripe (abscess)', standard Malay *bantut* '(to) retard'. The form also occurs as Tausug *bantut* 'impotent, castrated, transsexual' and Cebuano *bantut* 'effeminate'.
- (23) Sabah Malay *cacak* 'gecko'; Brunei Malay *cacak*, Banjar *cacak*, Delang *caca?* Tamuan *caca?* idem, but standard Malay *cicak* idem. Fluctuation between *cicak* ~ *cacak* also occurs in Chamic languages (Thurgood 1999: 314), pushing the doublet back to Proto-Malayo-Chamic.
- (24) Sabah Malay (*inda?* 'no(t), not want'; Brunei Malay *inda* 'no(t)', Kutai (Kota Bangun variety) *inda* idem, Banjar *indah* 'not want', but standard Malay *tidak* 'no(t)'. The form also occurs as Minangkabau *inda?* 'no(t)'.
- (25) Sabah Malay *kamah* 'dirty'; Brunei Malay *kamah*, Iban *kamah*, Sarawak Malay *kamah* idem, but standard Malay *kumuh* idem. The form *kamah* also occurs in Seraway (Helfrich 1904: 64), Lampung (Junaiyah 1985: 125) and Aslian languages on the Malay Peninsula (Benjamin 1976: 105), suggesting a wider distribution in Malay varieties in the past. Adelaar (1985: 154) reconstructs *kamah* ~ *kumuh* 'dirty' for Proto-Malayic.
- (26) Sabah Malay *lawà* 'attractive'; Brunei Malay *lawà*, Kedayatn *lawà* idem, also understood in the Malay Peninsula, but standard Malay *cantik* idem.
- (27) Sabah Malay *lilipan* 'centipede'; Banjar *halilipan* (various dialects: *lilipan* ~ *lalipan*; Kawi 2002: 413); Tamuan *lalipan*, Mualang *nyelipan* idem, but standard Malay *lipan* idem; compare PMP **qalu-hipan* idem.
- (28) Sabah Malay *limpang* 'to lay down'; Brunei Malay *limpang* idem, Iban *limpang* 'go aside, turn aside, go past, go over', Mualang *limpang* 'turn off to another direction; change the subject'; Berau Malay *ba-lempang* ~ *ba-limpang* 'to lay down' (Kurniawati et al. 2002: 75); compare standard Malay *gə-limpang-an* 'sprawled everywhere', but *baring* 'to lay down'. The form also occurs as Mukah *ba-limpang* 'to lay down' and Murut *bə-limpang* idem (Nothofer 1987 Appendix).
- (29) Sabah Malay *limpas* 'to go past'; Brunei Malay *limpas*, Banjar *limpas* idem, Iban *limpas* 'overshoot, go past, go over'; compare standard Malay *limpas* 'after', but *lalu* 'to go past'.
- (30) Sabah Malay *mangkali* 'maybe'; Brunei Malay *mangkali*, Iban *mangkali* idem; compare Ambon Malay *mangkali*, Kupang Malay *mangkali* idem, but standard Malay *barangkali* idem.

- (31) Sabah Malay *palui* 'stupid'; Brunei Malay *palui*, Iban *palui* idem, but standard Malay *bodoh* idem; compare Kadazandusun *palui* 'naughty', PMP **paluy* 'taboo' (Adelaar 2005c: 367).
- (32) Sabah Malay *siring* 'side'; Brunei Malay *siring*, Banjar *siring*, Berau Malay *siring* idem, compare standard Malay *sə-iring* 'alongside', but *təpi* 'side'.
- (33) Sabah Malay *tapuk* 'to hide'; Brunei Malay *tapuk*, Berau Malay *ba-tapuk*, Sarawak Malay (*bə-*)*tapok*, Iban *tapok* idem (Collins 1991: 691-692); Mualang *tapuk*, Kedayatn *tapuk* idem, compare West Coast Bajau *tapuk* idem, but standard Malay *bər-səmbunyi* idem.
- (34) Sabah Malay *tutuk* 'to pound'; Berau Malay *tut:uk*, Sarawak Malay *nutok*, Iban *tutok* idem (Collins 1991: 691-692), Mualang *tutuk* 'to pound, peck'; compare PMP **tuktuk* 'to pound', but standard Malay *tumbuk* idem; Adelaar (1985: 157) reconstructs **tumbuk* ~ **tutuk* idem for Proto-Malayic.

Sabah Malay is further characterized by ongoing lexical influence from Bahasa Indonesia. Many Sabahans argue that their dialect is closer to Bahasa Indonesia than to Peninsular Malay, although such statements are fuelled by regional pride and not based on scientific research. Nevertheless, Sabah's great number of legal and illegal Indonesian labourers, geographical proximity to Kalimantan and porous borders help us explain why Sabah Malay exhibits several lexical items common in Indonesia but rarely used in Malaysia. Many of these Indonesian loanwords are also used by elderly speakers (Kartini Abd. Wahab, personal communication), suggesting that lexical borrowing from Bahasa Indonesia is not a completely recent phenomenon. I have encountered the following Indonesian loanwords in Sabah Malay:

(35) SM	Bahasa Indonesia		BM
amar	hamər (from Dutch)	'hammer'	tukul
alis mata	alis mata	'eyebrows'	kəning
bikin	bikin	'to make'	buat
bilang	bilang	'to say'	cakap
bisuk	besok	'tomorrow'	esok
ciwik	cewek	'girl'	awek
garubak	gərɔbak	'pushcart'	kəreta sorong
girigitan	grəgətan	'infuriated'	gəram
injak	injak	'to step on'	pijak
jam	jam	'o'clock'	pukul
jandila	jəndela	'window'	tingkap
juruk	jorok	'filthy'	cəmar
kamarin	kəmarin	'yesterday'	səmalam, kəlmarin
kasi	kasih	'to give'	bagi
lamari	ləmari	'closet'	almari
mau	mau	'to want'	nak
mangarti	məngərti	'to understand'	faham
pantat	pantat	'buttocks'	punggung
pariuk	pəriuk	'cooking pot'	kuali
piring	piring	'plate'	pinggan
salana	cəlana	'trousers'	səluar
sapi	sapi	'cow'	ləmbu

siang	siang	'afternoon'	təŋgahari
silaka	cəlakə	'misfortune'	malang

Since most speakers use Sabah Malay as their second or third language, they tend to borrow extensively from their native tongues. Some of these words have subsequently been picked up by speakers from other linguistic backgrounds. Examples of loanwords adopted into mainstream Sabah Malay are *bayau* 'buddy' from West Coast Bajau *bayau* 'brother-in-law', *pilak* 'immigrant from the Philippines' from Tausug *pilak* 'money' and *anu* 'so-and-so' used in the same meaning in several regional languages. Furthermore, speakers of Sabah Malay tend to use *hari satu*, *hari dua*, *hari tiga*, etcetera. 'day one, day two, day three' for the days of the week, which is also done in several regional languages, compare Kadazandusun *tadau koiso?*, *tadau koduwo*, *tadau kotolu*, etcetera. The standard Malay names of the days are derived from the Arabic numerals: *hari Isnin* 'Monday', *hari Səlasa* 'Tuesday', *hari Rabu* 'Wednesday', etcetera.

As Sabah's largest indigenous language, Kadazandusun has made various lexical contributions to Sabah Malay. Unfortunately, most of the Dusunic languages and dialects are poorly described, making it very difficult to pinpoint the exact origins of these loanwords. The list given in (36) contains words that according to consultants originate from Kadazandusun, although I have not been able to find most of them in the dictionaries available to me (Antonissen 1958; Lasimbang 2009). Many of the words listed carry a particularly negative semantic load, which might be the reason that Sabah Malay speakers felt the need to adopt them from another language. Compare the following examples:

(36)	balabak	'scrotum'
	bobot	'vagina'
	gutuk	'stupid'
	kogutan	'to have a hangover'
	lugai-lugai	'to walk around aimlessly'
	moginum	'to consume alcoholic drinks'
	mulau	'crazy'
	pusas	'snacks'
	tontolou	'penis'

Sabah Malay has also borrowed from Southern Chinese dialects, although most speakers do not realize that the relevant words are loanwords. "The words *kaw* (< Hokkien 厚 *kau*³²) 'thick', *nai* (< Cantonese 奶 *nai*²², Hakka *nai*¹¹) 'milk' and *ping* (< Hokkien and Cantonese 冰 *ping*¹¹) 'ice' have become accustomed in coffee stalls across Sabah, regardless of the ethnicity of the staff and customers. Thus, people will order *kupi kaw* 'strong (thick) coffee', *kupi nai* 'coffee with milk' or *kupi ping* 'coffee with ice cubes'." The Sabah Malay word for 'spoon' is *cikang* ~ *sikang* (< Cantonese 匙羹 *chi*¹² *kang*¹¹) instead of standard Malay *sudu*. Waitresses are addressed as *amoi*

(< Hakka 阿妹 *a¹ moi³* ‘younger sister’). Other Chinese loanwords are *cincai* ‘unorganized’ (< Hokkien 清采 *chin³ chai²* idem), *ngam* ‘accurate’ (< Cantonese 啱 *ngaam¹* idem), *ngam-ngam* ‘just (now)’ (< Cantonese 啱啱 *ngaam¹ ngaam¹* idem) and *siuk* ‘excellent’ (< Hokkien 惜 *siok⁴* ‘to cherish’). From English, Sabah Malay borrowed several lexical items rarely used in standard Malay,⁹ such as *aisbuk* ‘icebox’, *bagon* ‘Volkswagen’, *buk* ‘book’, *bul* ‘ball’, *gibap* ‘to give up’, *gohed* (< go ahead) ‘to go forward (vehicles)’, *gostan* ~ *gustan* (< go astern) ‘to go reverse (vehicles)’, *grabal* ‘gravel’, *karan* (< current) ‘electricity’, *insum* ‘handsome’, *kimbit* ‘camp bed’, *pemes* ‘famous’, *sikul* ‘school’ and *sot* (< short) ‘electrical malfunction (short circuit); crazy’.

4 MORPHOSYNTAX

This section examines several distinctive morphosyntactic features of Sabah Malay, giving especial attention to periphrastic causatives, head-final alternatives and discourse particles. Before elaborating on these topics, we will briefly examine the personal pronouns and the markers *suda* and *ada*. As is the case in many colloquial Malay varieties, the Sabah Malay personal pronouns are pluralized by the addition of *urang* ‘person, people’. Compare the paradigm in (37):

(37)	1s	aku
	1s _{polite}	sia
	2s	kau ~ ko
	2s _{polite}	kita
	3s	dia
	1pi	kita-urang
	1pe	sia-urang, kami-urang
	2p	kamurang (< kamu-urang)
	3p	dia-urang ~ durang

We will now move on to the use of the marker *suda*. In modern standard Malay, the equivalent *sudah* ‘already’ tends to be placed clause-initially. There are exceptions to this rule: Van Minde and Tjia (2002: 285-286) show examples from modern Malay literature where the particle “moved to the right”. The authors elaborate on this atypical post-clausal use and compare examples from the Ambon Malay vernacular, in which *suda* is frequently placed after the clause it modifies. Contributing to this discussion from a Sabah Malay perspective, we may note a preference for post-clausal use of *suda*. This is illustrated in examples (38) and (39):

(38)	Buli nampak suda	tu	bukit.
	can see	already	DEM hill
	‘You can see that hill already.’		

⁹ However, most of these English loanwords occur in Brunei Malay (Azam Othman, personal communication) and presumably also in other dialects.

- (39) Aku lapar suda bah ni.
 1s hungry already DP DEM
 'I'm already hungry.'

The marker *ada* functions as an existential verb 'there is'. In several Malay dialects, it also functions as an aspect marker. Adelaar (2005a: 213) demonstrates how *ada* marks progressive aspect in pidgin-derived Malay varieties (for which standard Malay uses the marker *səɗang*). In colloquial Peninsular Malay, the marker *ada* appears to function as a perfective aspect marker (for example, *ada jatuh* 'has fallen'), but this use has, to the best of my knowledge, not satisfactorily been described. In Sabah Malay, too, *ada* marks perfective aspect. This is illustrated in examples (40) to (43):

- (40) a. Ko ada nampak ka tu munyit?
 2s PST see DP DEM monkey
 'Have you seen that monkey?'
 b. Ada
 PST
 'I have.'
- (41) Siapa ada pigi Kuala Abai?
 Who PST go Kuala Abai
 'Who has gone to Kuala Abai?'
- (42) Aku ada pukul dia.
 1s PST hit 3s
 'I've beaten him up.'
- (43) Butul ka si Tom ada bilang bagitu?
 really DP ART Tom PST say like.that
 'Did Tom really say that?'

4.1 PERIPHRASTIC CAUSATIVES

Colloquial Peninsular Malay and pidgin-derived Malay varieties typically employ verb serialisation to express a causative construction (Koh 1990: 288 and further; Adelaar 2005a: 215-216), while standard Malay uses the verbal suffix *-kan* for this purpose. Sabah Malay displays a preference for verb serialisation, in which the auxiliary verbs *kasi* 'give', *bikin* 'make' and *kana* 'hit' are placed before the modified verb or adjective. This is illustrated in examples (44) to (48)

- (44) Kadang-kadang kalo satu urang dia kasi mati dulu?
 sometimes if one person 3s AUX dead first
 baru angkat.
 then lift
 'Sometimes when there's only one person he kills it [a big fish] first before pulling it [on board].'

- (45) Kalo ada urang mati? ramai-ramai kasi kubur dia.
if EXIST person dead RED-many AUX grave 3s
'If somebody dies there are many people to bury him/her.'
- (46) Itu untuk bikin takut-takut budak saja?.
DEM to AUX RED-scared kid just
'That's just to scare children.'
- (47) Bikin panas bah kan anjing ni.
AUX hot DP DP dog DEM
'This dog is so annoying.'
- (48) Tu anak nangis bah kana mara mama dia.
DEM child cry DP AUX angry mother 3s
'The child cried when his mother got angry at him.'

4.2 HEAD-FINAL ALTERNATIVES

In a pioneering study on colloquial Peninsular Malay, Koh (1990: 180) describes three atypical constructions not used in standard Malay:

- DEM-N: the position of the demonstrative before the noun or noun phrase. In standard Malay, the demonstrative (*ini, itu*) always follows the noun or noun phrase it modifies; placed before the noun or noun phrase it constitutes an independent segment.
- N-punya-N: the use of *punya* to indicate possession. In standard Malay, possession is expressed by a possessive enclitic or personal pronoun or noun (phrase) positioned directly after the possessed item.
- MOD-punya-N: the use of *punya* as a relative marker to connect a modifying phrase with a modified noun. In standard Malay, the relative marker *yang* is inserted between the modified noun and the modifying phrase (N-yang-MOD).

These head-final constructions are traditionally attributed to "low Malay" or "Bazaar Malay" – umbrella terms denoting the simplified Malay varieties used by second-language speakers – but Koh (1990: 180-181) also encounters them in colloquial Peninsular Malay used by native speakers. These constructions are prevalent in Sabah Malay and will be addressed one by one. To start with the position of the demonstratives, Sabah Malay displays a preference of DEM-N constructions:

- (49) Macamana mau kasi mati tu giuk?
how want AUX dead DEM worm
'How would you kill those worms?'
- (50) Sia nampak tu urang bikin pusas tu munyit.
1s see DEM man AUX snack DEM monkey
'I saw those people make a snack of that monkey [eat that monkey as a snack].'

- (51) Padi mata aku kana tu bawang.
 burn eye 1s AUX DEM onion
 'My eyes burn because of that onion.'

In some Malay varieties, such as the dialect of Southwest Maluku, noun-initial versus noun-final position of demonstratives clearly marks a grammatical difference (Van Engelenhoven 2008). This does not seem to be the case in Sabah Malay, in which fluctuation between DEM-N and N-DEM constructions appear to be accounted for by diglossia with standard Malay. My Sabah Malay corpus also contains examples of N-DEM and DEM-N-DEM constructions, given respectively in (52) and (53):

- (52) Burung tu nda? pandai tarabang.
 bird DEM NEG able fly
 'That bird cannot fly.'
- (53) Bukan gia pemes tu urang tu.
 NEG DP famous DEM person DEM
 'That person is not that famous.'

Possession in Sabah Malay can be expressed in two ways, both frequently used: 1) the possessor directly follows the possessed item or 2) the marker *punya* is inserted between the possessor and the possessed item. Constructions of the first type are reflected in examples (54) and (55):

- (54) Manggigit gia anjing ko tu?
 bite DP dog 2s DEM
 'Does that dog of yours bite?'
- (55) Aku nampak hari tu muka dia lawa bah.
 1s see day DEM face 3s pretty DP
 'I saw that her face looked pretty that day.'

Examples (56) to (58) illustrate the use of the possessive marker *punya*, which is inserted between the possessor and the possessed item:

- (56) Macam urang tu buli baca sia punya utak ...
 as.if person DEM can read 1s POS brain
 'As if that person could read my mind ...'
- (57) Ko punya bos cuti.
 2s POS boss vacation
 'Your boss is on vacation.'
- (58) Sabah punya makanan nyaman bah.
 Sabah POS food nice DP
 'Sabah's food is nice.'

The third head-final construction in Sabah Malay is of the MOD-punya-N type. As mentioned above, standard Malay uses the relative particle *yang* inserted between the modified noun and the subordinate clause to form a relative noun phrase. This construction is also used in Sabah Malay. In addition, Sabah Malay can form a relative noun phrase with the marker *punya*, which is inserted between the subordinate clause and the modified noun. This is illustrated in examples (59) to (61):

- (59) Ko ni paling bida punya muka.
 2s DEM most ugly POS face
 'You have the ugliest face.'
- (60) Palui punya urang datang sini bikin gadu.
 stupid POS person come here AUX fight
 'Stupid people, coming here to make trouble ...'
- (61) Apa punya urang tinggal sini rampuk urang punya tana?
 what POS person live here steal people POS land
 'What kind of person would live here and steal people's land?'

Not preceded by a modifying phrase, the marker *punya* is used to emphasize the adjective it modifies. This construction is not uncommon in other Malay varieties, such as Kupang Malay (Hein Steinhauer, personal communication). Compare the examples (62) and (63):

- (62) Punya bida pacar ko.
 POS ugly date 2s
 'You're date is so ugly.'
- (63) Uina, punya basar tu pukuk langsung!
 INT POS big DEM tree langsung
 'Wow, that langsung tree is huge!'

4.3 DISCOURSE PARTICLES

"Discourse particles" or "discourse markers" are small, morphologically invariable lexical items that do not bear obvious meaning. Their exact definition is still a matter of controversy, but a useful theoretical overview is provided in Jucker and Ziv (1998). Discourse particles are found cross-linguistically in informal speech and contain pragmatic information on the utterances they modify and the relation between the speech participants. Examples in English include 'you know', 'isn't it' and 'actually'. There is not much information on discourse particles in Malay varieties, although an increasing amount of literature is gradually improving this situation (Goddard 1994; Wouk 1998, 1999, 2001; Ewing 2005; Stoel 2005; Sneddon 2006).

The two most frequently used discourse particles in Sabah Malay are *gia* and *bah*. Their major functions are addressed in what follows. The particle *gia* probably originates from Kadazandusun, in which it conveys

slight remonstrance (Antonissen 1958: 39, 132). In Sabah Malay, the particle is only used by ethnic Kadazandusuns (Kartini Abd. Wahab, personal communication). In a question, *gia* is placed after the question word to convey emphasis and urge the speech participant(s) to answer what is being asked. Compare examples (64) to (66):

- (64) Kanapa gia ko ni?
 why DP 2S DEM
 'What are you on about?'
- (65) Sapa gia nama ko punya kawan?
 who DP name 2S POS friend
 'What's the name of your friend again?'
- (66) Macamana gia mo buka tu air?
 how DP want open DEM water
 'How do you make the water come out?'

Interestingly, this usage of *gia* corresponds to *kan* in colloquial Brunei Malay (but not in standard Malay), in which the above examples would be (63) "Kenapa kan ko ani?", (64) "Siapa kan nama kawan mu atu?" and (65) "Macamana kan buka aing ani?" (Azam Othman, personal communication). Moving on to the function of *gia* in an imperative or negative imperative structure, the particle is used to invite the listener to do something. This can be seen in examples (67) to (69):

- (67) Cuba gia bali dulu.
 try DP buy first
 'Please buy it first.'
- (68) Ko saja gia ni!
 2s just DP DEM
 'You go on then!'
- (69) Kalau dia tidur jangan gia kasi guyang.
 if 3s sleep VET DP AUX move
 'If he sleeps, don't move him.'

In this function, the particle *gia* is equivalent to *tah* in colloquial Brunei Malay, in which the above examples would be (66) "Cuba tah bali dulu.", (67) "Kau tah saja!" and (68) "Kalau ia tidur jangan tah diguyang." (Azam Othman, personal communication). In a declarative utterance, the particle *gia* conveys emphasis on the segment it follows, urging the listener to believe that what is being said is true. This is illustrated in examples (70) to (72):

- (70) Lama lagi gia tu kalo mo tunggu.
 long more DP DEM if want wait
 'It will take very long to wait.'

- (71) Lawa gia bini ko!
 pretty DP wife 2s
 'Your wife is so pretty!'
- (72) Ada juga gia yang nda? suka.
 EXIST also DP REL NEG like
 'Also, there are those who don't like it.'

Having explored the major functions of *gia*, we now move on to the particle *bah*. This particle is found in several of Borneo's regional languages, including West Coast Bajau (Miller 2006: 2 note 5), Kadazandusun (Antonissen 1958: 39), Timugon Murut (Nathesan 1993: 455), Belangin (Adelaar 2006: 78), Bisaya and Lun Bawang (Ozóg and Martin 1996: 249 note 4). It is also used in Brunei Malay and Bruneian English (Ozóg and Martin 1996). Although the use and functions of *bah* in Brunei Malay have not sufficiently been studied to allow a detailed comparison, the particle seems to occur more often and in more different ways in Sabah Malay, in which it is "over-used" in the ears of Bruneians (Azam Othman, personal communication). Indeed, in Sabah one is likely to overhear entire telephone conversations filled with *bah*. As opposed to *gia*, the particle *bah* can constitute intonational segments by itself. It is used as an agreement marker, a politeness marker, an emphatic marker, a question marker, an exhortative marker, an interruptive marker and a way to end a conversation. All of these functions will be discussed in what follows. Foremost, *bah* marks agreement with other speech participants, as illustrated in examples (73) and (74):

- (73) Bah, jumpa lagi bisuk!
 DP meet again tomorrow
 'All right, we'll see each other tomorrow.'
- (74) a. Jangan kasi kluar tu anjing
 VET AUX go.out DEM dog
 'Don't let the dogs escape.'
- b. Bah!
 DP
 'Okay!'

The particle also functions as a politeness marker, softening the assertive force of an utterance and making it less abrupt. In this function, it often urges other speech participants to show solidarity with what has been said. This is illustrated in examples (75) to (78):

- (75) Ko pigi mana bah?
 2s go where DP
 'Where are you going, then?'
- (76) Kasian bah urang Indun di Sabah.
 too.bad DP person Indonesia LOC Sabah
 'Those poor Indonesians in Sabah ...'

- (77) Sia nda? mo cari gadu butul-butul bah.
 1S NEG want search fight RED-really DP
 'I wasn't looking for an actual fight, you know.'
- (78) Siaurang nda? ada bah.
 1P NEG EXIST DP
 'We don't happen to have any.'

The particle *bah* also functions as an emphatic marker, adding mild stress to the phrase it follows. This use is partly comparable to standard Malay *lah* and Sabah Malay *gia*,¹⁰ although the latter is predominantly used by people with a Kadazandusun-speaking background, whereas *bah* is used by all Sabahans. The emphatic function of *bah* is illustrated in examples (79) to (83):

- (79) Iya bah, buli bah kalo ko!
 AFF DP can DP if 2s
 'Sure, if it's you it's no problem!'
- (80) Sabab tu sia suka Sabah bah.
 reason DEM 1s like Sabah DP
 'That's why I really like Sabah.'
- (81) Tambirang bah durang ni ging.
 punch DP 3P DEM gang
 'They are lying, guys.'
- (82) Mimang palui bah ko ni.
 really stupid DP 2s DEM
 'You're really stupid.'
- (83) Bida bah kalo buyuk, bikin panas ja? ni!
 ugly DP if cheat AUX hot only DEM
 'Cheating is bad, it only makes people angry!'

As has been demonstrated for *gia*, the discourse particle *bah* can be added to a question to urge a response. This is illustrated in examples (84) to (86):

- (84) Butul bah?
 really DP
 'Is it really?'
- (85) Ko tu mo ikut mangkali bah?
 2s DEM want join maybe DP
 'Maybe you would like to join?'

¹⁰ See Goddard (1994) on the use of *lah* in Malay. This particle is known and used in Sabah Malay, as can be seen in example (86), but does not occur as frequently as in Peninsular Malay varieties.

- (86) Nam blas la bah, buli la bah?
 sixteen DP DP can DP DP
 'Sixteen please, would that be okay?'

The particle is also used as an exhortation, inviting or encouraging the other speech participants to do something. In this function, it can be placed before and/or after the phrase it modifies or constitute a separate segment. This is illustrated in examples (87) to (89):

- (87) Bah, sini ko bah kalo brani.
 DP here 2s DP if brave
 'Okay, come here if you dare.'
- (88) Datang bah lagi di sini.
 Come DP again in here
 'Please visit again.'
- (89) Bah!
 DP
 'Please!/Yes!'

As an independent segment positioned utterance-initially, *bah* is used to attract the attention of the other speech participants and to interrupt a conversation in order to provide comment on it or introduce a new topic. This is illustrated in examples (90) and (91):

- (90) a. Ui, basar-nya pantat dia.
 INT big-DEF bottom 3s
 'Wow, her bottom looks so big.'
- b. Bah, apalagi pakai calana pendek.
 DP moreover wear trousers short
 'Yeah, especially when she's wearing short pants.'
- (91) Bah, ko bagus-bagus sarita ...
 DP 2s RED-beautiful tell
 'Okay, you tell it beautifully ...'

The last function of the discourse particle *bah* is to indicate that the speaker intends to end a conversation. Neglecting to use *bah* when doing so is considered impolite. This is illustrated in examples (92) and (93):

- (92) Bah, jalan dulu bah.
 DP walk first DP
 'Well, I'm off then.'
- (93) a. Bah, pigi la? ko dulu.
 DP go DP 2s first
 'Now then, off you go.'
- b. Bah!
 DP
 'Bye!/Okay! [to indicate leave-taking]'

Given their partly overlapping functions, people from a Kadazandusun-speaking background tend to use the discourse particles *bah* and *gia* in a supplementary way, in which the former precedes the latter. This is illustrated in example (94) and (95):

(94) Jangan bah gia.
 VET DP DP
 'Don't be like that.'

(95) Dari mana bah gia ko ni?
 from where DP DP 2s DEM
 'Where on earth did you come from?'

5 DISCUSSION

This study has highlighted several distinctive features that set Sabah Malay apart from other Malay varieties. Based on the presented lexical data, Sabah Malay appears to be closely related to Brunei Malay and other dialects spoken in Borneo. It is also characterized by lexical influence from Sabah's indigenous languages, Bahasa Malaysia, Bahasa Indonesia, southern Chinese dialects and English, making it a unique variety in its own right. The historical development of Sabah Malay opens up questions about the existing sociolinguistic classifications of Malay varieties. During the colonial era, the distinction between "low Malay" or "bazaar Malay" versus "high Malay" prevailed in the Dutch East Indies and the Malay Peninsula. The former roughly denoted the parlance of non-Malays, whereas the latter referred to literary Malay and "pure" Malay speech. A nuance in this high-low distinction was introduced by Hassan (1969), who distinguishes bazaar Malay, literary Malay and colloquial Malay (*Bahasa Melayu Lisan*). The latter refers to vernacular, non-literary varieties spoken among Malays. An essentially similar distinction is made by Adelaar and Prentice (1996) and Adelaar (2005a), who use the terms 1) literary Malay, 2) trade-Malay or pidgin-derived Malay and 3) vernacular Malay.

This study on Sabah Malay offers several new perspectives on Malay dialectology in a more general context. Usually, the addition of *orang* to plural personal pronouns, the use of *ada* as an aspect marker and the occurrence of periphrastic causatives and head-final alternatives are attributed to pidgin-derived Malay varieties, such as Ambon Malay, Sri Lanka Malay and Cocos Malay. The prevalence of this "package" of features across the pidgin-derived Malay varieties is strongly suggestive of a historically widespread "trade Malay". If this hypothesis is true, such a trade vernacular almost certainly underwent extensive influence from southern Chinese dialects.¹¹ In a next stage, several pidgin-derived Malay varieties developed a vast core of mother

¹¹ In a similar way, Chinese merchants were of key significance in shaping the idiom of "Bazaar Malay" spoken in the Malay Peninsula. It is well documented that most of the distinctive features of that variety are directly adopted from Hokkien (Shellabear 1913; Lim 1981; Pakir 1986; Adelaar and Prentice 1996: 674-675).

tongue speakers and started to behave sociolinguistically as vernaculars instead of *lingua franca* (Adelaar 2005a:204). While the interrelatedness and shared history of Baba Malay, Sri Lanka Malay and eastern Indonesian Malay varieties remains slightly controversial, the development of Sabah Malay is less obscure and somewhat different. It seems that upon its introduction in Sabah, an offshoot from the Brunei Malay vernacular underwent regional influences and developed into an interethnic means of communication with “pidgin features”. A similar development seems to have taken place in Bacan Malay (Collins 1983b) and comparative data on that variety would be welcome for future research. We are also in need of more data on Malay and “Malayic Dayak” varieties spoken in Borneo, so that we may passionately pursue our attempts at charting the borders of Malay dialectology on this island.

REFERENCES

- Abdul Djebar Hapip. 1977. *Kamus Banjar-Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Adelaar, Alexander K. 1985. *Proto-Malayic; The reconstruction of its phonology and parts of its lexicon and morphology*. PhD thesis, Leiden University.
- Adelaar, Alexander K. 1996. “Endangered Malay isolects; The case of Salako, Cocos Malay, and Sri Lanka Malay”. Paper presented at *International Workshop on South-East Asian Studies: Endangered languages and literatures of South-East Asia*, Leiden, 9-13 December.
- Adelaar, Alexander K. 2005a. “Structural diversity in the Malayic subgroup”, in: Alexander Adelaar and Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, pp. 202-226. London and New York: Routledge.
- Adelaar, Alexander K. 2005b. *Salako or Bademeà; Sketch grammar, texts and lexicon of a Kanayatn dialect in West Borneo*. Wiesbaden: Harrassowitz Verlag.
- Adelaar, Alexander K. 2005c. “Malayo-Sumbawan”, *Oceanic Linguistics* 44/2: 357-388.
- Adelaar, Alexander K. 2006. “Where does Belangin belong?”, in: Fritz Schulze and Holger Warnk (eds), *Insular Southeast Asia; Linguistic and cultural studies in honour of Bernd Nothofer*, pp. 65-84. Wiesbaden: Harrassowitz Verlag.
- Adelaar, Alexander K. and D.J. Prentice. 1996. “Malay: its history, role and spread”, in: S.A. Wurm, P. Mühlhäusler, and D. Tryon (eds), *Atlas of languages of intercultural communication in the Pacific, Asia and the Americas*, pp. 673-693. Berlin and New York: Mouton de Gruyter.
- Antonissen, A. 1958. *Kadazan-English and English-Kadazan dictionary*. Canberra: Government Printing Office.
- Benjamin, Geoffrey. 1976. “Austroasiatic subgroupings and prehistory in the Malay Peninsula”, *Oceanic Linguistics Special Publications* 13: 37-128.
- Blust, Robert. 2006. “The linguistic macrohistory of the Philippines; Some speculations”. Paper presented at the 10th International Conference on Austronesian Linguistics (10-ICAL), Puerto Princesa City, Palawan, Philippines, 17-20 January.

- Broek, Jan O.M. 1962. "Place names in 16th and 17th century Borneo", *Imago Mundi* 16: 129-148.
- Chuchu, Haji Jaludin bin Haji. 2003. *Dialek Melayu Brunei dalam salasilah Bahasa Melayu Purba*. Bangi: Penerbit Universiti Kebangsaan Malaysia.
- Chuchu, Haji Jaludin bin Haji. 2005. *Ragam-ragam dialek Melayu Brunei*. Brunei Darussalam: Dewan Bahasa dan Pustaka.
- Chuchu, Haji Jaludin bin Haji. 2008. "Berau – Brunei: Jalinan hubungan utuh", in: Katharina E. Sukamto (ed.), *Kelana Bahana sang Bahasawan; Persembahan untuk Prof. Soenjono Dardjowidjojo, Ph.D. dalam rangka ulang tahunnya yang ke-70. Bagian 1*, pp. 139-160. Jakarta: Universitas Katolik Atma Jaya.
- Collins, James T. 1983a. *Dialek Ulu Terengganu*. Kuala Lumpur: Penerbit Universiti Kebangsaan Malaysia.
- Collins, James T. 1983b. "Penggolongan bahasa Bacan", *Nusantara* 10: 86-125.
- Collins, James T. 1987. *Dialek Melayu Sarawak*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Collins, James T. 1989. "Malay dialect research in Malaysia: the issue of perspective", *Bijdragen tot de Taal-, Land- en Volkenkunde* (BKI) 145/2-3: 235-264.
- Collins, James T. 1990. *Bibliografi dialek Melayu di Pulau Borneo*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- Collins, James T. 1991. "Rangkaian dialek Melayu di Pulau Borneo", *Jurnal Dewan Bahasa* 35/8: 687-696.
- Engelenhoven, Aone van. 2008. "Ini apel ni nya 'This here apple now' ; Deictics in the Malay speech of Southwest Malukan migrants in the Netherlands", *Wacana, Jurnal Ilmu Pengetahuan Budaya* Vol. 10 No. 1 (April): 90-116.
- Ewing, Michael C. 2005. "Colloquial Indonesian", in: Alexander Adelaar and Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, pp. 227-258. London and New York: Routledge.
- Goddard, Cliff. 1994. "The meaning of lah: Understanding 'emphasis' in Malay (Bahasa Melayu)", *Oceanic Linguistics* 33/1: 145-165.
- Hassan, Abdullah. 1969. "Bahasa Melayu Pasar di-Malaysia Barat", *Dewan Bahasa* XIII: 207-218.
- Haynes, H.S. 1900. "A list of Brunei-Malay words", *Journal of the Straits Branch of the Royal Asiatic Society* 34: 39-48.
- Helfrich, O.L. 1904. *Bijdragen tot de kennis van het Midden Maleisch (Běšěmahsch en Sěrawajsch dialect)*. Batavia: Landsdrukkerij. [Verhandelingen van het Bataviaasch Genootschap van Kunsten en Wetenschappen 53.]
- Hudson, Alfred B. 1967. *The Barito isolects of Borneo; A classification based on comparative reconstruction and lexicostatistics*. Ithaca, New York: Cornell University.
- Jucker, A. and Y. Ziv (eds). 1998. *Discourse markers; Description and theory*. Amsterdam and Philadelphia: Benjamins.
- Junaiyah, H.M. 1985. *Kamus Lampung-Indonesia*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.

- Kamus. 2006. *Kamus Kedayan-Melayu, Melayu-Kedayan*. Brunei Darussalam: Dewan Bahasa dan Pustaka.
- Kawi, Djantera. 2002. *Bahasa Banjar; Dialek dan subdialeknya*. PhD thesis, University of Indonesia.
- King, Julie K. and John Wayne King. 1983. *Languages of Sabah; A survey report*. Canberra: The Australian National University.
- Koh, Ann Sweesun. 1990. *Topics in colloquial Malay*. PhD thesis, University of Melbourne.
- Kurniawati, Wati, Non Martis, Buha Arintonang, and Hidayatul Astar. 2002. *Kosakata dasar Swadesh di Kabupaten Berau, Kotamadya Samarinda, dan Kotamadya Balikpapan*. Jakarta: Pusat Bahasa.
- Lasimbang, Rita. 2009. *English-Malay-Kadazandusun beginner's dictionary*. Kota Kinabalu: Kadazandusun Language Foundation.
- Lim, Sonny. 1981. *Baba Malay; The language of the 'Straits-born' Chinese*. Melbourne: Monash University.
- Mahdi, Waruno. 2005. "Old Malay", in: Alexander Adelaar and Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, pp. 182-201. London and New York: Routledge.
- Maxwell, Allen R., 1981-82. "The origin of the name 'Sabah'", *Sabah Society Journal* VII/2: 91-105.
- McDonnell, Bradley. 2008. "A conservative vowel phoneme inventory of Sumatra; The case of Besemah", *Oceanic Linguistics* 47/2: 409-432.
- Miller, Mark. 2006. *Learn to speak West Coast Bajau (Kota Belud dialect); A series of twenty-five language learning lessons*. Kota Kinabalu: Jabatan Muzium Sabah.
- Minde, D. van and J. Tjia. 2002. "Between perfect and perfective; The meaning and function of Ambonese Malay su and suda", *Bijdragen tot de Taal-, Land- en Volkenkunde (BKI)* 158/2: 283-303.
- Moody, David C. 1984. "The Malayic language family", in: Julie K. King and John Wayne King (eds), *Languages of Sabah, a survey report*, pp. 91-100. Canberra: Pacific Linguistics. [Pacific Linguistics C-78.]
- Nathesan, S. 1993. "Bahasa Melayu dialek Sabah; Satu tinjauan terhadap kelainannya", *Jurnal Dewan Bahasa* 37/5: 447-456.
- Nothofer, Bernd. 1987. "Bahasa-bahasa Austronesia di Brunei Darussalam". Paper presented at Seminar Antarabangsa; Bahasa, kesusasteraan dan kebudayaan Melayu, Bandar Seri Begawan, 3-8 August.
- Ozóg, A. Conrad K. and Peter W. Martin. 1996. "The bah particle in Brunei English", in: Peter W. Martin, A. Conrad K. Ozóg and Gloria Poedjosoedarmo (eds), *Language use and language change in Brunei Darussalam*, pp. 236-249. Athens, Ohio: Ohio University Centre for International Studies.
- Pakir, Anne Geok-In Sim. 1986. *A linguistic investigation of Baba Malay*. PhD thesis, University of Hawai'i, Honolulu.
- Prentice, D.J. 1969. "A wordlist for use in Borneo", in: D.J. Prentice, *Papers in Borneo linguistics*, pp. 1-8. Canberra: The Australian National University. [Pacific Linguistics A-20.]

- Shellabear, W.G. 1913. "Baba Malay; An introduction to the language of the Straits-born Chinese", *Journal of the Straits Branch of the Royal Asiatic Society* 65: 49-63.
- Sin, Chong. 2009. "Dialek-dialek Melayu di Lembah Baram", *Sari; International Journal of the Malay World and Civilisation* 27/2: 59-71.
- Sneddon, James Neil. 2006. *Colloquial Jakartan Indonesian*. Canberra: Australian National University. [Pacific Linguistics 581.]
- Stoel, Ruben. 2005. *Focus in Manado Malay; Grammar, particles, and intonation*. Leiden: CNWS Publications.
- Suryadikara, Fudiat, Durdje Durasid, Djantera Kawi, and Abdurrahman Ismail. 1979. *Bahasa Kutai*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Sutlive, Vinson and Joanne Sutlive. 1994. *A handy reference dictionary of Iban and English*. Kuching: Tun Jugah Foundation.
- Thurgood, Graham. 1999. *From ancient Cham to modern dialects; Two thousand years of language contact and change*. Honolulu: University of Hawai'i Press.
- Tjia, Johnny. 2007. *A grammar of Mualang; An Ibanic Language of Western Kalimantan, Indonesia*. PhD thesis, Leiden University.
- Wong, Jane Kon Ling. 2000. *The Sabah Malay dialect; Phonological structures and social functions*. Kota Kinabalu: University Malaysia Sabah.
- Wouk, Fay. 1998. "Solidarity in Indonesian conversation; The discourse marker *kan*", *Multilingua* 17: 381-408.
- Wouk, Fay. 1999. "Gender and the use of pragmatic particles in Indonesian", *Journal of Sociolinguistics* 3: 194-200.
- Wouk, Fay. 2001. "Solidarity in Indonesian conversation; The discourse marker *ya*", *Journal of Pragmatics* 33: 171-191.
- Zainal Abidin Bakar. 1993. "Pengaruh dialek tempatan dalam penggunaan Bahasa Melayu baku", *Jurnal Dewan Bahasa* 37/12: 1128-1135.