

A GRAMMAR OF BABA MALAY
WITH SOCIOPHONETIC CONSIDERATIONS

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF
THE UNIVERSITY OF HAWAII AT MĀNOA IN PARTIAL FULFILMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

LINGUISTICS

DECEMBER 2014

By

Nala Huiying Lee

Dissertation Committee:

Lyle Campbell, Chairperson
Andrea Berez
Robert Blust
Katie Drager
Barbara Watson Andaya

Copyright © 2014
by
Nala Huiying Lee

Dedicated to the memory of
Lee Tai San alias Yap Kim Seng,
my Peranakan grandfather
and
Hung Kok Cheong,
my Hockchia grandfather

ACKNOWLEDGMENTS

I had looked forward to many things. For example, being able to say, yes there is a doctor on board the plane and could I give you an aspirin? (Just kidding). Chief among which, I had always looked forward to writing this page, knowing that it would mark a particular milestone in my dissertation writing. Yet, this has been the most difficult page I have ever had to write in my thus-far short academic career, not because there is no one to thank, but because I am afraid I might have left you out, and if I do thank you here, there is a high chance that even the most profuse expression of gratitude falls short. Do bear with me as I attempt to do justice to all the kindness I have received.

This dissertation was made possible, and due in no small part, to the encouragement and support of many people. My gratitude goes out to my key language consultants who have been nothing but generous with their time and language expertise. Thank you Baba Peter Wee, president of the Peranakan Association Singapore and proprietor of Katong Antique House, for the exceptional company and for encouraging my love of books and all things old; Baba Victor Goh, cultural advisor to the Gunong Sayang Association, for being ever so patient and for sharing with me your passion for the Peranakan culture; my dearest *mamah* ‘grandmother’ Nyonya Jane Quek, for all the chuckles and for spoiling me with good old home-cooked Peranakan *lauk-lauk* ‘dishes’; Baba Albert Ku and wife, Aunty Lydia, for welcoming me to Malacca and for treating me like your own; and my own Grand Uncle Chan, for inspiring me with your genteel nature and for sharing with me Lee family stories from *dulu-dulu* ‘long long ago’. I am also thankful to all other members of the Associations who have extended their generosity to me in some way or other.

At the University of Hawai‘i, I am indebted to my dissertation committee. I wish to thank my chairperson, Lyle Campbell, for encouraging me to pursue a grammar, for constantly being my voice of reason when all I wanted were the moon and the stars (and at times Jupiter and Mars), for having a big heart and for lending me the perfect space to work in, for his superhuman response time, and for making my dissertation stronger just by asking the right questions; Andrea Berez for sharing her expertise and enthusiasm on matters of language documentation, for encouraging me to explore different technologies to keep my dissertation organized and on-track,

and for her good humour and well-timed wit; Robert Blust for being one of the first people to welcome me to Hawai‘i, for making the Austronesian world such an interesting one to be in, for combing through my references in every submission with a fine-toothed comb, so that my bibliographies became models of exception, and for encouraging me to challenge myself every step of the way; Katie Drager for introducing me to sociophonetics, for being patient with me when statistics got me (not the other way around), and for her warmth and energy all around; and Barbara Watson Andaya for her kindness, for coming on board so readily as an external member, and for taking a genuine interest in my work.

I am also grateful to classmates and friends at the University. Thank you Katie Gao (née Butler) for the late nights at Sinclair Library, for making sure I had enough food and coffee to get through the weeks of my comprehensive exams, for all the times it got tough, and for all the good ones too; Sunju Kim for equally good company, for the evenings at the coffeeshop, and for teaching me random Korean phrases; Carolina Aragon for keeping me upbeat, for making me *doce de leite*, and for being my constant gym buddy; John Van Way for being an all-round excellent colleague and co-writer, for creating the Language Endangerment Index with me, and for keeping my competitive edge sharp; Tamami Harada for her selfless friendship, her generosity, and for keeping me sane over the weekends. I also wish to thank members of the sociolinguistics interest group, In-group, and Da Pidgin Coup for making me feel at home with my work in Hawai‘i. Special thanks goes to Kent Sakoda for all the talk story time and for feeding me candy and *zabon*. I am also grateful to Jennifer Kanda and Nora Lum. I am forever in awe at their efficiency and adeptness, and how much they know everything there is to know about what form there is to fill in, or when I have to turn it what.

At the National University of Singapore, I am grateful to current head of department, John Richardson, and my previous head of department, Lionel Wee for all their encouragement in my academic career; to my mentor, Bao Zhiming for always finding time for me and for all his sagely advice; to Mie Hiramoto, the elder sister I never had, for all the times I panicked and needed a listening ear, and for showing me that people who live with the same philosophies as I do can come out just fine. At the Nanyang Technological University, I owe much to past colleagues, especially Ng Bee Chin and Tan Ying Ying. They were among the first people to stand by me.

To Mama, thank you for sharing your love of different cultures and your curiosity of the world with me. To Papa, thank you for years of ferrying my friends and me around in your car, for tirelessly standing by me in your quiet ways. To Juan, if not for you, I don't think I would have ever been able to come do this and feel good about it. Thank you for being my amazing sister. To Shawn, thanks for the laughs and welcome to the family. To my in-laws, you are probably the best ones there are out there. Thanks for supporting the kind of lifestyle that I lead and the work that I do. To Iswen, my partner in life, this journey would have been so different without you, and we did it, husband. I cannot wait to explore the world with you. Woof to Ruffy and Snowy, who can't read this, without whom this dissertation might have been completed some weeks earlier.

This research was supported by funding provided by the National University of Singapore Overseas Graduate Scholarship, the University of Hawai'i Department of Linguistics Graduate Assistantship, Achievement Scholarship, and Endowment Fund, as well as the University of Hawai'i Graduate Students Organization Award.

ABSTRACT

Baba Malay, the home language of Peranakans, was formed via early intermarriage between Hokkien-speaking male traders and Malay-speaking indigenous women in the Malay Peninsula. The language is endangered, with less than 1,000 speakers in Singapore, and less than 1,000 speakers in Malacca, Malaysia. This dissertation describes the language's sociohistorical background, its current circumstances of endangerment, and provides information regarding the phonology, parts of speech, and syntax of Baba Malay as it is spoken in Singapore. The language has 19 consonants and 8 vowels, of which [ɛ] occurs only in a refined style of speaking. Acoustic investigation of the vowel system shows that that [ɛ] is falling out of use, especially among less proficient speakers. A matched guise task is conducted to show the associations that listeners form between this changing variant and the speakers who produce this variant. Results show that younger listeners in particular perceived forms with [ɛ] as being more emblematic of the Peranakan culture and community than corresponding coarse speech forms that do not use [ɛ]. This is consequential for language change. Results from a post-matched guise survey also indicate that Peranakans are very concerned about language loss. Beyond this extended phonological investigation, the language's basic clausal word order is Subject Verb Object, where Subject is optional. Topicalization also occurs frequently in the language. The basic phrasal word orders in Baba Malay are Noun Adjective, Genitive Noun, and Preposition Noun. Relative clauses occur prenominal and postnominal. Other than presenting a traditional description of Singapore Baba Malay, this grammar also highlights differences between Singapore Baba Malay and Malacca Baba Malay, and addresses whether Baba Malay is a genetic dialect of Malay, a mixed language, or a creole. Comparison between Singapore Baba Malay and Malacca Baba Malay shows that Malacca Baba Malay is more influenced by standard Malay, particularly where lexicon is concerned. Investigation into the issue of classification shows that BM should be classified as a creole. This dissertation includes vocabulary and texts; the audio files associated with these texts are archived at Kaipuleohone.

TABLE OF CONTENTS

Acknowledgments	iv
Abstract.....	vii
Table of Contents	viii
List of Tables	xv
List of Figures	xviii
1. Introduction to Baba Malay, its speakers and this dissertation	1
1.1 The formation of Baba Malay	1
1.2 The Peranakans	3
1.2.1 Geographical location	9
1.2.2 Dialects and language environments.....	10
1.2.3 Speaker numbers.....	13
1.2.4 Language endangerment index	14
1.3. Language documentation	20
1.3.1 Community and participants.....	23
1.3.2 The scope of this dissertation	25
2. An overview of Baba Malay typology and this grammar	35
2.1 An overview of Baba Malay typology	35
2.1.1 Sound system.....	35
2.1.2 Basic morphology	36
2.1.3 Basic syntax.....	37
2.2 A user guide to this grammar	41
2.2.1 Features	41
2.2.2 Structure of the grammar.....	42
2.2.3 List of linguistic abbreviations used	42
2.2.4 Example sentences	44
2.2.5 Transcription.....	45
3. Phonetics and phonology	47
3.1 Phoneme inventory	47
3.1.1 Consonants	47
3.1.1.1 Minimal and near-minimal pairs of consonants.....	48
3.1.1.2 Bilabials	51

3.1.1.3 Alveolars.....	54
3.1.1.4 Post-alveolars.....	59
3.1.1.5 Velars.....	62
3.1.1.6 Glottals	64
3.1.1.7 Glides.....	66
3.1.2 Vowels.....	67
3.1.2.1 Minimal pairs of vowels	68
3.1.2.2 Monophthongs	70
3.1.2.3 Diphthongs.....	76
3.1.2.4 Vowel sequences across syllables	77
3.2 Phonotactics.....	78
3.2.1 Resyllabification and reduplication.....	80
3.3 Phonological rules.....	83
3.3.1 Syllable-final velar plosive to glottal stop rule.....	88
3.3.2 Deletion of word-initial <i>h</i> rule	88
3.3.3 Metathesis of <i>ə</i> and <i>r</i> rule	89
3.3.4 Assimilation to <i>u</i> before <i>lw</i> rule.....	90
3.3.5 Vowel-raising to <i>e</i> and final <i>l, r, s</i> loss in refined style rules.....	90
3.3.6 Rule ordering	92
3.3.6.1 Vowel-raising, then loss of final <i>l, r, and s</i> rule	92
3.3.6.2 <i>u</i> -assimilation, then vowel-raising, followed by loss of final <i>l, r, and s</i> rule	92
3.4 Fast speech	93
3.5 Writing system.....	95
3.6 Stress and intonation.....	100
3.6.1 Word stress and pitch	100
3.6.2 Tones on Hokkien-derived words	106
3.6.3 Sentence intonation	112
3.7 Variation and change	117
3.7.1 Sociophonetic investigation.....	125
3.7.1.1 Stimuli	127
3.7.1.2 Procedure	131
3.7.1.3 Participants.....	134
3.7.2 Results	136
3.7.2.1 Is this person a speaker of BM?	137

3.7.2.2 Is this person likely to take part in activities at the Peranakan Association or Gunong Sayang Association?	141
3.7.2.3 Is this person well-educated?	144
3.7.3 Implications of sociophonetics experiment	150
4. Parts of speech	153
4.1 Nouns	154
4.1.1 Nominal morphology	155
4.1.1.1 Derivational noun morphology	156
4.1.1.2 Nominal compounds	157
4.1.2 Noun classifiers and partitives	159
4.1.3 Pronouns	161
4.1.3.1 Personal pronouns	161
4.1.3.2 Reflexives	163
4.1.3.3 Reciprocals	165
4.1.3.4 Interrogative pronouns	166
4.2 Verbs	168
4.2.1 Verbal morphology	168
4.2.1.1 Derivational verb morphology	169
4.2.1.2 Non-productive verb morphology	172
4.2.1.3 Verbal compounds	176
4.3 Adjectives	178
4.3.1 Adjectival morphology	179
4.4 Adverbs	179
4.4.1 Adverbial morphology	180
4.4.2 Interrogative adverbs	182
4.5 Determiners	183
4.5.1 Person marker	183
4.5.2 Demonstratives	184
4.5.3 Numerals	185
4.5.3.1 Other numerical expressions	187
4.5.4 Quantifiers	188
4.6 Prepositions	189
4.7 Conjunctions	196
4.8 Discourse elements	198
4.8.1 Interjections	198

4.8.2 Particles	200
5. Syntax	203
5.1 Noun phrases	203
5.1.1 Genitive	205
5.1.2 Noun phrases with determiners.....	207
5.1.2.1 Noun phrases with demonstratives	208
5.1.2.2 Noun phrases with person marker	210
5.1.2.3 Noun phrases with numerals, noun classifiers and partitives.....	211
5.1.2.4 Noun phrases with quantifiers	213
5.1.3 Noun phrases with adjectival modifiers	214
5.1.4 Negation of noun phrases	216
5.1.5 Order of elements in noun phrases.....	217
5.2 Verb phrases	217
5.2.1 Copula constructions	219
5.2.2 Modality	220
5.2.3 Passivization	223
5.2.4 Ditransitive, causative and benefactive constructions.....	226
5.2.4.1 Ditransitive constructions	226
5.2.4.2 Causative and benefactive constructions	227
5.2.5 Tense and aspect	229
5.2.5.1 Future tense.....	230
5.2.5.2 Perfective aspect.....	232
5.2.5.3 Recent perfect aspect.....	233
5.2.5.4 Progressive aspect	233
5.2.5.5 Habitual aspect.....	234
5.2.5.6 Experiential perfect aspect.....	235
5.2.5.7 Tentative aspect.....	235
5.2.6 Serial verb constructions	236
5.2.7 Verb phrases with adverbial modifiers.....	238
5.2.8 Negation of verb phrases	240
5.2.9 Order of elements in verb phrases.....	242
5.3 Adjectival phrases.....	242
5.3.1 Comparatives	244
5.3.2 Comparison of equality	245
5.3.3 Comparison of similarity.....	246

5.3.4 Superlatives	247
5.3.5 Excessive degree.....	247
5.3.6 Adjectival phrases with adverbial modifiers	248
5.3.7 Order of elements in adjectival phrases	249
5.4 Adverbial phrases	250
5.4.1 Temporal adverbial phrases.....	251
5.4.2 Location adverbial phrases	251
5.4.3 Manner adverbial phrases.....	252
5.4.4 Order of elements in adjectival phrases	253
5.5 Summary of word order at the phrase level.....	253
5.6 Clauses	255
5.6.1 Word order at the clause level	255
5.6.2 Grammatical relations and alignment	256
5.6.3 Relative clauses	258
5.6.4 Complement clauses	262
5.6.5 Direct and indirect speech	263
5.6.6 Adverbial clauses	264
5.6.6.1 Temporal adverbial clauses.....	265
5.6.6.2 Location adverbial clauses.....	267
5.6.6.3 Manner adverbial clauses.....	268
5.6.7 Conditionals.....	268
5.6.8 Conjunctions.....	269
5.6.8.1 Coordinating conjunctions	269
5.6.8.2 Subordinating conjunctions	272
5.6.9 Topicalization	274
5.6.10 Questions	276
5.6.10.1 Content questions	276
5.6.10.2 Tag questions	283
5.6.11 Imperatives	288
6. Differences between Singapore Baba Malay and Malacca Baba Malay	290
6.1 Phonetic and phonological differences	290
6.1.1 Vowel /ɔ/	290
6.1.2 /e/, /o/ versus /aj/, /aw/	295
6.2 Morphological difference.....	297
6.2.1 <i>Ke-</i> -an nominalizing circumfix	297

6.2.2 Other affixes	298
6.3 Syntactic differences.....	301
6.3.1 Noun phrases	301
6.3.1.1 Demonstrative determiners	301
6.3.1.2 Personal pronouns	303
6.3.1.3 Relative clauses	305
6.3.2 Verb phrase.....	306
6.3.2.1 Progressive aspect	306
6.3.2.2 Perfective aspect.....	307
6.3.3 Adjectival and adverbial phrases	308
6.3.3.1 <i>Daripada</i> comparatives	309
6.3.3.2 <i>Dengan</i> ‘with’ adverbial phrase	309
6.3.4 Conjunctions.....	310
6.3.4.1 <i>Dan/sama</i> ‘and’ coordinating conjunction.....	310
6.3.4.2 <i>Atau/ka</i> ‘or’ coordinating conjunction.....	311
7. What type of language is BM?	313
7.1 Classification	314
7.2 BM as a “dialect” of Malay	316
7.2.1 Thurgood (1998)	317
7.2.2 Tan (1980)	320
7.2.3 Pakir (1986)	321
7.3 BM as a mixed language	323
7.3.1 Sociohistorical factors	328
7.3.2 Structural factors.....	331
7.4 BM as a creole	343
7.4.1 Lim (1991, 1988)	345
7.4.2 Ansaldo and Matthews (1999) and Ansaldo et al. (2007)	353
7.4.3 Shih (2009)	357
7.4.4 Creole theories and BM.....	359
7.4.4.1 The universalists.....	361
7.4.4.2 The substratists.....	365
7.4.4.3 The superstratists.....	375
7.5 Conclusion.....	379
References.....	382

Appendix A: Word lists	396
Swadesh 100 word list	396
Kinship terms	399
Expressions for day, month and time	403
Appendix B: Texts	407
Transcription conventions for texts	407
Pear story with Peter Wee	408
Turnip story with Victor	427
Conversation with Aunty Jane	437
Pantuns by Albert Ku	465
Appendix C: Vocabulary	470
Singapore Baba Malay-English	470
English-Singapore Baba Malay	490
Appendix D: Sociophonetics	510
Post-listening questionnaire	510
Responses to fillers	512

LIST OF TABLES

Table 1: Language Endangerment Index- Intergenerational transmission	15
Table 2: Language Endangerment Index- Absolute number of speakers	16
Table 3: Language Endangerment Index- Speaker numbers	17
Table 4: Language Endangerment Index- Domains of use	18
Table 5: Language Endangerment Index and Level of Certainty Index	20
Table 6: Personal pronouns in Baba Malay, Malay and Hokkien (Lim 1988: 27).....	30
Table 7: Consonant chart of Singapore Baba Malay	48
Table 8: Bilabial consonants by position	51
Table 9: Alveolar consonants by position.....	55
Table 10: Post-alveolar consonants by position	60
Table 11: Velar consonants by position.....	62
Table 12: Glottal consonants by position	64
Table 13: Glides by position	66
Table 14: Vowel chart of Singapore Baba Malay	68
Table 15: Vowels and formant values for a proficient speaker of Singapore Baba Malay.....	72
Table 16: Distinctive features of vowels and glides in Singapore Baba Malay.....	85
Table 17: Distinctive features of consonants in Singapore Baba Malay (not including glides)	86
Table 18: Writing system of Baba Malay used by this grammar	96
Table 19: List of matched guise tokens	128
Table 20: List of Baba Malay fillers.....	130
Table 21: List of clearly non-Baba Malay fillers	130
Table 22: Demographics of perception experiment participants.....	135
Table 23: Coefficients of fixed effects; larger negative coefficients indicate a lower likelihood of responding ‘yes’ to the question “Is this person a speaker of BM?” when listening to a COARSE guise	139
Table 24: Coefficients of fixed effects; larger negative coefficients indicate a lower likelihood of responding ‘yes’ to the question “Is this person likely to take part in activities at the Peranakan Association or Gunong Sayang Association?”	143
Table 25: Coefficients of fixed effects; larger negative coefficients indicate a lower likelihood of responding ‘yes’ to the question “Is this person well-educated?”	148
Table 26: List of some nominal compounds in Baba Malay	157

Table 27: Noun classifiers in Baba Malay.....	159
Table 28: Mass classifiers in Baba Malay	160
Table 29: Personal pronouns in Baba Malay	162
Table 30: Interrogative pronouns in Baba Malay.....	167
Table 31: List of some verbal compounds in Baba Malay	176
Table 32: Interrogative adverbs in Baba Malay	182
Table 33: Numerals derived from Malay.....	186
Table 34: Numerals derived from Hokkien	187
Table 35: List of prepositions in Baba Malay.....	190
Table 36: List of conjunctions in Baba Malay	197
Table 37: List of common interjections in Baba Malay	199
Table 38: List of commonly used adverbs that modify verb phrases and their distribution	239
Table 39: List of phrases, their word orders and general tendencies	254
Table 40: List of interrogative pronouns and adverbs in Baba Malay	277
Table 41: Vowel chart of Malacca Baba Malay.....	293
Table 42: Consonant chart of Malacca Baba Malay.....	294
Table 43: Personal pronouns in Baba Malay (Table 29)	305
Table 44: List of first 20 words on Swadesh basic word list	332
Table 45: Key roles in a Peranakan wedding.....	334
Table 46: Personal pronouns in Baba Malay, Hokkien, and Malay	336
Table 47: Tense and aspect of Baba Malay, Hokkien, and Malay	339
Table 48: Relativization patterns in Baba Malay, Hokkien, and Malay	342
Table 49: Consonant charts of Hokkien and Malay (Lim 1988).....	346
Table 50: Consonant chart of Malacca Baba Malay (Lim 1988)	348
Table 51: Vowel charts of Hokkien and Malay (Lim 1988).....	348
Table 52: Vowel chart of Malacca Baba Malay (Lim 1988)	349
Table 53: Relativization patterns in Baba Malay, Hokkien, and Malay (Table 48).....	373
Table 54: Swadesh 100 word list in Baba Malay.....	396
Table 55: Kinship terms in Baba Malay	399
Table 56: Days of the week in Baba Malay.....	403

Table 57: Months in Baba Malay	404
Table 58: Special occasions on the lunar calendar in Baba Malay	405
Table 59: General time expression in Baba Malay.....	406
Table 60: Specific time in Baba Malay	406
Table 61: Transcription conventions for texts.....	407

LIST OF FIGURES

Figure 1: Waveform and spectrogram of /p/ in [panas] ‘hot’	53
Figure 2: Waveform and spectrogram of /b/ in [bakol] ‘basket’	53
Figure 3: Waveform and spectrogram of /m/ in [sama] ‘same, and, with’	54
Figure 4: Waveform and spectrogram of /t/ in [tareʔ] ‘pull’	56
Figure 5: Waveform and spectrogram of /d/ in [datan] ‘come’	56
Figure 6: Waveform and spectrogram of /s/ in [sama] ‘same, and, with’	57
Figure 7: Waveform and spectrogram of /n/ in [mana] ‘where, which’	57
Figure 8: Waveform and spectrogram of /l/ in [labilabi] ‘tortoise’	58
Figure 9: Waveform and spectrogram of /l/ in [bakol] ‘basket’	58
Figure 10: Waveform and spectrogram of /r/ in [garan] ‘fierce’	59
Figure 11: Waveform and spectrogram of /tʃ/ in [tʃarita] ‘story’	60
Figure 12: Waveform and spectrogram of /dʒ/ in [dʒaga] ‘guard, take care’	61
Figure 13: Waveform and spectrogram of /ŋ/ in [manaʔ] ‘many’	61
Figure 14: Waveform and spectrogram of /k/ in [kaki] ‘leg’	62
Figure 15: Waveform and spectrogram of /g/ in [garan] ‘fierce’	63
Figure 16: Waveform and spectrogram of [ŋ] in [garan] ‘fierce’	63
Figure 17: Waveform and spectrogram of /h/ in [habis] ‘finish’	65
Figure 18: Waveform and spectrogram of /ʔ/ in [budaʔ] ‘child’	65
Figure 19: Waveform and spectrogram of /w/ in [wajan] ‘play (performance)’	66
Figure 20: Waveform and spectrogram of /j/ in [wajan] ‘play (performance)’	67
Figure 21: Waveform and spectrogram of /i/ in [abi] ‘then’	71
Figure 22: Vowel space of a Singapore Baba Malay speaker	76
Figure 23: Vowel space of a more proficient Singapore Baba Malay speaker (Figure 22)	119
Figure 24: Vowel space of a less proficient Singapore Baba Malay speaker	120
Figure 25: Normalized vowel plots of six Singapore Baba Malay speakers	121
Figure 26: Responses to “Is this person a speaker of Baba Malay?” with participant’s age group as a predictor	138
Figure 27: Yes responses to “Is this person likely to take part in activities at the Peranakan Association or Gunong Sayang Association?” with participant’s age group as a predictor	142

Figure 28: Yes responses to “Is this person well-educated?” with participant’s age group as a predictor	145
Figure 29: Yes responses to “Is this person well-educated?” with participant’s origin as a predictor	146
Figure 30: Yes responses to “Is this person well-educated?” with participant’s cultural association membership status as a predictor	147
Figure 31: Tense and aspect system of Baba Malay	230
Figure 32: Vowel space of a Malacca Baba Malay speaker	291
Figure 33: Vowel space of a Singapore Baba Malay speaker (Figure 22).....	292
Figure 34: Responses to fillers for the question ‘Is this person speaking BM?’	512
Figure 35: Responses to fillers for the question “Does this person have many close friends who are not Peranakan?’	512
Figure 36: Responses to fillers for the question ‘Is this person likely to take part in activities at the Peranakan Association or the Gunong Sayang Association?.....	513
Figure 37: Responses to fillers for the question ‘Is this person well-educated?’	513

1. Introduction to Baba Malay, its speakers and this dissertation

The purpose of this dissertation is to provide a grammar of Baba Malay as spoken in Singapore. Other key components to documenting this highly endangered language include a sociophonetic focus on a language change in progress, observations of the differences between the varieties of Baba Malay that are spoken in Singapore and Malacca, as well an investigation of the type of language that Baba Malay is, whether it is a genetically related “dialect” of Malay, a mixed language, or a creole. For details of these goals, please see section 1.3.2.

1.1 The formation of Baba Malay

Baba Malay (ISO 639-3: mbf), also known as *Peranakan* to its speakers, is a contact language with two main component languages. While the term *Peranakan* itself is an endonym, *Baba Malay* is the name that researchers have given to the language, *baba* referring to the males of the community (see section 1.2). Most of the words in Baba Malay are of Malay (ISO 639-3: zlm) origin, while it is disputed whether the grammar is derived from Hokkien, also known as Fukien, Fujian, or Southern Min (ISO 639-3: nan) (Holm 1989; Lee 2009; Shih 2009) or if Baba Malay is a dialect of Malay (Pakir 1986; Thurgood 1998).¹ Regardless, it is generally agreed that Baba Malay was formed via Chinese-Malay intermarriage.

Based on Chinese imperial records, Chinese contact with the Malay Peninsula is said to have begun around the 3rd century AD, continuing into the 5th AD (Thurgood 1998). Chinese traders en route to India would have had to travel via the Malay Archipelago. The ports in the Malacca Strait area were geographically advantaged, being at the receiving end of both northeast and southwest monsoons. The northeast monsoon aided vessels in their journey from China to the Malacca Strait, where they awaited changes in the wind direction to continue to their journeys or to return home (Andaya and Andaya 2001). The development of Chinese settlements southwards

¹ While some have stated that Baba Malay was influenced by colloquial Malay rather than by the standard variety (Tan 1993; Ansaldo and Matthews 1999), others have stated that the language was more likely influenced by more standard Malay (Lee 2009). The Hokkien dialect that most probably contributed to the formation of Baba Malay is that of Zhangzhou and Quanzhou, where most Singapore and Malacca Hokkiens originally came from. More discussion on this topic ensues in Chapter 7: What type of language is Baba Malay?

from China began as early as the 15th century AD—at least dates have been documented for the travels of the Chinese admiral Zheng He, who led trading exchanges in the South China Sea, Java Sea and Malacca² Strait between 1405 to 1433 (Wade 1994; Wang 1964; Widodo 2002/2003). In particular, Chinese settlements were built in Malacca, and it became a highly favored trading post of the Chinese. At that time, Malacca was a kingdom in its own right (Windstedt 1948), and it notably became the first nation to receive an official inscription on a tablet from a Chinese emperor of the Ming dynasty, Yong Le (Wang 1964; Widodo 2002/2003).

During his voyaging years, the Chinese admiral Zheng He made a total of five trips to Malacca. After Zheng He's death in 1433, a scholar on his ship by the name of Fei Hsin wrote in *Hsing-ch'a sheng-lan*³ 'The Overall Survey of the Star Raft' that besides dark-skinned people, he had also seen fairer-looking people of Chinese descent in Malacca (Fei 1436). A century later, in 1537, a Chinese traveler by the name of Hwang Chung wrote in his travel journal, *Hai yu* 'News from the Ocean', that the Chinese in Malacca ate pork, lived in hotels, and had female slaves who served them food and drinks (Groeneveldt 1876). These narratives complement the account that many of the Chinese men who had come to trade at the new port in Malacca did not return to China. Most Chinese immigrants were men, as it was rare for women to make these voyages out of China (Skinner 1996). Eventually, these Hokkien-speaking men who mostly originated from the Zhangzhou and Quanzhou⁴ regions in the Fujian province on the south-east coast of China, married indigenous Malay-speaking spouses.⁵ Some were said to have married slaves of Batak, Balinese, and Javanese origin (Purcell 1980). By 1750, it was reported by the then Governor of Malacca, Balthasar Bort that the Chinese population in Malacca had grown to 2161, a fifth of the total population (Purcell 1980), and that the number of China-born Chinese only contributed

² In Malay, Malacca is Melaka orthographically.

³ The Wade-Giles system of romanization was used by the translator of this edition, J.V.G. Mills.

⁴ With the exception of referring to the romanized forms used by other authors who have used Romanization systems such as the Wades-Giles, the Hanyu Pinyin system is used for representing Mandarin in this dissertation.

⁵ I hesitate to use the term 'indigenous Malay', which is a highly contested term. For example, according to Article 160 of the Constitution of Malaysia, which has been in effect since the country's independence day in 1957, a Malay is a person who is Muslim, speaks the Malay language, conforms to Malay customs, and was before the time of independence born in Malaysia or in Singapore, or born of at least one parent who was born in Malaysia or in Singapore, or was living in Malaysia or Singapore at the time of independence, or who is a descendent of such a person (The Commissioner of Law Revision 2010). This is in effect, different from identifying Malay-ness as an ethnicity. See Barnard (2004) for further explications on the complexities in defining the Malay identity.

minimally to this figure (Skinner 1996). Descendants of this community are referred to as ‘Straits-born Chinese’, or *Peranakan*, and it is their language, *Baba Malay*, that this dissertation is concerned with.

1.2 The Peranakans

It is important to note that while Chinese-Malay intermarriages still currently occur, these intermarriages result in neither Peranakan ethnicity nor culture. Besides having developed their own home language that they refer to as a *patois*, the Peranakans also created their own unique culture, which is a blend of Chinese, local, and Western customs and traditions. Mostly Chinese, the Peranakan culture is also influenced by Malay and Western practices (Wee 2013).

Traditionally, they follow Chinese religious practices, such as praying to *ti gong* ‘sky god’, and the *datok dapor* ‘deity of the stove’, in addition to ancestral worship⁶. Malay influence on the other hand, is seen in how the Peranakan language resembles Malay, hence its name *Baba Malay* (henceforth ‘BM’). Where dress is concerned, the Peranakan men traditionally wore the *baju lokchuan*, a Chinese attire made of silk, while the females wore the *baju panjang*, a long blouse, over a *sarong*, reminiscent of what Malay and other local women would wear. Eventually, the men began to prefer English suits due to colonial influence, and the women favored the *sarong kebaya*⁷, the *kebaya* being a short jacket made of European material such as voile and often elaborately embroidered by hand around the edges with Chinese motifs, such as phoenixes and dragons. The *kebaya* is fastened in the front by three brooches linked together by a common chain. This set of brooches, unique to the culture, is called the *kerosang*. Women also wore beaded shoes called the *kasot manek*. While most Peranakans no longer wear traditional outfits on a daily basis, the women do wear the *kebaya* for special occasions, such as formal dinners, while the men have begun wearing batik shirts, similar to their counterparts in Southeast Asia. The picture on the following page was taken in the 1920s, and it shows a Peranakan woman

⁶ The Peranakan Chinese were never Muslims, even though the group was formed via Chinese intermarriage with Malay speakers.

⁷ In the past, the *kebaya* was known as the *baju nona*, literally the ‘clothes (of a) Nyonya’. This term is no longer commonly used by Peranakans today (Peter Wee, p.c.).

dressed in a *baju panjang* fastened together with a set of *kerosang*, worn over a *sarong*, and a Peranakan man dressed in an English suit.



Photograph of a Peranakan couple from the 1920s.⁸

⁸ The subjects in the picture are the researcher's great-grandmother and step-great-grandfather. The picture was taken at a now-defunct photography studio, Wong Fong, which used to be at 809 North Bridge Road, Singapore.

Where cooking is concerned, Peranakan cooking is a fusion of Chinese, local, and Western ingredients. The *iték tim* ‘duck soup’ for example, comprises Chinese preserved vegetables, the Southeast Asian tamarind, and Western brandy, among other ingredients (Wee 2013). Western influence brought about particularly by British colonialism cannot be underestimated. With the British settlement of Penang, Malacca, and Singapore in the nineteenth century, Peranakans were learning English (thus differentiating themselves from other Chinese migrants), and identifying themselves as the King’s Chinese (Hardwick 2008). While wealthy families sent their children to schools in England, Peranakan men were also employed as middlemen who mediated between the migrants and the English colonial administrators. Eventually, a group of Peranakans set up the Straits Chinese British Association in 1900, the name of the association reflecting the Peranakans’ ties with the British administrators. Post- World War II, the British completely removed themselves from Malaya in 1963, and only then was the association renamed the Singapore Chinese Peranakan Association. It was later renamed the Peranakan Association Singapore in 1966, and the association’s name has since stayed the same (Wee 2013).

Peranakan has many denotations. The word *Peranakan* is composed of a person prefix, *per-*, the morpheme for child, *anak*, and a nominalizer, *-an*. Both affixes are of Malay origin. While the nominalizer *-an* is a productive suffix in BM, the person prefix *per-* is not used to create new words in the language. As a single word, *Peranakan* denotes ‘descendants’ and also ‘a mother’s womb’⁹. In the Malay Peninsula, *Peranakan* can be used to indicate a person born of a Malay mother and a foreign father, and it is also commonly taken to mean ‘locally-born’, as with Peranakan Yahudi (locally-born Jews), and Peranakan Jawi (locally-born Malay-speaking Muslims of non-Malay and Malay ancestry, such as Arab and South Asian).^{10, 11} While the term

⁹ It is also not implausible that *per-* was actually derived from middle voice prefix *ber-*, which may combine with *anak* ‘child’ and nominalizer *-an* to indicate ‘womb’.

¹⁰ It is curious that Peranakan Jawi and Peranakan Yahudi are most likely exonyms, even though Peranakan is an endonym to the Peranakan Chinese, who use it to refer to their community and the language that they speak. On a separate note, there is only a handful of Peranakan Jawi and Peranakan Yahudi left in Singapore and Malaysia.

¹¹ The Peranakan Jawi have to be distinguished from a separate group known as Chitty. The Chitty are locally-born Tamils; some of whom have mixed ancestries. Traditionally Hindu, they share many practices with the Peranakan Chinese, such as their manner of dress (see page 3). The Chitty are said to speak Baba Malay, but with Tamil loanwords—this requires further investigation.

Peranakan Cina¹² (locally-born Chinese) was more popular in the past, the Peranakan Chinese are now commonly referred to simply as *Peranakan* by the speakers themselves and by people outside the community. Pakir (1986:23) postulates that *Peranakan*, which is supposedly a more generic term by nature, is used to refer to the Peranakan Chinese due to the significantly larger numbers of this particular community.

The males in the Peranakan community are referred to as *baba*, while the females are referred to as *nyonya*. The word *baba* is of unclear etymology. Pakir (1986:23) mentions JD Vaughan (1879)'s explanation of the term's origin, that "the term *baba* was used by the natives of Bengal to designate the children of Europeans and was probably introduced by the Indian convicts at Penang to describe Chinese children", but postulates that the word could have been borrowed from Turkish, in which it means 'father' (1986:25). In a different vein, Khoo (1996) postulates that *baba* is a loanword borrowed from Persian by Malay speakers. Used as an honorific solely for grandparents, the term was supposedly brought to the Malay Peninsula by Hindustani speakers such as vendors and traders. It then became part of the vernacular. In light of this disparity in etymology, a simpler, more direct explanation of the term *baba* can be proposed. The term could have come from Hokkien, which the Chinese traders would originally have spoken. In Hokkien, one addresses one's father as [pàpǎ] with an unaspirated [p], a low falling tone on the first syllable, and a rising tone on the second syllable. The children of the Chinese-Malay intermarriages may have used the Hokkien term while addressing their father, which is entirely plausible since most of the kinship terms are derived from Hokkien (See Pakir 1986, Shih 2009, and Appendix A: Kinship terms). The term [papá] evolved into [baba], as unaspirated [p] is aurally close to [b], and Peranakans call their fathers [bapa?]. Some Malacca Peranakans call their fathers [baba].

Nyonya, which refers to a female Peranakan, appears to have a slightly clearer etymology. Pakir (1986) suggests that the word derives from the Indonesian word *nona*, which means 'lady'. On a related but different note, Thurgood (1998:90) states that *nyonya* is a Javanese term, corresponding to "English *Mrs.*, German *Frau*, or Polish *Pani*". To uncover the actual

¹² The conventional orthography for this term is from Malay, where *cina* is pronounced as [tʃina]

etymology, it is necessary to look at the common influences that underlie both Indonesian and Javanese forms. Pakir (1986:25) questions if *nyonya* has a Dutch or Portuguese origin. While there was considerable Dutch influence in the region due to that fact that Malacca and parts of Indonesia (in particular Java) formed part of the Dutch East Indies administration from the 17th century to the early 19th century, there is no similar form for *nyonya* in Dutch that means ‘lady’, ‘female’, or an equivalent. In addition, it is said that Malacca under the Dutch rule attracted few Chinese immigrants, “so that the emergent society had a long period of incubation with relatively little incorporation of new Chinese “blood”” (Skinner 1996:57). It is likely that the Peranakan community would have been stable for a period of time, and that essential terms in the language would already have been developed prior to this. On the other hand, Portuguese has a comparable form, *dona*. The dates of the first Chinese settlement in Malacca and the Portuguese colonization of the place also agree with this account. Portugal had colonized Malacca between 1511 and 1641, arriving a century after Chinese traders had likely begun to settle in it. Within a century, the Chinese-Malay intermarriages that led to the development of the Peranakan culture and identity would have taken place, and it would have been possible for the Peranakans to borrow the word for ‘female’ from the then-dominant Portuguese administrators. Remarkably, the word *nona* from that era is still preserved in the region through a popular Papia Kristang (an endangered Portuguese-based creole spoken in Malacca) song called *Jingli Nona* (‘Dancing Girl’), demonstrating how salient the word *dona* or *nona* could have been in Malacca, so much that the Peranakans would have come to use *nyonya* to mean ‘female Peranakan’.

Other essential terms associated with the Peranakan identity are *peranakan jati*, *peranakan chelop*, *embok-embok*, *embok jantan*, and *bibik*. Both *peranakan jati* and *peranakan chelop* relate to lineage. *Jati* means ‘teak’ in Malay, and a *peranakan jati* is one who is considered to be true-blue Peranakan, or one whose parents are both Peranakan. *Chelop* means ‘to dip in dye’ in Malay¹³, and a *peranakan chelop* is ironically, no longer “pure”, since their ancestors or parents have married outside the community. These concepts are paradoxical, considering that the Peranakan community itself arose out of intermarriages across ethnic lines. However, in the

¹³ Written *celup* in Malay orthography, but *chelop* in the writing system used by this grammar. See section 3.5.

years after the formation of the Peranakan identity, marriages within the community were preferred (Clammer 1980, Tan 1979) and the Peranakans differentiated themselves from the new Chinese migrants or *sinkék*. The word *sinkék* is derived from Hokkien to mean ‘newcomers’. These marriages within the community became rare after the Japanese invasion of the Malay Archipelago (1942-1945), and consequently, there are not many Peranakans left who can claim to be *peranakan jati*.

The term *embok-embok* is close in nature to that of *peranakan jati*. *Embok-embok* in Peranakan refers to elderly Peranakan women who are highly traditional, and fastidious in maintaining customs. Pakir (1986:23) cites Horne (1974), pointing out that *embok-embok* could have originated from Javanese, in which [əmbəʔ] would mean ‘mother’, or be a ‘term of address for an older woman, especially of the servant class’. Pakir (1986:23) also cites Ikranagara (1980) who states that Melayu Betawi, which is spoken in Jakarta, uses [mboʔ- mboʔ] to mean ‘older woman’. As before, it is important to ask if this word can be traced further back to an earlier existence in Hokkien, since the Chinese from Fujian were once dominant in the region. There exists in Hokkien a word that refers to ‘grandaunt’—[pǒ], with a rising tone. This word would appear as [ŋpǒ] in the Zhangzhou variety of Hokkien, Zhangzhou being where most of these Chinese immigrants originated from. [ŋ] is a vocative particle in the Zhangzhou variety of Hokkien, used commonly in kinship terms (Freedman 1979). The meaning of [ŋpǒ], which denotes grandaunt, and connotes a much older lady, is semantically congruous with those of the Peranakan, Javanese, and Melayu Betawi terms. In Peranakan, it is interesting to note that *embok-embok* can also mean ‘highly traditional Peranakan elders’ collectively. However, the term is never used to refer to elderly Peranakan men. This is perhaps due to the fact that the original terms from which *embok-embok* was coined refers to ‘grandmother’ or ‘elderly woman’ in Hokkien. The male counterpart of the female *embok-embok* is the *embok-jantan*, wherein *jantan* means ‘male’ in Malay.¹⁴

¹⁴ Whereas *jantan* ‘male’ is used when referring to animals, it can be used for human beings in BM. Interestingly, the term for female animals in Malay, *betina*, is not used to refer to human beings. Instead, *perompuan* ‘female, woman’ is used.

The last term that indicates Peranakan identity is *bibik*. This refers to ‘a mature Peranakan woman’. It is said that this word may have been derived from Hindustani, and has a similar meaning of ‘aunt’ in Indonesian (Pakir 1986:25). There is no direct male equivalent to *bibik*. At the time of writing, the terms *embok-embok* and *embok-jantan* are falling out of use with members of the community. Only the term *bibik* remains highly used. This may be due to the fact that *embok-embok* refers to not just an elderly Peranakan person, but also to a very traditional person who still follows the Peranakan customs and traditions. By definition then, there are only a handful of Peranakans who can be considered to be *embok-embok*, and the word is not often used.

1.2.1 Geographical location

Although the Peranakans first developed their identity, language, and culture in Malacca (2.2000° N, 102.2510° E), many later moved to Penang (5.4000° N, 100.2333° E) and Singapore (1.3667° N, 103.7500° E). While Penang is 475 km (295 miles) north of Malacca, Singapore is 232 km (144 miles) south of Malacca. Penang is about 1,048km², Malacca is about 1,664 km², and Singapore is about 710 km². Both Penang and Malacca are states in current-day Malaysia, and Singapore is a nation state on its own. The local names for Penang and Malacca in Malay are *Pulau Pinang* ‘Island Penang’ and *Melaka*. Penang and Malacca are positioned on the west coast of the Malay Peninsula, while Singapore occupies the southern-most tip of this peninsula. All three locations run along the Malacca Strait, which was made prosperous by trading ships that plied these waters. Being the main channel that links the Indian Ocean and the Pacific Ocean, the Malacca Strait is still a vital shipping channel today.

Malacca flourished as a trading port under the Portuguese rule between 1511 and 1641. In contrast, it was not as successful under the Dutch rule between 1641 and 1825. The Dutch preferred to use Batavia, present-day Jakarta, as their main center of economic activities, and had only occupied Malacca to prevent other European powers from occupying it (De Witt 2008). Thus, a considerable number of Peranakans moved to Penang after the British annexed it in 1786, for Penang was thriving as a trading port at the expense of Malacca. In a letter dated 1

February, 1787, founder of Penang, Captain Francis Light, wrote that “[d]id not the Dutch keep a strict watch over the Chinese, most of them would leave Malacca” (Purcell 1967: 244).

Following that, a sizeable number of Peranakans also moved to Singapore at the beginning of the 19th century, shortly after the founding of Singapore by Sir Thomas Stamford Raffles in 1819 (Skinner 1996).

In addition to the Peranakans of Malacca, Penang, and Singapore, there is also a Peranakan population in Java, Indonesia (7.5028° S, 111.2631° E). This population is not directly related to the community that first developed in Malacca. It is said that even though Chinese immigrants arrived in Java prior to the 18th century, they had mostly assimilated to indigenous society. This Peranakan community, with its own practises and language, only developed in Java during the 18th century (Skinner 1996). The language of this community is also different from BM. It is called Peranakan Indonesian (ISO 639-3: pea), and is based on Indonesian (ISO 639-3: ind) and Javanese (ISO 639-3: jav), with Mandarin elements as compared to BM, which instead has Hokkien elements (Lewis et al. 2013).

1.2.2 Dialects and language environments

There are two varieties of BM spoken today. The Peranakans from Singapore speak a different variety from the Peranakans from Malacca. Interestingly, while it has been noted that there is a population of Peranakans in Penang who originated from Malacca, this group of Peranakans does not speak BM. Instead, they speak Penang Hokkien. Skinner (1996) postulates three reasons for this difference. First, he states that there was already a stable settlement of married Hokkien speakers in Penang preceding the arrival of the Peranakans from Malacca, whereas there was no such settlement in Singapore. Next, the trading patterns of Penang and Singapore differed.

As Skinner (1996:58) reports it:

Penang served as entrepot for Medan in northeastern Sumatra and for Phuket in southern Thailand, and in both of those ports Hokkien speakers predominated among Chinese traders, whereas Singapore was the entrepot for ports along the north coast of Java, where Peranakans predominated.

Lastly, he states that there was considerable intermarriage between the Peranakans of Penang with Hokkien families in Sumatra and southern Thailand, whereas Singapore Peranakans intermarried with Peranakan families in Java. Some of these explanations appear to be more plausible than others. While it may be less accurate to suppose that the Singapore Peranakans maintained BM due to their links with Java Peranakans, as the Java Peranakans speak a different language than the Singapore Peranakans, it appears that the Penang Peranakans assimilated to Hokkien due to their strong ties with Hokkien speakers, and that they already had models in the form of other Hokkien families they could emulate. It is important to note that although Peranakan culture is a mixture of Chinese, indigenous, and Western ways, it is still predominantly Chinese. For example, regardless of whether Peranakans speak Penang Hokkien or BM, they maintain Chinese kinship terms (See Tan 2001 for kinship terms in Penang Hokkien and Appendix A for kinship terms in BM).

Malacca and Singapore Baba Malay differ mainly because of the other languages that they came in contact with. In Malacca, Malay is the language of administration and mainstream education, and it is spoken by the majority (Tan 2001). To a lesser extent, English, Hokkien, and Cantonese are spoken. It is interesting to note that whereas Cantonese is more widely spoken in other Malaysian states such as the capital state of Kuala Lumpur, Hokkien is more extensively used in Malacca and Penang (Chen 2003). English is also spoken by the younger generation. It is becoming more prominent in education, as it is increasingly used in private schools as a medium of education (David and Govindasamy 2005). In addition to this, Malaysia has its own colloquial English variety, more commonly known as “Manglish” (Lee 1998). Malay is, however, still the most dominant language in Malaysia.

In Singapore, English is the majority language. Used for administrative purposes and as the main medium of education, English is designated as an official language together with Malay,

Mandarin, and Tamil. Previously, since the period of British settlement, most schools taught in one of the four official languages, but this changed in 1966 when the bilingual language education policy was implemented. The language policy for education in Singapore categorizes students by broad ethnic groups, and assigns students from each ethnic group a “mother-tongue” that they have to learn as an academic subject, in addition to English. Pakir (1994) calls this “English-knowing bilingualism”. With this rule, those who are classified to be “Chinese”, regardless of whether they are Hokkien, Teochew, or Cantonese, are required to learn Mandarin. “Indians” are to learn Tamil, and “Malays” are to learn the Malay language. In addition, Malay has an additional status of being the national language, and it is the language of Singapore’s national anthem and pledge. However, it is not widely used outside the Malay community, its status as a national language notwithstanding. Other languages spoken in Singapore with no official status include Colloquial Singapore English (commonly known as “Singlish”), Bazaar Malay (a pidgin used in the market place by older Singaporeans), and other Chinese varieties such as Hokkien, Teochew, Cantonese, and Hakka. The 2010 Singapore census of households shows that English and Mandarin were most frequently spoken at home. Among the Chinese, Hokkien was the most commonly spoken Chinese “dialect”, while Malay was preferred at home among the Malays (Singapore 2011).

The two varieties of BM are influenced by their different environments. It appears at first glance, that Singapore Baba Malay (SBM) has more Hokkien lexical items than Malacca Baba Malay (MBM), which seems more Malay. While *huahi* can be used to express the notion of ‘happy’ in SBM, MBM users will use *gembira* or *suka hati* and not *huahi*. The words *gembira* ‘happy’ and *suka hati*, literally meaning ‘like heart’, are derived from Malay, while their counterpart *huahi* is derived from Hokkien. Similarly, other words such as like *riyang*, meaning ‘lively’, are used in MBM, but are deemed by SBM speakers as being ‘too Malay’. SBM speakers use the Hokkien-derived term *laujiet* to mean ‘lively’ instead. Syntactically, the two varieties also differ from one another. It is possible to construct phrases with both sequences of ‘Noun Determiner’, and ‘Determiner Noun’ in SBM, whereas the sequence ‘Determiner Noun’ is not usually found in current-day MBM. In Malay, only the sequence ‘Noun Determiner’ is permitted, while in Hokkien, the opposite is true.

There has been little research done on the differences between SBM and MBM. While Pakir (1986) and Lee (1999) explicitly state that their fieldwork was conducted in Singapore, and Sharif (1981) and Lim (1981, 1988) state that their fieldwork took place in Malacca, the two varieties are often conflated together. Pakir (1986) suggests towards the end of her dissertation, that comparisons between the two dialects could be carried out in the future.

1.2.3 Speaker numbers

The 2010 census of Malaysia shows that Malacca had a population of 788,706 as of 2010 (Malaysia 2011). In 2007, 32% of the total population in Malacca was Chinese, while the Malays outnumbered them at 57%. In Singapore, as of 2012, there were 5.31 million people. The Chinese formed the majority at 74%, followed by Malays at 13%, and Indians at 9.2% (Singapore 2011). These numbers however shed no light on the current Peranakan population and BM speaker numbers, as Peranakans are subsumed under ‘Chinese’ for the purpose of administration in both Malaysia and Singapore. In addition, while Singapore’s population census does report language trends such as the language most often used at home, it only includes a set list of languages such as English, Mandarin, Malay, Tamil, Hokkien, Teochew, Cantonese, “Other Chinese dialects”, “Other Indian languages”, and “Others” (Singapore 2011). Presumably, BM would be included in the count under “Others”.

In 1986, based on a 1957 census report of Singapore that stated that 2.1% of 442,707 people had “Malaysian dialects” as their mother tongue, Pakir (1986) reported that there was possibly an ethnic population of 10,000 Peranakans in Singapore. Pakir (1986) further estimated there to be at least 5000 speakers. This number comprises both fluent and semi-fluent speakers.

Monolinguals were also reported to be at least 70 years old. Twenty-eight years on, BM has become a moribund language that is no longer being learnt by younger generations. Speakers themselves often cite that there are less than 1,000 people in Singapore who can speak the language fluently. Based on Pakir’s (1986) numbers, all surviving monolingual speakers should be in their late 90s at the very least. The situation in Malacca appears to be equally bleak. In 2006, it was reported that there were 2,000 Peranakans in Malaysia (Salleh 2006). There are no official numbers for how many of them speak BM. Given that the language is moribund and that

speakers under the age of 50 are rare, assuming equal numbers of Peranakans below and above the age of 50, it can be estimated that there are less than 1,000 speakers in Malaysia.

1.2.4 Language endangerment index

Once said to be “the dialect of commerce and the lingua franca of the Straits Settlements” (Tan 1988:121), BM has now become a home language that is endangered (Chia 1983, Lau 1984, Pakir 1986). This is due to the fact that other languages have taken over the home domain. In Malacca, languages that are spoken at home in place of MBM include English, Hokkien, and to a lesser extent, Malay. In Singapore, languages that are spoken in place of SBM are predominantly English and Mandarin.

There are different methods of assessing language vitality. For example, there is the 8-level Graded Intergenerational Disruption Scale (Fishman 1991), the Expanded Intergenerational Disruption Scale (Lewis and Simons 2010) and UNESCO has also proposed nine factors for evaluating language vitality (2003). One of the latest scales to date is the Language Endangerment Index (Van Way and Lee 2013) that has been proposed for assessing levels of endangerment in the *Catalogue of Endangered Languages* (Endangered Languages 2012).¹⁵ Most of these scales emphasize similar factors that indicate endangerment. Van Way and Lee (2013) propose that some of the most important factors include intergenerational transmission (as underscored by Fishman 1991, Lewis and Simons 2009 in their scales that measure intergenerational disruption), absolute number of speakers, speaker number trends, and domains of use.

Van Way and Lee’s (2013) endangerment index is presented in the following tables. The numbers (1-5) in the top rows represent the score that is to be given to the language if it matches the description in the column below. Being the most important indicator of language vitality, the score given for intergenerational transmission is multiplied by two, while the other scores are not multiplied. The scores for all separate indicators are then added up, and converted into a percentage out of a total possible score of 25. The vitality score based on these percentages can

¹⁵ www.endangeredlanguages.com

then be interpreted with Table 5: Language Endangerment Index – Interpretation. Note that in this table, there is also a certainty score in addition to the vitality score. This is a unique feature of Van Way and Lee’s (2013) Language Endangerment Index. This score indicates how reliable the vitality score is, by summing up the total possible scores that a language can receive based on the number of indicators that are used. If there is not enough information for a particular indicator, that indicator does not have to be used, and a score of zero will be assigned instead of five, which is the total possible score for each indicator (except for intergenerational transmission, for which a language receives a possible total score of 10 points instead of 5).

5	4	3	2	1	0
<i>Critically Endangered</i>	<i>Severely Endangered</i>	<i>Endangered</i>	<i>Threatened</i>	<i>Vulnerable</i>	<i>Safe</i>
Few speakers, all elderly	Many of the grandparent generation speak the language.	Some adults of child-bearing age know the language, but do not speak it to children.	Most adults of child-bearing age speak the language.	Most adults and some children are speakers.	All community members/ members of the ethnic group speak the language.

Table 1: Language Endangerment Index – Intergenerational transmission

Where intergenerational transmission is concerned, BM is “severely endangered”. A severely endangered language is one that is spoken by many from the grandparents’ generation. A severely endangered language on the index is more in danger of dying out than an “endangered language”, which is spoken by some adults of child-bearing age, but not spoken to children. Both MBM and SBM are rarely spoken by adults of child-bearing age, and would be better classified as being “severely endangered”, being moribund and not learnt by children. At the same time, it

cannot yet be classified as being “critically endangered”, as not all speakers are elderly, some are middle-aged. The language scores an 8 in this respect (4x2).

5 <i>Critically Endangered</i>	4 <i>Severely Endangered</i>	3 <i>Endangered</i>	2 <i>Threatened</i>	1 <i>Vulnerable</i>	0 <i>Safe</i>
1-9 speakers	10-99 speakers	100-999 speakers	1000-9999 speakers	10,000-99,999 speakers	100,000+ speakers

Table 2: Language Endangerment Index – Absolute number of speakers

In terms of absolute number of speakers, having less than 1,000 speakers, both MBM and SBM are classified as being “endangered” on Van Way and Lee’s index (2013). The language scores a 3 for absolute number of speakers.

5	4	3	2	1	0
<i>Critically Endangered</i>	<i>Severely Endangered</i>	<i>Endangered</i>	<i>Threatened</i>	<i>Vulnerable</i>	<i>Safe</i>
A small percentage of community members or members of the ethnic group speak the language; the rate of language shift is very high.	Fewer than half of community members or members of the ethnic group speak the language; the rate of language shift is accelerated	About half of community members or members of the ethnic group speak the language; the rate of language shift is frequent but not rapidly accelerating	A majority of community members or members of the ethnic group speak the language; the number of speakers is gradually diminishing.	Most community members or members of the ethnic group are speakers; speaker numbers are diminishing, but at a slow rate.	Almost all community members or members of the ethnic group speak the language; speaker numbers are stable or increasing.

Table 3: Language Endangerment Index – Speaker number trends

Besides absolute numbers, Van Way and Lee (2013) emphasize the importance of looking at general trends in speaker numbers. These varieties are not commonly spoken by those under the age of 50. Assuming that there are community members equally distributed over an age range of 0-90, SBM and MBM can be categorized on the index as being “severely endangered” since fewer than half of the community members speak the language. The rate of language shift is accelerated in these instances. The language scores a 4 in this instance.

5 <i>Critically Endangered</i>	4 <i>Severely Endangered</i>	3 <i>Endangered</i>	2 <i>Threatened</i>	1 <i>Vulnerable</i>	0 <i>Safe</i>
Used only in very few domains, (for example, restricted to ceremonies, to few specific domestic activities); a majority of speakers supports language shift; no institutional support.	The language is being replaced even in the home; some speakers may value their language while the majority supports language shift; very limited institutional support, if any.	Used mainly in the home; some speakers may value their language but many are indifferent or support language shift; no literacy or education programs exist for the language; Government encourages shift to the majority language; there is little outside institutional support.	Used in non-official domains; shares usage in social domains with other languages; most value their language but some are indifferent; education and literacy programs are rarely embraced by the community; government has no explicit policy regarding minority languages, though some outside institutions support the languages.	Used in all domains except official ones (i.e., government and workplace); nearly all speakers value their language and are positive about using it (prestigious); education and literacy in the language is available, but only valued by some; government and other institutional support for use in non-official domains.	Used in government, mass media, education and the workplace; most speakers value their language and are enthusiastic about promoting it; education and literacy in the language are valued by most community members; government and other institutions support the language for use in all domains.

Table 4: Language Endangerment Index – Domains of use

In addition to intergenerational transmission and speaker numbers, it is also important to consider the language’s domains of use. On Van Way and Lee’s (2013) index, both MBM and SBM are “severely endangered”, as they are being replaced even in the home. In both Malacca and Singapore, the varieties are being replaced by English to a large extent. As mentioned earlier, other languages spoken at home in Malacca include Hokkien, and to a lesser extent, Malay; and Mandarin is spoken in the Singapore Peranakan home as well. It is valuable to note

that the Peranakans were among the first in the Malay archipelago to embrace the English medium education that began with British colonialization (Lim 2010). In addition, it is important to remember that the community had positive relations with the British administrators, to the extent that they were referred to as the “King’s Chinese” during the colonial period (Lim 2010). It appears that since then, Peranakan families have been more inclined towards speaking English. In addition, although the Peranakan language is highly valued particularly by those involved in cultural and clan activities, the majority would rather have their children speak English at home for the purpose of social advancement, as the mastery of English is conflated with better academic performance, and better job opportunities. Chua (2003:71) states that “weakness in English language skills are not only detrimental to educational achievements, but also incur serious economic disadvantages” in the Singapore context. There is very limited institutional support for BM in both Malacca and Singapore, and there is no official support for the language in Malacca. The situation is harsher in Singapore, as the use of non- official languages is forbidden on the national free-to-air radio and television stations (Wee 2010). Although the Peranakan Museum was established in Singapore in 2008 by the National Heritage Board, the focus of the museum is the Peranakan material culture, and not its non-material culture, which includes its language. Limited institutional support comes from the various Peranakan associations. BM is often used on promotional material for Peranakan events, although these are limited to short *pantun* ‘traditional poem’, and short paragraphs, since many community members do not understand the language. At the time of writing, there are no BM language classes available anywhere. The language thus scores a 4 on this scale.

<i>Vitality</i>	<i>Certainty</i>
100-81% = Critically Endangered	0-10 = Uncertain
80-61% = Severely Endangered	15-20 = Fairly Certain
60-41% = Endangered	25 = Certain
40-21% = Threatened	
20-1% = Vulnerable	
0% = Safe	

Table 5: Language Endangerment Index and Level of Certainty Index

BM scores $(8+3+4+4)/25 \times 100 = 76\%$ on the Language Endangerment Index, and the language is considered as being “severely endangered”. This can be said with a high level of certainty, since all four indicators have been used to derive the vitality score.

1.3 Language documentation

This dissertation is based on an ongoing language documentation project that began in June 2012. Both SBM and MBM are being documented as part of this project, even though this dissertation focuses primarily on the SBM variety.

Essentially, language documentation emerged as a response towards current concerns regarding language vitality. Krauss (1992) predicted that as much as 90% of the world’s languages may become “doomed” (moribund or extinct) by the end of this century. Repeating Krauss’

predictions, UNESCO (2003) states that at least 50% of the world's languages are losing speakers, and that by the end of the 21st century, about 90% of the languages may be replaced by dominant languages. Since then, more recent numbers have also been made available. Campbell et al. (2013) show that 3230 of the world's languages are endangered by the careful criteria of the *Catalogue of Endangered Languages* (Endangered Languages 2012), 45% of the world's languages, and on average, only 0.36 of a language becomes extinct per month. Essentially, this means that one language dies every 3 months, or 4.3 languages per year. The fact that Krauss' initial estimates were too high is corroborated by Simons and Lewis (2013), who based their findings based on information in the latest edition of *Ethnologue* (Lewis et al. 2013). Simons and Lewis (2013) state that in more urbanized areas such as Australia, Canada, and the United States, 75% of the languages that were in use in 1950 are now extinct or moribund, while in less urbanized areas such as sub-Saharan Africa, less than 10% of the languages are extinct or moribund. Overall, Simon and Lewis (2013) estimate that 19% of the world's living languages are moribund. Regardless of how high or low these numbers are, there should be no doubt that language loss is highly consequential.

Crystal (2000) states the importance of language in expressing identity, and in encapsulating its speakers' history, for as Emerson once said, "Language is the archives of history" (1855:26). Languages also contribute to the sum of human knowledge (Crystal 2000), and the loss of languages and their cultures is accompanied by the irreparable loss of diverse and interesting intellectual wealth (Hale 1992). With specific regard to BM, the language expresses and represents a particular ethnic identity. Arguably, the language is a key external marker of the Peranakan identity in this current era. Most Peranakans look Chinese, having long been assimilated into the Chinese population through marriage. Nyonyas no longer wear their traditional outfit, the *sarong kebaya* on a daily basis. Neither do the Babas go about their regular day dressed any differently from other people in a modern-day society. Essentially, it is difficult to identify a person on the street as being Peranakan unless the person speaks BM. Beyond being an ethnic marker, BM also captures the cultural knowledge of the Peranakans. Without the language, it is impossible to talk about the traditional ceremonies that take place in a Peranakan household, such as *lapchai* 'gift exchange ceremony' and *chiutau* 'coming-of-age ceremony',

which are just two of the many ceremonies conducted for marriage. In addition to being a marker of identity and a cultural repository, BM is linguistically interesting. It is a contact language that has both Austronesian and Sino-Tibetan components. As such, it also combines non-tone and tone languages. Furthermore, while it is regarded to be a creole by some (Lim 1981, 1988, Ansaldo and Matthews 1999, Shih 2009), it has more complex features that are not traditionally associated with creoles, such as relative clauses to mention one example (see section 5.6.3). Also with regard to its likely creole-hood, Ansaldo et al. (2007) identify correctly, that while most typical creoles were formed against a background of unequal power between two groups, BM had its roots in the intermarriage of ethnicities of relatively equal power. These are but just some of the attributes that make BM unique as a language, and that gives further impetus to document it.

Language documentation has been positioned to address the concerns above. However, there are two quite separate views about what language documentation should encompass. There are researchers such as Himmelmann, who state that language documentation is “concerned with the methods, tools, and theoretical underpinnings for compiling a representative and lasting multipurpose record of a natural language or one of its varieties” (2006: 5). In the same vein, Austin states that “[t]he role of data in language documentation is rather different from the way that data is traditionally treated in language description” (2006:86). A more extreme version of this view is that of Woodbury, who proposes that language documentation is “direct representation of naturally occurring discourse is the primary project, while description and analysis are contingent, emergent byproducts which grow alongside primary documentation but are always changeable and parasitic on it” (2003:39). While researchers like Himmelmann may prefer to leave the task of description out of a language documentation project until the task of language documentation is fully complete (see Himmelmann 2006), it is clear that a language may never be satisfactorily documented (in their sense of documentation) to its fullest extent. In addition, Rhodes et al. (2006) state that description and analysis is a crucial step in accounting for how adequately a language is documented. In this sense, grammars and dictionaries appear to be basic requirements of language documentation projects – this is the view held by most scholars who work on the indigenous languages of North America and by older scholars in other

regions. This dissertation hence adopts the view that language documentation should encompass both documentation and analysis.

This dissertation, which aims to present a descriptive grammar of BM, is designed to be part of a wider documentation project that will ideally encompass also the creation of an annotated corpus and a bilingual BM- English, English-BM dictionary.¹⁶ In particular, a well-annotated corpus can be used by community members in the development of pedagogical materials, or serve as a record of the culture. For researchers, material in a well-annotated corpus can be used to answer different questions about the nature of BM, and/or contact languages, as well as to verify or reject the claims of other researchers with regard to these subject matters among others. All material that result from language documentation efforts will be archived in Kaipuleohone,¹⁷ the University of Hawai‘i Digital Ethnographic Archive, that is managed by the Department of Linguistics. The university and department have a strong focus on research in the Asia-Pacific region. Kaipuleohone also conforms to international archiving standards for digital archives. By default, all material in this archive will be openly accessible after an initial period of five years, unless they are deemed to require restricted-access by the language consultants. It cannot be reiterated often enough how crucial it is to archive documentation materials. This will contribute towards preserving linguistic and cultural diversity, BM being a repository of culture for its speech community.

1.3.1 Community and participants

With regard to language documentation, I have been working with speakers of both SBM and MBM. Among my key consultants are three SBM speakers (two males and one female) and one MBM speakers. Being well-connected with Peranakan culture scene or with family members who still speak the language, these speakers also serve as my contact point to other speakers. The ages of my four key consultants range from 60 to 80, and they all use BM on an everyday basis. According to Pakir (1986), monolinguals were in their 70s in 1986, and should be in their 90s

¹⁶ With due respect to researchers who have previously worked on BM, it should be noted that the extensiveness of this current project was made possible by technological advances, such as digital audio recording, digital archiving, and a range of transcription and annotation software, that were not available in the past.

¹⁷ www.kaipuleohone.org

currently, but none could be found. Interviews have also been conducted with less proficient speakers, who claim that they used to speak the language natively, but who no longer use the language regularly. This allows me to examine and establish the extent of variation among these different sorts of speakers.

Among the four key speakers that I have been working with is the president of the Peranakan Association of Singapore, Baba Peter Wee, who is in his 60s. The Peranakan Association of Singapore is a cultural association for the ethnic group established in 1900 with links to similar associations in Malacca, Penang, and also Sydney, where a number of Peranakans have migrated to in recent years¹⁸. It is headquartered informally at the Katong Antique House that Baba Peter Wee owns. The Katong Antique House, which is part shop and part private museum, is located along East Coast Road, a traditional Peranakan enclave. Recording sessions with Baba Peter Wee take place here. He also occasionally receives Peranakan visitors who speak BM. With informed consent, some of the conversations that take place at the Katong Antique House have also been recorded and archived.

Another location where recording sessions take place is the Gunong Sayang¹⁹ Association clubhouse situated at Geylang. The Gunong Sayang Association is a social club established in 1910 that aims to create an awareness of Peranakan culture by promoting *dondang sayang* ‘art of singing Malay poetry’, *wayang Peranakan* ‘Peranakan theater’, music, and dance.²⁰ Besides eliciting data from Baba Victor Goh, the club’s cultural advisor (who is also in his 60s), I record conversations that take place at the dinner table, and apart from attending events organized by both associations, I attend weekly singing rehearsals at the clubhouse. In particular, I have formed a close bond with the Gunong Sayang singers, a group of mostly female singers between the ages of 50 to 80. Under the direction of Baba Victor Goh, they rehearse every Monday evening at the Gunong Sayang clubhouse, and put on several public and private performances every year. Although this is a *language* documentation project, cultural documentation is equally important, since language and culture are highly intertwined. Hence, my corpus will include

¹⁸ This information has been derived from the Peranakan Association website (www.peranakan.org.sg).

¹⁹ The words *Gunong Sayang* translate directly into ‘mountain love’ to mean ‘mountain of love’.

²⁰ This information has been derived from a National Library Board of Singapore website (http://infopedia.nl.sg/articles/SIP_1383_2010-05-24.html)

recordings of cultural performances, in addition to recordings of interview sessions. Besides documenting cultural activities, I believe that spending time intensively interacting with a group of Peranakans who have come together for the purpose of carrying out a cultural activity, can better inform my language documentation efforts. Also, by attending cultural events, and interacting with active community members, I can create awareness for my project, as well as awareness about the endangerment of BM.

The only female key speaker that I am working with is Bibik Jane Quek, who is 80 at the time of writing. Besides being a singer and committee member of the Gunong Sayang Association, Bibik Jane Quek is known for her Peranakan home-cooking in Singapore. Recording sessions with her usually take place at her home. All three key speakers mentioned above speak SBM. The key speaker that speaks MBM is Baba Albert Ku, who is also in his 60s. Recordings with him take place mostly at his home in Malacca. A retiree, he speaks the language with his family, and he is at proficient at writing *pantun* (traditional poetry) in BM.

All of the key speakers in this project speak BM natively, and still use the language on a daily basis. They have provided informed consent for their data to be recorded and archived in Kaipuleohone, so that it can be used for this dissertation and subsequent publications on BM.

1.3.2 The scope of this dissertation

This dissertation is a grammar of BM as it is spoken in Singapore, although it makes observations where SBM differs from MBM. It is based mainly on the material that has been collected for the language documentation project. It differs from other grammars in that:

- a) It incorporates sociophonetics as a means of documenting the associations that people form between phonetic variants that are changing due to language endangerment and the speakers who produce these variants.

The term ‘sociophonetics’ has been used in connection with research concerning variation, change, sociologically-informed fieldwork, speaking style, as well as with other topics as disparate as the phonological relationship between liquids, descriptive accounts of Albanian and

Cocos Malay, loanword pronunciation, conversation analysis, methods for developing large corpora, and psycholinguistic experiments on information processing (Foulkes et al. 2010). For the purposes of this dissertation, a narrower sense of the term ‘sociophonetics’ is used. Here, the term ‘sociophonetics’ refers to “the study of socially conditioned phonetic variation in speech” (Hay and Drager 2007:90). Hay and Drager state that sociophonetics usually involves “using increasingly sophisticated phonetic analysis to show that very fine phonetic detail is used in constructing social identity,” “using ethnographic approaches to break free of traditional social categories,” and “using experimental techniques to probe listener’s use of sociophonetic detail in speech perception” (2007:89).

There are different reasons for incorporating sociophonetics into a language documentation project and grammar. This project identifies four ways in which language documentation and sociophonetics can complement one other. First, to fully document a language, it is also necessary to document and describe how the language is used, and how it varies according to social contexts. In addition, variation in structure is still part of structure. Variation in structure needs to be described and explained, and the explanation may most likely be found in social conditioning of language variants. Second, language documentation emerged as a reaction to the fact that languages were being lost at an alarming rate. A documentation description captures a snapshot of the language, and preserves it, but the reality is that this snapshot is seldom that of a stable system. Moribund languages are said to be often less normative and more variable (Dorian 2001). It has been demonstrated that in minority language communities, variation can occur in phonetics and phonology, and in syntax, morphology, and even morphophonology when a language’s viability is being threatened by a more dominant language (see Stanford and Preston 2009). An endangered language community thus may be ideal for a language variation study. In view of that, this current project approaches variation via sociophonetic perception methods. Third, an extensive corpus of good quality sound files provides not only valuable production data from languages other than English, on which sociophonetics has been largely focused (Foulkes et al. 2010). This also provides a wealth of naturally occurring tokens, which may be preferred for their naturalness in perception experiments. Labov et al. (1991) in particular advocate against using unnatural, isolated, synthetic tokens. Fourth, linguists are usually expected to incorporate

aspects of ethnography into the language documentation project. Hill states, “Documentary linguists need to be ethnographers, because they venture into communities that may have very different forms of language use from those of the communities in which they were socialized as human beings or trained as scholars” (2006:113). This is compatible with the notion of ethnographic approaches in sociophonetics, wherein researchers are able to “break free from using traditional social categories that may not be relevant for a particular group of speakers, and to investigate in depth the social meaning of particular variants” (Hay and Drager 2007:89).²¹ Although the goals of language documentation and sociophonetics tend to be different, the basic stances are the same. With careful planning, it is possible and optimal for a language documentation and sociophonetics to mutually exist within a single project.

Although this is not a language revitalization project, I hope that by documenting the associations that people form between phonetic variants that are changing due to language endangerment and the speakers who produce these variants, I can create a greater awareness of language endangerment in the Peranakan community and increased enthusiasm for attempting to revitalize the language. The sociophonetics experiment conducted during the course of fieldwork is discussed in section 3.7.1. Besides being an experiment in itself for incorporating sociophonetics into language documentation work, this grammar can be used to shed insight on issues such as:

b) How is SBM different from MBM?

As mentioned earlier (see section 1.2.2), researchers usually regard BM as a single entity, using the same label ‘Baba Malay’ for both varieties, spoken in Singapore and Malaysia. Again, the only references to particular locations were made by Pakir (1986), who states that her dissertation describes BM as it was spoken in Singapore, and Lim (1988), who mentions that his fieldwork took place in Malacca. Based on materials collected for the archive, I will be able to

²¹ Note that the notions of ethnography in sociophonetics and language documentation differ in some sense. Ethnography in sociophonetics allows for more open-ended investigation than immediately focusing on a few key variables, while ethnography in language documentation takes into account culture when it impinges on the language structure. These also differ drastically from the ethnography of communication, which usually refers to a qualitative analysis of speech in interaction within the context of ethnography, as well as from ethnography in anthropology, which requires a holistic description of society and culture.

evaluate how these two varieties differ. Broad generalizations that can be made are that SBM utilizes more forms that are derived from Hokkien, while MBM utilizes more forms that are derived from Malay. For example, where the vowel inventory is concerned, [ɔ] exists in SBM, but is not observed to occur in MBM. Differences are also clear where function words are concerned. An example of such a lexical difference is the use of conjunctions. SBM speakers use [sa.ma] to mean ‘and’ and ‘with’, whereas in addition to [sa.ma], MBM speakers use both Malay forms [dan] to indicate ‘and’ and [dəŋan] to indicate ‘with’, the basic lexical meaning of [sa.ma] being ‘same’ in Malay. With a more extensive corpus, it will be possible to establish how the two dialects differ. This dissertation, being mainly a grammar of SBM, makes observations of certain differences between SBM and MBM in chapter 6.

Besides exploring language change through sociophonetics, as well as the differences between SBM and MBM, this grammar also address:

- c) Whether or not Baba Malay is a creole language, a dialect of Malay, or some other sort of contact language

With regard to the classification of BM, researchers are divided into two main camps. There are those who regard BM as a dialect of Malay (Pakir 1986, Thurgood 1998), and those who regard BM as a language that has undergone creole formation (Lim 1981, Lim 1988, Ansaldo and Matthews 1999, Ansaldo et al. 2007, Shih 2009).

Thurgood’s (1998) dissertation focuses on 19th century BM. Based on her analysis of old BM texts, Thurgood (1998:245) concludes that the language is a Malay variety that has undergone shift, rather than a creole. For Pakir (1986: 116, 117), BM is a variety of Malay that has borrowed Hokkien loanwords in domains such as the household; terms of address and reference; celebration, customs, beliefs and religion; and clothing and personal ornaments, in addition to others. Pakir (1996:207) states, for example, that the phone [ɔ] only occurs in Hokkien loanwords, and that all syntactic structures in BM that resemble Hokkien structures, are also found in other dialects of Malay, such as Bazaar Malay, Jakarta Malay and Ambonese. Also, Pakir (1986:210) points out cases wherein these links appear to be less strong—not all of these structures and their functions match up to either that of Hokkien or Malay. Therefore, Pakir

(1986) views BM as a unique dialect of Malay. Note that the findings of this current grammar do not support the notion that BM is a genetic dialect of Malay.

In response to Pakir (1986), Ansaldo and Matthews (1999) state that the fact that other varieties of Malay have many of these BM features by no means invalidates the arguments for the Hokkien substrate, since there are other ways through which these features could have entered these other Malay dialects. They point out that Bazaar Malay and Jakarta Malay especially have substantial Chinese language contact influence. Besides suggesting that there could have been direct Chinese influence through immigration across the Malay Peninsula and archipelago, and possibly “restructuring under conditions of language contact” (Ansaldo and Matthews 1999:43), the authors also point to Lim’s (1981:31) suggestion that there could have been an “antecedent lingua franca spoken by and between different groups over several centuries, underlying both Baba and Bazaar Malay.” The rest of Ansaldo and Matthews’ structural arguments are based mostly on data provided by a play. They investigate a passive construction, an attributive construction, and the tense-aspect system of BM, stating that these have similar counterparts in Hokkien.

Based on his fieldwork in Malacca, Lim (1981, 1988) recognizes the structural ties between BM and Hokkien. For example, Lim points out that most of the pronominal system in BM is derived from Hokkien (see table 6). For example, second person and third person plurals appear to have been calqued from Hokkien forms using a combination of singular form and the Malay word *orang*; *lâng* in Hokkien means ‘people’.²² Note that an expanded version is presented in table 46.

²² It is not implausible that the phonetic similarity of Malay *orang* and Hokkien *lâng* may have facilitated the adoption of the calqued forms for second person and third person plurals in BM, comprising Malay word *orang* ‘person’. See tables 6 and 46.

	Baba Malay	Malay	Hokkien
<i>I</i>	gua ²³	saya/ aku	gua
<i>you</i>	lu	kamu/awak/engkau	lú/ lí
<i>he/she</i>	dia	dia	I
<i>we</i>	kita	kita (inclusive) kami (exclusive)	gún-lâng (inclusive) lán-lâng (exclusive)
<i>you (pl)</i>	lu-orang	kamu/awak/engkau	lín-lâng
<i>they</i>	dia-orang ²⁴	mereka	ín-lâng

Table 6: Personal pronouns in Baba Malay, Malay and Hokkien, replicated from Lim (1988: 27)

Similarly, Shih (2009) attempts to establish BM as a creole, based on observations of its phonology, morphology, lexicon, and syntax. While comparisons of its phonological system with that of Hokkien and Malay appear to be less conclusive, she states that there is less affixation in BM than in Malay, and the use of the ‘number classifier noun’ structure appears to have been

²³ My Singapore Baba Malay informants and Malacca Baba Malay consultants state that *saya* is used in Baba Malay to mean first person singular as well, and that it is more formal than *gua*. This is also observed by Chia (1983). The word *saya* ‘I (humilific)’ is borrowed from Sanskrit, not only in Malay, but in a few other languages of the region. Whereas Marsden (1812) notes that *saya* is humilific in Early Modern Malay, which may have influenced BM’s structure, *saya* is used as a more formal first person pronoun in current day Malay and Indonesian, the informal first person pronoun being *aku*.

²⁴ In my collected data, *dia* can be used interchangeably with *dia-orang* to represent the third person plural.

derived from Hokkien. In addition, kinship and cultural terms have also been derived mainly from Hokkien. With regard to syntax, Shih (2009) compares BM *kasi*, with Hokkien *hoo*, which has the literal meaning of ‘to give’. She demonstrates that there are overlaps in their functions, among which are their uses in causative constructions. Examples (1) and (2) demonstrate two types of causative constructions using *kasi*.

- (1) *tak orang mo senang senang kasi lu tumbok*
 NEG person want free free let 2.SG punch
 ‘No one would volunteer to be hit by you.’
 (Shih 2009: 82)
- (2) *dia tak dani, dia kasi saya susah-hati*
 3.SG NEG present 3.SG cause 1.SG worry
 ‘His absence frets my heart (troubles me a lot).’
 (Shih 2009:83)

In example (1), *kasi* expresses ‘to let’; in example (2), *kasi* means ‘to cause’. Similar constructions can be found in Hokkien.

- (3) *I hoo gua lai*
 3.SG let 1.SG come
 ‘he let me come.’
 (Shih 2009:79)
- (4) *i bo lai hoo gua jin huan.lo*
 3.SG NEG come cause 1.SG very worried
 ‘that he doesn’t come troubles me a lot.’
 (Shih2009:80)

While in (3) *hoo* means ‘to let’, in (4), it means ‘to cause’. These are similar in structures to examples (1) and (2). Based on these examples that appear to be substrate transfers, Shih (2009) concludes that creolization must have taken place.

It is useful to consider the different views in the literatures about creole formation, to see if any of these may contribute to a resolution of this controversy. Thus far, the evidence above appears to be congruent with the notion of substrate transfer or relexification (Lefebvre 1998, Muysken 1981, Siegel 1999), where the functions and or syntax of the substrate language are adopted into the creole language with forms from the lexifier language. A more detailed description of the

differences between substrate transfer and relexification are found in section 7.4.4.2.

Nevertheless, Pakir (1986) had pointed out that there is not a complete point-by-point match between the grammars of Hokkien and BM. A universalist point of view may also be inadequate — Bickerton's (1984) list of 15 features that are presumably found in all creoles makes no predictions of more complex material, such as relative clauses (see section 5.6.3). Neither does the universalist approach acknowledge that features can be retained from the substrate languages as seen from examples (1) to (4).

Would BM, for example, be a feasible creole according to Mufwene's views of founder effect and language ecology? The founder effect principle is largely applied to plantation creoles, and it asserts that the structural features of creoles have been predetermined to a large extent (but not exclusively) by characteristics of the vernaculars spoken by the populations that founded the colonies in which they developed (Mufwene 1996). According to Mufwene (1996: 85), an ecology is:

the ethnographic setting in which the lexifier (the displaced population) has come into contact with diverse languages (other populations) whose structural features (genes) enter into competition with its own features.

The features may then be adapted to suit the ecology, so that features that are compatible with the features of the other languages in the ecology are more likely to be maintained than features that are not. Structurally, for this to be a plausible account of BM, it must be shown that basilectalization had taken place, so that the creole became more and more divergent from the initial lexifier language. Basilectalization in this instance, refers to the creole becoming more like its substrate language in terms of structure. However, this begs another question. Why is Malay the lexifier language of BM and not Hokkien? Should the Malay-speaking population then be treated as the founder population, rather than the Hokkien traders? In relation to this, as mentioned earlier, Ansaldo et al (2007) identify correctly, that while most typical creoles were formed against a background of unequal power between two groups, BM had its roots in the intermarriage of ethnicities of relatively equal power. If Malay should indeed be considered the language of the founder population (possibly since they were in Southeast Asia prior to the

Hokkien traders), it will be necessary to consider how the language ecology may have affected the selection of features. When the Chinese traders arrived after the beginning of the 15th century, the Malay language was already a dominant lingua franca in the region, having been established as one during the Srivijaya period (from about the 7th to the 13th century A.D.) (Hall 2011).²⁵ In addition, the founder effect may partially explain for why there is so little Portuguese influence in BM, even though the Portuguese came into power following their conquest of Malacca in 1511, and bearing in mind that the Portuguese language was also a dominant trade language in Asia at that time. By the time of the Portuguese conquest of Malacca, the Chinese traders had already arrived and were already exposed to the language of the founder population, Malay. It is also relevant to note that the Portuguese never forced their language onto the local population, instead they were said to respect the culture and traditions of the Malays, the Tamil Chetty, Chinese, and other communities (Muzzi 2002:31). Taken as a whole, these factors may explain for why there is not more Portuguese influence on BM, aside from some loanwords. With regard to the influence of languages brought in through colonialism, it should be highlighted that even though Malacca fell to the Dutch later in the 1641, the Dutch language was never a dominant trade language in the region (Borschberg 2010). The two dominant languages that existed at the same time were Malay and Portuguese, although it was noted by different individuals that Malay was the more significant of the two (Borschberg 2010). This language ecology appears to be largely maintained (with the exception of Portuguese and Dutch). While Dutch is not currently spoken in Singapore and Malacca, there is a small community of Portuguese creole speakers left in Malacca. The name of this creole is *Papia Kristang* ‘Christian language’. Currently, with Singapore situated among Malay-speaking neighbors (Indonesia, Malaysia, and Brunei), and Malacca being part of Malaysia, it may be no surprise as to why on the surface, BM still appears to be most congruent with Malay.

The founder effect and the notion of a linguistic ecology are useful, but some questions remain unanswered. For example, why do some structures appear to be caused by substrate transfer, such as the causative patterns observed in examples (1) to (4)? Why are forms that are Hokkien

²⁵ The Srivijaya kingdom began in the late 7th century A.D, became politically dominant in the 9th to 10th century, and then declined into the 13th century.

in origin used by BM speakers for particular domains such as with kinship terminology (See Appendix A)? While the structure of BM cannot be completely explained by these theories, their pertinence could rest in their aggregate contribution to explaining BM's origins and development. Based on the findings of this grammar, this dissertation aims to shed light on whether or not BM can be classified as a creole language, a dialect of Malay, or some other sort of contact language (see chapter 7). I will argue for the conclusion that BM is to be classified as a creole.

2. An overview of Baba Malay typology and this grammar

2.1 An overview of Baba Malay typology

This is a brief typological overview of BM, surveying its basic characteristics. It includes notes on its vowel and consonant inventories, morphology, word order, alignment, and case.

2.1.1 Sound system

BM has eight phonemic vowels; its vowel inventory is as follows:

i		u
e	ə	o
ɛ		ɔ
	a	

Its 19 phonemic consonants are as follows:

		Labial	Alveolar	Post-alveolar	Velar	Glottal
plosives	voiceless	p	t		k	ʔ
	voiced	b	d		g	
affricates	voiceless			tʃ		
	voiced			dʒ		
fricatives	voiceless		s			h
nasals		m	n	ɲ	ŋ	
lateral			l			
flap			r			

glides: w, j

2.1.2 Basic morphology

There is some affixation in BM. All of the affixes are derived from Malay, in which the corresponding affixes are obligatory (Marsden 1812)²⁶. Most of these affixes are optional and may not be meaningful to BM speakers. For example, in Malay, the prefix *me-* is used as an active marker. But in BM, the use of this marker is not mandatory, and it makes no difference in meaning to the speakers. This is demonstrated below by examples (5) and (6), where *me-* is non-existent in the former example, but is attached to *nangis* ‘cry’ in the latter. Alternatively, *menangis* can be used in the context of (5) and *nangis* can be used in the context of (6) to express the same notions.

- (5) *Téngok, bapak dia mati, dia tak nangis.*
see father 3.SG die 3.SG NEG cry.
‘See, his father passed away, (and) he did not cry.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-021, 00:32:44.0-00:32:46.3)
- (6) *Dia kuluardari rumah menangis.*
3.SG go.out from house cry
‘She left the house crying.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:04:18.3-00:04:20.9)

Very few affixes are obligatory. Two affixes that are more commonly used are the prefixes *ter-* ‘accidental, movement’, as well as the suffix *-kan* ‘transitive’. The two types of *ter-* are analysed as the same prefix, since they are semantically related, as they both involve movement in general. The different usages of *ter-* are shown in examples (7) and (8). In example (9), the verb *jatoh* ‘to fall’ is intransitive, while in (10), it is made transitive with the affixation of *-kan*.

- (7) *Ter-masok kat dalam ayé.*
ACD-enter PREP inside water
‘Fell into water.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:14:33.2-00:16:34.8)

²⁶ Crucially, Early Modern Malay contributed to the formation of Baba Malay, rather than Malay as it is currently spoken; Baba Malay in the present day being a moribund language.

- (8) *Ter-balék-kan itu ikan pun boléh lah.*
 MVT-turn.over-TR that fish also can EMP
 ‘Turn over that fish is also okay.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:10:32.3-00:10:33.8)
- (9) *Budak jatoh,*
 child fall
 ‘The child falls.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:46.6-00:03:50.7)
- (10) *Jatoh-kan dia,*
 fall-TR 3.SG
 ‘drop it.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:59.2-00:01:02.5)

2.1.3 Basic syntax

The basic clausal level word order in BM is SVO (Subject Verb Object). This is demonstrated by the following examples.

- (11) *Itu kucing makan ikan.*
 that cat eat fish
 ‘That cat eats fish.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:00:35.9-00:00:37.7)
- (12) *Gua jatoh-kan gua mia bég*
 1.SG fall-TR 1.SG POSS bag
 ‘I dropped my bag.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:15:37.8-00:15:41.9)
- (13) *Dia kata buah berangan*
 3.SG say CLF.FRUIT chestnut
 ‘He said chestnuts.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:00:49.2-00:00:51.2)

The subject is not always obligatory. This happens when the unexpressed subject is known to participants in the conversation. In (14), the speaker is asked what she is doing, to which she responds *kopék bawang* ‘peeling onions’. In (15), the speaker was speaking about what old

people do to stay healthy, one of the suggestions she gives is *makan ikan banyak* ‘eat a lot of fish’, the unexpressed subject here being old people.

- (14) *Kopék bawang lah.*
Peel onion EMP
‘Peeling onions.’
(Kim Choo, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:12:49.8-00:12:51.2)

- (15) *Makan ikan banyak*
Eat fish many
‘Eat a lot of fish.’
(Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:03:44.4-00:03:46.2)

At the phrasal level, BM has word orders of ADJ N (Adjective Noun) (not basic) *and* N ADJ, GEN (Genitive Noun) *and* NG (not basic), as well as PREP N (Preposition Noun). Relative clauses occur both prenominally and postnominally. Example (16) shows the ADJ N word order, while example (17) shows the N ADJ word order. Although the ADJ N order is a very commonly used structure, a relativizer has to be used, and it cannot be considered as basic as N ADJ.

- (16) *betol mia cherita*²⁷
real REL story
‘story that is real.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:03:24.8-00:03:26.0)

- (17) *Idong panjang.*
nose long
‘Long nose.’
(Lillian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:00:13.7-00:00:15.0)

Example (18) demonstrates the GEN N word order. The N GEN word order is not basic since it requires the use of a pronoun. The possessor in the N GEN construction has to be a pronoun. This is shown in example (19).

²⁷ *Punya*, which can be shortened to *nia and mia*, is a possessive marker that can also be analysed as a relative clause marker. This phenomenon is described in section 5.6.3

- (18) *Peter punya bapak,*
 Peter REL father
 ‘Peter’s father,’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:45:01.7-00:45:06.9)
- (19) *Badan lu*
 body 2.SG
 ‘Your body.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:11:33.1-00:11:36.1)

Example (20) illustrates the PREP N word order.

- (20) *Kebun dekat Europe ini.*
 garden PREP this
 ‘This garden is in Europe.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:32.8-00:00:34.6)

Example (21) shows prenominal relativization, while example (21) shows postnominal relativization. Both are equally common in BM. For restrictions on relativization and more details, please refer to section 5.6.3

- (21) [*Anak perempuan nia*] *satu*
 child female REL ONE
 ‘The one that is a girl.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:44.8-00:00:47.0)
- (22) *Ini semua dia mia kawan [nang²⁸jaga kambing]*
 This all 3.SG REL friend REL guard sheep
 ‘These (are) all his friends that guard the sheep.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:05:58.2-00:06:00.8)

For discussion on other word-order tendencies within the different types of phrases, please refer to sections 5.1.5, 5.2.9, 5.3.7, 5.4.4 and 5.5.

In terms of alignment and case, the subject NP is not marked differently from the object NP. The subject of the clause mostly precedes the verb. BM is a nominative-accusative language. The

²⁸ Also *yang*.

subject of the intransitive verb is treated equivalently to the agent of the transitive verb. Both precede verb phrase. Example (11) is replicated as (24).

(23) *Budak tu senyum.*
child that smile
'That child smiles.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:35.9-00:00:37.7)

(24) *Itu kucing makan ikan.*
that cat eat fish
'That cat eats fish.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:37:23.3-00:37:25.5)

In addition, it is possible for the object to undergo passivization with the use of the adversative passive markers *kasi* (meanings: 'give', 'cause', 'let') and *kena* (other meaning: non-volitional 'subjected to', see section 5.2.3) where the object is promoted to subject. Note that the logical subject is not expressed overtly when *kena* is used. For more discussion on passivization, please refer to section 5.2.3. Example (26) is the passive version of example (25), while example (28) is the passive statement that corresponds to example (27). (27) replicates (24).

(25) *Mary bunoh dia*
kill 3.SG
'Mary killed him.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:39:17.1-00:39:18.8)

(26) *Dia kasi Mary bunoh.*
3.SG PASS kill
'He was killed by Mary.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:40:01.8-00:40:14.9)

(27) *Itu kucing makan ikan.*
that cat eat fish
'That cat eats fish.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:37:23.3-00:37:25.5)

(28) *Itu ikan sudah kena makan.*
that fish already PASS eat
'That fish has already been eaten.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:38:17.8-00:38:19.9)

In topicalization, logical object is also fronted, without passivization. This topicalization construction is highly productive in BM, and hence worthy of mention. It has been observed by Lee (1999) in SBM, and Lim (1988) in MBM. This generates sentences such as (29). More information on topicalization can be found in section 5.6.9.

- (29) *Itu ikan sudah kucing makan*
 that fish already cat eat
 ‘That fish, the cat already ate.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:37.37.1-00:39:39.8)

2.2 A user guide to this grammar

This grammar was written to be of use to as wide an audience as possible. Where possible, I have avoided using any theoretical framework that would require specific training to understand. Some basic background in linguistics is however necessary, since this grammar uses basic grammatical concepts. These basic grammatical concepts have to be used so that the language can be adequately described, and so that researchers can make comparisons to other languages where relevant. This grammar has also adopted the use of some features that are intended at making it more user-friendly.

2.2.1 Features

These user-friendly features are (A) morpheme-by-morpheme glosses; (B) a list of linguistic abbreviations used; (C) a Baba Malay-English glossary and an English-Baba Malay glossary featured in Appendix C; as well as (D) reference to where each example can be found in the relevant archive, following Berez (2011). Example (30) illustrates this.

- (30) *Dia chakap sama dia.*
 3.SG speak with 3.SG
 ‘He speaks with him.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:02:44.4-00:02:46.0)

Kaipuleohone is hosted online at scholarspace.manoa.hawaii.edu, and if a reader wanted to search for example (30) within the archived file, she or he would have the information that this example uttered by the speaker Peter Wee, can be found on the site in the file labeled NL1-042,

and more specifically, this utterance can be heard between 00:02:44.4-02:46.0. Appropriate citation will also be provided in the rare case that an example is derived from a previous publication.

2.2.2 Structure of the grammar

This grammar is organized as follows. Chapter 3 describes the phonetics and phonology of BM; chapter 4 describes the different parts of speech; and chapter 5 provides information concerning the syntax of the language. These chapters deal with the more traditional aspects of the language's grammar, except the inclusion of the results of the sociophonetics investigation in Chapter 3 that sheds light on an ongoing change to the vowel inventory. The dissertation then continues in a less traditional way, addressing in Chapter 6 the differences between SBM and MBM, and in Chapter 7, whether BM is a genetic dialect of Malay, or if it is a mixed language or a creole. Appended at the end of this dissertation is a collection of texts, vocabulary lists, indexes and bibliography.

2.2.3 List of linguistic abbreviations used

The following is a list of abbreviations used within the glosses of examples. Leipzig glossing conventions are used where possible.

ACD	Accidental marker
ADJ	Adjective
ADV	Adverb
AdvP	Adverbial phrase
Aux	Auxiliary verb
AP	Adjective phrase
CLF	Noun classifier
COP	Copula
COMP	Complement

CONF	Confirmative particle
DEM	Demonstrative
EMP	Emphatic particle
EXCLAM	Exclamation
EXIST	Existential marker
GEN	Genitive
HAB	Habitual marker
MVT	Movement marker
N	Noun
NEG	Negation marker
NMZ	Nominalizer
NP	Noun phrase
NUM	Numeral
PFV	Perfective marker
PL	Plural
POSS	Possessive marker
PREP	Preposition
PROG	Progressive marker
Q	Question particle
REL	Relativizer
SG	Singular
TR	Transitive marker
V	Verb
VP	Verb phrase

- 1 First person
- 2 Second person
- 3 Third person

2.2.4 Example sentences

Example sentences are mostly taken from material I have recorded during fieldwork that was carried out between June 2012 to February 2014. More than 80 hours of SBM data and about 12 hours of MBM data were collected. These sound files and selected transcripts are hosted at the above-mentioned Kaipuleohone²⁹ archive. Example sentences are taken from both naturally occurring and elicited texts. While it is always ideal to use naturalistic data, some grammatical constructions are not common in daily speech, for example, noun classifiers and relative clauses. Where elicited data is used, checks were made with other consultants to ensure that the utterances are acceptable in the language. In addition, these examples also come from a range of different genres, such as conversation and narratives among others. While texts collected include *pantun* ‘traditional poems’, these are generally not used for grammatical examples, as these *pantun* are highly stylized, and do not reflect ordinary language use.³⁰ Some *pantun* have been included in the selection of texts that accompany this grammar.

Again, the source of examples or place in the archive is indicated following the example, as explained in section 2.2.1. Similarly, examples are duly acknowledged in parentheses following them, if they are derived from some other source.

²⁹ www.kaipuleohone.org

³⁰ A *pantun* is a traditional Malay verse form with an *abab* rhyme system. The significant lines in a *pantun* are the third and fourth lines, in which the main message of the *pantun* is embedded. *Pantun* can be sung to a tune called *Dondang Sayang* ‘melody (of) love’, and singers used to exchange verses they would compose on the spot (Chia 1994: 70). As compared to the Malay *pantun*, the subject of the Peranakan *pantun* is often a matter that the Peranakan community can relate to. It is interesting that even in Peranakan *pantun*, the Malay language is followed as much as possible (Chia 1994: 70). Singing *pantun* to *Dondang Sayang* is now an endangered art form, but still performed by groups such as the Gunong Sayang Association in Singapore.

2.2.5 Transcription

The transcription in this grammar is based mostly on Du Bois et al. (1992)'s conventions for discourse transcription. Speech streams are segmented by intonation unit, which is defined by Du Bois et al. as “one spurt of speech” (1992:16) or “a stretch of speech uttered under a single coherent intonation contour” (1992:21). However, when presented in this grammar to demonstrate a particular grammatical pattern or a well-formed phrase or sentence, I have taken the liberty of conflating intonation units for ease of interpretation. Thus, an example such as (31), which has two intonation units, would be produced in the grammar as (32) in a section on transitivity.

(31) *Anjing tutop-kan.*
dog close-TR
‘The dog turned off’

Itu aloji.
that alarm.clock
‘That alarm clock.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:25.6-00:00:29.9)

(32) *Anjing tutop-kan itu aloji.*
dog close-TR that alarm.clock
‘The dog turned off that alarm clock.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:25.6-00:00:29.9)

Du Bois et al.'s (1992) conventions allow for a range of transcription delicacy, from a very broad level one to their very detailed level seven that would include multimodal information. For the purpose of this grammar, a level three transcription that includes information on what the boundaries are will suffice. A period at the end of an utterance indicates finality, a comma signals continuity (that the speaker has more to say), whereas a question mark signals appeal, that the speaker is seeking a response from her or his interlocutor. These are dependent on intonation (see section 3.6.3). Incontrovertibly, the function of an utterance, whether it is a statement or a question, is important in understanding the basic structure of a language. The broader level one and two transcription systems do not include this information. Level four is

more detailed than necessary for this grammar, including information such as whether a speaker is saying an utterance with a smile. Transcription conventions for understanding the appended full transcripts at the end of this grammar are included with the transcripts in Appendix B. Transcription is orthographic, adapted in part from Malay spelling conventions in Singapore (see section 3.5), rather than phonetic. A solely phonetic transcript might be alienating to any non-linguists who are interested in the transcripts.

3. Phonetics and Phonology

At its core, the sound system of BM must take into account the fact that BM is a contact language. While the main lexifier of BM is Malay, BM also derives a significant amount of its lexicon from Hokkien. To a lesser extent, it is observed to have borrowed words from English, Portuguese, Dutch, and Tamil, among other languages (Pakir 1986, Shellabear 1913). This chapter explains how BM has its own phonological system, different from the Malay that is spoken in Singapore and Malaysia, and incorporating words from Hokkien and other languages based on its own unique phonological template.

3.1 Phoneme inventory

The phoneme inventory of BM accounts for words of Malay origin, words of Hokkien origin, and words of other origins that have long existed in the BM speakers' repertoire. This inventory comprises 19 consonants and 8 vowels that can be combined to form diphthongs and long vowels. For the reader's convenience, the inventory shown in section 2.1.1 is replicated here as table 7. Differences between this phoneme inventory and that of Malay, as well as of Hokkien, are discussed in section 7.4.1.

3.1.1 Consonants

BM has the following consonants:

		Labial	Alveolar	Post- alveolar	Velar	Glottal
plosives	voiceless	p	t		k	ʔ
	voiced	b	d		g	
affricates	voiceless			tʃ		
	voiced			dʒ		
fricatives	voiceless		s			h
nasals		m	n	ɲ	ŋ	
lateral			l			
flap			ɾ			

glides: w, j

Table 7: Consonant chart of Singapore Baba Malay

3.1.1.1 Minimal and near-minimal pairs of consonants

Minimal pairs can be used to establish the phonemic status of the consonants that have been discussed. Some minimal pairs for the consonant phonemes of the language are as follows.

- (33) /b/ : /p/
[baji] ‘similar to’ : [paji] ‘morning’
- (34) /b/ : /w/
[bajaŋ] ‘shadow’ : [wajaŋ] ‘play’
- (35) /p/ : /w/
[paʔ] ‘male fellow’ : [waʔ] ‘elderly Malay person’
- (36) /m/ : /w/
[majaŋ] ‘palm blossom’ : [wajaŋ] ‘play’
- (37) /m/ : /n/
[mja] ‘life’ : [nja] ‘contracted form of possessor and relativizer *punya*’
- (38) /m/ : /p/
[mati] ‘die’ : [pati] ‘first cream of coconut’
- (39) /d/ : /t/
[da.own] ‘leaf’ – [ta.own] ‘year’
- (40) /d/ : /r/
[dzari] ‘finger’ – [dzadi] ‘happen, become’
- (41) /t/ : /tʃ/
[tareʔ] ‘pull’ – [tʃareʔ] ‘search’
- (42) /d/ : /dʒ/
[dari] ‘from’ – [dʒari] ‘finger’
- (43) /tʃ/ : /dʒ/
[tʃam] ‘observe’ – [dʒam] ‘time, hour’
- (44) /n/ : /l/
[nama] ‘name’ – [lama] ‘long (describing time)’
- (45) /r/ : /l/
[rupa] ‘appearance’ – [lupa] ‘forget’

- (46) /r/ : /n/
[kuraŋ] ‘less’ – [kunaŋ] ‘under the influence of black magic’
- (47) /n/ : /ŋ/
[piŋgan] ‘plate’ – [piŋgaŋ] ‘waist’
- (48) /n/ : /ɲ/
/na/ ‘here you go’ – [ɲa] ‘mother’
- (49) /dʒ/ : /g/
[dʒantaŋ] ‘heart’ – [gantaŋ] ‘hang’
- (50) /k/ : /g/
[kaja] ‘rich’ – [gaja] ‘splurge’
- (51) /k/ : /tʃ/
[kjam] ‘stingy’ – [tʃjam] – ‘divination stick’
- (52) /g/ : /ŋ/
[gaga] ‘daring’ – [ŋaŋa] ‘mouth agape’
- (53) /ŋ/ : /k/
[pəŋat] ‘sweet dessert of thick coconut milk with banana chunks’ – [pəkət] ‘thick (describing liquid)’
- (54) /ʔ/ : /ŋ/
[pulaʔ] ‘instead’ – [pulaŋ] ‘return’
- (55) /s/ : /h/
[so] ‘burn’ – [ho] ‘good’
- (56) /h/ : /ŋ/
[hantoʔ] ‘bang’ – [ŋantoʔ] ‘sleepy’
- (57) /ʔ/ : /k/
[baʔu] ‘smell’ – [baku] ‘standard’

A near minimal pair that may establish that /ʔ/ and /g/ are different is:

- (58) /ʔ/ : /g/

[baʎu] ‘smell’ – [bagus] ‘good’

3.1.1.2 Bilabials

There are three bilabials in BM, the voiceless plosive /p/, voiced plosive /b/, and the nasal /m/. The following provides the environments in which they are found. The items in grey are words of Hokkien origin. Note that Hokkien is an isolating language (Wright 1983), meaning that while it lacks inflectional morphology, it has many compound words comprising more than one syllable (See Bodman 1955 and 1958 for examples). Hokkien compound words are however taken directly into BM, and BM speakers who mostly do not speak Hokkien, are unable to identify the original individual components that make up these compound words. These words are often incorporated into BM as whole, monomorphemic items that cannot be further broken down, abiding by the phonotactics of BM words (see section 3.2 for discussion on phonotactics).

	Initial position	Medial position	Final position
p	[pokoʔ] ‘tree’	[kəpeʔ] ‘pinch’	[tʃukop] ‘enough’
	[peʔpeʔ] ‘father’s elder brother’	[peʔ.peʔ] ³¹ ‘father’s elder brother’	[tʃap] ‘ten’
b	[bawaŋ] ‘onion’	[tʃobeʔ] ‘long-jawed’	[dʒawab] ‘answer’
	[bapaʔ] ‘father’	[awban] ‘selfish’	[tʃanab] ‘decorative altar stand’
m	[masaʔ] ‘cook’	[gəmoʔ] ‘fat’	[malam] ‘night’
	[mja] ‘life’	[chut.mja] ‘famous’	[geʔsim] ‘unhappy’

Table 8: Bilabial consonants by position

³¹ Examples such as [ʔ.p] in [peʔ.peʔ] and [t.m] in [chut.mja] should not be regarded as true consonant clusters, as the consonants that appear adjacent to each other belong to separate syllables.

While the consonant [f] is noted to occur in [maʔaf] ‘forgiveness’, which is of Arabic origin, it only occurs in this one word, and may be an instance of code-switching, since the expression [ampun] is more commonly used to denote ‘forgiveness’ in BM.

The following are acoustic representations of these bilabials. The purpose of acoustic analysis is two-fold. In addition to being more precise than a purely auditory analysis, an acoustic analysis can form the basis of comparison in the future if changes occur to particular phonemes. All waveforms and spectrograms are generated using computer software Praat version 5.3.59 (Boersma and Weenink 2013). All recordings in this section are made with using a Zoom H4 recorder paired with either a Shure WH30XLR cardioid headset microphone or a Countryman E6 XLR omnidirectional earpiece microphone. For the purpose of making the characteristics of these phones comparable, speech from the same speaker is utilized. The words used in these instances are derived from naturally occurring forms in narratives including story-telling and conversational data. While phoneticians have been, and still are debating over whether it is preferable to use word-list data or naturally occurring data, this grammar maintains that it is important to understand language as a natural occurrence. Hence word-list data is not used. Although naturally occurring forms are used, it is possible to choose tokens whose acoustic characteristics of the phones in question can appear distinctly. For example, it may be more useful to look at plosives in the word initial position, so that the waveform captures information such as whether there is aspiration or not, whereas a nasal’s own formants show up better on the spectrogram when the nasal is between vowels. Tokens in stressed positions are also preferred, so that acoustic information on in the waveforms and spectrograms appear clearer. In general, BM is syllable-timed instead of stress-timed (see section 3.6), and thus in most instances, syllables may receive equal stress in a multi-syllabic word. In addition, it must be noted that these words are extracted from language documentation data that are recorded in natural environments, and there may be some background noise in some of the clips. The formant settings used for this section are – Maximum formant: 5000Hz, number of formants: 5, window length: 5 milliseconds, dynamic range: 40dB, dot size of 2.0mm. To facilitate comparison, all waveforms and spectrograms presented are normalized to 0.5 seconds in length. Formant dots are only used on spectrograms to illustrate nasal formants.

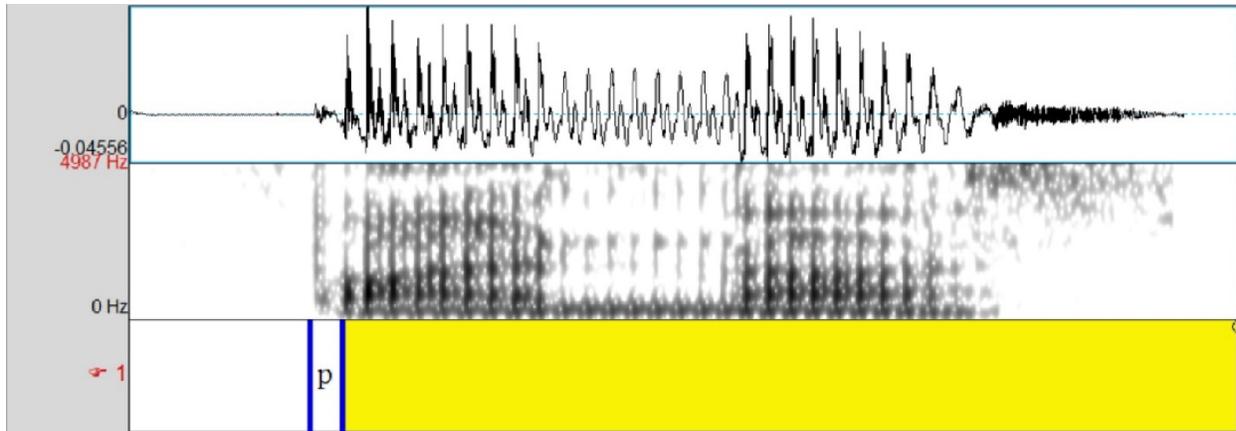


Figure 1: Waveform and spectrogram of /p/ in [panas] ‘hot’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:26.7-00:02:27.2)

Notice from the waveform that /p/ in BM has very little aspiration and the vowel begins quickly after a complete closure. The plosive is also voiceless as indicated by the lack of a voice bar in the corresponding area on the spectrogram. The locus of F2 and F3 (second and third formants) in the following vowel are also relatively low, as characteristic of preceding bilabials (Ladefoged and Johnson 2011).

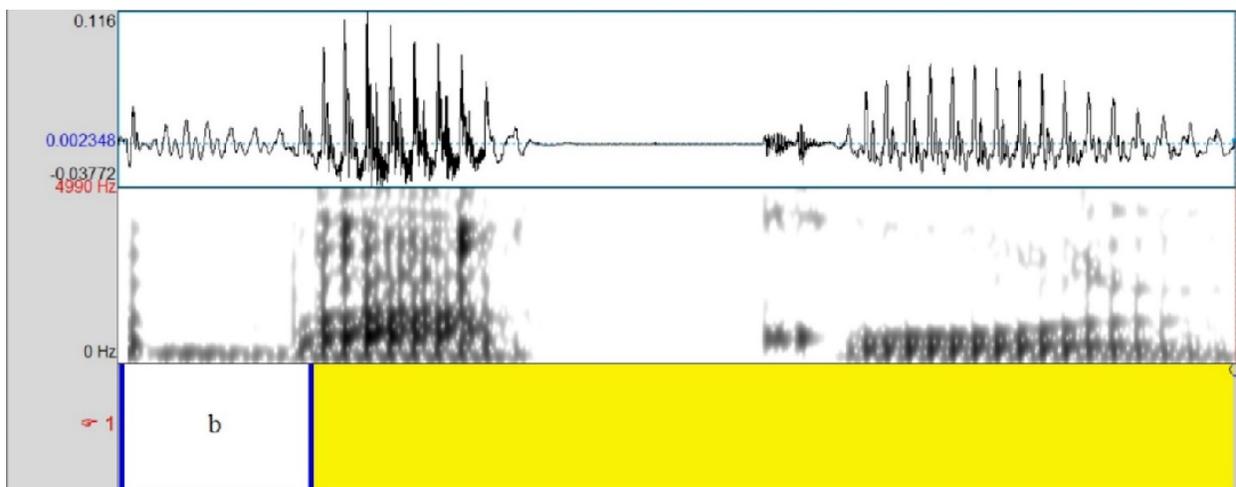


Figure 2: Waveform and spectrogram of /b/ in [bakol] ‘basket’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:54.4-00:00:54.9)

The corresponding voiced bilabial plosive of /p/ is /b/, the main difference being that there is no aspiration at all in this instance. Voicing is present during the closure period and there is a negative voice onset time. As expected, the locus of both F2 and F3 in the following vowel are also relatively low for this bilabial.

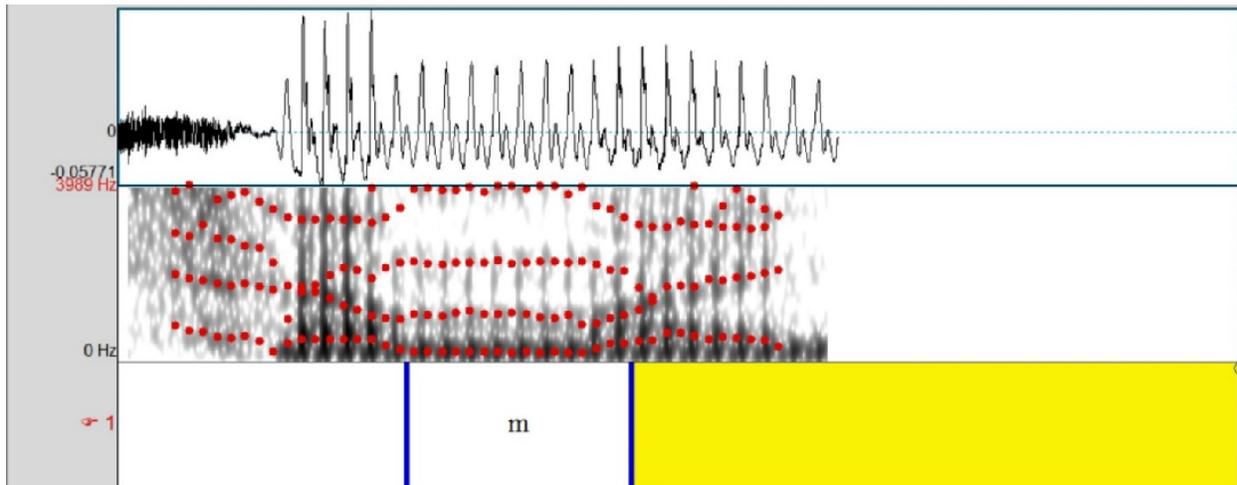


Figure 3: Waveform and spectrogram of /m/ in [sama] ‘same, and, with’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:01:43.0-00:01:43.3)

As characteristic for /m/ there is energy at the base at about 240Hz (hertz), at about 1000Hz, and at about 2300Hz. Again, the locus of F1 and F2 are relatively low for /m/ as expected, since it is a bilabial.

3.1.1.3 Alveolars

The alveolars found in BM are the voiceless plosive /t/, the voiced plosive /d/, voiceless fricative /s/, nasal /n/, lateral /l/ and flap /ɾ/. The following are environments in which they can be found.

	Initial position	Medial position	Final position
t	[taw] ‘know’	[rəti] ‘understand’	[lipat] ‘fold’
	[tu] ‘cupboard’	[kam.tioʔ] ‘develop an illness’	[tʃwe.it] ‘first day of the lunar month’
d	[da.own] ‘leaves’	[bodoh] ‘stupid’	
	[djam] ‘quiet’	[djam.djam] ‘quietly’	
s	[sədap] ‘delicious’	[kasi] ‘give, let, cause, PASS’	[ləkas] ‘quick’
	[sehjit] ‘birthday’	[tʃut.si] ‘birth’	
n	[naŋis] ‘cry’	[mənaŋ] ‘win’	[ta.own] ‘year’
	[njo] ‘mother-in-law’	[anchəŋ] ‘blessing of the marital bed’	[tʃin] ‘close to one another’
l	[lipat] ‘fold’	[kalu] ‘if’	[dəgil] ‘stubborn’
	[lotʃeŋ] ‘bell’	[haw.lam] ‘male mourner’	
r	[rəti] ‘understand’	[buoʔ] ‘ugly’	[kasar] ‘coarse’

Table 9: Alveolar consonants by position

The following are waveforms and spectrograms of BM alveolars.

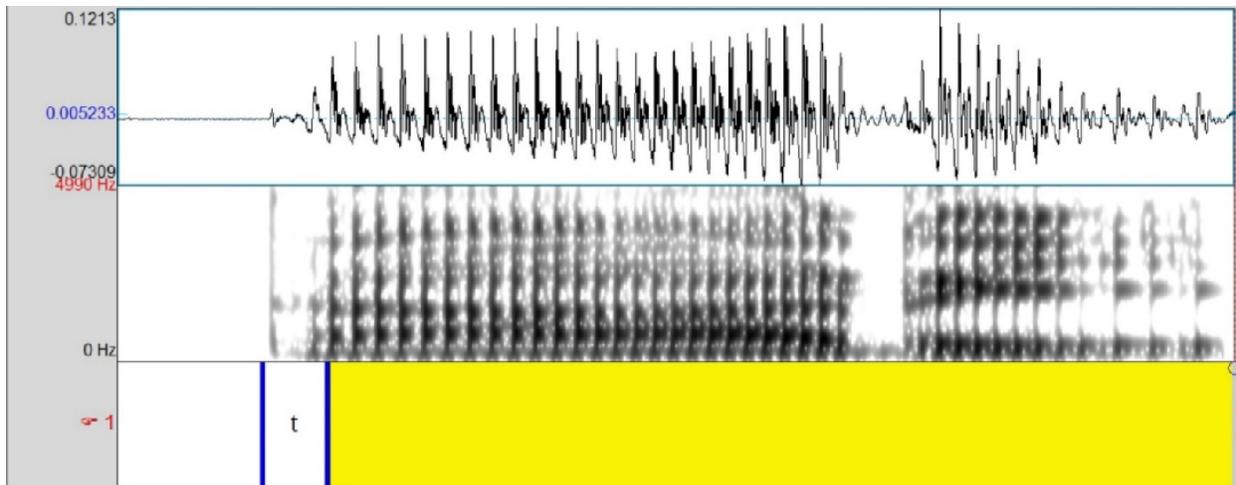


Figure 4: Waveform and spectrogram of /t/ in [tareʔ] ‘pull’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:54.7-00:01:55.2)

There is slight aspiration after the closure made for /t/ in the above instance, and minimal voicing. The locus of F2 is about 1700Hz, which characterizes alveolars (Ladefoged and Johnson 2011).

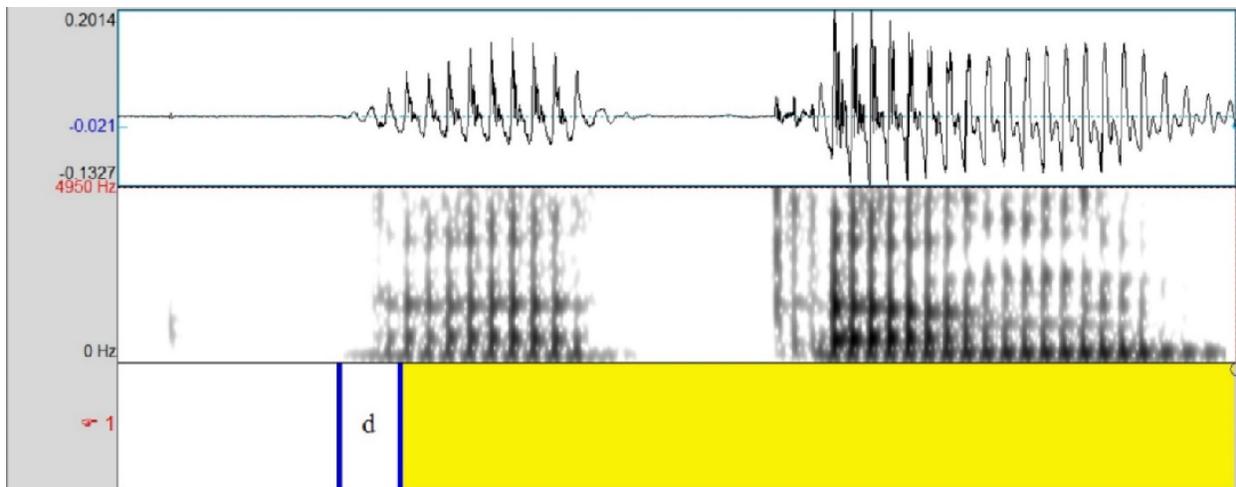


Figure 5: Waveform and spectrogram of /d/ in [datan] ‘come’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:14.5-00:02:15.0)

In this position, /t/ and /d/ do not look very different. Yet as compared to /t/, no aspiration is involved at all when producing /d/, and there a voice bar appears clearly in the spectrogram.

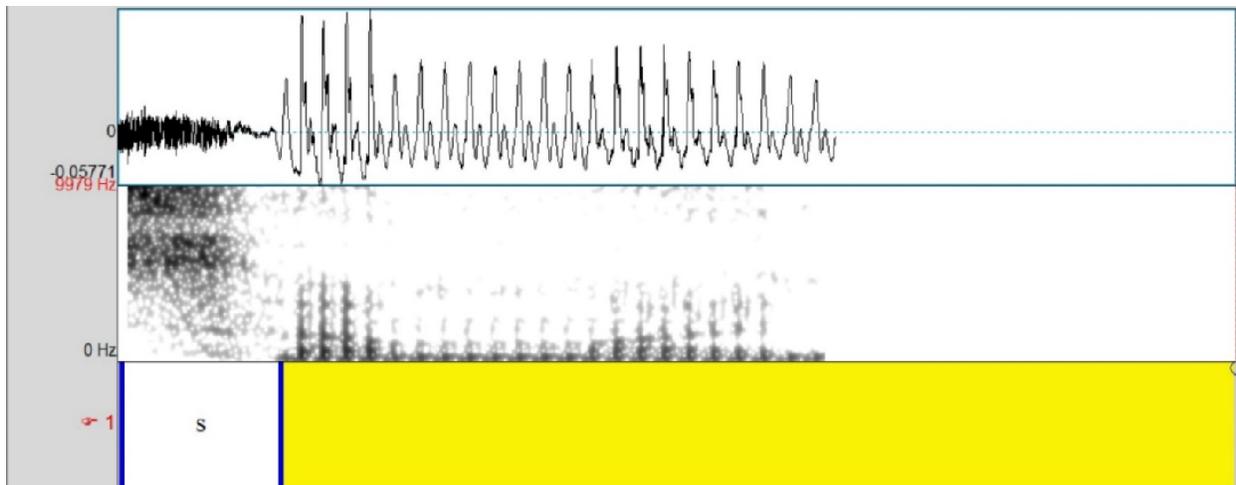


Figure 6: Waveform and spectrogram of /s/ in [sama] ‘same, and, with’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:01:43.0-00:01:43.3)

The /s/ fricative has a lot of energy in the spectrogram above 5200 Hz until 10,000 Hz or so. The fact that there is energy concentrated in the higher range (just under 10,000Hz) may indicate that /s/ is more fronted than typical alveolar sibilants. Frication noise can also be seen in the waveform. This range for the alveolar sibilant can be contrasted with the range expected for the palatoalveolar sibilant, which does not show energy above 7000Hz (Evers et al. 1998: 348).

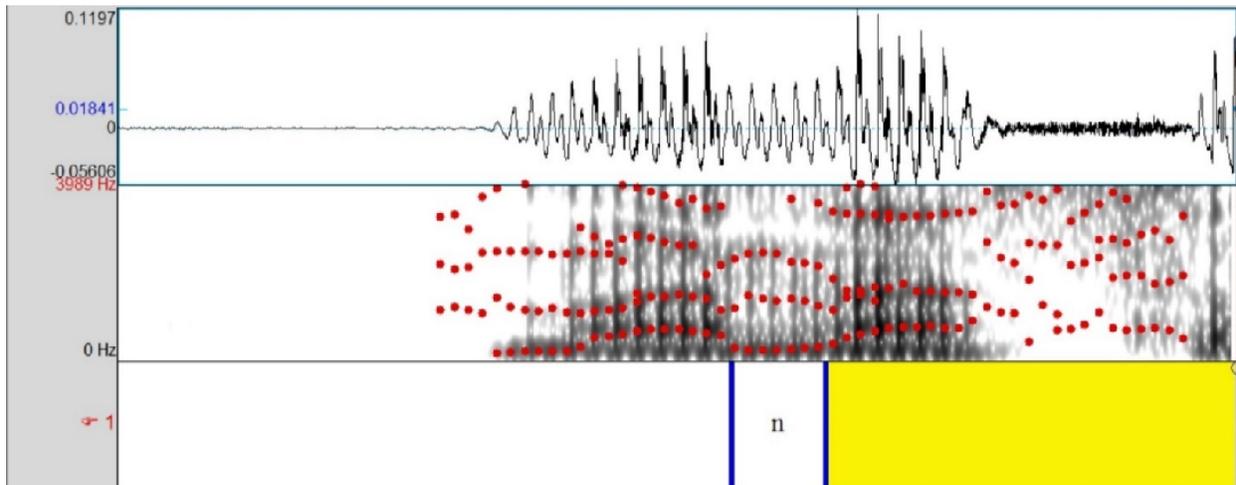


Figure 7: Waveform and spectrogram of /n/ in [mana] ‘where, which’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:26.3-00:01:26.8)

As with nasals at the initial position, faint formants can be seen, and there is energy at about 250Hz, 1300Hz and 2400Hz in this case. Formants are weak at the nasal murmur portion, nasal murmur comprising extra resonances that arise due to the nasal passage being used in addition to the oral tract. The nasal formant occurs weakly at about 300Hz.

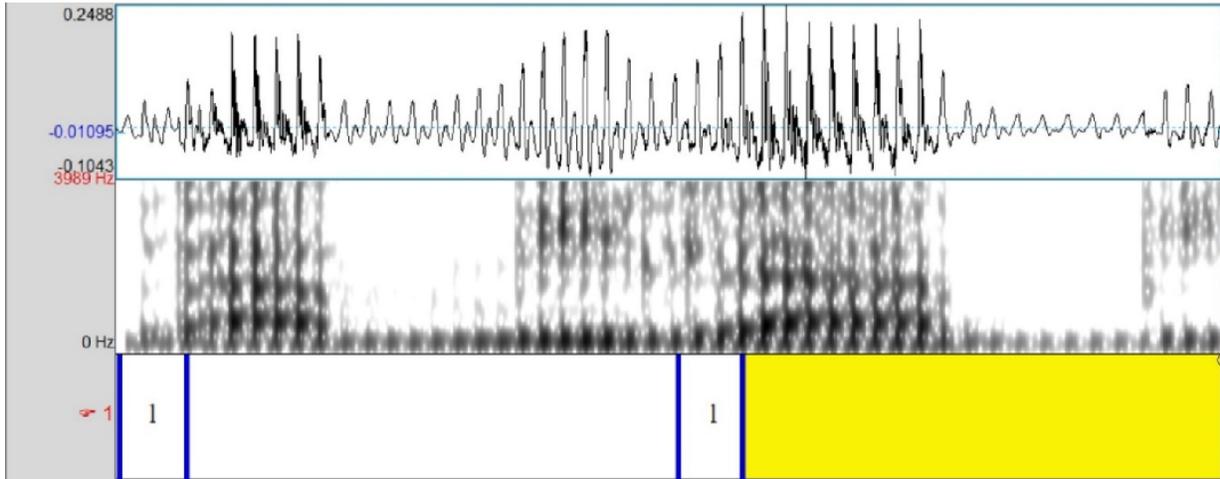


Figure 8: Waveform and spectrogram of /l/ in [labilabi] ‘tortoise’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:00:15.6-00:00:16.1)

When in an initial position, clear /l/ has a second formant at about 1400Hz.

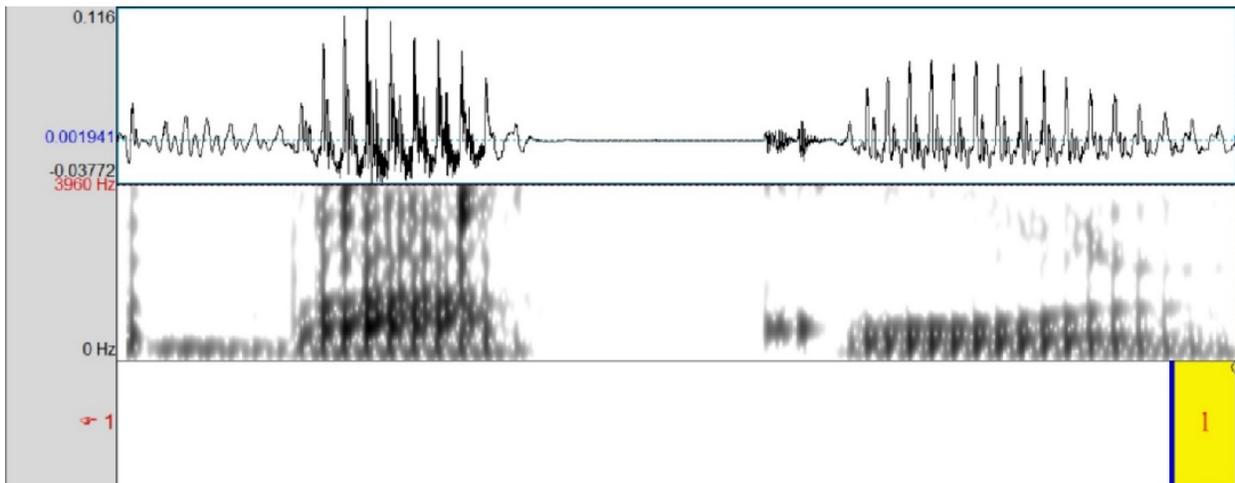


Figure 9: Waveform and spectrogram of /l/ in [bakol]

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:54.4-00:00:55.0)

There is no dark l in final position. Instead, /l/ appears to be vocalized where it occurs.

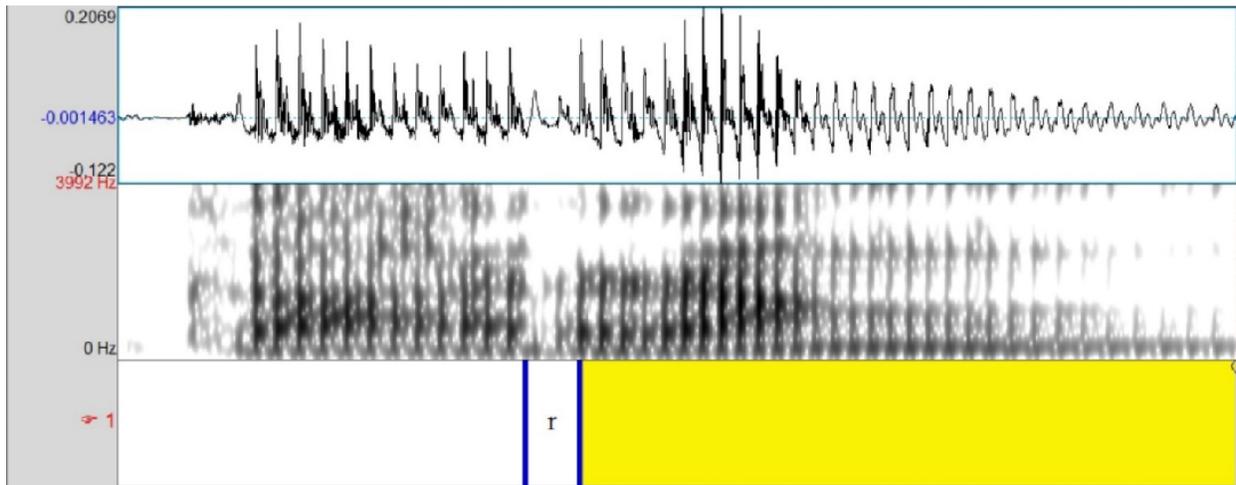


Figure 10: Waveform and spectrogram of /r/ in [garaŋ] ‘fierce’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:06.4-00:02:07.0)

There is very little rhoticity or r-coloring to /r/, with the relatively high third formant value of about 1600Hz (Ladefoged 2003).

3.1.1.4 Post-alveolars

The three post-alveolar affricates in the language are voiceless affricate /tʃ/, voiced affricate /dʒ/, and nasal /ɲ/.

	Initial position	Medial position	Final position
tʃ	[tʃam.por] ‘mix’	[bun.tʃit] ‘distended stomach’	
	[tʃaj.ki] ‘banner’	[lap.tʃaj] ‘wedding gift exchange ceremony’	
dʒ	[dʒum.pa] ‘meet’	[badʒu] ‘clothes’	
	[dʒi] ‘two’	[peʔ.dʒi] ‘eight characters for Chinese horoscope’	
ɲ	[ɲoɲa] ‘Peranakan lady’	[ɲoɲa] ‘Peranakan lady’	

Table 10: Post-alveolar consonants by position

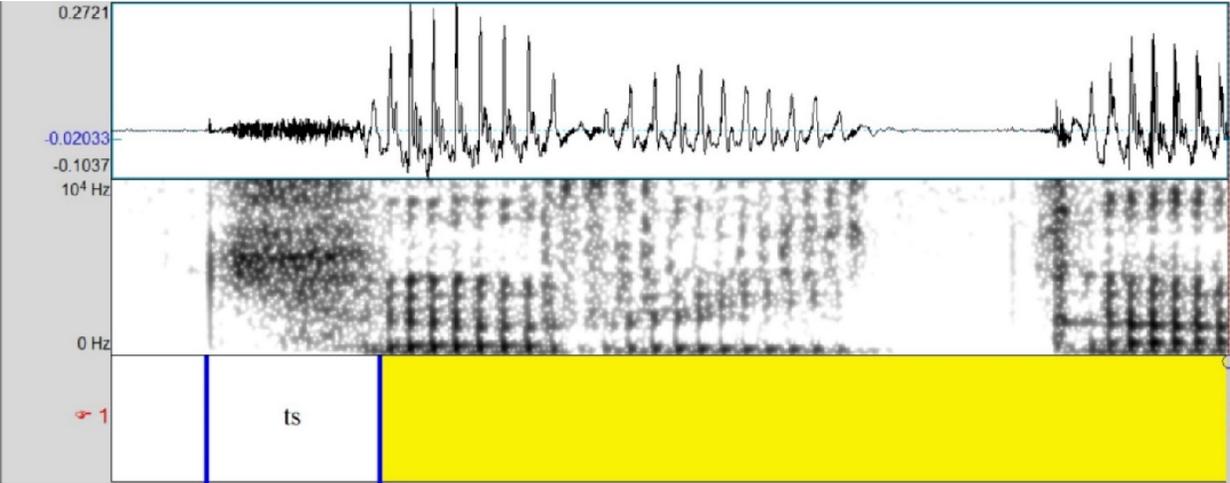


Figure 11: Waveform and spectrogram of /tʃ/ in [tʃərɪta] ‘story’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:00:18.3-00:00:19.2)

In the waveform above, there is clear frication that occurs after the closure with the production of affricate /tʃ/.

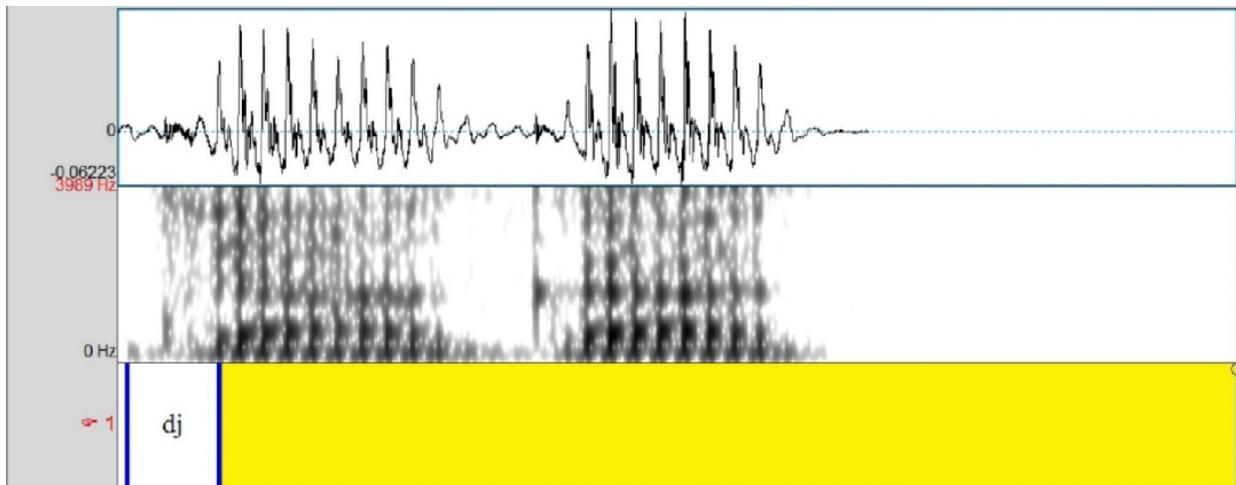


Figure 12: Waveform and spectrogram of /dʒ/ in [dʒaga] ‘guard, take care’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:04:06.3-00:04:06.6)

There is some prevoicing in the instance of /dʒ/, and there is clear frication in the waveform, albeit lesser than involved in the production of /tʃ/.

In addition to affricates, there is also nasal /ŋ/ at the post-alveolar place of articulation.

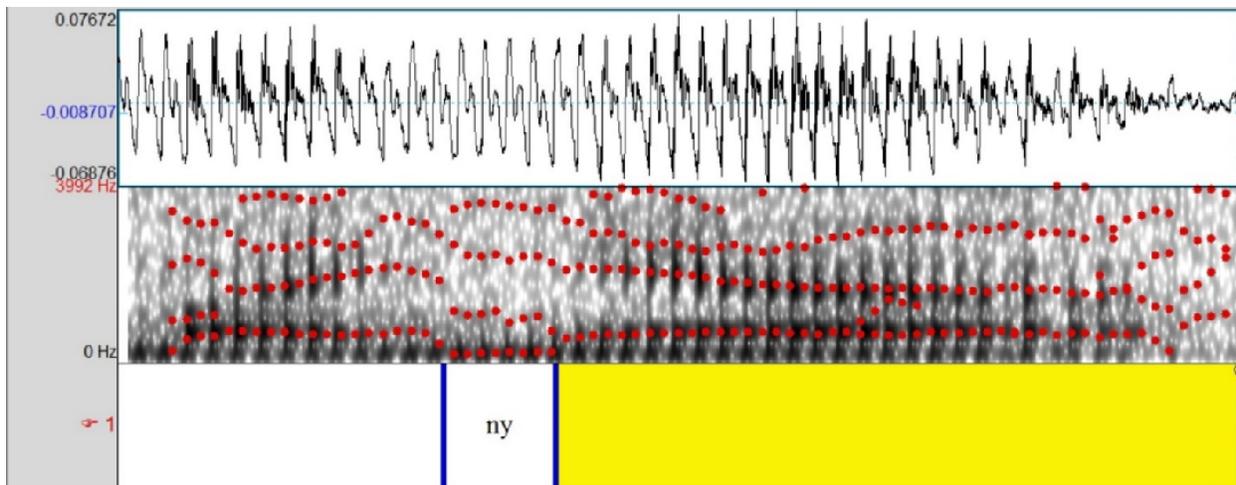


Figure 13: Waveform and spectrogram of /ŋ/ in [maŋaʔ] ‘many’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:19:41.5-00:19:42.4)

As is common with nasals universally, some formants can be seen between vowels for /ŋ/ in BM. Energy is concentrated at around 200Hz, 900Hz and about 2400Hz.

3.1.1.5 Velars

The velars in BM are voiceless plosive /k/, voiced plosive /g/, and nasal /ŋ/.

	Initial position	Medial position	Final position
k	[kalu] ‘if’	[ləkas] ‘quick’	
	[ko] ‘paternal aunt’	[laŋ.kéʔ] ‘guest’	
g	[gəmoʔ] ‘fat’	[paŋi] ‘morning’	
	[gwa] ‘1.SG’	[luŋ.gweʔ] ‘Chinese intercalary month’	
ŋ	[ŋantoʔ] ‘sleepy’	[naŋis] ‘cry’	[biŋ] ‘noisy’
	[ŋe.ŋe] ‘obstinate’	[laŋ.kéʔ] ‘guest’	[lotʃeŋ] ‘bell’

Table 11: Velar consonants by position

The following are acoustic representations of these velars.

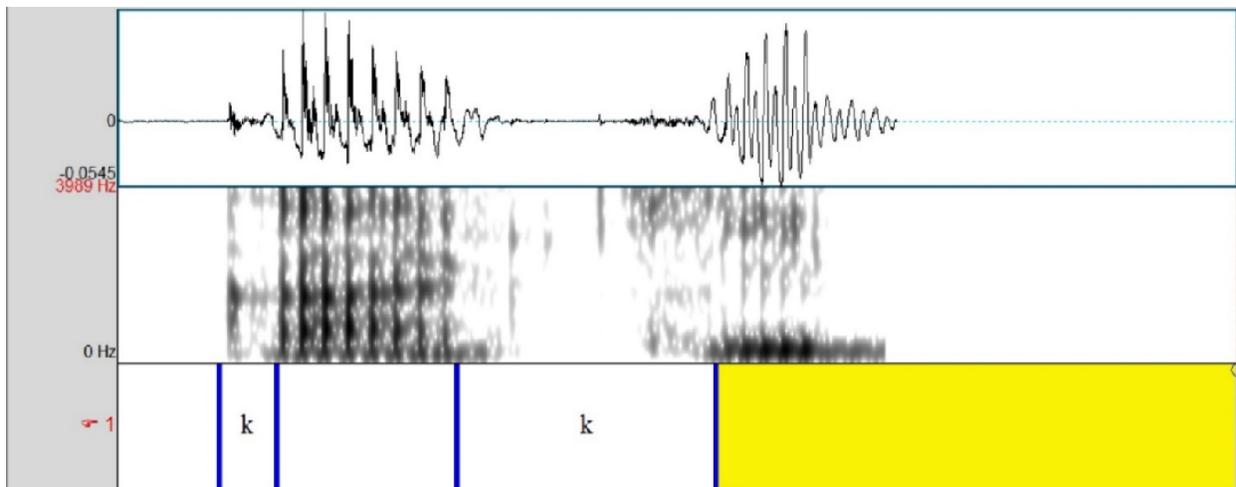


Figure 14: Waveform and spectrogram of /k/ in [kaki] ‘leg’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:06.4-00:02:07.0)

There is minimal aspiration following closure as observed from both initial and medial /k/ on the waveform. The velar pinch is more obvious with the intervocalic /k/. The velar pinch is where F2 and F3 appear to converge on the spectrogram.

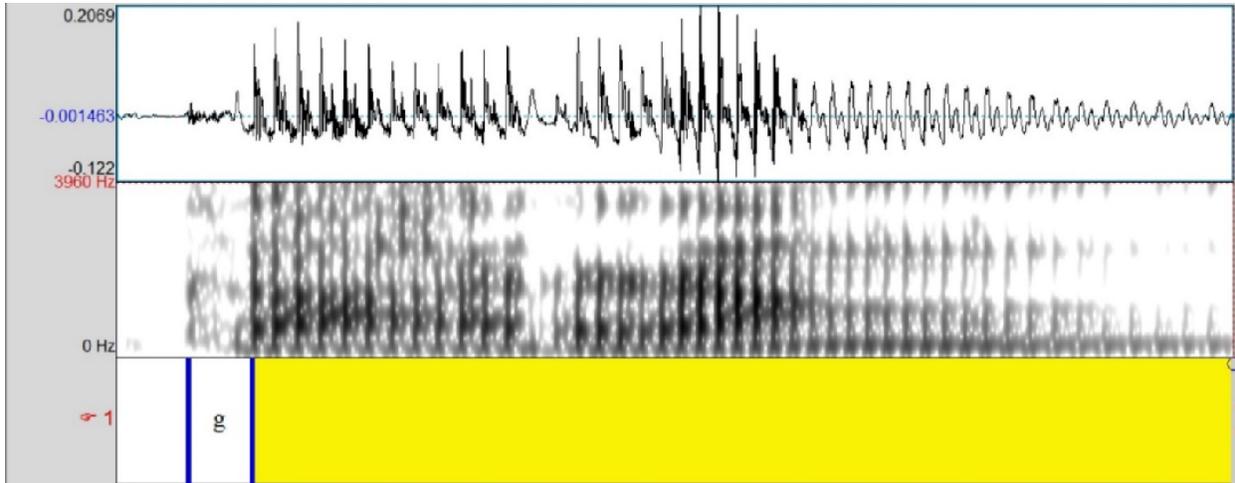


Figure 15: Waveform and spectrogram of /g/ in [garaŋ] ‘fierce’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:06.4-00:02:07.0)

It is interesting to note that /k/ and /g/ do not look acoustically very different, except for slightly more voicing associated with /g/. Similar to /k/, there also appears to be some aspiration where /g/ is concerned. The same spectrogram also demonstrates the acoustic characteristics of /ŋ/

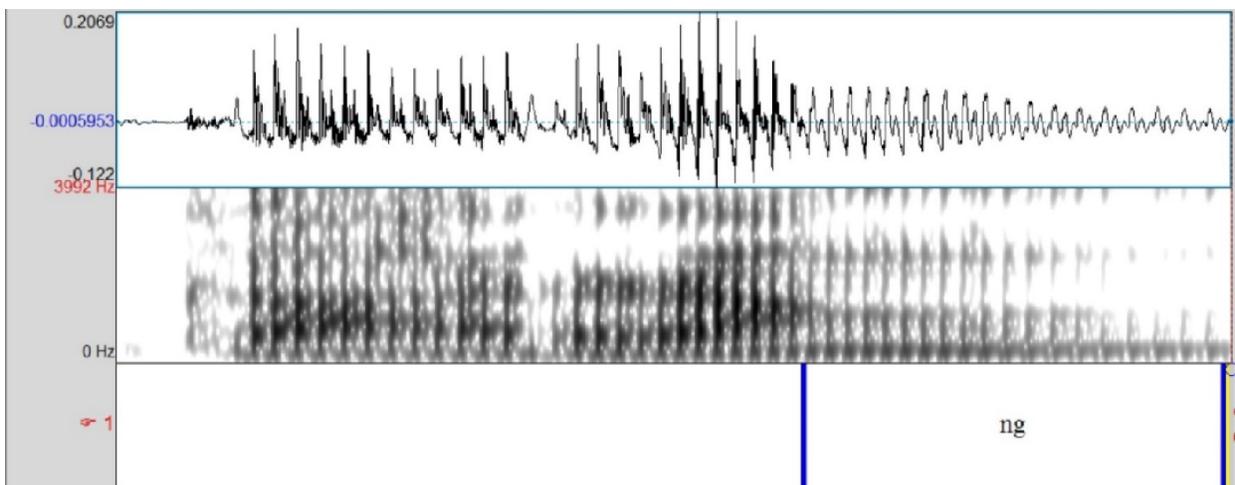


Figure 16: Waveform and spectrogram of [ŋ] in [garaŋ] ‘fierce’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:05:07.8-00:05:08.1)

Weak formants, which are characteristic of nasals, can be seen in the region of /η/. The velar pinch at 1200Hz shows that this is a velar nasal, and not an alveolar nasal.

3.1.1.6 Glottals

There are two glottals in BM, these being the voiceless plosive /ʔ/³², and the voiceless fricative /h/.

	Initial position	Medial position	Final position
ʔ		[a.ʔus] ‘thirsty’	[kəpeʔ] ‘pinch’
		[geʔ.sim] ‘unhappy’	[sin.keʔ] ‘newcomer’
h	[habis] ‘finish’	[dahi] ‘forehead’	[mən.tah] ‘raw’
	[hwan.tjoʔ] ‘meet ill spiritual forces’	[lihaj] ‘cunning’	

Table 12: Glottal consonants by position

The following are acoustic representations of these glottals.

³² It is worth noting that words with final consonants such as glottal stops, have been derived from both Malay and Hokkien. For example, [gəmoʔ] ‘fat’ is derived from Malay, while [peʔpeʔ] is derived from Hokkien.

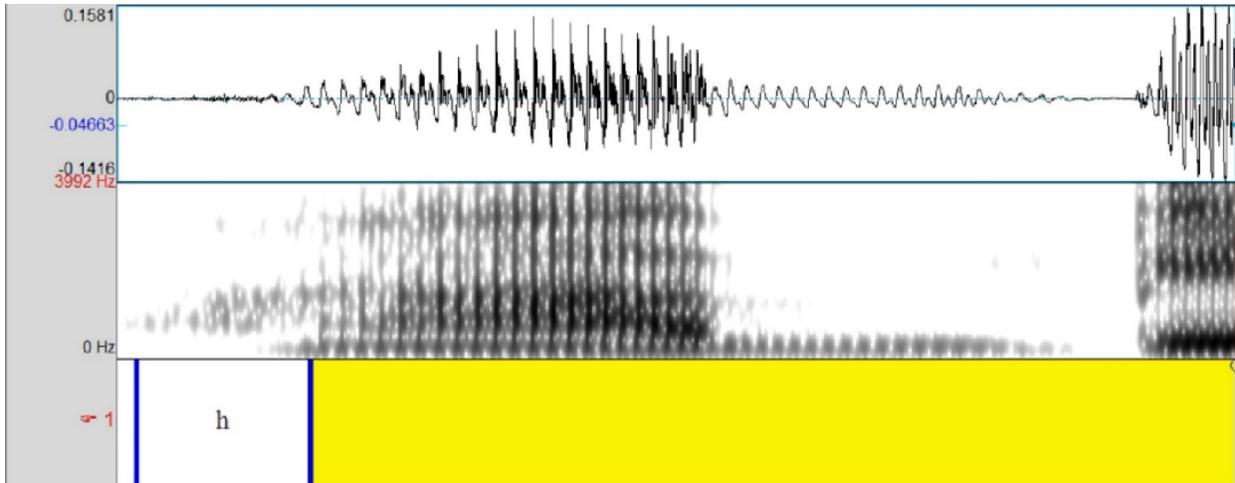


Figure 17: Waveform and spectrogram of /h/ in [habis] ‘finish’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:41.7-00:03:42.4)

Even though there is very little striation in the spectrogram due to its voiceless quality, glottal frication can be observed in the spectrogram with [h].

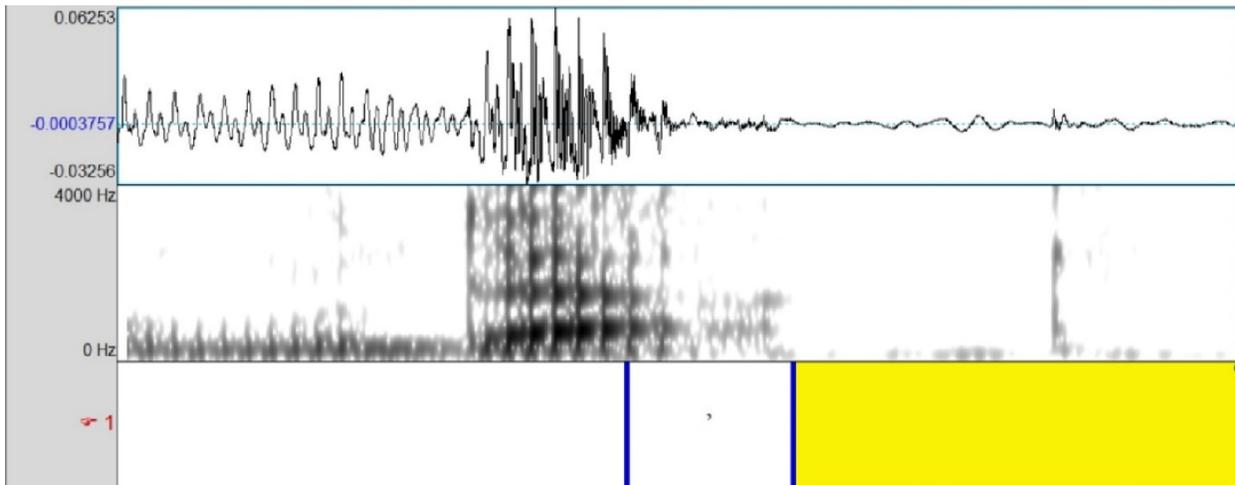


Figure 18: Waveform and spectrogram of /ʔ/ in [budaʔ] ‘child’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:54.1-00:005:54.6)

While the glottal plosive shows up more clearly between vowels, it is rare to find one in this position in BM. A word-final glottal stop is therefore used, and glottal frication shows up at the position of plosive /ʔ/.

3.1.1.7 Glides

There are two glides, the labiovelar approximant /w/ and the palatal approximant /j/.

	Initial position	Medial position	Final position
w	[waŋi] ‘fragrant’	[bawaŋ] ‘onion’	[taʋ] ‘know
	[wiɔ] ‘Lunar New year reunion dinner’	[kawiŋ] ‘marry’	[hoɾ.paʋ] ‘purse’
j	[jaŋ] ‘REL’	[səmajaŋ] ‘pray’	[paŋdaʋ] ‘clever’
	[jeŋ.tʃi] ‘rouge’	[tʃaʋ.jeŋ] ‘jelly’	[jaʋ.gwaʋ] ‘demon’

Table 13: Glides by position

These glides are represented acoustically in the following.

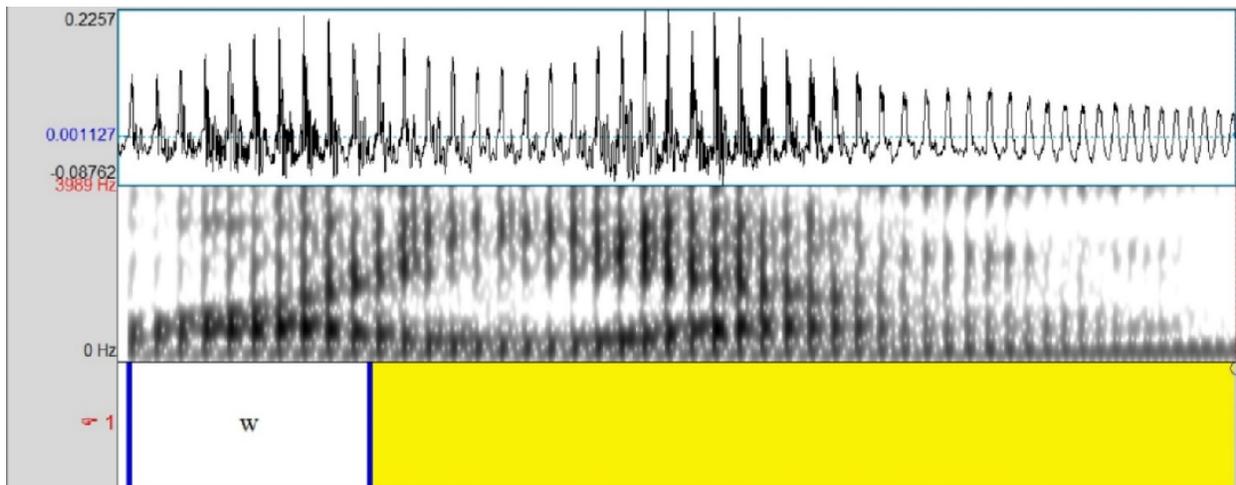


Figure 19: Waveform and spectrogram of /w/ in [wajaŋ] ‘play (performance)’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:11:22.4-00:11:23.0)

The labiovelar approximant /w/ appears to have formants but no steady state. There is a gradual dip in its F2, but it does not descend below 700Hz and based on that, it can be concluded that the lips were not very rounded (Ladefoged 2003). The acoustic characteristics of /j/ can be observed on the same spectrogram. This is replicated below for the convenience of readers.

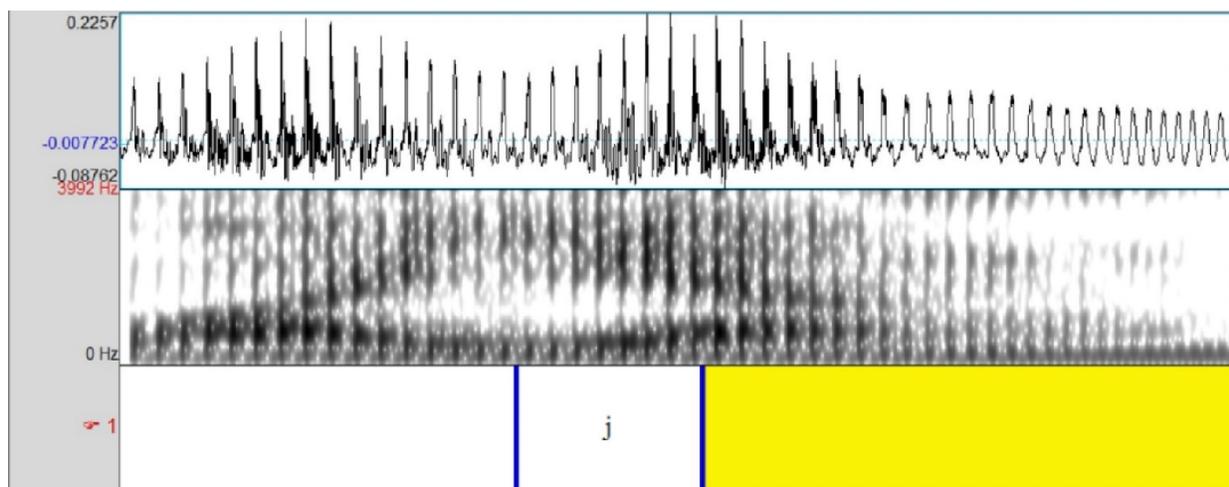


Figure 20: Waveform and spectrogram of /j/ in [wajan] ‘play (performance)’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:11:22.4-00:11:23.0)

Using the same word as above, [wajan], the palatal approximant [j] in this instance also shows vowel-like qualities except that it has no steady state. The tongue is at its highest front position at about 2100Hz, and is retracted and lowered as F2 falls.

3.1.2 Vowels

The 8 phonetic vowels are represented in the vowel chart below:

	front	central	back
	non-rounded	non-rounded	rounded
close	i		u
close-mid	e	ə	o
open-mid	ɛ		ɔ
open		a	

Table 14: Vowel chart of Singapore Baba Malay

Out of these vowels, /ɛ/ and /a/ are contrasting phonemes, and /a/ becomes [ɛ] before /l/, /r/ and /s/ (see section 3.3.5). Both do also form minimal pairs (see section 3.1.2.1).

3.1.2.1 Minimal pairs of vowels

The minimal pairs for the following vowel phonemes of the language are as follows:

- (59) /i/ : /e/
 [tapi] ‘but’ : [tape] ‘fermented rice dessert’
- (60) /i/ : /ə/
 [siram] ‘flush’ : [səram] ‘frightening’
- (61) /i/ : /u/
 [abi] ‘then, but’ : [abu] ‘ash’

- (62) /u/ : /ə/
[kunaŋ] ‘under the influence of black magic’ : [kənaŋ] ‘reminisce’
- (63) /u/ : /o/
[ku] ‘maternal uncle’ : [ko] ‘paternal aunt’
- (64) /o/ : /ə/
[kɔpɛʔ] ‘peel’ : [kəpɛʔ] ‘pinch’
- (65) /o/ : /a/
[ikot] ‘follow’ : [ikat] ‘tie’
- (66) /o/ : /ɔ/
[toʔ] ‘table’ : [tɔʔ] ‘poisonous, evil’
- (67) /ɔ/ : /ə/
[əŋ] ‘prosperous’ : [əŋ] ‘honorific prefix for familial relations’
- (68) /ɔ/ : /a/
[gantɔŋ] ‘hang’ : [gantəŋ] ‘a cylindrical measure of one gallon of rice’
- (69) /a/ : /ɛ/
[kena] ‘PASS’ : [kenɛ] ‘know (refined)’
- (70) /a/ : /ə/
[basi] ‘stale’ : [bəsi] ‘iron’
- (71) /a/ : /e/
[tʃobaʔ] ‘taste’ : [tʃobeʔ] ‘long-jawed’
- (72) /e/ : /ə/
[bedeʔ] ‘tell a lie’ : [bədeʔ] ‘pummel’

3.1.2.2 Monophthongs

Acoustic measurements are made of the vowels of the same proficient speaker whose consonants were acoustically represented in 3.1.1. Again, the data comes from naturally occurring forms in narratives including story-telling and conversational data, and tokens are chosen if their relevant acoustic characteristics appear distinctly. 10 tokens are measured for each vowel, and the resultant vowel plot in F1-F2 space is presented at the end of this section. It has been proposed that that vowel formants are acoustic correlates of vowel features, and are more representative of vowels rather than their articulatory properties. This is due to idiosyncratic differences in vowel articulation as well as inconsistencies between patterns of linguistic vowel height and frontness and measured tongue height and frontness during vowel production (Johnson 2012, Johnson et al. 1993, Ladefoged et al. 1972).

Formant values are read manually at the midpoint of the steady state of the vowel, as shown in figure 21. The settings used for this analysis of this speaker are– Maximum formant: 5000Hz, number of formants: 5, window length: 5 milliseconds, dynamic range: 40dB, dot size of 2.0mm. F1 and F2 values are derived for all vowels. F1 corresponds approximately to vowel height while F2 corresponds roughly to vowel frontness. The higher F1 is, the lower the vowel, the higher F2 is, the more front the vowel. F3 is primarily used for differentiating between vowels that are only distinguishable by lip-rounding (Ladefoged 2003), for example, the difference between [i] and [y]. There are neither front rounded vowels nor back unrounded vowels in BM, hence it is not necessary to measure F3. Analysis is carried out in Praat (Boersma and Weenink 2013). Again, where possible, the same speaker's vowels are presented. Tokens are not arbitrarily selected. Rather vowels are selected with the following preferred characteristics: Vowels in stressed positions are preferred, since it is difficult to read the formants of an unstressed vowel off a spectrogram. For the same reason, creaky voice data is discarded. In addition, it is preferable to use vowels that are not surrounded by consonants, as the formant frequencies of vowels can be affected by consonants in their immediate environment. This ensures that the measurements derived are purely that of the vowel, and not that of a consonant's effect on a vowel. In addition, if there are preceding or following consonant, /h/ initial and glottal stop finals are preferred since

these do not affect formant values as much as other consonants. For example, laterals lower the following formant values. However, due to the fact that many tokens are preceded by labials, these can be used consistently for different vowel measurements. Tokens outside of these environments are used only where there are not enough tokens to constitute the 10 required for each vowel. In addition, it is necessary to remember that these clips are taken from language documentation data. This inevitably means that there is some noise in the clips, but nothing that should obscure the vowel being analysed. The following figure demonstrates how the formants are measured for individual tokens.

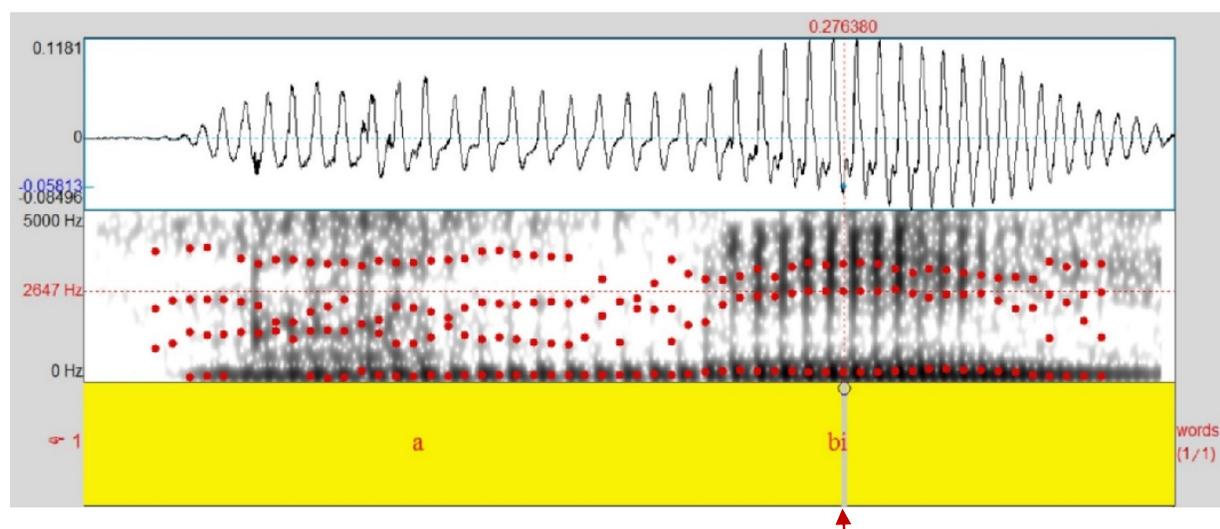


Figure 21: Waveform and spectrogram of /i/ in [abi] ‘then’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-30, 00:03:12.4-00:03:13.0)

In the spectrogram generated by Praat, the dots correspond to where the darkest horizontal stripes are. These dark horizontal stripes represent areas of energy, and also the formants that are to be measured. F1 is indicated by the lowest stripe, and F2 by the one above it. The arrow indicates where the vowel /i/’s formants is measured, at the midpoint of the vowel’s steady state. The measurements along the left axis of the spectrogram shows that the value of /i/’s F2 is 2647Hz (F1 whose value is not seen here is 282Hz). The relatively low F1 value and the relatively high F2 value show that /i/ is a high front vowel.

The following table shows the different F1 and F2 values generated for the different vowels. Where relevant, the preceding context across the word boundary is provided, since this potentially influences the formant values. This is important especially when spontaneous speech is being analysed. Individual spectrograms are not provided since formant values may vary widely between speakers and tokens.

No.	Vowel	Word(s)	Environment	F1	F2
1	i	labilabi	b_	286	2145
2	i	labilabi	b_	267	2165
3	i	habi	b_	286	2242
4	i	tapi	p_	286	2203
5	i	labilabi	b_	267	2300
6	i	labilabi	b_	267	2397
7	i	habi	b_	305	2591
8	i	təpi	p_	267	2203
9	i	babi	b_	286	2242
10	i	abi ilaŋ	i#_	267	2397
11	e	belo?	b_	402	2010
12	e	pile	l_	441	2203
13	e	pəte?	t_?	402	2107
14	e	tare?	r_?	402	2029
15	e	tfare?	r_?	421	2107
16	e	teŋo?	t_ŋ	383	2010

17	e	habes	b_s	383	2145
18	e	sampe	p_	421	2436
19	e	habes	b_s	344	2262
20	e	belo?	b_	402	2010
21	ε	laŋε	ŋ_	538	2107
22	ε	laŋε	ŋ_	538	2087
23	ε	kəre	r_	615	2242
24	ε	sε	s_	538	2149
25	ε	bərgətε	t_	538	2203
26	ε	kərdzε	dʒ_	557	2184
27	ε	dəŋε	ŋ_	576	2107
28	ε	hospitε	t_	596	1971
29	ε	hospitε	t_	615	2087
30	ε	gambe	b_	596	1855
31	ə	pəte?	p_	383	1448
32	ə	pəte?	p_	441	1526
33	ə	tərlaŋε	t_r	480	1545
34	ə	tərlaŋε	t_r	460	1545
35	ə	məntah	m_n	499	1506
36	ə	tərkəle?kəle?	t_r	441	1603
37	ə	tərkəle?kəle?	k_	480	1546
38	ə	tərkəle?kəle?	k_	383	1564
39	ə	təpi	t_	421	1564

40	ə	kəna	k_	480	1661
41	a	mata	m_	789	1255
42	a	mari	m_	770	1545
43	a	maso?	m_	789	1448
44	a	mja apa	a#_	770	1506
45	a	apa	p_	634	1255
46	a	babi	b_	809	1390
47	a	taw apa	w#_	731	1216
48	a	apa	p_	751	1235
49	a	sudah apa	h#_	673	1351
50	a	apa	p_	770	1371
51	u	kəbun	b_n	267	809
52	u	kəbun	b_n	247	906
53	u	buda?	b_	286	903
54	u	mulot	m_	247	925
55	u	mulot	m_	286	906
56	u	bukit	b_	247	925
57	u	buda?	b_	286	925
58	u	buda?	b_	286	867
59	u	buda?	b_	247	867
60	u	bulan	b_	305	886
61	o	poko?	p_	408	698
62	o	mo	m_	383	777

63	o	mo	m_	363	777
64	o	mo	m_	344	750
65	o	lotʃeŋ	l_	354	830
66	o	dʒatoh	t_h	407	803
67	o	toloŋ	t_	407	1015
68	o	mo	m_	354	750
69	o	mo	m_	354	803
70	o	katʃo	tʃ_	381	1094
71	ɔ	ala ɔraŋ	a_	468	964
72	ɔ	lambəŋ	b_ŋ	538	698
73	ɔ	teŋɔʔ	ŋ_ʔ	539	936
74	ɔ	lambəŋ	b_ŋ	518	645
75	ɔ	tarɔʔ	r_ʔ	513	1041
76	ɔ	lɔmpat	l_m	518	1068
77	ɔ	təŋkat	t_ŋ	538	990
78	ɔ	teŋɔʔ	ŋ_ʔ	576	1068
79	ɔ	təŋ	t_ŋ	513	989
80	ɔ	bəgi ɔraŋ	i#_	500	1068

Table 15: Vowels and formant values for a proficient speaker of Singapore Baba Malay

That [ɔ] exists in words derived from Malay is worthy of mention. This contrasts with Pakir's (1986:56) observations that [ɔ] exists only in loanwords. It is also important to note that there are [ɔ] vowels that are undergoing change, so a word such as [ɔraŋ] 'person' may be pronounced as [oraŋ] by others (see section 3.7 for more information regarding change in the vowel space).

The F1 and F2 values derived are presented in figure 22, which is a vowel plot with F1 values on the Y-axis and F2 values on the X axis.

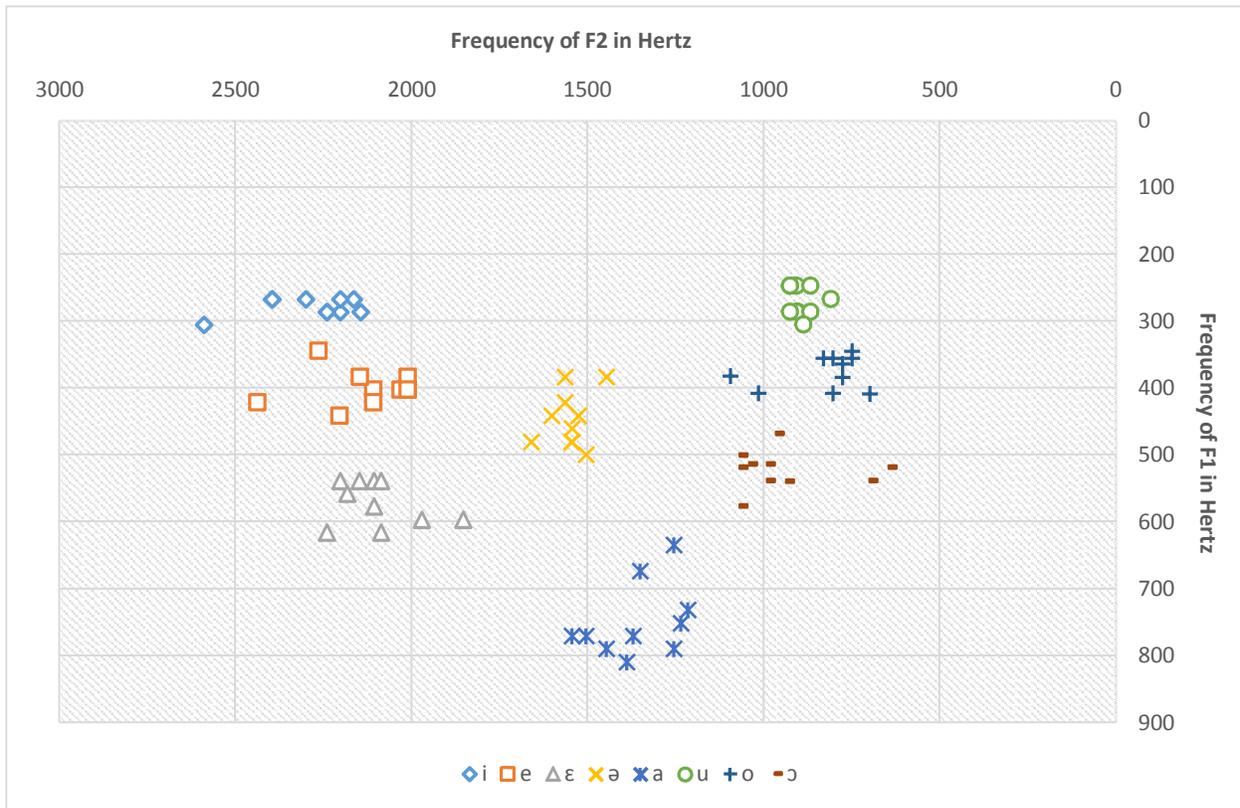


Figure 22: Vowel space of a Singapore Baba Malay speaker

The speaker whose vowels are represented in figure 22 is a proficient speaker, relatively conservative (See section 3.7), and the above chart is a good representation of the general vowel space in BM. While the above is an individual’s vowel chart, and may involve some level of idiosyncrasy, the vowel spaces of six speakers are contrasted in section 3.7.

3.1.2.3 Diphthongs

Diphthongs here refer from sequences of vowel and glide that occur within the same syllable, as compared to sequences of vocoids that occur broken across two separate syllables (see section 3.1.2.4). These have their own special category, as they can be distinguished from other consonant clusters that exist as a result of metathesis and fast speech (see section 3.2). This is in

fact similar to the treatment of diphthongs in Austronesian, where diphthongs are viewed to be combinations of vowel and semivowel (Dempwolff 1934-1938, Blust 1998), and such a treatment extends to many other languages. Diphthongs in BM can be divided into two types, those that only exist in words derived from Hokkien, and those that exist in words of Malay and Hokkien origins, among others. Words that are derived from Hokkien are marked with (H) in the following list. Note that these words may have been phonologically adapted into BM (see sections 3.3 and 3.3) and sound different from their original H counterparts.

(73) [aj]

Examples: [bə.kə.laj] ‘quarrel’, [tʃaj.təŋ] ‘nunnery that serves vegetarian food (H)’

(74) [aw]

Examples: [taw] ‘know’, [ɔ:paw] ‘purse (H)’

(75) [oj]

Examples: [amboj] ‘exclamation of surprise’, [bojʔ] ‘socks (H)’

(76) [ow]

Examples: [da.own] ‘leaves’, [bow] ‘NEG (H)’

3.1.2.4 Vowel sequences across syllables

The [a.o] vowel sequence occurs broken across two different syllables. These cannot be analyzed as diphthongs comprising vowel and glide, since each syllable should have its own nucleus, and each vowel in one of these vowel sequences form the nucleus of the different syllables. Note that there is no glottal stop in between [a] and [o] in these instances.

(77) [ao]

Examples: [da.own] ‘leaves’, [ta.own] ‘years’, [ga.o] ‘mix’

3.2 Phonotactics

The BM syllable structure is (C)(C)V(C)(C) with some restrictions. Onsets and codas are optional. Thus, syllables are of the type V, CV, CCV, VC, CVC, CCVC, and CVCC. However, no examples are found with VCC and CCVCC. The consonant clusters that occur within the same syllable have a glide as one of the consonants in the cluster. This is demonstrated by examples such as the CCV syllable [mwi] in [mwi.laŋ] ‘matchmaker’, the CCVC syllable [bwaŋ] ‘throw’, and the CVCC syllable [bajʔ] ‘good’. Other consonant clusters within the same syllable occur as a product of local metathesis or fast speech. For example, [kr] in [krə.dʒar] ‘work’, [gr] in [grə.taʔ] ‘threat’, and [tr] in [trə.baŋ] ‘fly’ are derived from [kər.dʒar]³³, [gər.taʔ], and [tər.baŋ] respectively. An example of a consonant cluster in fast speech occurs when [ə] is reduced, as in [bla.tʃu] for [bə.la.tʃu] ‘unbleached cotton outfit used for mourning’. In all these instances of underlying consonant clusters in a single syllable, at least one of the consonants is a glide [j, w] or a liquid [l, r]. In general, the glottal stop [ʔ] cannot occur in the onset of the first syllable, while affricates [tʃ] and [dʒ] cannot occur in the coda. The following are permutations found in BM, as well as the restrictions found on each permutation. Syllable boundaries are marked with ‘.’ Words of Hokkien origin are followed by (H).

(78) V

Examples: [a.deʔ] ‘sibling’, [o.pan] ‘freckles (H)’

All vowels can occur as the obligatory, individually-occurring nucleus that comprises the entire syllable.

(79) CV

Examples: [da.pat] ‘receive’, [kweh.ji] ‘glutinous rice balls (H)’

There are no restrictions on V in this sort of syllable, and the only restriction on C is that it cannot be a glottal stop. Glottal stops are only allowed in the onset of the second syllable, such as

³³ Note that whereas the BM versions of work are [kər. ʒar], [krə.ʒar] (coarse), and [krə.ʒɛ], the Malay version is [ker.ja].

in [ba.ʔu] ‘smell’ and in the word [pe.ʔe] ‘good character’, which is borrowed from Arabic, [fi.ʔil].

(80) CCV

Examples: [trə.ban] ‘fly (after metathesis), [mwi.lang] ‘matchmaker (H)’

There are no restrictions on V, whereas the restriction for all consonant clusters is that one of the consonants must be a glide, or a liquid in the case of metathesis or fast speech. In the case of CCV, the second consonant has to be a glide or a liquid. The glottal stop should also not occur in the onset of the first syllable.

(81) VC

Examples: [aŋ.kat] ‘lift, hold, pick up, carry’, [aw.ban] ‘selfish (H)’

Both glides and full consonants, except for affricates, can occur in the C position in VC. Note that it is more common for [ŋ] to follow a vowel as a coda than to precede it as an onset.

(82) CVC

Example: [ti.mun] ‘cucumber’, [joʔ.hun] ‘Chinese medicinal powder (H)’

The glottal stop is not permitted in the onset of the first syllable, while affricates are not permitted in the coda. An example of a glottal stop occurring in the second syllable is the word [a.ʔus] ‘thirsty’.

(83) CCVC

Example: [pə.rjoʔ]³⁴ ‘cooking pot’, [hwat] ‘expand (H)’

Again, the glottal stop is not permitted in the onset of the first syllable, while affricates are not permitted in the coda. The second consonant in the CCVC consonant cluster has to be a glide.

³⁴ Not [per.joʔ]

(84) CVCC

Examples: [bajʔ] ‘good’³⁵, [bojʔ] ‘socks (H)’

In CVCC syllables, the penultimate consonant must be a glide. Similarly, the glottal stop is not permitted in the onset of the first syllable, while affricates are not permitted in the coda.

Words are usually between one to three syllables in length, with two syllable words being the most common in the corpus used. All possible syllabic permutations are found in words of Malay origin, as well as words of Hokkien origin. Words borrowed from other languages into BM also follow the above phonotactic constraints. For example, Dutch [lamp] for ‘lamp’ is borrowed into BM as [lampu] as there are no consonant cluster codas in BM that do not comprise a glide or a liquid, and the Portuguese word [padri] is borrowed into the language as [padəri] ‘priest’, as consonant clusters with liquids are not common, unless as a product of metathesis or fast speech.

3.2.1 Resyllabification and reduplication

Other processes that concern syllable shape are resyllabification and reduplication.

Resyllabification also occurs with the use of suffixes. For example, when nominalizer /-an/ is attached to /ma.nak/ ‘many’ to derive a noun out of an adjective, the new word is /ma.na.kan/, and /k/ is no longer the coda of the penultimate syllable but it is the onset of the ultimate syllable. This is not a very productive process because the use of nominalizer /-an/ is optional, and so is the use of most affixes except for the transitive suffix /-kan/ (see section 4.2.1.1), which has no effects on resyllabification since it has its own onset. Reduplication is also not always productive in BM. Words that appear reduplicated may not be recognised as being composed of a shorter component that has meaning by itself, that is repeated in the reduplicated form by speakers, as with the following list of words.

³⁵ Originally from Dutch, borrowed into Malay.

- (85) labi-labi ‘turtle’
- (86) kupu-kupu ‘butterfly’
- (87) kura-kura ‘tortoise’
- (88) antiŋ-antiŋ ‘earrings’
- (89) gədəbaʔ-gədəbuʔ ‘onomatopoeia: thudding of the heart’
- (90) gədəbaŋ-gədəboŋ ‘onomatopoeia: loud noises’
- (91) embɛ-embɛ ‘half-cooked’

In the above examples, none of the supposed individual (unreduplicated) “components” are meaningful to BM speakers. *Labi* in (85) *labi-labi* ‘turtle’ does not mean a singular turtle, assuming that reduplication functions as a plural marker as with many other varieties of Malay, such as with Manado Malay, Ambon Malay, Kupang Malay (Paauw 2009), Bahasa Melayu, and Bahasa Indonesia, to name a few. In fact, on its own, *labi* does not appear to be accepted as a word. In varieties such as Bahasa Melayu and Bahasa Indonesia, *kupu* means ‘equal in social or familial position’, but is in no way related to (86) *kupukupu* ‘butterfly’. Similarly *kura* means ‘spleen’ in Bahasa Melayu and Bahasa Indonesia, but this is unrelated to (87) *kurakura* ‘tortoise’. The word *antiŋ* means ‘weight on a scale’ in the same language, and the word for earrings (88) *antiŋ-antiŋ* may have come from the fact that the long earrings worn might have looked like the weights on traditional balancing scales, but the word *antiŋ* does not exist in BM and any such historical connection there may be is unknown to speakers. Where (89) is concerned, the components *gədəbaʔ* and *gədəbuʔ* have no meaning on their own in BM. In related varieties such as Bahasa Melayu and Bahasa Indonesia, *gədəbuʔ* is the onomatopoeic sound for stamping and pounding. These seemingly reduplicated words in examples (85) to (89) appear to have been directly derived from other Malay varieties, in which these words exist. More discussion on which variety of Malay may have constituted the lexicon of BM ensues in Chapter 7.

There are also instances of these “reduplications” in BM that have not been derived directly from any source. In (90), *gədəbaŋ* and *gədəboŋ* have no individual meanings in BM, but *gədəbam* is the onomatopoeic sound made by falling things in Bahasa Indonesia. BM speakers might have analogized and innovated *gədəbaŋ-gədəboŋ* as an onomatopoeia for loud noises themselves,

following the template for *gədəbaʔ-gədəbuʔ*, in which the first component ends with an open low vowel followed by a velar, and the second component ends with a rounded, back and high vowel also followed by a velar. In other cases such as (91), neither component nor “reduplicated” word is known to exist in any other associated variety of Malay, and *embe-embe* ‘half-cooked’ appears to be unique to BM (*embe* on its own does not exist as a word).

In other instances, words are reduplicated meaningfully, some more productively than others:

- (92) *peʔ-peʔ* ‘father’s elder brother’ / *dʒi peʔ* ‘father’s second eldest brother’
- (93) *kim-kim* ‘mother’s brother’ wife / *tua kim* ‘mother’s eldest brother’s wife’
- (94) *adeʔ bəradeʔ* ‘siblings’ / *adeʔ* ‘sibling’
- (95) *rumah rumah* ‘houses’ / *rumah* ‘house’
- (96) *pəlan pəlan* ‘slowly’ / *pəlan* ‘slow’
- (97) *ləkas ləkas* ‘quickly’ / *ləkas* ‘quick’
- (98) *tauwε tauwε* ‘somewhat tasteless’ / *tauwε* ‘tasteless’
- (99) *asien asin* ‘somewhat salty’ / *asin* ‘salty’
- (100) *dʒalan dʒalan* ‘to take a walk’ / *jalan* ‘walk’
- (101) *matʃam matʃam* ‘like this and that’ / *matʃam* ‘seems, like, like this’

There are five patterns of reduplication observed in the data above. (93) and (94) are examples of kinship terms in BM that have been derived from Hokkien. The individual components on their own have the same meanings as the reduplicated components, for example, *peʔ* also means ‘father’s elder brother’, same as *peʔ-peʔ*, but these stems are never used on their own in BM. Rather, they are combined with Hokkien numerals to indicate the position of this relative in relation to oneself, thus *dʒi peʔ* would mean ‘father’s second eldest brother’, *dʒi* meaning ‘two’ (see Appendix A: Kinship terms).

Whereas reduplication for plurals is common in other varieties of Malay such as Bahasa Melayu and Bahasa Indonesia, reduplication for plurals is not common in BM. Plurals in BM are usually indicated by adding the word *manak* before the noun. Reduplicated plurals found in the corpus include (94) *adeʔ bəradeʔ* ‘siblings’, reduplicated from *adeʔ* ‘sibling’ and, and *rumah-rumah* ‘houses’, reduplicated from (95) *rumah* ‘house’.

Examples (96) to (101) are more productive and typical in BM. Adjectives can be reduplicated to become adverbs, as with example (96), where *pəlan* is ‘slow’ and *pəlan pəlan* ‘slowly’, and example (97), where *ləkas* is ‘quick’ and *ləkas ləkas* means ‘quickly’. Adjectives can also be reduplicated to express tentativeness or moderateness, as with (98) and (99), where *tauwe* means ‘tasteless’ and *tauwe tauwe* can mean ‘somewhat tasteless’, and *asin* means ‘salty’ whereas *asien asin* can mean ‘somewhat salty’. Similarly, tentative can be expressed by reduplicating other word classes. On its own, *dzalan* would mean ‘walk’ but (100) *dzalan dzalan* means ‘to take a walk’ or to stroll with no general goal or purpose (see section 5.2.5.7), and *matfam* means ‘seems, like, like this’ whereas (101) *matfam matfam* indicates ‘like this and that’, and nothing in general.

Phonologically, some of these reduplications appear to be interesting, for example, *ade? bərade?* ‘siblings’ and ‘*asien asin*’ but none of these patterns are productive in BM. Monosyllabic, bisyllabic and trisyllabic roots can all be reduplicated, with the monosyllabic forms being derived entirely from Hokkien.

3.3 Phonological rules

The following are phonological rules in BM, some of them being optional. It is important to note that these rules generate surface forms from underlying BM forms, and to not assume that the underlying forms will coincide with equivalent Malay words. BM is not a daughter or sister language of Malay (in the historical sense), but a contact language (see section 7.2). More importantly, BM is a separate, distinct language in its own right, and its phonology must be analyzed accordingly, as that of a distinct language. It is therefore not technically necessary nor appropriate to use the phonologies of either Malay or Hokkien as preliminary points of comparison for BM phonological facts. However, there are several systematic and non-systematic differences that should be noted.

Where words end with /aj/ and /aw/ word-finally in Malay, they end with /e/ and /o/ respectively in BM. These close-mid monophthongs have the same degree of frontness and backness as the semivowels /j/ and /w/ in the diphthongs /aj/ and /aw/. Thus, forms like /kədəj/ ‘shop’, /pandaj/ ‘clever’, /halaw/ ‘chase away’ and /pisaw/ ‘knife’ in Malay correspond to /kədə/, /pandə/, /halo/ and /piso/ in BM. This phenomenon does not constitute a phonological monophthongization rule in BM (where the diphthongs /ai/ and /au/ become close-mid monophthongs [e] and [o] with similar degrees of frontness and backness as the high vowels /i/ and /u/ in the diphthongs /ai/ and /au/) since there is no surface alternation that shows up with /ai/ or /au/. One exception in BM is the word /kalu/ ‘if’, which corresponds to Malay /kalaw/.

In a separate phenomenon, where vowel sequences of /a/ and /i/, and /a/ and /u/ are mediated with /h/ in Malay, they occur without /h/ as /aj/ and /aw/ sequences in BM. For example, /dzahit/ ‘sew’, /pahit/ ‘bitter’ and /tahu/ ‘know’ in Malay correspond to /dzajt/, /pajt/ and /taw/ in BM. Again, this does not constitute an intervocalic h-deletion rule and a vowel to glide rule (/ai/ to [aj] and /au/ to [aw]) in BM, since there are no surface alternations that show up with /ahi/ and /ahu/ in BM itself. An exception to this phenomenon is the word /mo/ ‘want’ in BM (instead of /mau/), which corresponds to /mahu’ in Malay.

In addition to understanding these systematic differences between BM and Malay, it is interesting to note how some BM words are different from equivalents in Malay in non-systematic ways. For example, BM /bawak/ corresponds to Malay /bawa/ ‘bring’, BM /məsti/ to Malay /misti/ ‘must’, BM /amek/ to Malay /ambil/ ‘take’, BM /tʃarek/ to Malay /tʃari/ ‘find’, and BM /muŋka/ to Malay /muka/ ‘face’. The forms of words such as these are just idiosyncratically different in BM, and there should be no attempt to derive such BM words from the Malay forms with which they are compared, but by the phonological rules that are necessary and appropriate for BM phonology.

The following tables provides distinctive features of individual segments, as distinctive features are used in this section. Inasmuch as is possible, non-redundant features are used.

	i	e	ɛ	a	ə	o	ɔ	u	j	w
syllabic	+	+	+	+	+	+	+	+	-	-
high	+	-	-	-	-	-	-	+	+	+
low	-	-	-	+	-	-	+	-	-	-
back	-	-	-	+	+	+	+	+	-	+
round	-	-	-	-	-	+	+	+	-	+
tense	+	+	-	+	-	+	-	+	-	-
approximant	-	-	-	-	-	-	-	-	+	+

Table 16: Distinctive features of vowels and glides in Singapore Baba Malay

In a strict view, where economy is valued, BM would not require all these features. For example, the feature [approximant] is redundant in this table of distinctive features of vowels and glides, as glides can be specified as [-syllabic]. However, since this feature captures a natural class of glides and liquids (which are represented in table 16), it is included both tables 16 and 17.

	p	b	m	t	d	s	n	l	r
syllabic	-	-	-	-	-	-	-	-	-
voice	-	+	+	-	+	-	+	+	+
back	-	-	-	-	-	-	-	-	-
coronal	-	-	-	+	+	+	+	+	+
anterior	+	+	+	+	+	+	+	+	+
lateral	-	-	-	-	-	-	-	+	-
nasal	-	-	+	-	-	-	+	-	-
continuant	-	-	-	-	-	+	-	+	+
approximant	-	-	-	-	-	-	-	+	+

	tʃ	dʒ	ɲ	k	g	ŋ	ʔ	h	
syllabic	-	-	-	-	-	-	-	-	
voice	-	+	+	-	+	+	-	-	
back	-	-	-	+	+	+	-	-	
coronal	+	+	+	-	-	-	-	-	
anterior	-	-	-	-	-	-	-	-	
lateral	-	-	-	-	-	-	-	-	
nasal	-	-	+	-	-	+	-	-	
continuant	-	-	-	-	-	-	-	+	
approximant	-	-	-	-	-	-	-	-	

Table 17: Distinctive features of consonants in Singapore Baba Malay (not including glides)

The following phonological rules are found in BM.

3.3.1 Syllable-final velar plosive to glottal stop rule

$$(102) \begin{bmatrix} - \text{syllabic} \\ + \text{back} \\ - \text{nasal} \end{bmatrix} \rightarrow \begin{bmatrix} - \text{back} \\ - \text{coronal} \\ - \text{anterior} \\ - \text{continuant} \end{bmatrix} / _ \$$$

$$g \rightarrow ? / _ \$$$

$$k \rightarrow ? / _ \$$$

In syllable-final position, velar plosives become a glottal stop obligatorily. Otherwise, it surfaces as /k/ or /g/. For example, /k/ remains /k/ in [maŋak-an] ‘many-NMZ’ and /g/ remains /g/ in [gasa?] ‘guess’, where syllable-final velar plosives becomes [?].

Underlying forms	beg ‘bag’	pekpek ‘father’s elder brother’	masok ‘enter, put in’	gasak ‘guess’	maŋak-an ‘many- NMZ’	bikin ‘make, do’
syllable-final plosive to glottal stop rule	be?	pe?pe?	maso?	gasa?	maŋakan	bikin

3.3.2 Deletion of word-initial *h* rule

$$(103) \begin{bmatrix} - \text{coronal} \\ + \text{continuant} \end{bmatrix} \rightarrow (\emptyset) / \# _$$

$$h \rightarrow (\emptyset) / \# _$$

Word initially, /h/ is deleted optionally.

Underlying forms	haʔus 'thirsty, worn-out'	hudzan 'rain'	hitam 'black'	mahal 'expensive'	dʒahat 'evil'	bohɔŋ 'lie'
word initial h-deletion rule	aʔus	udʒan	itam	mahal	dʒahat	bohɔŋ

This rule does not apply to all Hokkien-derived words with initial [h], for example, [hawlam] for 'female mourners', [hauli] 'male mourners', and [huahi] 'happy'. For words derived from Malay, the rule is optional; speakers produce both forms with and without [h] (speakers can produce both [udʒan] and [hudʒan] for 'rain'), although the general consensus is that the forms that lack [h] are more BM, since the forms with [h] exist in the standard Bahasa Melayu as well.

3.3.3 Metathesis of ə and r rule

$$(104) \begin{bmatrix} + \text{syllabic} \\ - \text{high} \\ - \text{low} \\ + \text{back} \\ - \text{round} \end{bmatrix} \begin{bmatrix} - \text{syllabic} \\ - \text{lateral} \\ + \text{approximant} \end{bmatrix} \Rightarrow \left(\begin{bmatrix} - \text{syllabic} \\ - \text{lateral} \\ + \text{approximant} \end{bmatrix} \begin{bmatrix} + \text{syllabic} \\ - \text{high} \\ - \text{low} \\ + \text{back} \\ - \text{round} \end{bmatrix} \right) / _ [-\text{syllabic}]$$

$$\text{ər} \Rightarrow (\text{rə}) / _ \text{C}$$

Metathesis occurs as /ər/ becomes [rə]. This only occurs before a consonant. It is also reasonable to assume that laterals following /ə/ might cause metathesis too, since both /r/ and /l/ are liquids. However, there are no instances of /l/ followed immediately by a consonant in BM, and hence it is not necessary to further define the liquid by including [-lateral]. This rule is optional, although the word for work, [krədʒa] appears consistently as such in the corpus. /ər/ is postulated as the underlying segment and not [rə], because when the segment /ər/ occurs in more general conditions elsewhere, such as in bərat 'heavy'.

Underlying forms	gərdʒa 'church'	kərdʒa 'work'	bərsi 'clean'	tərbəŋ 'fly'	bərat 'heavy'	pəletʃok 'twist foot'
ər to rə - metathesis rule	grədʒa	krədʒa	brəsi	trəbəŋ	bərat	pəletʃok

3.3.4 Assimilation to *u* before *lw* rule

$$(105) \left[\begin{array}{c} + \text{syllabic} \\ \end{array} \right] \rightarrow \left(\left[\begin{array}{c} + \text{high} \\ + \text{round} \end{array} \right] \right) / _ \left[\begin{array}{c} + \text{lateral} \\ \end{array} \right] \left[\begin{array}{c} - \text{syllabic} \\ + \text{round} \end{array} \right]$$

$$V \rightarrow (u) / _ lw$$

Optionally, vowels before /lw/ can assimilate towards the rounded, back and high /w/. This anticipatory assimilation is triggered specifically by a subsequent /lw/ and not by an intervening lateral, /lj/ or any IV environment.

Underlying forms	səlwar 'pants'	kəlwar 'go out, take out'	dilwar 'outside'	kalu 'if'	halja 'ginger'	katʃua 'cockroach'
u- assimilation rule	sulwar	kulwar	dulwar	kalu	halia	katʃua

3.3.5 Vowel-raising to *ɛ* and final *l*, *r*, *s* loss in refined style rules

Word-finally, /al/, /as/, and /ar/ optionally become [ɛ], as noted by Pakir (1986). Whether /al/, /as/, and /ar/ becomes [ɛ] depends however, on whether the speaker intends to come across as being 'refined' or 'coarse'. More details are provided in section on 3.7 on variation and change.

There are two rules that have to be posited for refined forms, since it is phonologically more plausible and preferable for phonological rules to change only a single segment at a time. These rules are identified in (106) and (107).

$$(106) \begin{bmatrix} + \text{syllabic} \\ + \text{low} \\ - \text{round} \end{bmatrix} \rightarrow \left(\begin{bmatrix} + \text{syllabic} \\ - \text{back} \\ - \text{tense} \end{bmatrix} \right) / _ \begin{bmatrix} + \text{anterior} \\ + \text{continuant} \end{bmatrix} \#$$

$$a \rightarrow (\epsilon) / _l \#$$

$$a \rightarrow (\epsilon) / _r \#$$

$$a \rightarrow (\epsilon) / _s \#$$

In the vowel raising rule, /a/ becomes [ε] when followed by /l/, /r/, and /s/, although this only occurs for refined forms.

Underlying forms	kenal 'know (a person)'	tampal 'mend'	panas 'hot'	nanas 'pineapple'	benar 'logical'	kasar 'coarse'
vowel raising rule	kəneɪ	tampeɪ	paneɪ	naneɪ	bəneɪ	kaseɪ

$$(107) \begin{bmatrix} + \text{anterior} \\ + \text{continuant} \end{bmatrix} \rightarrow \emptyset / \begin{bmatrix} + \text{syllabic} \\ - \text{back} \\ - \text{tense} \end{bmatrix} _$$

$$l \rightarrow \emptyset / \epsilon _ _$$

$$r \rightarrow \emptyset / \epsilon _ _$$

$$s \rightarrow \emptyset / \epsilon _ _$$

The vowel raising rule feeds the rule that deletes final /l/, /r/ and /s/ that follow [ε].

Underlying forms	kenal 'know (a person)'	tampal 'mend'	panas 'hot'	nanas 'pineapple'	benar 'logical'	kasar 'coarse'
final r, l, s - deletion rule	kəne	tampe	pane	nane	bəne	kase

There is idiosyncratic variation when it comes to the vowel [ɛ]. The transitive marker is usually pronounced [kan]. However, it is expressed as [kɛn] by an individual as a refined version of the transitive marker. Thus, the speaker produces [bɔrsi -kɛn] as a refined version of ‘clean -TRAN’ and [dʒato -kɛn] as a refined version of ‘fall -TRAN’, with the notion meaning to drop something. He deems the [kan] version as being coarse, and avoids it, opting to use the refined [kɛn] version mostly.

3.3.6 Rule ordering

There are two sets of rule ordering.

3.3.6.1 Vowel-raising, then loss of final *l*, *r*, and *s* rule

The vowel raising rule feeds the final *l*, *r*, *s* -deletion rule, since the final *l*, *r*, *s* segments are only deleted when they follow vowel [ɛ], which is created from the vowel raising rule.

Underlying forms	nanas ‘pineapple’	pasar ‘market’	kenal ‘know’
vowel raising rule	nanɛs	pasɛr	kenɛl
final <i>l</i> , <i>r</i> , <i>s</i> -deletion rule	nanɛ	pasɛ	kenɛ
Surface forms	nanɛ	pasɛ	kenɛ

The feeding order would be disrupted if the rules are applied in the wrong sequence, and there would be no change to the surface form of words such as /nanas/ ‘pineapple’, /pasar/ ‘market’, and /kenal/ ‘know’.

3.3.6.2 *u*-assimilation, then vowel-raising, followed by loss of final *l*, *r*, and *s* rule

Vowel-raising and loss of final *l*, *r*, and *s* following raised vowel takes place only after assimilation to *u*.

Underlying forms	kəlwar ‘go out, take out’	səlwar ‘pants’
u-assimilation rule (Surface forms)	kulwar	sulwar
vowel-raising rule	kulwɛr	sulwɛr
final r, l, s -deletion rule	kulwɛ	sulwɛ
Surface forms	kulwɛ	sulwɛ

This is due to the fact that non-refined forms are also observed to undergo optional u-assimilation, so that forms such as [sulwar] and [kulwar] exist in addition to [səlwar] and [kəlwar] as well for non-refined forms. Vowel raising followed by the loss of final l, r, and s are optional rules that makes forms such as [səlwar] and [kəlwar] refined. Otherwise, surface forms such as [kəlwɛr] for ‘go out, take out’ and [səlwɛr] are expected.

3.4 Fast speech

In fast speech, speakers tend to use contracted lexical items, and omit certain vowels, and join up words, so that the original word boundaries are lost. Fast speech phenomena are presented fairly extensively by Pakir (1986).

Certain contracted forms that are not phonologically rule-governed are *swa* for *sudah* ‘already’, *kat* for *dəkat* ‘near, at’, *sapa* for *sjapa* ‘who’, *pi* for *pərgi* ‘go’, *ni* for *ini* ‘this’, and *tu* for ‘itu’ ‘that’. These contracted forms are usually function words or words that are very commonly used, so that very little context is required to understand what these words refer to. Among these contracted forms, *dʒoraŋ* is an interesting contraction for *dia-orang* meaning ‘they’. The word initial *d* appears to be undergoing palatalization, brought about by anticipatory articulation before *i* with the sequence *jao* reducing to just *o*. This is not a commonly-used contracted form

though, and in fact, whether it should be represented in writing for the stage has been a source of dispute among playwrights.³⁶

The vowel ə is also often reduced in fast speech in BM (Pakir 1986). This occurs usually on the first syllable, so that *səkali* ‘very’ becomes *skali*, *sərono* becomes *srono* ‘proper’, *kəriŋ* ‘dry’ becomes *kriŋ*, *bərapa* ‘how many’ becomes *brapa*, *tʃəlop* ‘dip in dye’ becomes *tʃlop*. In particular, there is a preference for ə to be omitted if it is followed by a liquid, but this is not always the case, as with *səkali* and *skali*. This appears to be connected to the optional /əɾ/ to [ɾə] metathesis rule, although the metathesis rule has no effect on the number of syllables. For example, *bərsi* ends up optionally as *brəsi* ‘clean’, both forms having the same number of syllables. In the instances of ə reduction in fast speech, the number of syllables is reduced. For example, whereas there are two syllables in *kəriŋ* ‘dry’, there is only one syllable in *kriŋ*.

The other phenomenon that is common in fast speech is that word-finally, liquids and glottal fricative can be omitted. For example, in their coarse forms, *kapal* ‘ship’ becomes *kapa* in fast speech, and *bəladʒar* ‘study’ becomes *bəladʒa*. Non-coarse forms are affected too. For example, *dapor* ‘kitchen’ becomes *dapo*, *pe’el* ‘good character’ becomes *pe’e*, and *buah* becomes *bua* ‘fruit’. There appears to be connected to the rule that deletes final /ɾ/, /l/, /s/ after raising to /ɛ/ in refined speech. The difference between these two phenomena is that final /ɾ/, /l/, /s/ is mandatorily deleted after raising to /ɛ/ in refined speech, whereas word-final liquids and glottal fricative are not mandatorily deleted in fast speech.

Fast speech can also compound words together, common ones being *taʔ-a* and *taʔ-da* for *taʔ ada* ‘there isn’t’, *taʔ-pa* for *taʔ apa* ‘it is nothing’, *toʔ-sa* for *toʔ-usa* ‘there is no need’, and *niari* for ‘this *ini ari* ‘this day’. These are also recorded in Pakir (1986). She has more examples of these that show up in her data than in the archive data used. She states that “disyllabic structures are the favoured output” (Pakir 1986:80).

³⁶ Information is recorded in an interview with Victor (oai:scholarspace.manoa.hawaii.edu: NL1-015, 00:00:00-00:03:43.0)

3.5 Writing system

As a non-official language that is not taught in schools, BM has no official writing system of its own. There are however individual systems used by playwrights who write Peranakan plays known as *wayang Peranakan*, or authors who write about Peranakan subject matters. These systems are closely modelled on the Malay writing system used in Singapore, Malaysia, and Brunei. Until the early 1900s, Malay was written using the Jawi script, modified for use from the original Arabic system (Omar 1989, Omar 1993). Only in 1904 was the writing system reformulated and romanized by Richard James Wilkinson, a Malay scholar and member of the colonial administration. This was the system adopted in Malaysia, Singapore, and Brunei. In fact, Wilkinson systematized the writing system predominantly based on the English writing system, whereas the romanized writing system in Indonesia introduced by Charles Adriaan Van Ophuijsen was influenced by Dutch (Omar 1989, Omar 1993). The standard Malay orthography was then revamped several times, with key changes taking place in 1972, during which administrators attempted to unify the spelling systems of Malay in Malaysia and Singapore with the one in Indonesia.

The system that this grammar uses is closest to the one utilized by William Gwee, a Peranakan community member in Singapore who has published a BM dictionary (Gwee 2006) and a compendium of BM sayings (Gwee 1998), the reasons being that his works may be familiar to the Peranakan community, and that the system he uses is logical and mostly easy to follow for general readers. The following table shows the phonetic values, writing system used for this grammar adapted from Gwee (2006) and the corresponding equivalents used in modern Malay orthography. The sense of the word ‘orthography’ as it is used here refers to standard usage.

<i>IPA</i>	<i>Writing system of this grammar (adapted in part from Gwee 2006)</i>	<i>Modern Malay orthography</i>
a	<i>a</i>	<i>a</i>
i	<i>i</i>	<i>i</i>
e	<i>é</i> ³⁷ except syllable-finally where it is represented by <i>ay</i>	<i>e</i> , except before a coda where it is represented by <i>i</i> . ³⁸
ɛ	<i>air</i>	
ə	<i>e</i>	<i>e</i>
u	<i>u</i>	<i>u</i>
o	<i>o</i>	<i>o</i> , except before a coda where it is represented by <i>u</i> .
ɔ	<i>or</i>	

³⁷ Differs from Gwee (2006), where both [e] and [ə] are represented by *e*.

³⁸ Vowel lowering takes place in Malay before a consonant coda, so that /i/ becomes [e], and /u/ becomes [o]. See Onn (1980); Teoh (1994). There is no evidence that this takes place in BM.

p	<i>p</i>	<i>p</i>
b	<i>b</i>	<i>b</i>
m	<i>m</i>	<i>m</i>
t	<i>t</i>	<i>t</i>
d	<i>d</i>	<i>d</i>
s	<i>s</i>	<i>s</i>
n	<i>n</i>	<i>n</i>
l	<i>l</i>	<i>l</i>
r	<i>r</i>	<i>r</i>
tʃ	<i>ch</i>	<i>c</i>
dʒ	<i>j</i>	<i>j</i>

ɲ	<i>ny</i>	<i>ny</i>
k	<i>k</i>	<i>k</i> , except for words borrowed from Arabic that originally contained velar fricative, these are spelled <i>kh</i> .
g	<i>g</i>	<i>g</i>
ŋ	<i>ng</i>	<i>ng</i>
ʔ	ʔ, except in word-final position, <i>k</i> or <i>g</i>	
h	<i>h</i> ,	<i>h</i>
w	<i>w</i> , except in non-onset position, <i>u</i>	<i>w</i>
j	<i>y</i> , except in non-onset position, <i>i</i>	<i>y</i>

Table 18: Writing system of Baba Malay used by this grammar

There are several points to note:

The sound [ɛ] is represented by *air* as the vowel sounds similar to the vowel in the English word ‘air’. Singaporeans’ pronunciation of this English word is similar to that of the Received Pronunciation rendition. This would be familiar to most literate Peranakans who would also be exposed to English, be it in Singapore in Malaysia. In fact, it should be recalled that the romanized Malay script itself was formulated based on English in these areas (Omar 1989, Omar 1993).

The two sounds [e] and [ə] are represented by the same symbol, *e* in Malay and in Gwee’s writings (Gwee 1998, Gwee 2006). Whereas [ə] is still represented as *e* in this grammar, [e] is represented by *é* since it is not predictable where [ə] occurs, or where [e] occurs. For example, [bedeʔ] ‘tell a lie’ and [bədəʔ] ‘pummel’ form a minimal pair in BM. It is also more economical to mark [e] with an extra diacritic as opposed to [ə] since [ə] occurs much more frequently than [e] does.

However, the vowel sound [e] is written *ay*, such as with *sampay* for [sampe], meaning ‘until, reach’. This marks a salient difference between BM and Malay. BM [sampe] would correspond to Malay [sampaj]. Similarly, while [ɔ] does not exist in Malay, [o] and [ɔ] are represented by *o* and *or* respectively, so as to be consistent with Gwee (2006).

Diphthong combinations are represented using *i* for [j] and *u* for [w] when [j] or [w] are not found in the onset. For example, [ja] is written as *ia* and [wa] can be written as *ua*. For example, [sjut] ‘singe’ is written as *siut* and [kwat] ‘strong’ is represented as *kuat*. In onset positions, the glide [j] is represented as *y* and [w] is represented as *w*. For example, the word [wa.jaŋ] ‘play’ is written as *wayang*. This follows Gwee (2006).

Although word-final glottal plosives manifests as [ʔ], these are still represented as *k* or *g*, since it should be relatively easy to apply the word-final plosive to glottal stop rule for readers. Word-final *k* and *g* are also used by Gwee (2006).

For readers familiar with Malay, it is interesting to note that the sounds [f] and [z] do not exist in BM, even though they exist in Malay.

Whereas Hokkien is a tone language, BM words of Hokkien origin do not have consistent tones on them and are thus not marked for tonal contrasts. Note that they do however receive regular lexical stress (see section 3.6.2).

3.6 Stress and intonation

In general, BM is a syllable-timed system, where the duration of each syllable is approximately equal. This is not surprising, considering that both its lexifier language, Malay (Wan 2012) as well as its substrate language, Hokkien are syllable-timed (Hung 1996) instead of stressed-timed.

3.6.1 Word stress and pitch

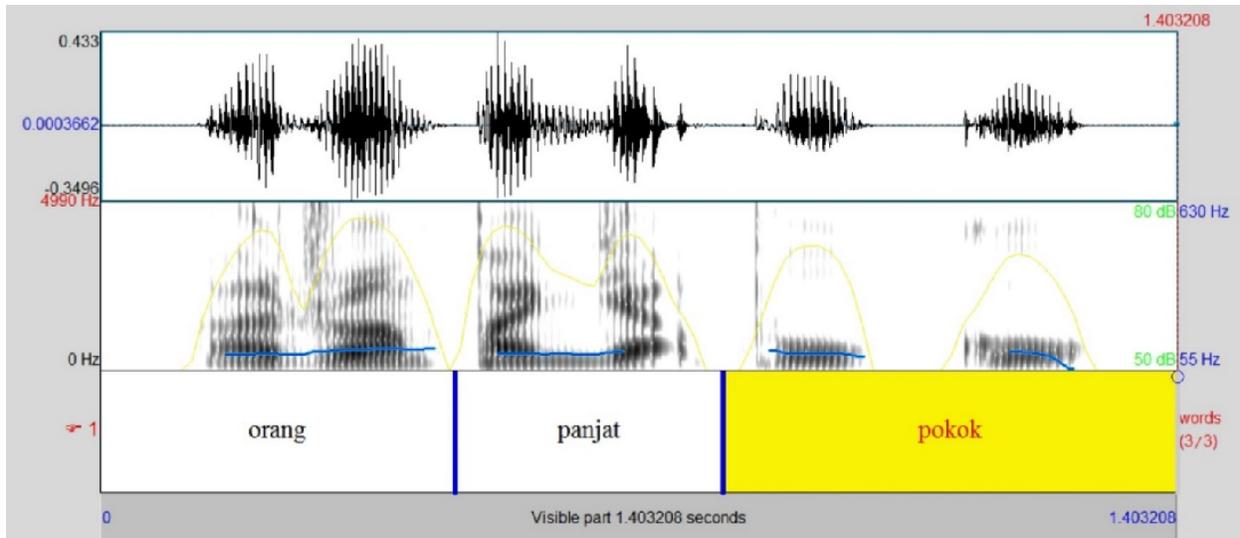
While the lexical stress system is not based on duration, pitch (and to some extent intensity) is used for the purpose of lexical stress. All three factors discussed are common features that are said to “conspire in varying degrees in many languages to give some syllables prominence when compared with other syllables” (Cruttenden 1997:7). In BM, pitch is the most accurate indicator of stress. The ultimate syllable tends to receive the most stress. Wee (2000) characterizes this as a ‘step-up progression’. However, when there are three syllables, the final syllable receives the most stress, followed by the first syllable. Lexical stress allocation begins from the final syllable, and stress alternates between syllables. Words with four syllables are rare, but when they occur, they follow the general lexical stress template, where the ultimate syllable receives the most stress, the penultimate syllable is not stressed, and the second syllable is stressed. This differs from Wee’s (2000) findings, where step-up progression is also observed for words that are more than two syllables long

The following waveform and spectrogram captures natural speech comprising words of two syllables. There are two sets of non-striated contour lines on the spectrogram. The lower line represents pitch while the upper line represents intensity.³⁹ The dynamic range on each spectrogram in this section is set at 40dB so that pitch and intensity contours can be seen more clearly. For the same purpose, pitch range and intensity range are adjusted for individual

³⁹ On a color copy of this dissertation, the lower line indicating pitch is blue, while the upper line indicating intensity is yellow.

spectrograms, so that none of the relevant contours are cut off. Otherwise, the settings in Praat for this section are – Maximum formant: 5000Hz,⁴⁰ number of formants: 5, window length: 5 milliseconds. Words are orthographically transcribed in this section.

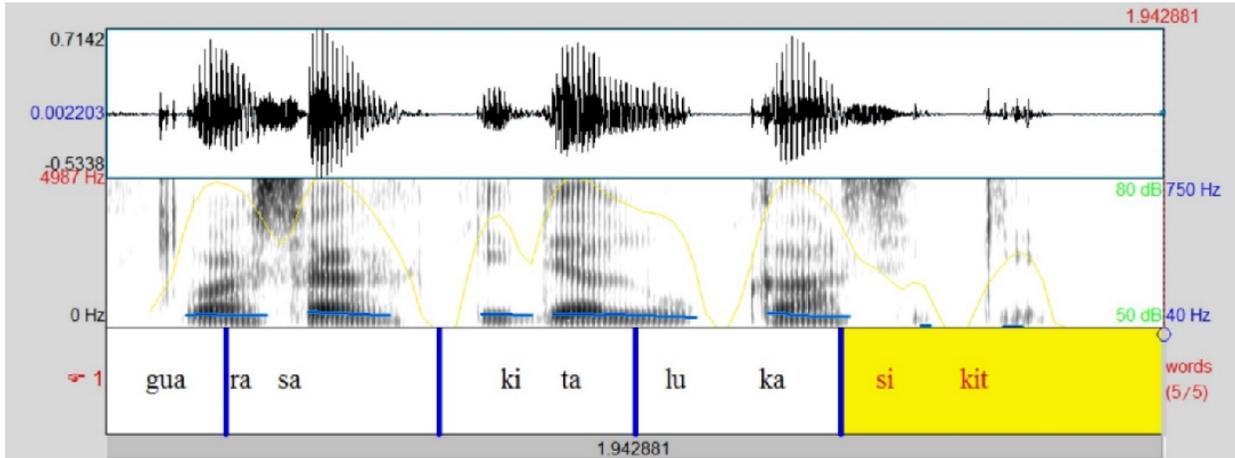
- (108) *Orang panjat pokok.*
 person climbs tree
 A person climbs a tree.
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:00:15.6-00:00:17.6)



Featuring a series of three two syllable words, example (108) shows that each syllable within each word is equivalently the length of the other word (as with syllable-timed systems), and each syllable within each word also has almost similar intensity patterns as the other. Besides a slight drop in pitch towards the end of the *pokok* ‘tree’ which seems to indicate sentence finality, the other disyllabic items show a slight step-up in pitch within the word. These can be represented as *o. 'rang* ‘person’ and *pan. 'jat* ‘climb’ respectively, where ' represents primary stress. In this sequence of two-syllable words, stress can be said to fall on the final syllable since the accompanying higher pitch indicates more prominence. Note that while (108) shows examples of closed syllables, (109) demonstrates that the same word stress pattern occurs in open syllables.

⁴⁰ This setting is also suitable for the female speaker whose voice is featured in this section, as she has a low voice.

- (109) *Gua rasa kita luka sikit.*
 1.SG think 1.PL wound little
 ‘I think we are a little wounded.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:03:42.5-00:03:44.9)

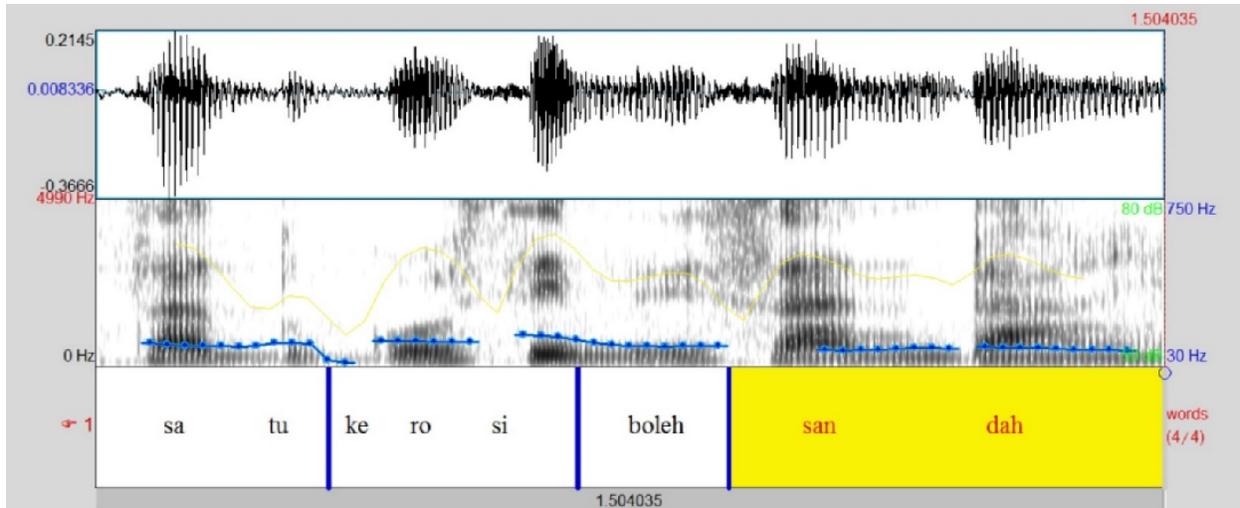


The disyllabic words in (109) can be represented as *ra*. 'sa 'think', *ki*. 'ta '1.PL' and *lu*. 'ka. In addition to step-up progression in pitch, the second syllable in each word is also accompanied by higher intensity. Thus, in disyllabic words featuring both open and closed syllables, lexical stress always occurs on the ultimate syllable.

The following two examples demonstrate lexical stress in trisyllabic words.

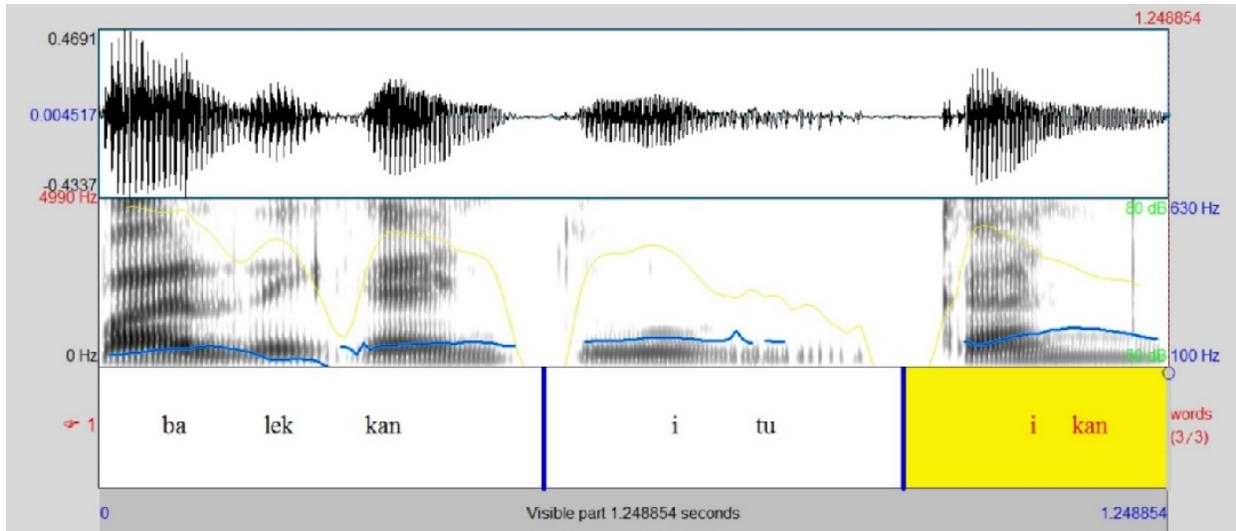
(110) *Satu kerosi boléh sandah.*
one chair can lean
'one chair (that) can lean.'

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:05:46.6-00:05:48.1)



In (110), the trisyllabic word for chair can be represented as *,ke.ro.'si* 'chair', where *,* indicates secondary stress and *'* indicates primary stress, with pitch being the most accurate indicator of stress. The following is another example of a trisyllabic word.

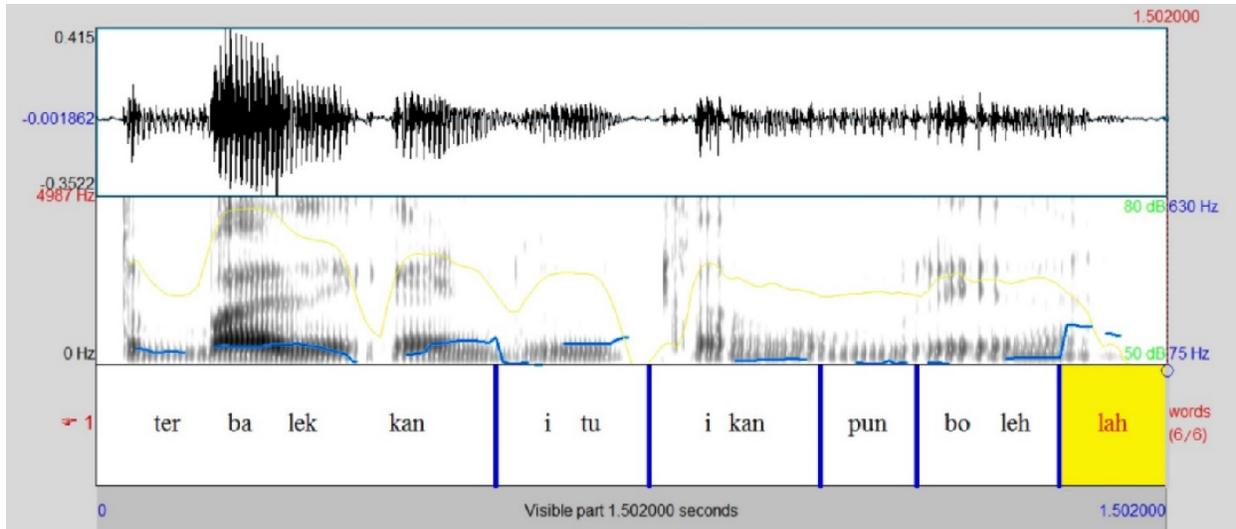
- (111) *Balék-kan itu ikan,*
 turn.over-TR that fish
 ‘Turn over that fish,’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:10:30.7-00:10:32.3)



In example (111), ‘turn over-TR’ can be represented as *ba.lék. 'kan*, since pitch is the most prominent on the ultimate syllable followed by the first syllable. Hence in trisyllabic BM words, primary stress falls on the ultimate syllable, followed by the first syllable.

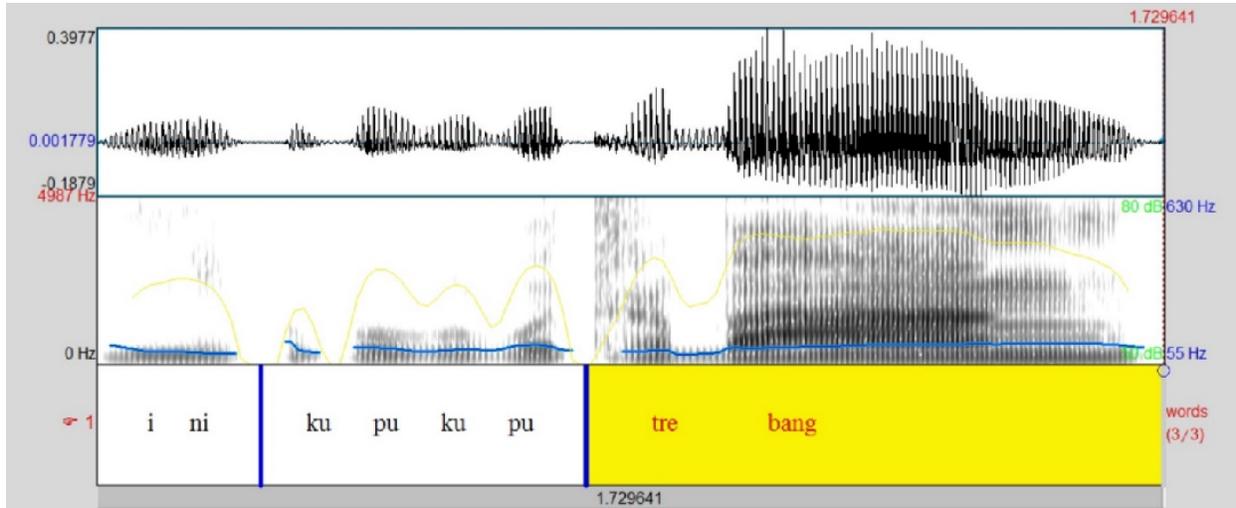
Words of four syllables are rare in BM, but they follow the template where the ultimate syllable receives primary stress, and the second syllable receives secondary stress. Example (111) can be compared to (112), where the prefix *ter-* ‘MVT’ is attached to *ba.lék. 'kan* ‘turn over-TR’. The word now comprises of four syllables instead of three. Example (8) is analyzed here as (112).

- (112) *Ter-balék-kan itu ikan pun boléh lah*
 MVT-turn.over-TR that fish also can EMP
 ‘Turn over that fish is also okay.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:10:32.3-00:10:33.8)



With an additional prefix, the four syllable word can be expressed as *ter*, *.ba.lék*, *'kan* ‘MVT-turn.over-TR’. The same syllables receive the same stress as *.ba.lék*, *'kan* ‘turn over-TR’. This can be explained by the fact that stress allocation begins from the ultimate syllable; the ultimate syllable receives the primary stress, while the alternate syllable receives the secondary stress. Lexical stress is mainly indicated by pitch. (113) is another example of lexical stress allocation on a word comprising four syllables.

- (113) *Ini kupu-kupu terbang,*
 this butterfly fly
 ‘This butterfly flies,’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:01:37.5-00:01:39.2)



The word for butterfly can be represented as ku .pu.ku.'pu, based on the pitch patterns, with the stress alternating between syllables.

3.6.2 Tones on Hokkien-derived words

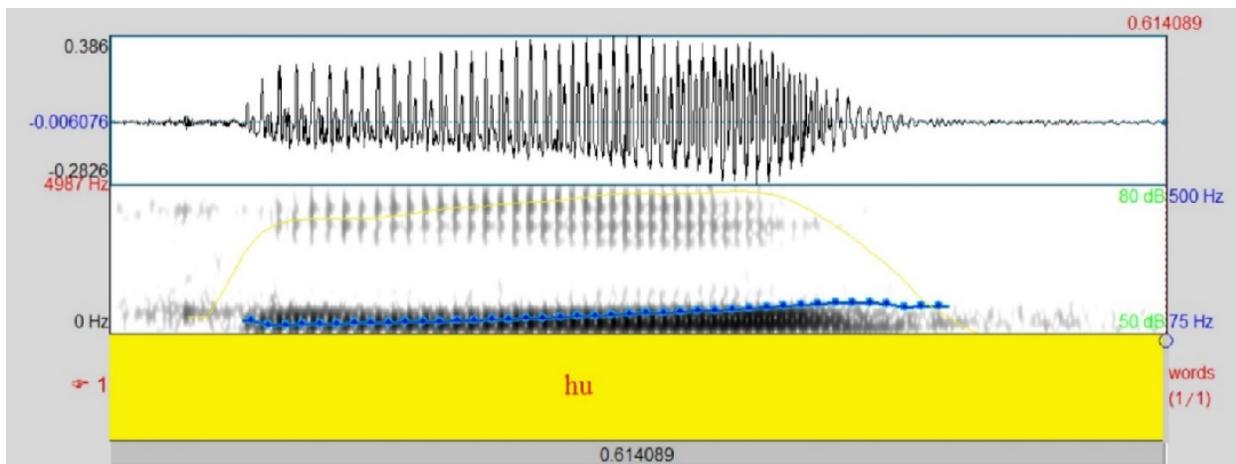
One of the more interesting questions regarding pitch is whether or not tones in words derived from Hokkien have been preserved. Hokkien is a tone language with seven contrastive tones, a tone language being one where changes in pitch can change the meaning of the word (Yip 2002). Pakir (1986) states that words derived from Hokkien show no tonal distinctions.⁴¹ This is partially true where the current corpus is concerned. On monosyllabic items, BM speakers appear to produce no contrastive tone in comparison with Hokkien speakers, and on disyllabic items, BM speakers produce pitch on these Hokkien-derived lexical items that are unlike Hokkien tones, but consistent with general BM patterns for other words. These are essentially

⁴¹ Pakir (1986) treats BM as a Malay dialect that has borrowed extensively from Hokkien.

step-up progressions like the ones recorded in the previous section. Trisyllabic words from Hokkien are rare.

The following is an example of a monosyllabic word as it would have been produced in Hokkien. The speaker speaks BM, but is aware that Hokkien words sound different in BM than the corresponding form in Hokkien itself sounds.

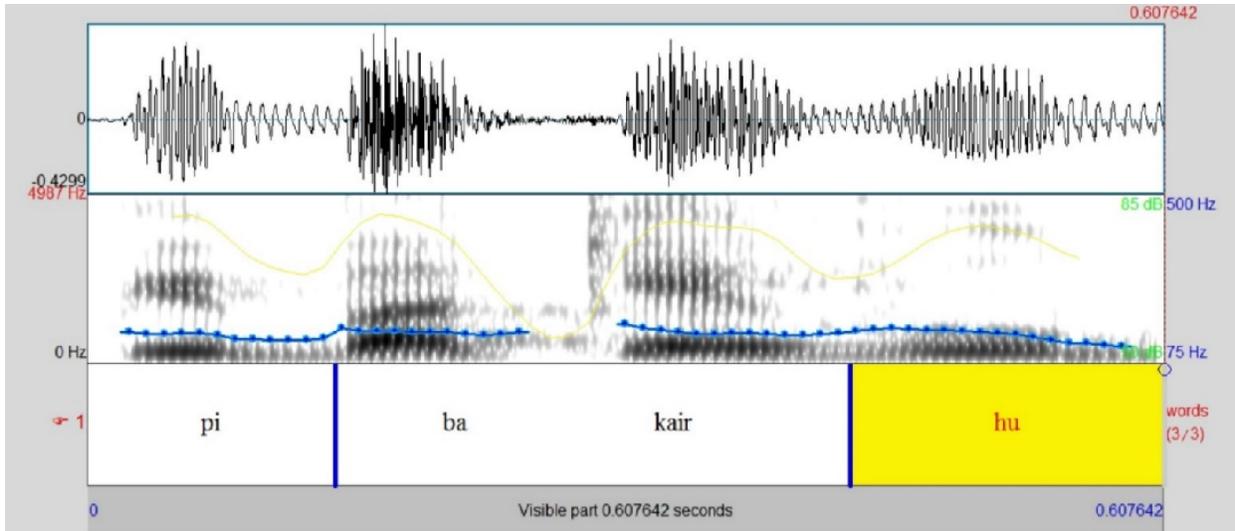
- (114) *Hu*
amulet
'amulet'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-169, 00:06:25.0-00:06:25.5)



In example (114), the speaker states that *hu13*⁴² is usually pronounced with a rising tone in Hokkien. In example (115), the BM speaker uses *hu* in a BM sentence. There is no rise in pitch accompanying *hu*. In fact, pitch falls slightly owing to *hu*'s sentence-final position.

⁴² Chao (1930)'s system is used to annotate tones here, as is common in describing tones in Chinese languages. The numerals 1-5 are used, and the first number in the transcription indicates the starting point of the tone, while the second number indicates the ending point of the tone. 1 represents the lowest pitch, and 5 the highest.

- (115) *Pi bakair hu.*
 go burn amulet
 'Go burn an amulet.
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-169, 00:06:26.7-00:06:27.3)



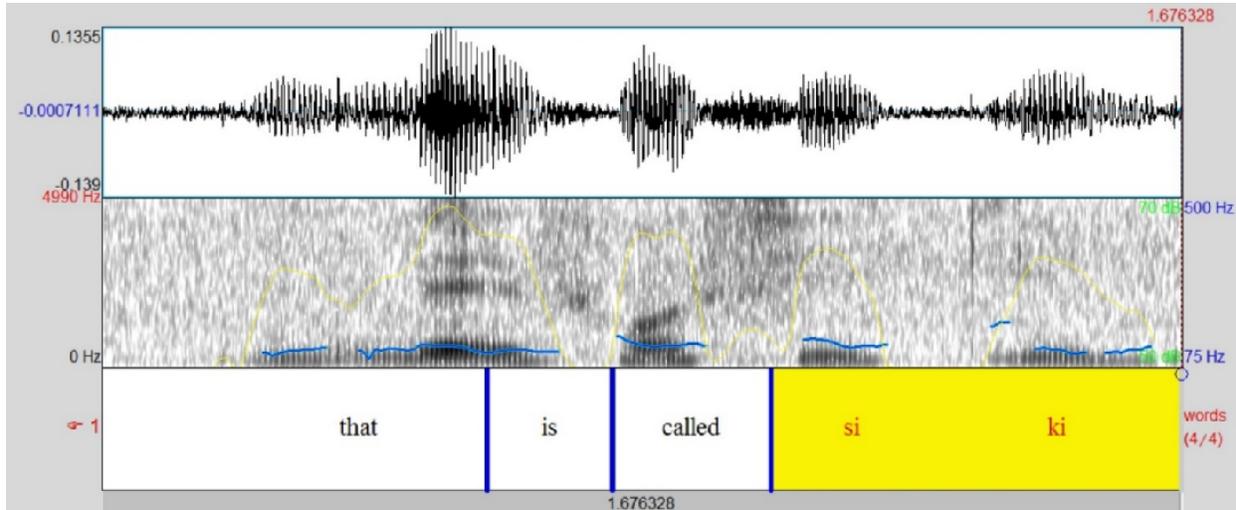
While monosyllabic items appear to lack tone, disyllabic items are accompanied by different pitches in BM from those in Hokkien. The following is taken from a speaker of both BM and Hokkien who differentiates between the two languages.

(116) *That is called siki.*

death.anniversary

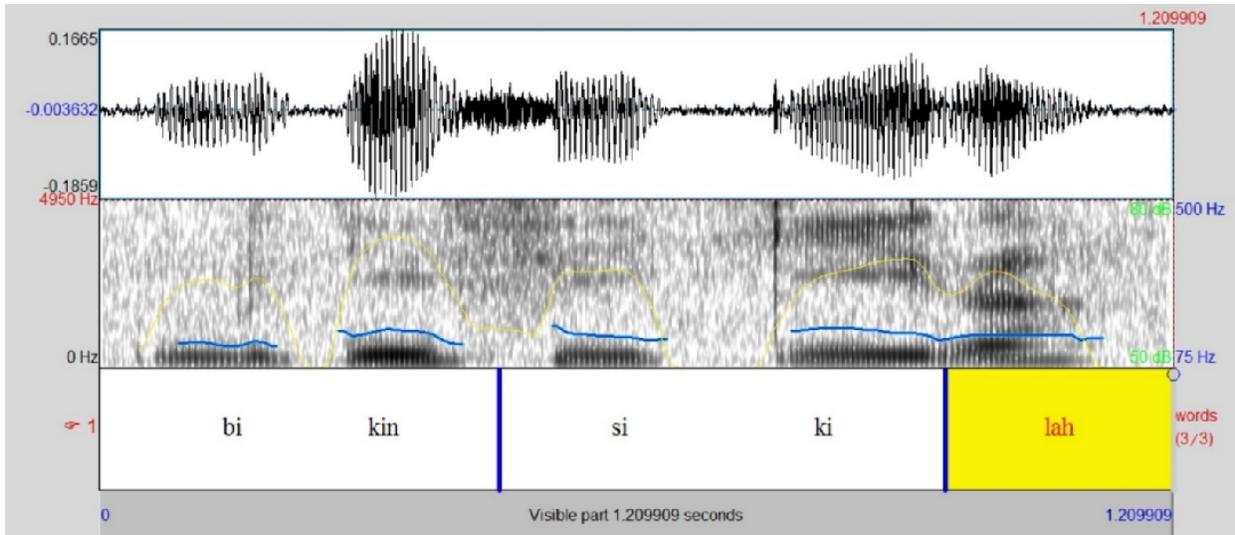
‘That is called death anniversary.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-173, 00:08:01.1-00:08:02.6)



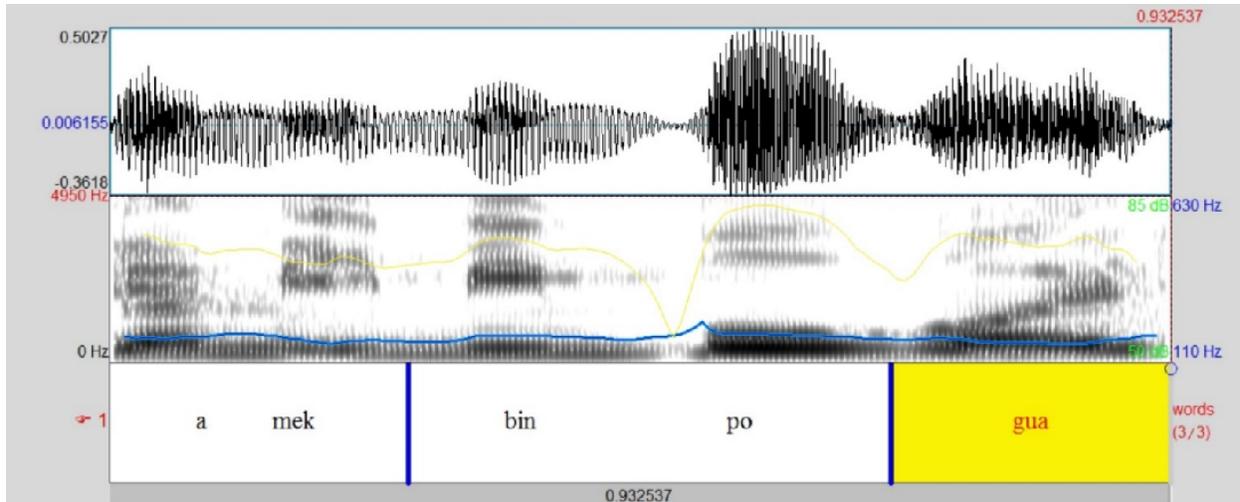
In example (116), the speaker says that in Hokkien, ‘death anniversary’ is pronounced *si53 ki32*, whereas in example (117), when the same speaker is speaking BM, there is a step-up progression in *si'ki*, where the second syllable ends slightly on a higher pitch than the first.

(117) *Bikin siki lah*
 make death.anniversary EMP
 ‘Commemorate death anniversary.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-173, 00:08:04.6-00:08:05.8)



The following is another example taken from a different speaker, who also says that Hokkien as spoken by the Chinese is different from Hokkien as spoken by BM speakers.

- (119) *amék binpo gua,*
 take handkerchief 1.SG
 ‘Take my handkerchief,’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-171, 00:34:34.1-00:34:35.2)



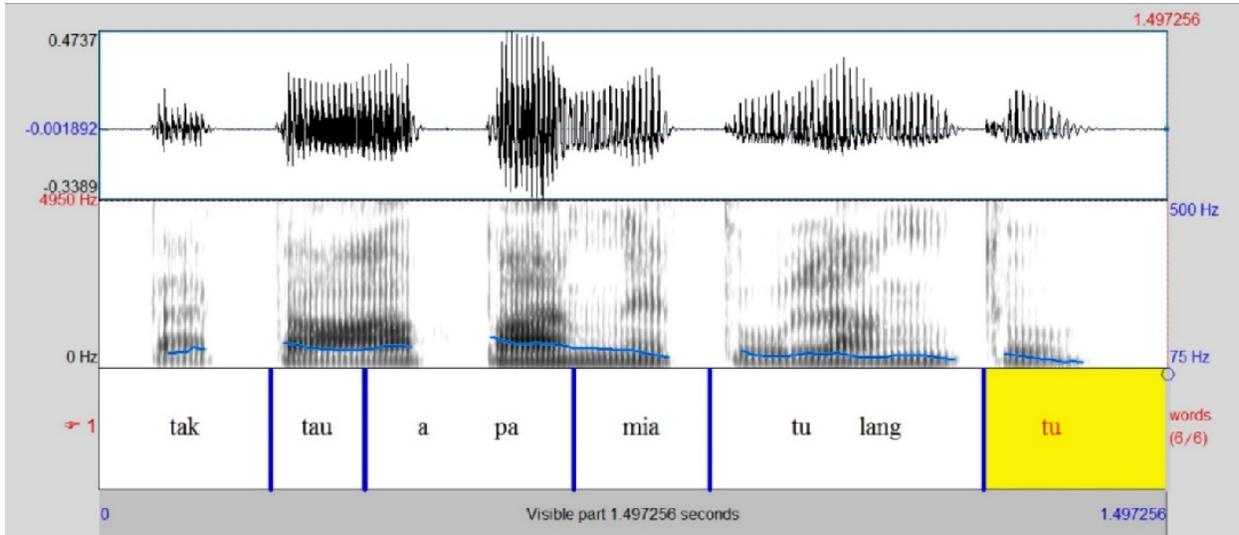
Although there is a slight up-step progression accompanying these words, there is technically no tone on these words, since tones are defined as what differentiates one lexical item from another (Yip 2002).

3.6.3 Sentence intonation

There are four main types of sentential intonation that may be observed. These are found in declaratives, clauses signaling continuity, content questions, and tag questions.

Declarative sentences end with a fall in the overall pitch contour of the sentence, as also observed by Wee (2000). This is demonstrated in example (120), and also in earlier examples such as (108) and (110). Only pitch contours (and not intensity contours) are shown in the spectrograms for this section.

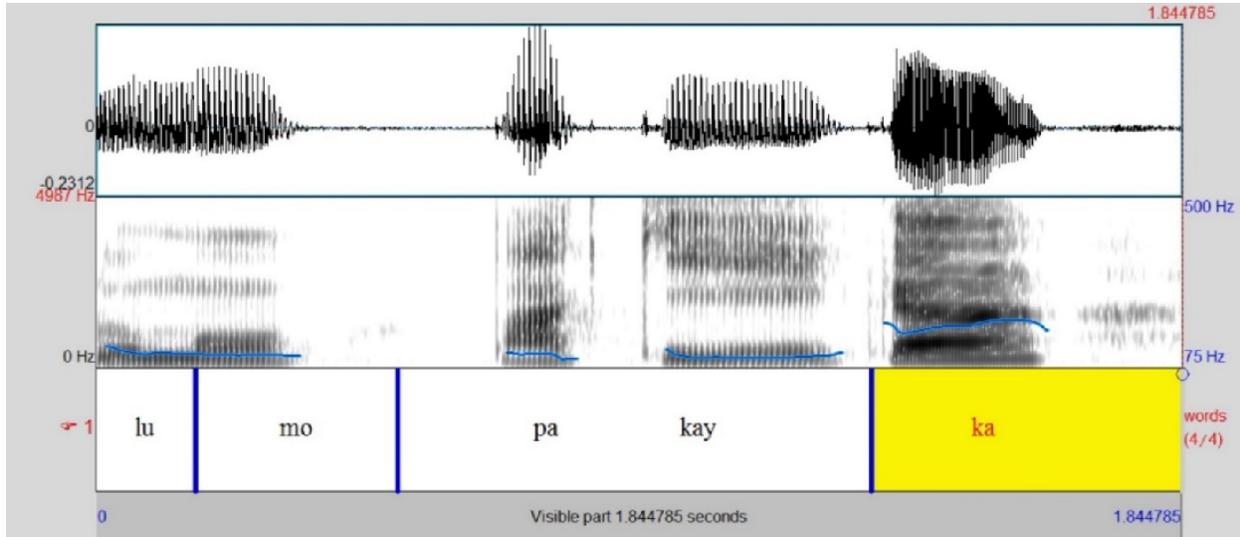
- (120) *Tak tau apa mia tulang itu.*
 NEG know what REL bone that
 ‘Don’t know what bone that is.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:40.8-00:02:42.5)



With a different pattern, BM speakers signal continuity, that they have more to say on the subject matter, even though they may have come to a pause. The pitch of the sentence rises towards the end of the clause, as shown in example (121).

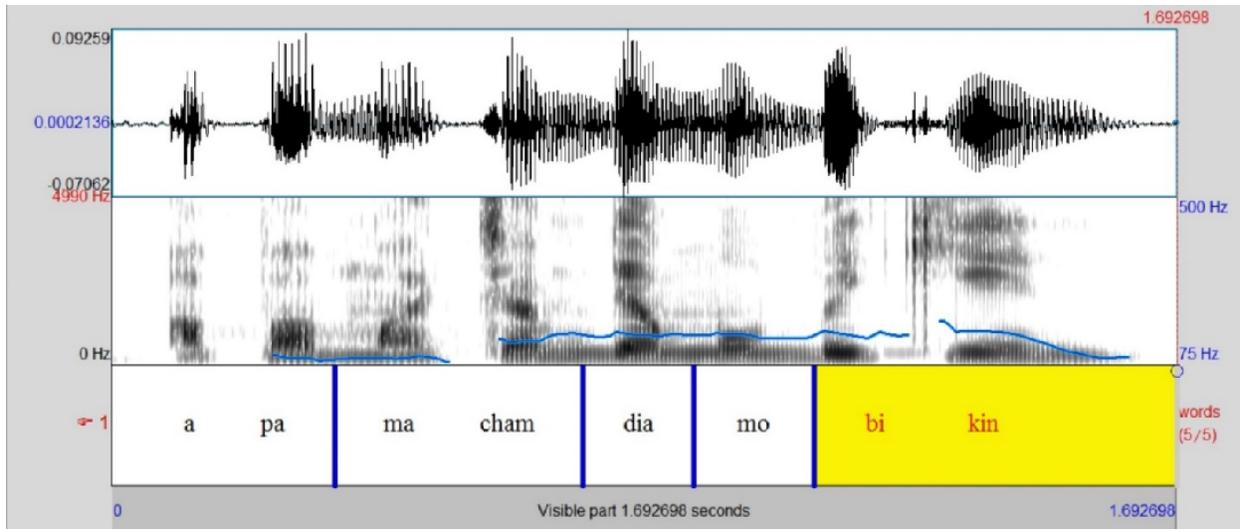
(121) *Lu mo pakay ka,*
2.SG want use or
'You want to use (it) or,'

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-170, 00:04:10.5-00:04:12.4)



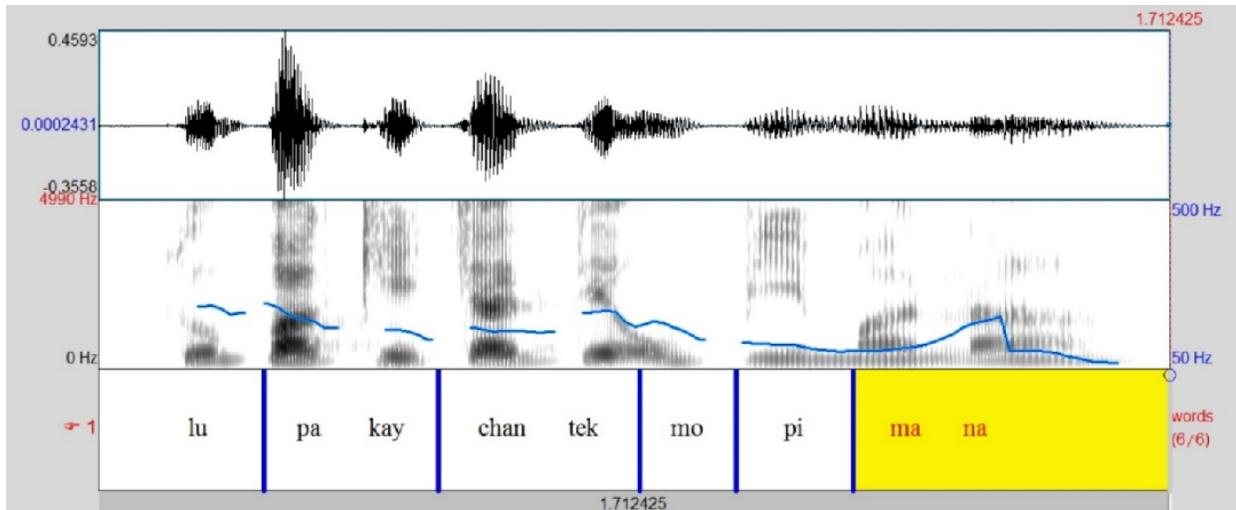
There are two separate intonation patterns that accompany content questions, depending on the position of the question word. When the question word occurs at the beginning of the utterance, pitch rises at the beginning and falls towards the end, as with example (122). This is similar to the observations of Wee (2000) that a rise-fall intonation accompanies interrogatives in BM.

- (122) *Apa macham dia mo bikin?*
 What like.that 3.SG want do
 ‘How does she/he want to do this?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-37, 00:42:37.6-00:42:39.3)



Conversely, when the question word is at the end of the utterance, pitch rises towards the end of it, as with example (123).

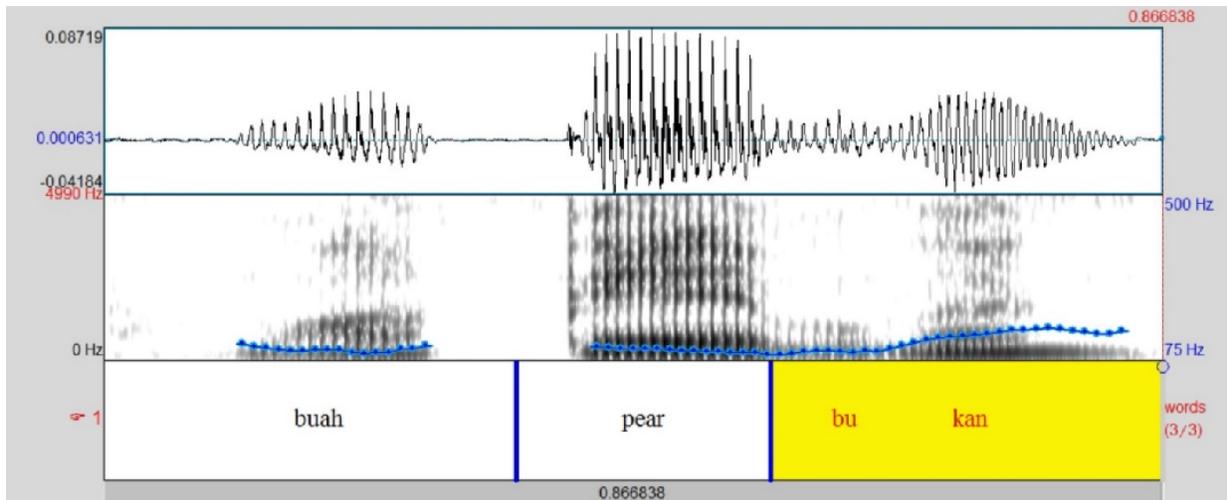
- (123) *Lu pakay chanték mo pi mana?*
 2.SG wear beautiful want go where
 ‘Where do you want to go dressed (so) beautifully?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-37, 00:07:04.0-00:07:05.9)



Similarly, tag questions have a rise that come towards the end of question, as also observed by Wee (2000)⁴³. This is demonstrated by example (124).

⁴³ Wee (2000)'s term for tag questions is yes-no questions.

- (124) *Buah pear bukan?*
 CLF.fruit no
 'Is this pear?'



Hence, there are a few intonation contours at the sentential level in BM. Declaratives fall in pitch towards the end of the utterance, a final rise can signal continuity, and content questions are accompanied by rise-fall when the question word occurs at the beginning of the interrogative. When the question word occurs at the end of the interrogative, or when a tag question is asked, pitch tends to rise at the end of the interrogative.

3.7 Variation and change

Tsunoda (2005) reports that the loss of phonetic information is common in situations of language decay. Even though phonological variation takes place when languages are not threatened, it has been demonstrated in minority communities that phonetic or phonological variation may become more apparent when the viability of the language is threatened (Babel 2009, Brunelle 2009, Campbell and Muntzel 1989). This is true of BM, where the number of phonemic contrasts is diminished as speakers become less proficient in BM, and more proficient in other dominant languages such as English and Malay. In fact, the vowel space appears to be contracting, so that it appears to be more similar to that of Standard Malay. Clynes and Deterding (2011) provide the only phonetic account of Standard Malay, this being a variety spoken in Brunei. Crucially, they

state that the Standard Malay of Brunei is similar to varieties spoken in Singapore and Malaysia, and different from the variety that is spoken in Indonesia. In the absence of information from Singapore, it is reasonable to assume that the vowel system of the variety in Singapore is comparable to the one that Clynes and Deterding (2011) describe. The essential difference between their vowel plot of Standard Malay and that of BM, is that open-mid front and back vowels, [ɛ] and [ɔ] that are found in BM do not occur in Standard Malay.

The vowels of six speakers are plotted below. These speakers are between the ages of 60 to 80, four of them are males and two of them are females. All of these speakers stated that they grew up with BM as their first language. In addition, all of them are bilingual in English, and profess to have different levels of proficiency in BM. Four speakers stated that they are proficient in BM, while two stated that they have become less proficient in the language as English is now their dominant language.⁴⁴

Ten tokens of each vowel were extracted from natural occurring speech (not wordlist data) and their formants measured in Praat, following the methodology in section 3.1.2.2. The only difference is that the settings used for this analysis are the same as the ones used earlier for the analysis of vowels, except that the maximum formant for is set to 5500Hz for one female speaker,⁴⁵ while it is set to 5000Hz for the rest of the speakers (including one female who has a low voice).⁴⁶ Where possible, ten [ɛ] and ten [ɔ] tokens are measured for each speaker, but where speakers only seem to produce [e] and not [ɛ], or [o] and not [ɔ], twenty [e] or [o] tokens are measured instead. Where vowels appeared to overlap, participants' beliefs were used to identify if a vowel was [e] or [ɛ], and [o] or [ɔ].⁴⁷

For the purpose of a quick general comparison, the following two vowel charts in figures 23 and 24 show the vowel spaces of a more proficient speaker, and a less proficient speaker respectively, the difference being that the more proficient speaker appears to have open-mid

⁴⁴ The speakers have not been named here, so as to protect their identities, especially since language attrition is being discussed.

⁴⁵ Speaker B in figure 25.

⁴⁶ Speaker A in figure 25.

⁴⁷ This was done due to the fact that consultants differed in how they felt some words should be pronounced. For example, [ɔraŋ] versus [oraŋ] for 'people'.

front and back vowels where the less proficient speaker produces none. Both speakers are males in their sixties. While the more proficient speaker still uses BM daily, the less proficient speaker has stopped using BM as much. This speaker used to speak BM with his parents who passed away over ten years before the time of recording. Figure 23 is from Figure 22.

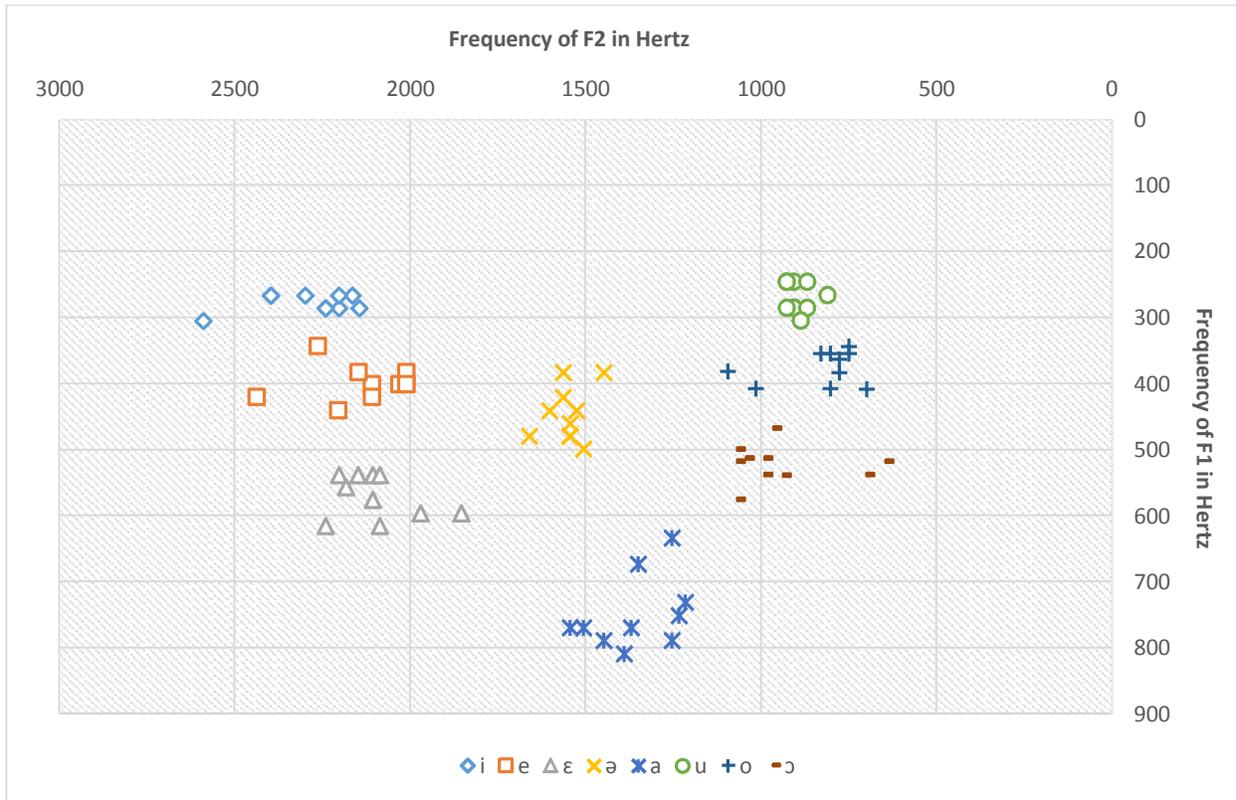


Figure 23: Vowel space of a more proficient Singapore Baba Malay speaker

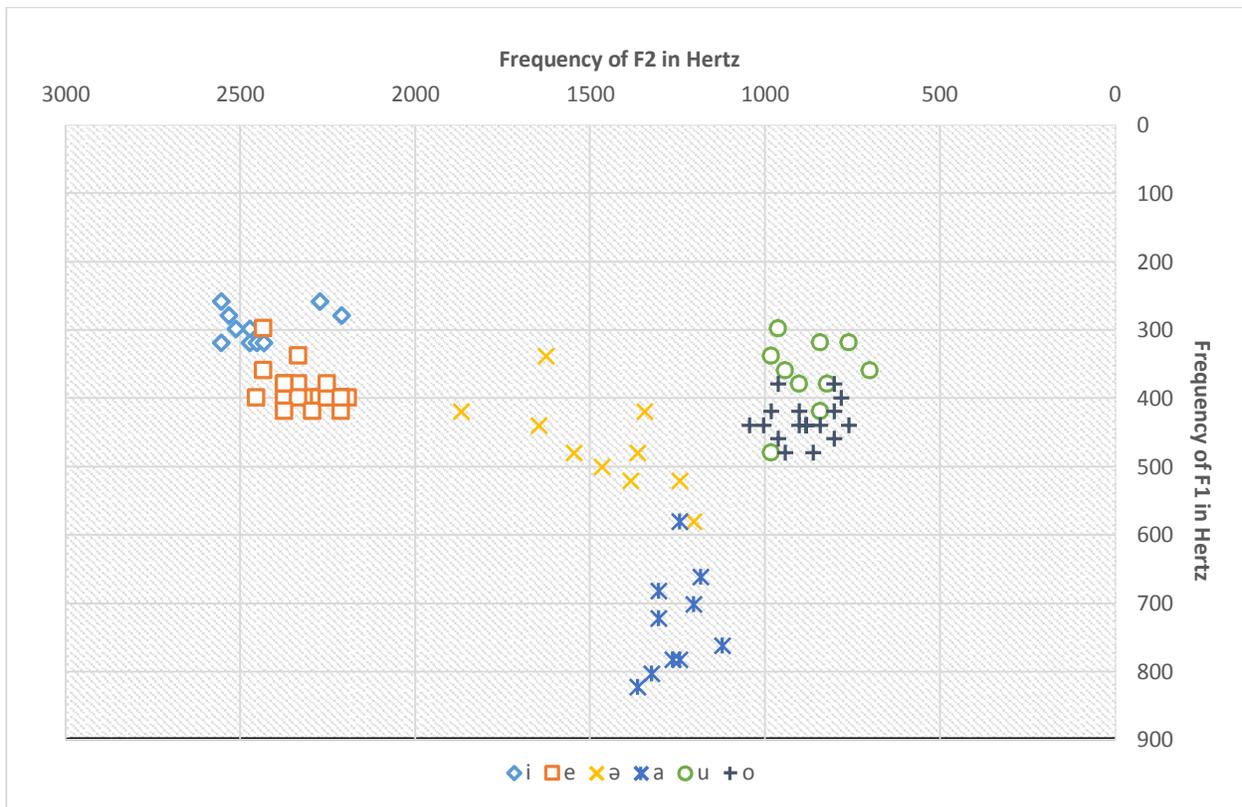


Figure 24: Vowel space of a less proficient BM speaker

While the above figures are provided for purpose of quick illustration, vowel normalization is generally required when comparing vowel spaces across different speakers, since the length and shapes of the vowel tracts of different speakers vary, especially when comparing across different genders. The Lobanov method is utilized in this instance (see Lobanov 1971).⁴⁸ The following are the resultant vowel plots.

⁴⁸ There are vowel intrinsic and vowel extrinsic methods of normalization (Ainsworth 1975, Nearey 1989). Vowel intrinsic methods use only acoustic information contained within a single vowel to normalize that vowel token, while vowel extrinsic methods rely on information distributed across more than one vowel of the speaker. Vowel extrinsic methods such as the Lobanov have been shown to perform better than vowel intrinsic methods. The Lobanov method is represented by the following formula $F_{n[V]}^N = (F_{n[V]} - MEAN_n) / S_n$, where $F_{n[V]}^N$ is the normalized value for $F_{n[V]}$ (formant n of vowel V), $MEAN_n$ is the mean value for formant n for the speaker, and S_n is the standard deviation for the speaker's formant n .

The vowel plot shown is obtained through vowel normalization and plotting suite, NORM (Thomas and Kendall 2014).

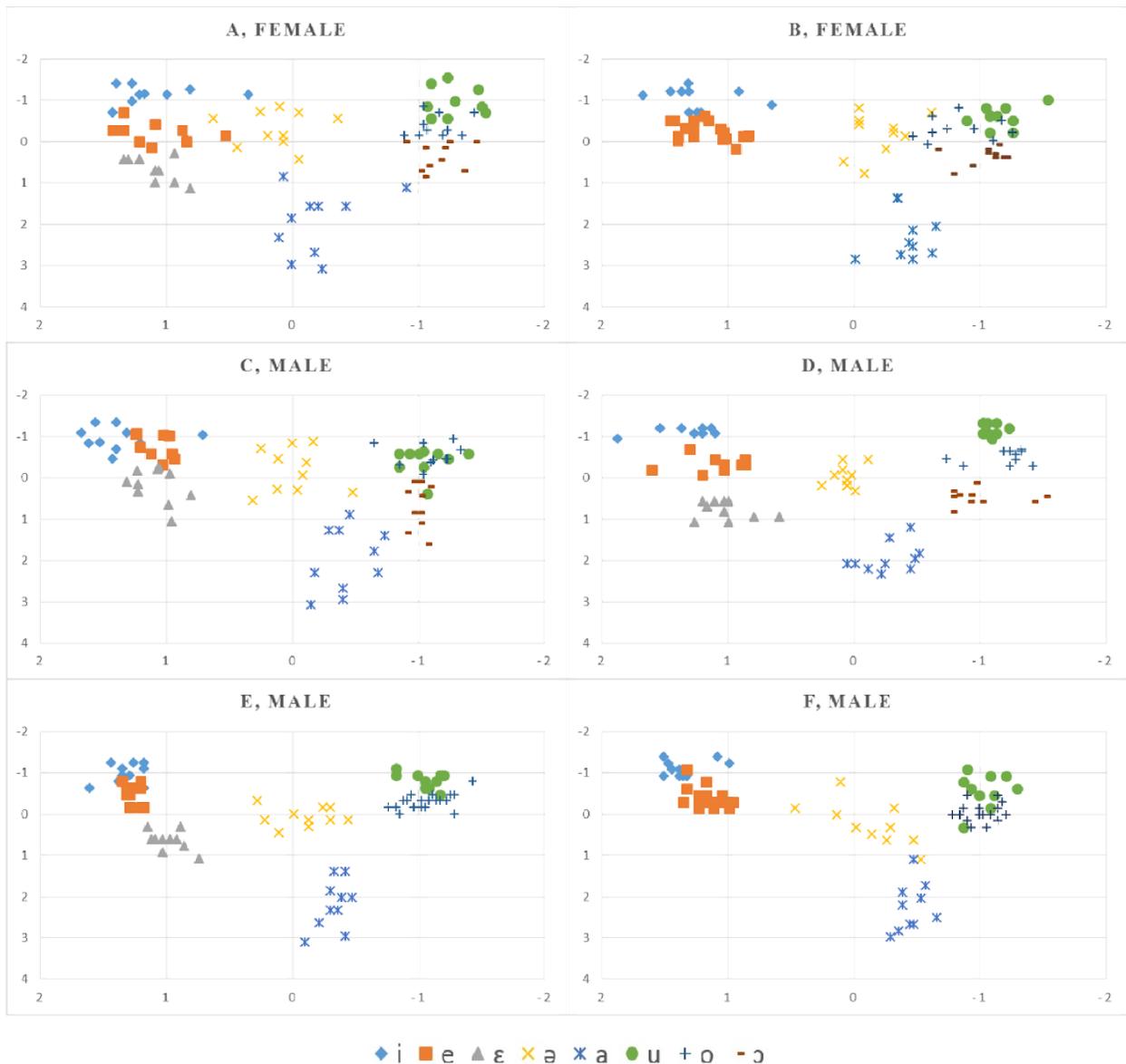


Figure 25: Normalized vowel plots of six Singapore Baba Malay speakers⁴⁹

According to the plots, two out of six speakers, B and F, do not produce anything in the region of the open-mid front vowel [ɛ], while only speakers A, B, C and D produce tokens in the region of open-mid back vowel [ɔ]. It is not surprising that speakers B and F state that they are no longer

⁴⁹ The normalized values shown are obtained through vowel normalization and plotting suite, NORM (Thomas and Kendall 2014).

proficient in BM, and that English is now their dominant language. Although speaker E states that he is a proficient BM speaker, he does not produce [ɔ] tokens. All other speakers state that they are proficient in BM. There is ongoing change, as two vowels are on the verge of being lost. The emerging vowel plot can be compared to the one that Pakir (1986) describes, wherein [ɛ] and [ɔ] can both be distinguished from other vowels. It is also noteworthy that vowels [u], [o], and [ɔ] all appear to be in the process of merging.

It is interesting to recall that [ɛ] only occurs in refined speech, corresponding to words in coarse speech that end with phonological /aɪ/, /aɪ/ or /as/. That [ɛ] is not produced as often as /aɪ/, /aɪ/ or /as/ shows that BM speakers are producing less ‘refined’ speech, and using more ‘coarse’ speech instead. Words with the /aɪ/, /aɪ/, and /as/ endings are acknowledged by speakers to be *kasar* [kasar] or *kasair* [kaseɪ], literally meaning ‘coarse’. Their [ɛ] counterparts are *halus* or *alus*, the literal meaning being ‘refined’. Previously unnoted by other researchers, when speakers engage either form, it has little to do with the level of familiarity the speaker has with the interlocutor, their social statuses or the social circumstances surrounding language use, but rather, that the speaker simply wants to be perceived as being more refined (hence the use of the terms *halus*⁵⁰ ‘refined’ and *kasar* ‘coarse’, indicating an elevated versus colloquial style, is more appropriate, rather than *formal* and *informal*).

That speakers cite [ɛ] as being *alus* ‘refined’, and /aɪ/, /aɪ/ or /as/ as being *kasair* ‘coarse’⁵¹, may be rooted in [ɛ]’s association with urbanity, and some sense of sophistication and elegance. This is different from findings that urban forms can be associated with toughness (Eckert 1988, Lawson 2014). While Lim (1988) does not label [ɛ] as a refined speech form, he associates it to urbanity in Malacca, and essentially, exposure to Hokkien speakers. Lim (1988) states that BM speakers in Malacca were more likely to produce [ɛ] if they lived nearer the city center, since this is where early Chinese settlers would have established themselves, and [ɛ] exists in Hokkien (but not in Malay).⁵² In contrast, if the BM speakers lived in the more rural areas, they were less likely to come in contact with Hokkien speakers, and less likely to produce [ɛ]. Where Singapore

⁵⁰ Also occurs as *halus* (after initial h- deletion)

⁵¹ Ironically, this is the refined form of the word *kasar*, meaning ‘coarse’.

⁵² However, note that /ɛ/ does not exist in any of the Hokkien-derived words in BM.

is concerned, there are no records that such a division existed in the past, but it is not an implausible situation. However, noting that Singapore is small at 710 square kilometers (50 kilometers (31 miles) from east to west, and 26 kilometers (16 miles) from north to south), and that it is currently Chinese-dominant at 74 percent (Singapore 2011), a present-day division based on amount of contact with the Chinese is not perceptible. In fact, it is also observed that all Malacca speakers of BM who were interviewed for this project use both [ɛ] and non-[ɛ] forms regardless of how close or far away from the city they live. This may be due to urbanization within Malacca. The Chinese have spread out beyond the city center, and commuting has been made more convenient since the 80s when Lim carried out his research, thus enabling Peranakans to come in contact with others they may not necessarily live close to and vice versa. Instead, speakers of MBM also state that [ɛ] forms are *alus* ‘refined’, while non-[ɛ] forms are *kasair* ‘coarse’.

The only other BM items that have been described as *halus* ‘refined’ or *kasar* ‘coarse’ are the first person singular pronouns *saya* (refined) and *gua* (coarse) (see section 7.3.2). This can be compared to Standard Malay, which involves a *halus* ‘refined’ and *kasar* ‘coarse’ distinction in some vocabulary. Whereas the refined and coarse forms in BM are almost always derived from the same stems (with the exception of refined *saya* and coarse *gua* for first person singular), the refined/coarse distinction in Standard Malay involves unrelated stems. For example, *isteri* is the refined form for ‘wife’ in Standard Malay, whereas *bini* is the coarse form for the same word. Similarly, *suami* is the refined form for ‘husband’ in Standard Malay, while the coarse form is *laki*. It is highly interesting that the coarse forms *bini* ‘wife’ and *laki* ‘wife’ are the ones that have entered BM as neutral forms. In addition, the refined/coarse distinction in BM can be compared with the Ngoko/Krama system in Javanese, Ngoko being the register used for speaking to social equals or for speaking to someone who is lower on the social hierarchy than oneself, whereas Krama is the register that is used for speaking to social superiors (Poedjosoedarmo 1968). While the Ngoko/Krama system does not involve a refined/coarse distinction, some Ngoko forms involve back vowels (historically conservative) where they would correspond to front vowels in the Krama form. For example, *agama* ‘religion’ is Ngoko while *agami* is Krama; *ajon* ‘to herd or tend livestock’ is Ngoko while *ajen* is Krama; and *bubar* ‘to disperse’ is Ngoko while *bibar* is

Krama.⁵³ In a way, this parallels the fact that the historically more conservative form in BM is the coarse form. However, in BM, there is no front/back vowel distinction involved in the refined/coarse distinction, and there is no way of distinguishing between refined/coarse forms if the coarse stem does not end in /ar/, /al/ or /as/ (as compared to the Ngoko/Krama system which also uses unrelated stems). Nothing similar to the refined/coarse system of BM has been reported for other languages in the region or for other varieties of Malay.

On a separate note, that BM speakers are producing less [ɛ] forms in Singapore and also in Malacca may be due to increased contact with Standard Malay or Bahasa Melayu. Bahasa Melayu is endorsed as an official language in both countries, as compared to Hokkien, which is gradually losing younger speakers. Although it should be mentioned that most BM speakers in Singapore do not speak Malay, they exist in an ecology that comprises Malay, and may be influenced reflexively by the Malay that they hear in their environment. In particular, if they do not encounter BM as much as they do Malay (the Malay population is much bigger than the Peranakan population, and there are television programs and radio broadcasts that take place in Malay, whereas none takes place in BM), the BM vowel system may assimilate towards that of Malay, and [ɛ] forms are in danger of being lost.

With respect to [ɔ], two reasons may be postulated for its loss. One plausible reason is that phonological systems tend to not be completely symmetrical, having more front than back vowels, due to the narrowing of the vocal track towards the back of the mouth. BM is consistent with this, as speakers appear to produce fewer cases of [ɔ] than [ɛ]. This is especially evident with male speaker, E. Next, Pakir (1986) states that [ɔ] is only found in words of Hokkien origin. Noting that BM is endangered and in the event that it is decreolizing, [ɔ] may be lost simply because words of Hokkien origin are being used less frequently. This may also be caused by language contact influence from standard Malay. Still, this explanation is not entirely satisfactory because it has been observed that some speakers of BM produce [ɔ] in words of Malay origin (see section 3.1.2, where the speaker produces ɔraŋ for oraŋ ‘person’). Yet, the notions of decreolization or even language change due to contact with a more dominant language do not

⁵³ Robert Blust, p.c.

have to be abandoned. In fact, speakers are producing [o] where they would traditionally produce [ɔ]. For example, speakers may produce [binpo] ‘handkerchief’ where originally, the word is pronounced [binpɔ] in Hokkien. Language change in which the contact language becomes more like its lexifier language, due to contact influence may explain for the change in phonetic space, since Malay has the same vowels as BM, apart from [ɛ] and [ɔ].

3.7.1 Sociophonetic investigation

That BM speakers are producing more of the less refined speech is interesting. It is important to also note that the [ɛ] vowel only exists in refined speech, and that [ɛ] distinctively marks BM speech since it does not occur in Malay, and [ɛ] is constantly brought up by members of the community as an example of what differentiates BM from Malay. Note that even though [ɛ] also occurs in Hokkien, BM is much more similar to Malay than Hokkien, due to the substantial influence of the Malay lexicon, and on the surface, there is no concern that BM cannot be differentiated from Hokkien. Given these phenomena, I pose the question: what social characteristics will listeners attribute to speakers who produce either variant? Will members of the community associate the refined [ɛ] form with speaking BM, even though the form is less often produced than the coarse /ar/, /al/, and /as/ forms? A perception experiment was carried out to understand associations that BM speakers have between the variant that is changing due to language endangerment, and the speakers who produce these variants. Instead of relying on self-report measures that can generate skewed results, a carefully designed experiment that relies on the perception of phonetic variables can provide a more accurate view of language endangerment. Experiments such as the one conducted here allow researchers to approach attitudes that are below the level of consciousness, and these attitudes cannot be reported accurately using another method (Kristiansen 2010). These views may be particularly informative for researchers or community members who are interested in considering language attitudes for the revitalization prospects of BM (for example, what counts as BM, and what forms should be maintained).

For the above purposes, a matched guise experimental task was conducted. The matched-guise task is a method first pioneered by Lambert and his associates (Lambert 1967, Lambert et al. 1960). In their experiments, recordings of bilinguals speaking French or English were played to subjects who rated the speakers on various scales in order to compare their attitudes towards English and French speakers. The listeners did not recognize that the same individuals were speaking English and French in different stimuli, and so rated them differently for various traits depending on which language they were speaking (Lambert 1967, Lambert et al. 1960). Essentially, the matched guise technique enables the researcher to ask, what social characteristics do listeners attribute to speakers based only on hearing their voice (Drager 2013)?

A recent project featuring a modified version of the matched guise task is that of Campbell-Kibler (2007). Campbell-Kibler (2007) investigates the relationship between English variable (ING) and the Southern accent and the “gay accent”, as they are conceptualized in listener’s perceptions of spontaneous speech. The speech of 8 speakers was collected in sociolinguistic interviews. Speakers responded to digitally manipulated excerpts of 32 matched pairs differing in whether they were alveolar variant *-in* or velar variant *-ing*. Campbell-Kibler (2007) found that hearers associate the velar variant with education, intelligence, formality and articulateness, and the alveolar form with the lack of these characteristics. In addition, the listeners’ perceptions of these forms were also based on other characteristics attributed to the voice, whether they were Southern or gay. The alveolar variant increased the perceived strength of a Southern accent, dampening what was perceived as a gay and urban accent, and increasing perceptions of being rural and relatively uneducated. The conclusion that Campbell-Kibler (2007) draws is that accent is not perceived as a continuum as much as a multidimensional landscape surrounding a central norm. This means that, more often than not, a variant cannot be attributed one single social meaning, and an accent does not simply index more of a social meaning or less of it. Instead, there may be several related meanings that are associated with a central norm. In view of findings like this, it is important to consider the other indexed meanings that BM [ɛ] or /ar/, /al/, and /as/ forms might have besides their respective central meanings of being refined or coarse, such as whether these variants mark speakers as speaking BM or not speaking BM, or whether these variants mark speakers as being more educated, or less educated for that matter.

With regard to ongoing changes in a language, much like what is happening in BM, van Bezooijen (2005) uses the matched guise task to elicit attitudes towards the transmission of a linguistic change. van Bezooijen (2005) states that three variants of /r/ co-occur in northern Standard Dutch, spoken in the Netherlands. In addition to the older alveolar and uvular consonantal types of /r/, there is now an approximant type /r/, which is restricted to the syllable coda. Through interviews, van Bezooijen establishes that approximant /r/ is spreading from the west to other parts of the Netherlands. It is used more often by women than by men, and more often by children than by adults. Matched guises with different variants were provided by a single speaker. These guises were interspersed by filler recordings of four other speakers reading the same text. Listeners were not told about the purpose of the experiment, that the researcher wanted to establish how attractive the different variants of /r/ were. A seven point Likert scale was used for listeners from four different regions (30-40 listeners per region, equally distributed across sex and two age groups of older than 45, and younger than 45) of the Netherlands to rank their answers to questions, such as “*Would you like to talk like this yourself?*” and “*What do you think the social position is of this speaker?*” (van Bezooijen 2005:23). The finding was that listeners from the west found approximant /r/ more attractive than listeners from other regions. van Bezooijen (2005) states that this variant is associated with people in the west and with a high social position. Although van Bezooijen (2005), as well as Lambert et al. (1960) and Campbell-Kibler (2007) were not working on particularly marginalized communities, the matched guise method that they use can be replicated in an endangered language community experiencing variation, as with BM speakers.

3.7.1.1 Stimuli

Ideally, the technique utilizes pairs of tokens that are produced by the same speaker, but vary in some domain (Drager 2013). The pairs of tokens in this instance are pairs of words that differ mainly in whether the word is coarse or refined, the coarse version ending in /ar/ and the refined version ending with the [ɛ] variant. There are other coarse variants such as /al/ and /as/, but it is preferable to provide only one coarse form in this instance, so that the critical comparison is between the refined form and the coarse form, and not between the refined form and the different

underlying representations that comprise the coarse form. The experiment was conceived as part of a language documentation project, and naturally-occurring data that had been elicited during fieldwork was used. While it would have been ideal to utilise word pairs from multiple speakers, there is a small number of native speakers who naturally produce both forms, and it is a concern that these speakers would be aware of the goals of this perception experiment if they were required to participate in the perception experiment.

Hence, five pairs of tokens were extracted from the speech of two speakers (two from one speaker, and three from another). The five pairs of matched guises are presented in the list below.

No.	Target token	Gloss	Type	Speaker
1	[ba.kɑr]	burn	coarse	A
2	[ba.kɛ]	burn	refined	A
3	[bə.sɑr]	big	coarse	B
4	[bə.sɛ]	big	refined	B
5	[də.ŋɑr]	listen	coarse	A
6	[də.ŋɛ]	listen	refined	A
7	[ka.sɑr]	coarse	coarse	A
8	[ka.sɛ]	coarse	refined	A
9	[pa.sɑr]	market	coarse	B
10	[pa.sɛ]	market	refined	B

Table 19: List of matched guise tokens

Besides the 10 target tokens that comprise 5 target pairs, there were 30 filler tokens. It is important to utilize filler tokens, so that listeners do not focus too much on their responses to target tokens. It is also necessary to use a larger number of filler tokens and voices in this experiment, as this is a within-subjects experimental design. This means that participants hear and respond to the same voice in both target guises, as opposed to having one group of participants listen to one guise, and another group of participants listen to another guise. While a within-subjects design is more robust, since the subjects remain the same in this setup, it is necessary to use a larger number of fillers and voices to keep subjects from noticing that they are re-hearing the same target speakers (Drager 2013). As a result, within-subjects experiments tend to take longer than across-subjects experiments. However, this will not be a major concern, as the tokens are single words, instead of long sentences. The whole experiment takes no more than 20 minutes, even for listeners who take their time when responding. Tokens of single words are used instead of entire sentences. As these recordings are made in natural, non-studio environments, there is no way to control for any preceding or following content. Splicing the target token into a longer control sentence, following Campbell-Kibler (2007, 2009), resulted in clearly-manipulated, unnatural-sounding sentences, due to the clear difference in background noise in the recordings. In addition, when more words are used in a single token, there is a chance that the phonetic and intonational qualities of these other words will impinge on the perception of the phonological variation that this experiment is concerned with. Previous work on the perception of sociolinguistic variables has demonstrated that listeners extract social information from very small amounts of speech, including single words (Strand 2000, Rakerd and Plichta 2003, Koops 2011).

The 30 fillers include 6 fillers that are unambiguously native BM, and 24 fillers that are clearly unambiguously not native BM, shown in in tables 20 and 21 respectively. All fillers were spoken by different speakers. The 6 fillers that are clearly native BM comprise lexical items that occur in BM, but not in Malay. These fillers were extracted from fieldwork data. For the 24 fillers that can clearly not be identified as native BM, recordings were made of Singapore English words produced by Singapore English speakers, BM words produced by Americans, and English words produced by American speakers.

No.	Filler	Gloss
1	idong	nose
2	samplang	improper etiquette
3	léihéi	neck
4	kéwat	fussy
5	embok-embok	Traditional Peranakan
6	piso	knife

Table 20: List of Baba Malay fillers

No.	Filler	Gloss	Note
1	corrigh	correct	Singlish, Singlish speaker
2	dowan	don't want	Singlish, Singlish speaker
3	ex	expensive	Singlish, Singlish speaker
4	kiasu	afraid to lose	Singlish, Singlish speaker
5	laidat	like that	Singlish, Singlish speaker
6	sian	ennui	Singlish, Singlish speaker
7	suaku	mountain tortoise (having no experience of the world)	Singlish, Singlish speaker
8	tapau	to take away (food)/ fail an exam	Singlish, Singlish speaker
9	kuping	ear	Baba Malay, American speaker
10	kawin	marry	Baba Malay, American speaker
11	tuahia	big brother	Baba Malay, American speaker
12	tipu	to cheat	Baba Malay, American speaker
13	Mak	mother	Baba Malay, American speaker
14	huahi	happy	Baba Malay, American speaker
15	yipoh	mother's sister	Baba Malay, American speaker
16	menantu	daughter-in-law	Baba Malay, American speaker
17	sandwich	-	English, American speaker
18	umbrella	-	English, American speaker
19	coffee	-	English, American speaker

20	plant	-	English, American speaker
21	pencil	-	English, American speaker
22	teeth	-	English, American speaker
23	airplane	-	English, American speaker
24	ocean	-	English, American speaker

Table 21: List of clearly non- Baba Malay fillers

3.7.1.2 Procedure

While participants listen to matched-guises of words with either variants, the key question that will be asked is:

- I. Is the person a speaker of Baba Malay?

Binary yes-no responses are recorded. There are four possible outcomes. First, the two variants could both not be linked to BM speakers. This is the most unlikely outcome. Next, both variants could be associated with speakers of BM. In this case, no shift is observed. Third, the coarse variant may be linked to BM speakers, and this may provide an explanation for the continuation of a shift that will eventually eliminate [ɛ] from the phonological inventory. Fourth, the refined form [ɛ] will be linked to speaking BM, but not the coarse variants. During fieldwork, it has been noticed that younger speakers are not using the [ɛ] variant. One of the major possibilities is that younger, or less proficient speakers might identify the [ɛ] variant word as being spoken by a person who speaks BM, and state that the coarse variant word is not produced by a BM speaker, even if these younger, or less proficient speakers do not produce the [ɛ] variant form themselves. This can be attributed to two things. First, the younger, or less proficient speakers, may not know many other lexical items that differentiate BM speakers from Malay speakers. The only difference that they may be highly aware of is that [ɛ] occurs at the end of words in BM, where they would otherwise be /ar/, /al/, /as/ in Malay. This may lead them to identifying the [ɛ] variant as being BM, and the coarse variant as not being BM. The second reason, which is not completely unrelated to the first, is that the Peranakans want to differentiate themselves from the Malay community. Now that the culture and language are being threatened, the younger and/or less proficient speakers, who are experiencing the brunt of this loss, may want even more to

differentiate themselves from other ethnic communities. Crucially, other explanations may also be found in the answers to these questions that will also form part of the matched-guise task. These questions may help to disentangle the different possible interpretations of the coarse and refined guises.

- II. Does this person have many close friends who are not Peranakan?
- III. Is this person likely to take part in activities at the Peranakan Association or Gunong Sayang Association?
- IV. Is this person well-educated?

Binary yes-no responses are recorded for these questions as well. Question (II) is designed to test for the listener's perception of the speakers' orientation towards the Peranakan community. It is possible that the more closed the networks are, the more likely speakers come into contact with other speakers of BM (including fluent speakers), and know how to use both variants. Older, more proficient listeners may be more aware of this network effect than younger, less proficient listeners. However, it is rare in an exceptionally urban environment to have completely closed networks, as community members would interact with non-community members on a daily basis. It is therefore interesting to see if there is any correlation between speakers who retain [ɛ] and the listener's perception of how the speakers are oriented towards the Peranakan community. For example, the more close Peranakan friends a speaker has, the more oriented she or he is toward the Peranakan community, the more likely she or he may want to be distinguished from other communities. Thus, listeners might realize that speakers might only use the refined [ɛ] variant to distinguish themselves from the Malay community, since the coarse variant exists in Malay words from which BM words have been derived.

Related to (II), in addition to studying the listener's perception of the speaker's orientation, question (III) is conceived to uncover if there is any association between the [ɛ] form and speakers who are actively involved in Peranakan cultural activities. The outcomes could be similar to that of question (II). It is possible that the more involved speakers are in the Peranakan community, the more likely they are aware that there is a refined variant as well as a coarse one.

Again, older, more proficient listeners may be aware of this, as opposed to younger, less proficient listeners.

Question (IV) then, is designed to see if listeners perceive the [ɛ] form as being a reflection of higher education, and hence more prestigious. Recall that Lim (1988) had associated [ɛ] with urbanity. While levels of urbanity are now high in both Singapore and Malacca, it is still interesting to consider if speakers associate [ɛ] with a related aspect such as higher education levels.⁵⁴ Whether or not [ɛ] is associated with a higher level of education is also interesting, since BM has strong covert prestige as a language that signals exclusive group membership, and is regarded as a non-standard language in Singapore and Malaysia. While Peranakans are proud of their language, they commonly cite stories of being punished by their Malay teachers for using BM instead of standard Malay at school. At the same time, the Peranakans were the first to adopt English medium education in Singapore during the colonial era, and they were deemed to be a refined and wealthy class (Shellabear 1913). It is thus unclear if the listeners think of their language as being something that more educated people speak.

Other independent variables in this study are the listener's age, gender, proficiency in BM, and whether or not the listener belongs to a Peranakan cultural association. Also examined is whether there is an effect of languages that the listener speaks, other languages that the listener understands, their attitude toward the language, and their perception of the endangerment of BM. In particular, the listener's reported attitudes towards the language can be gauged through their responses to (i) to (ii) and the perceived endangerment of BM can be gauged through the questions shown in (iii) and (iv).

- i. If you have children, would you want your children to speak Baba Malay?
- ii. If a person does not speak Baba Malay, can that person be Peranakan?
- iii. Do you feel that Baba Malay is endangered?
- iv. Are you worried that Baba Malay is endangered?

⁵⁴ A higher level of education is associated with urbanity in general. While Singapore is completely urbanised, there is research specifically showing a general perception of lower levels of education at rural schools among rural school teachers themselves in Malaysia. See Othman and Muijs (2013) for example.

Yes-no responses to the above questions are collected in a post-listening questionnaire that also collects responses to questions regarding all other independent variables. A post-listening rather than pre-listening questionnaire is used since exposure to these questions could influence the listeners' responses on the matched guise task. The complete post-listening questionnaire that the listener completed has been appended in Appendix D. Reactions to both the matched-guise experiment and the questionnaire are recorded using a pen and paper method, instead of using computer experiment software, such as E-Prime and OpenSesame. Although there are clear advantages to using electronic means of recording reactions (for example, token randomization and ease of data processing), elderly participants feel more comfortable with a pen and paper task. Target tokens and fillers were randomized once, so that all participants hear the stimuli in a fixed order, but in an order that appeared random. The stimuli are portable sound files (.mp3), and each token is played twice. Participants pressed a button on the player (an iPhone 5) to listen to a new token. A pair of Sony dynamic stereo headphones (MDR-7506) was used for the task. A normalizing study was conducted with 10 other BM speakers who did not participate in this experiment, to ask if some of these words are 'more Baba Malay' than others. The BM fillers and tokens spoken by a BM speaker were more often identified as being Baba Malay than the fillers that were clearly not spoken by a BM speaker.⁵⁵

3.7.1.3 Participants

All participants learnt BM from birth, even though they state that they have different levels of proficiency. Participants were recruited through word of mouth, and as variation is expected between older and younger listeners, 20 older Peranakans and 20 younger Peranakans were recruited. Older listeners are defined as being 50 and above, and younger listeners are defined as being under 50. A group that was younger than this (18-25 year olds) did not take part since it is difficult to find young community members who are able to speak BM. Indicative of the lower numbers of BM-speaking youth, the estimated average age of active community members at the

⁵⁵ In terms of absolute responses, refined target tokens were identified as BM 50 times out of 50; coarse tokens were identified as BM 32 times out of 50; clearly BM fillers were identified as BM 52 times out of 60; Singapore English fillers were identified as BM 9 times out of 80; BM fillers spoken by Americans were identified as BM 7 times out of 80; and English fillers spoken by Americans were identified as BM 0 times out of 80.

various Peranakan cultural associations is 50. 20 per critical group, although small, is a natural limitation of working in an endangered language community. Across the older and younger groups, participants were matched for gender and whether they were from Singapore or Malacca. This means that among the 20 younger participants, there were 10 Singaporeans and 10 Malaccans. Among the 10 younger Singaporeans, there were 5 female participants as well as 5 male ones. Similarly, among the 10 younger Malaccans, there were 5 female and 5 male participants. The older participant group had the exact same composition. In addition, it is interesting to note if participants are members of cultural associations, since question (III) asks if they perceive the speaker to be a member of such an association. Half of the participants from the Singapore and half of the participants from the Malacca, matched across older and younger groups, are members of cultural associations. Among the 10 Singaporean participants who are members of Peranakan cultural associations, there were 3 older females, 2 older males, as well as 2 younger females, and 3 younger males. Among the 10 Malaccan participants who are members of similar associations, there were 3 older females, 2 older males, as well as 3 younger females and 2 younger males. The distribution of participant demographics is shown in table 22, indicating the number of participants in each group.

20 older participants								20 younger participants							
10 Singaporeans				10 Malaccans				10 Singaporeans				10 Malaccans			
5 females		5 males		5 females		5 males		5 females		5 males		5 females		5 males	
3	2	2	3	3	2	2	3	2	3	3	2	3	2	2	3
m	nm	m	nm	m	nm	m	nm	m	nm	m	nm	m	nm	m	nm

Table 22: Demographics of perception experiment participants, m indicating ‘association member’ and nm indicating ‘non-association member’

Note that due to difficulties with reading and or writing, the researcher read the survey to 8 participants and entered their responses on their behalf. These participants were the only eight participants who were in their 70s. All participants stated that they had no auditory impairment.

3.7.2 Results

In addition to reporting raw trends, statistical significance and the strength of any effects are determined by a binary mixed-effects logistic regression model, which was fit to the data by hand using the statistic tool R (R Development Core Team 2013) and *lme4* (Bates et al. 2013). The mixed effects model takes into account not only fixed effects (i.e. traditional independent variables or predicting factors), such as age group and gender, but also random effects such as individual listener's tendencies when responding. While simple linear and logistic models treat differences between responses from a single participant in the same way that they treat responses between participants, a random effects model assigns each participant an individual coefficient that can be scrutinized for anomalies. Where this experiment is concerned, binary yes-no responses are collected, and it is important to ascertain if individual listeners had a bias toward answering 'yes' or 'no', including participant as a random effect in the model can account for this bias, statistically.

Separate models were fit to the data for each of the matched guise questions (see section 3.7.1.2) for responses to only the target stimuli. Responses to fillers are attached in Appendix D.⁵⁶ As with simple logistic regression, a positive coefficient indicates a positive correlation between the dependent variable and the effect, while a negative coefficient indicates a negative correlation between the dependent variable and the effect. By default, the dependent variable is whether the listener responded 'yes' to the question being asked, so a larger coefficient indicates a higher likelihood of a 'yes' response. The random effects considered include listeners, individual item, and voice. The fixed effects considered include the guise (whether the token contained [ɛ] or /ar/), listener age group, listener gender, whether the listener is from Singapore or Malacca, her or his proficiency in BM, whether the listener belongs to a Peranakan cultural association, other languages that the listener speaks, other languages that the listener understands, as well as whether the listener would want her or his children to speak BM, whether the listener would consider a person to be Peranakan if the said person does not speak BM, whether the listener

⁵⁶ The clearly BM fillers were identified more consistently as being BM (194 times out of 240 instances), while the clearly non-BM fillers were usually identified as not being BM (924 times out of 960 instances). For a further breakdown of the non-BM fillers as well as responses to fillers on other questions, please refer to Appendix D.

feels that BM is endangered, and whether the listener is worried that BM is endangered (questions regarding these last four effects are included in the post-listening survey, see section 3.7.1.2). There are many predicting factors for a relatively small group of participants, so effects are tested separately, and only the relevant ones would be included in the final model. With regard to question (II) “Does this person have many close friends who are not Peranakan?”, no trend can be reported as the model shows no significant effect of any of the predicting factors tested. Responses to the other three questions are discussed in the subsections that follow.

3.7.2.1 Is this person a speaker of BM?

Where the main question was concerned, for all the 100 times the refined form was heard, the younger age group consistently responded that it was spoken by a BM speaker, whereas they identified the coarse form as being spoken by a BM speaker in only three instances. The older group performed very differently. While they also identified the speaker as a BM speaker for all 100 times the refined form was heard, they identified the speaker of the coarse form to be a BM speaker in 86 instances. These responses are presented in the chart in figure 26.

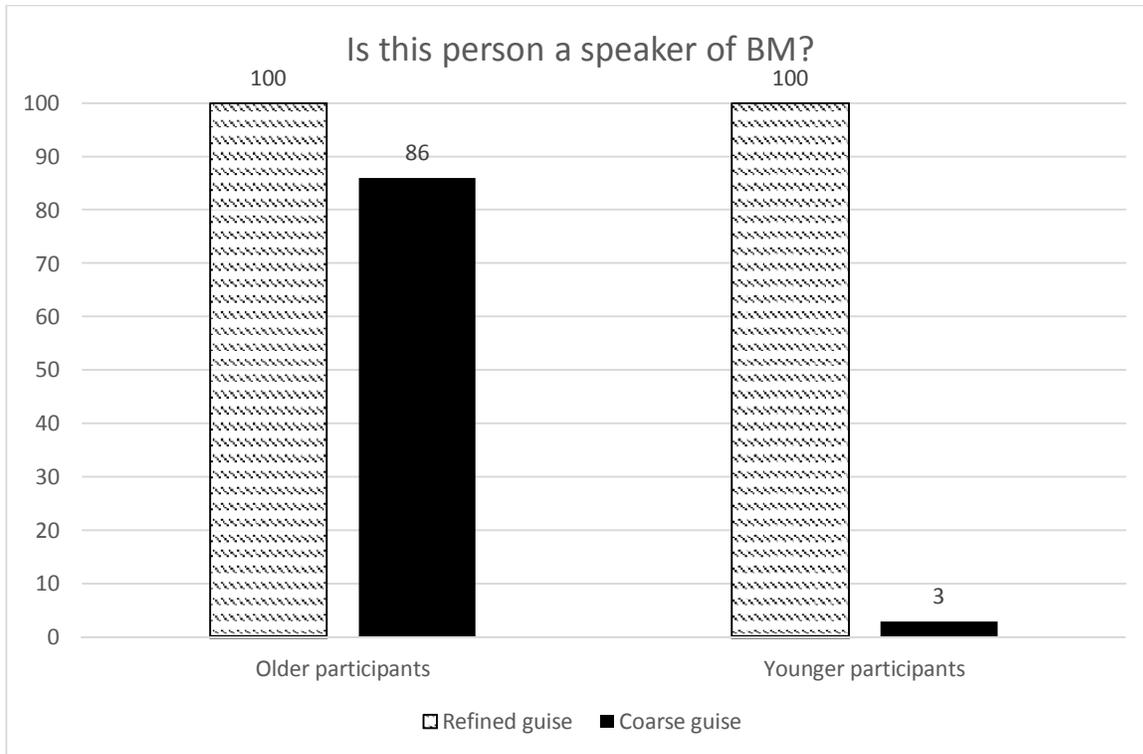


Figure 26: Responses to “Is this person a speaker of BM?” with participant’s age group as a predictor

It is not possible to generate a sound logistic regression model for the question “Is this person a speaker of BM” since there is perfect separation⁵⁷ — a refined guise completely predicts that the participant perceives the speaker as a speaker of BM. The analysis breaks down. The model has to be refit to only those responses in the coarse guise (i.e. only those tokens where not all participants perceived the speaker as someone who speaks BM). Only fixed effects that reached a significance level of $p < 0.05$ or smaller were included in this model.

⁵⁷ In the case of perfect separation, there is the option of using penalized regression (Firth 1993), but random effects cannot be included in such a model. It is important to consider random effects, hence penalized regression is not suitable for the purpose of this investigation.

	Estimate	Std. error	Z value	Pr(> z)
(Intercept)	4.729	1.436	3.293	0.0010
Age group=Young	-8.561	1.804	-4.745	<0.0001
Origin=Singapore	-2.800	1.105	-2.534	0.0113

Table 23: Coefficients of fixed effects; larger negative coefficients indicate a lower likelihood of responding ‘yes’ to the question “Is this person a speaker of BM?” when listening to a COARSE guise.

First, I step through how to interpret the model’s output and then, in the following paragraph, I discuss the results of the analysis. The model’s coefficients, listed in the ‘Estimate’ column, are negative for both predicting factors (age group and origin), indicating that coarse tokens were less likely to be perceived as being spoken by a BM speaker by younger listeners and listeners from Singapore compared with older listeners and listeners from Malacca. The coefficients are also indicative of effect size, with numbers that are further from zero being indicative of a larger effect. For listener age, the coefficient is -8.561 whereas it is -2.800 for origin; younger Peranakans are more likely to not identify the coarse form as being spoken by a BM speaker than older Peranakans, and this effect is stronger than the effect of origin. Significance level can be gleaned from column ‘Z value’, where numbers greater than 2 and smaller than -2 are significant, and the column with the p-value (labeled ‘Pr(>|z|)’), where numbers smaller than .05 are significant.

Participants who are younger (identified as being below the age of 50) are more likely to not respond ‘yes’ to the question “It this person a speaker of BM?” when they hear a coarse guise ($p < 0.0001$). This result is expected for two reasons as discussed in section 3.7.1.2. Younger speakers, being possibly less proficient, may not know many other lexical items that differentiate BM speakers from Malay speakers, the only difference they may be highly aware of, is the often-cited fact that [ɛ] can occur at the end of words in BM, where they would otherwise be /ar/, /al/, /as/ in Malay. Their responses show that they are unaware that /ar/, /al/, /as/ variants belong to BM in *kasar* ‘coarse’ forms. It is interesting that there is no significant effect of proficiency in

the mixed effects model. One would expect that older speakers are more proficient than younger speakers, and that this might have an effect on how they perceive these variants. Hence, a Kendall (1955) rank correlation test is run in R, using the package *kendall* (McLeod 2014). The Kendall rank correlation coefficient is able to evaluate the degree of similarity between two sets of scalar data. In effect, presented with (x_1, y_1) and (x_2, y_2) , the Kendall coefficient indicates how similar x_1 and x_2 , and y_1 and y_2 are. If the coefficient is zero, there is no correlation between these two items. When the coefficient is one, there is full correlation between the two items. The two items analyzed using Kendall' in this instance are age (grouped by tens), and proficiency (scale of one to five). Kendall's tau (τ) = 0.44, and the $z=3.2$, which is large enough to reject the null hypothesis that there is no correlation between the two items, at $p < 0.01$.⁵⁸ Hence there is a correlation between age and proficiency, and that the older participants are, the more likely they are to be more proficient.

Next, younger speakers who are experiencing the brunt of the loss of their culture and language may want to differentiate themselves even more from other ethnic communities such as the Malays. It is said that the third generation often expresses regret over the loss of their communities' languages and cultural practices (Stebbins 2004). In a production study on Cajun English within the St. Landry community, Dubois and Horvath (1998) demonstrate that the older generation use more dental variants than all others, the middle-aged dramatically decrease their use, but the younger generation shows a level of usage closer to the old generation.⁵⁹ Therefore, that the younger Peranakans do not identify the coarse form as being spoken by a BM speaker appears to be indicative of a desire to keep their language and culture distinct from others. This is an interesting phenomenon, considering that the refined form [ɛ] is falling out of use among less proficient BM speakers (see section 3.7).⁶⁰

⁵⁸ R states that exact p-values cannot be computed with ties, but the tau value still shows that there is a fair level of correlation.

⁵⁹ Except for young men in and women of all ages in open networks (See Dubois and Horvath 1998).

⁶⁰ Participants self-reported their proficiency on a scale of 1 to 5 in the post-listening survey, with 5 representing the highest level of proficiency and 1 representing the lowest level. The lowest level of proficiency reported was 3, while the highest was 5. In a separate mixed effects model (dependent variable: whether participants responded 'yes' to coarse form as being spoken by a BM speaker) that included proficiency as a fixed effect, the coefficient was - 5.498e-01 with a p value of 0.6705 for proficiency level 3. This indicates that there is a likelihood of less proficient

Interestingly and more unexpectedly, the fixed effect of origin has a significant effect on whether or not participants identified the coarse variant as being spoken by a BM speaker. Singaporeans (compared to Malaccans) were more likely to not identify the coarse guise as being spoken by a BM speaker. There is however no interaction between the fixed effects of age group and origin, meaning that, younger Singaporeans are not more likely to not identify the coarse guise as being spoken by a BM speaker, as compared to younger Malaccans. Returning to why Singaporean speakers may be more likely to not identify the coarse form as being spoken by a BM speaker in comparison to Malaccans, the reason may be highly political. Whereas Singaporeans have no incentives to identify themselves with other ethnic groups, Malaccans being Malaysians, may choose to align themselves with the Malays (the privileged *bumiputra*⁶¹), their country being pro-Malay when it comes to education, jobs, social and business opportunities (Holst 2012). Thus, Malaccan Peranakans may have less inclination to view themselves as being separate from the Malays, whereas Singaporean Peranakans are not disinclined towards disassociating themselves from the Malays. The fixed effect of origin is however less significant than that of age (at $p < 0.1$, compared to $p < 0.0001$).

3.7.2.2 Is this person likely to take part in activities at the Peranakan Association or Gunong Sayang Association?

The main predictor of whether a participant would respond ‘yes’ to the question is again, age. Out of all 100 instances that the refined guise was heard, it was identified by both older participants and younger participants as being spoken by a person who is likely to participate in activities at the Peranakan Association or the Gunong Sayang Association 96 percent of the time. When the coarse guise was heard, older participants identified it as being spoken by a person likely to participate in cultural association activities in 76 instances, whereas younger

participants identifying the coarse variant as being spoken by a BM speaker, although this is far from reaching significance. Future production work is required to test the interaction between proficiency and age group in particular.

⁶¹ In Sanskrit, *bhumi* means ‘earth’ and *putra* means ‘son’, the definition of *bumiputra* means ‘son of the earth’.

participants only did so for 2 instances.

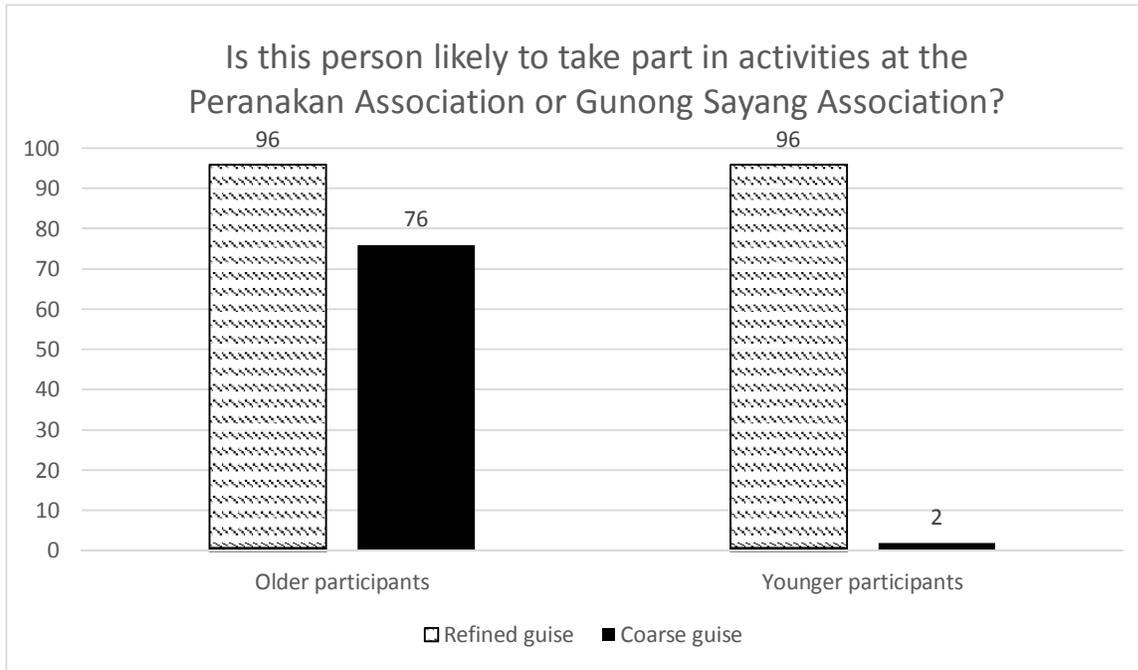


Figure 27: Yes responses to “Is this person likely to take part in activities at the Peranakan Association or Gunong Sayang Association?” with participant’s age group as a predictor

Compared to the question in section 3.7.2.1, there is no perfect separation. Not all refined guises are perceived as being spoken by a BM speaker. It is therefore possible to generate a binary mixed effects logistic model with the dependent variable being whether participants respond ‘yes’ to the question “Is this person likely to take part in activities at the Peranakan Association or Gunong Sayang Association?”. Because the model is fit to all of the data, guise can be included as a predicting factor in addition to the independent variables tested in section 3.7.2.1. Again, only fixed effects that reached a significance level of $p < 0.05$ or smaller were included in this model.

	Estimate	Std. error	Z value	Pr(> z)
(Intercept)	1.973	0.708	2.787	0.0053
Guise=Refined	2.952	0.802	3.682	0.0002
Age group=Young	-7.109	1.341	-5.302	<0.0001
Guise=Refined:Age group=Young	6.745	1.539	4.383	<0.0001

Table 24: Coefficients of fixed effects; larger negative coefficients indicate a lower likelihood of responding ‘yes’ to the question “Is this person likely to take part in activities at the Peranakan Association or Gunong Sayang Association?”

The positive intercept (1.973) indicates that people are likely to respond ‘yes’ to this question in general. The significant fixed effects on the other hand, include guise and age group. The positive coefficient for the refined guise factor, shows that participants are more likely to respond ‘yes’ to the question “Is this person likely to take part in activities at the Peranakan Association or Gunong Sayang Association?”, and this is significant at $p < 0.001$. The negative coefficient for the young age group factor suggests that the younger participants are more likely to not respond ‘yes’ in general, to the same question, although the interaction coefficient between refined guise and the younger age group is positive. This indicates that younger participants are more likely to respond ‘yes’ when listening to the refined guise, rather than when listening to the coarse guise. Both age group effect, and interaction effect between age group and guise are found to be highly significant at $p < 0.0001$.

Results show that the young participants shift the most between guises, the implication of which is that there is a change in progress with regard to the social meaning attributed to both guises. The assumption of the younger Peranakans appears to be that the refined form is more emblematic of the Peranakan culture and community than the coarse form, hence they identify the refined form as being spoken by a person who is likely to take part in activities at the Peranakan Association or Gunong Sayang Association, both being associations that emphasize Peranakan cultural activities. Again, this also shows that the younger listeners may be less

proficient in BM, and less aware that the coarse variant exists alongside the refined one. The older, more proficient speakers may be more aware of this. They do not identify the refined form as being more emblematic of their culture and community than the coarse form. There is no relation between the older age group and the refined guise; the older participants are not more likely to state that the refined guise is spoken by a person who is actively involved in cultural activities, although there is an overall main effect of guise that the older participants are more likely to state that the speaker takes part in cultural activities.

3.7.2.3 Is this person well-educated?

With regard to the question “Is this person well-educated?”, three main trends can be observed. The strongest trend is that of participant’s age group. In most instances, older participants perceived the speaker as being well-educated, regardless of which guise they heard, whereas when the refined guise was heard, younger participants perceived the speaker as being well-educated in 74 instances. This can be compared to their responses to the coarse guise. Younger participants perceived the speaker as being well-educated in 43 out of 100 instances that the coarse guise was heard. These raw responses are reflected in figure 28.

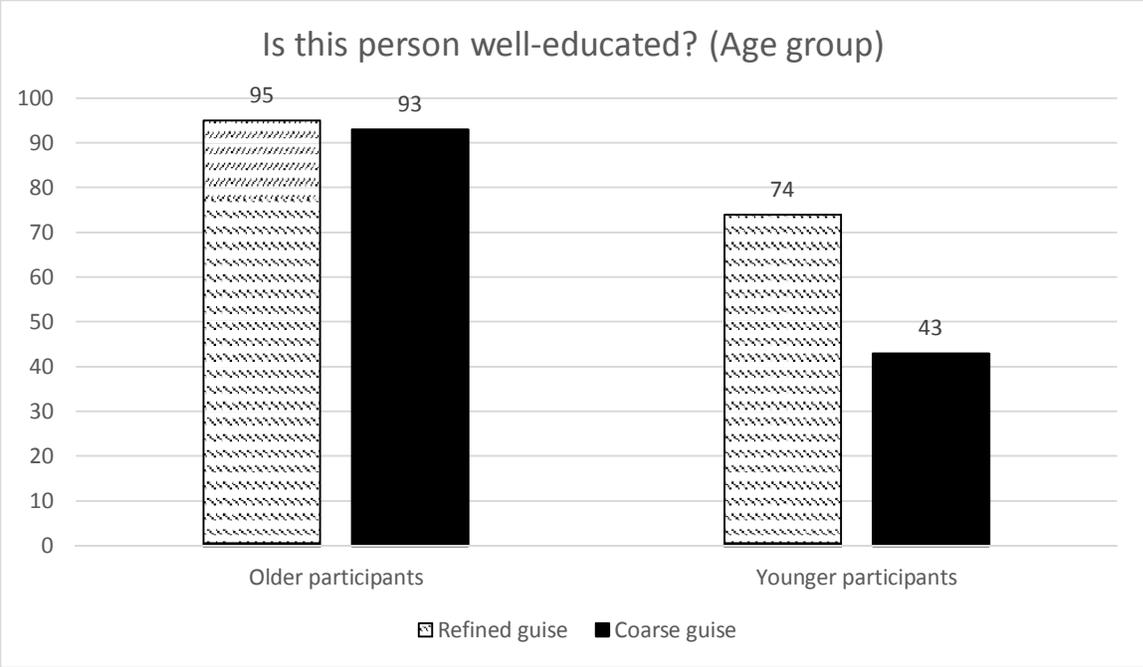


Figure 28: Yes responses to “Is this person well-educated?” with participant’s age group as a predictor

Participant’s origin is also an interesting predictor. Malaccan participants appeared to perceive the speaker as being well-educated regardless of guise. Singaporean participants perceived the speaker as being well-educated less often than the Malaccan speakers. Singaporean participants perceived the speaker as being well educated in 73 out of 100 instances in which the refined guise was heard, as compared to 52 out of 100 instances in which the coarse guise was heard. These raw responses are captured in figure 29.

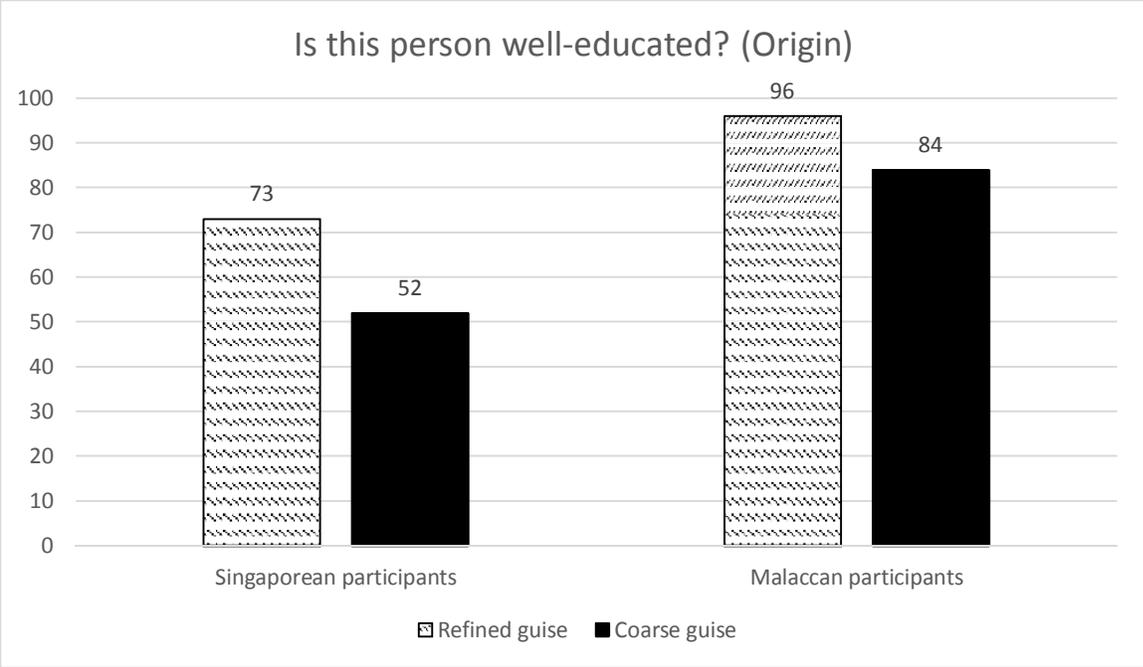


Figure 29: Yes responses to “Is this person well-educated?” with participant’s origin as a predictor

A third predictor that should be considered is whether or not the participant belongs to a Peranakan cultural association. Non-members appeared to perceive the speaker as being well-educated more often than association members, regardless of guise. Among association members, the refined guise was more often perceived as being well-educated than the coarse guise. In 100 instances that the refined guise was heard, association members perceived the speaker as being well-educated 77 times. In comparison, the same group perceived the coarse guise as being spoken by a speaker who is well-educated in 56 out of 100 instances that they heard. These raw responses are depicted in the chart in figure 30.

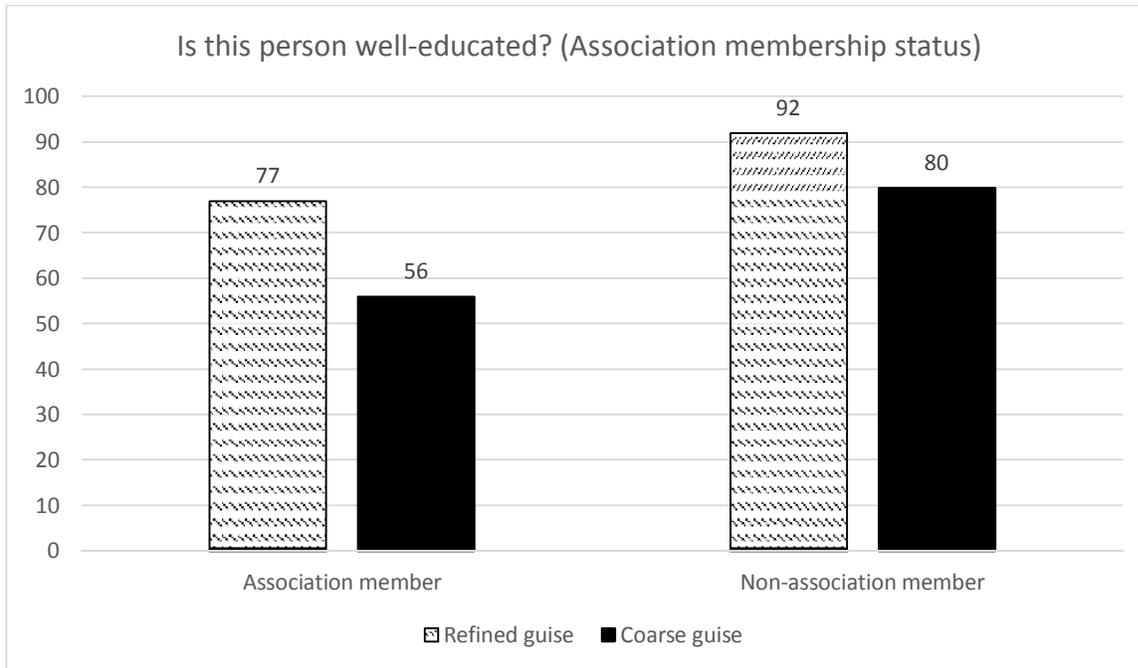


Figure 30: Yes responses to “Is this person well-educated?” with participant’s cultural association membership status as a predictor

A mixed model is generated for the question “Is this person well-educated?” Similar to the mixed effects logistic model in section 3.7.2.2, the dependent variable is whether participants respond ‘yes’ to the question being asked. Only fixed effects that are reached a level of significance of $p < 0.05$, either on their own or in interaction with other fixed effects, are included in this model.

	Estimate	Std. error	Z value	Pr(> z)
(Intercept)	10.703	2.519	4.249	<0.0001
Guise=Refined	0.511	0.716	0.713	0.4757
Age group=Young	- 7.289	1.768	-4.124	<0.0001
Origin=Singapore	-4.991	1.642	-3.039	0.0024
Association = Yes	-3.924	1.594	-2.462	0.0138
Guise=Refined:Age group = Young	3.932	1.158	3.396	0.0007

Table 25: Coefficients of fixed effects; larger negative coefficients indicate a lower likelihood of responding ‘yes’ to the question “Is this person well-educated?”

Fixed effects that are significant ($p < 0.05$) regardless of any interactions with other fixed effects, include age group, origin, and whether or not the participant belonged to an association. More specifically, younger participants, Singaporean participants, as well as those who belonged to a cultural association, were more likely to not respond ‘yes’ to the question “Is this person well-educated?” Of these effects, the most significant one is that of age ($p < 0.0001$), followed by the participant’s origin ($p < 0.01$), and finally that of association membership ($p < 0.05$).

It is then interesting to look at the relationships between age group, origin, membership, and guise, the question being whether or not younger participants, Singaporean participants, and participants who belong to Peranakan cultural associations are more likely to identify one guise as being more “well-educated” than another. However, only the interaction between young participants and the refined guise is significant ($p < 0.001$). Young participants are more likely to identify the refined guise as being spoken by a person who is well-educated. The other interactions have not been included in this final model presented, since those did not reach significance level of $p < 0.05$. The only substantial inference that can be made regarding participant’s origin and participant’s association membership status, is that Singaporean

participants and participants who belong to Peranakan cultural associations appear to be more selective as to who they perceive as being well-educated.

While there is no clear trend where the older participants are concerned, younger participants perceive speakers who sound “more BM” as being better educated. This is particularly interesting, considering that BM has a strong covert prestige, but not overt prestige. BM is not viewed as a language of particular prestige by people outside the community, it is not beneficial for upward social or economic mobility in the wider community, and it is regarded as a non-standard language in both Singapore and Malaysia. While their home language never gained any official status, the Peranakans have been known to place an emphasis on education. Before British trading began in the region, the male Peranakans were sent to China to be educated there (Ansaldo et al. 2007:208). When English medium education started in Singapore, the Peranakans were the first to adopt it (Shellabear 1913). In general, the Peranakans were perceived as “the best educated and wealthiest and most intelligent section of the Chinese community” (Nathan 1922:77). Since then, education has become widespread in both Singapore and Malaysia, and the Peranakans are no longer more educated than other ethnic groups. It is thus interesting that younger participants still associate the refined, “more BM” form with speakers who are better-educated, since it is common for them to receive the same level of education with their counterparts from other ethnic groups. On the other hand, it is curious that there is no identifiable trend of older participants perceiving the refined form as being more educated. There are two reasons for why this may be so. First, the refined form is not significantly “more BM” to the older participants than the coarse form (see section 3.7.2.1), and hence, there may be no reason for them to perceive the refined form as being more or less educated than the coarse form. Next, commonly cited experiences among older fieldwork consultants include being punished by Malay teachers for using BM instead of standard Malay. Punishments have since become less harsh⁶², and younger participants are less affected, especially in Singapore, since Singaporean Peranakans are now required to study Mandarin instead of Malay (as they are identified as

⁶² Older consultants (above the age of 60) shared classroom experiences with the researcher, and these include having been caned or fined for using BM in Malay language class; another consultant shared that his children were very recently told off by a Malay teacher for using BM in class, but they were not punished.

ethnically Chinese) (see Pakir 1994). If younger participants do not associate bad experiences in the classroom or in education with BM, they may not be as disinclined towards perceiving their language as being “more educated”. However, due to the fact that older participants often responded ‘yes’ to the question “Is this person well-educated?”, this interpretation may be less valid than the first one.

3.7.3 Implications of sociophonetics experiment

Several implications arise from the sociophonetics experiment.

The phenomenon of the loss of [ɛ] is somewhat paradoxical. Even though the data show change in apparent time for [ɛ] (refined form), where [ɛ] is not being passed on to the younger generation (see section 3.7), the sound has perceptual salience particularly for the younger Peranakans, who may not know other lexical terms that differentiate BM from Malay, or who may not know that /aɾ/, /aɭ/ and /as/ are coarse forms in BM (see section 3.7.2.1). Usage of the refined form also allows BM speakers to maintain a separate Peranakan identity. Facing the possible loss of their language, [ɛ] outrightly distinguishes BM speakers from Malay speakers, since the coarse forms /aɾ/, /aɭ/ and /as/ forms exist in Malay, but not the refined forms that utilize [ɛ]. In addition to perceiving the refined form as being “more BM” than the coarse form, younger Peranakans also perceive the refined form as being produced by Peranakans who are active in associations that emphasize cultural activities, and who are hence interested in keeping the culture and community alive.

On a separate but related note, the phenomenon that younger Peranakans have strong associations with a variant that they are not producing as often fits in with the notion of passive heritage speakers who are able to comprehend their language and participate in communication with other speakers, even though they may not speak the language itself (Valdés 2005: 419). While the younger participants in the current sociophonetic investigation may not be passive heritage speakers (since they are able to speak the language to some degree), what is noteworthy is the fact that these participants differ in the social meanings that they assign to the BM variants.

This could have consequences for language change in the future, where the two variants in question may lose their refined and coarse contrast and develop new fixed meanings.

The second paradox is the fact that a form associated with covert prestige can also reflect a high level of education, and hence better socioeconomic standing. Overt prestige, usually associated with higher socioeconomic standing, is supposedly characteristic of standard varieties (Labov 1972: 249), such as standard Malay. It is therefore interesting that the younger Peranakans perceive the “more BM” refined form as being spoken by well-educated speakers. Where language endangerment and revitalization is concerned, it is encouraging that the language bears no stigma for the younger generation, as well as the older generation who identify both forms as being spoken by a person who is well-educated.

Responses to other questions used as predicting factors for the sociophonetics experiment corroborate the potential for BM revitalization among the Peranakans. These include binary responses to the following questions (replicated from section 3.7.1.2).

- i. If you have children, would you want your children to speak Baba Malay?
- ii. If a person does not speak Baba Malay, can that person be Peranakan?
- iii. Do you feel that Baba Malay is endangered?
- iv. Are you worried that Baba Malay is endangered?

Responses to these questions did not affect the results to the matched guise tasks, but they show that the Peranakans are very concerned about language loss. In response to question (iii), 82.5% felt that BM is an endangered language. This includes responses from both the younger age group (80%) and the older age group (85%). All other questions yielded 100% to the same responses among both age groups (the lack of variation in responses also meant that these items could not be included as fixed effects in the models discussed in section 3.7.2.). While no participant felt that it was necessary for a person to speak BM in order to be Peranakan, they also responded that they would want their children to speak BM if they had children, and that they were worried about BM being endangered.

In general, the results from this experiment are encouraging for any potential revitalization efforts. Going forward, future research should investigate production more closely, especially the link between proficiency and the production of refined/coarse forms. With regard to perception itself, an interesting question is whether or not perceptions of refined/coarse forms will differ when the guises of older and younger speakers are contrasted, since a clear disparity in the perception of refined/coarse forms exists among younger and older Peranakans. Previous research has shown that social characteristics (including age) attributed to the speaker may influence the listener's perception of the speaker's variant (Drager 2011). It would be interesting to find out if listeners identify a variant as being refined more often when listening to an older speaker than a younger one. If they do, this may imply that listeners are more exposed to older speakers who utilize the refined form than younger ones. This would further corroborate the notion that language change is in process, and that [ε] may or may not fall out of use in the future.

4. Parts of speech

Since BM is mostly an isolating language, more so than Malay (which utilizes a number of affixes), word order is crucial in expressing grammatical relations. In view of this, the different parts of speech in BM can be identified appropriately by their individual distribution. This is more accurate than purely using semantics, whereby an action constitutes a verb, while a thing constitutes a noun, and a descriptive word constitutes an adjective. This is especially so when it is unclear semantically if the word is an action, a thing or a description. For example, the word *salah* ‘mistake, wrong’ can be both a noun and an adjective depending on how they are used. When used following a transitive verb, it is the object complement of the verb, when used immediately after a subject noun phrase, it can be assumed to most likely be an adjective (the copula verb is usually not expressed explicitly in BM).

- (125) *Gua bikin salah*
1.SG make **mistake**
‘I make mistake.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:19:04.0-00:19:05.8)

- (126) *Gua salah*
1.SG **wrong**
‘I (am) wrong.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:40:04.0-00:40:05.2)

Similarly, *marah* ‘angry’ can be used as an adjective or a transitive verb. After *jangan* ‘do not’, it is used as a verb (section 5.2.8) and where it precedes a noun phrase object complement as with (128), it is a transitive verb.

- (127) *Jangan marah*
do.not **angry**
‘Do not (be) angry.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:20:18.1-00:20:19.2)

- (128) *Dia-orang semua marah dia lah.*
3.PL all **angry** 3.SG EMP
‘They are all angry at him.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:04:37.9-00:04:40.4)

Hence, this grammar prefers to differentiate between parts of speech using distribution rather than semantic content. Note that square brackets [] are henceforth used to delineate meaning constituents that are relevant to the discussion, and no longer represent phonetic forms.

4.1 Nouns

Nouns in BM can be identified by the following distribution patterns. They can be used after noun classifiers (section 4.1.2), before and after demonstratives (section 4.5.2), and after numerals (section 4.5.3). The following are respective examples of these.

(129) *Satu batang payong*
 one CLF.long.thin umbrella
 ‘One umbrella’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-156, 00:05:43.5-00:05:45.0)

(130) *Budak ini*
 child this
 ‘this child.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:23.5-00:03:24.0)

(131) *Gua mo beli ni kayu.*
 1.SG want buy this wood
 ‘I want to buy this wood.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:02:23.3-00:02:24.8)

(132) *Ni tiga ékor babi kuluar rumah.*
 this three CLF.animal pig go.out house
 ‘These three pigs went out of the house.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:00:43.7-00:00:47.2)

Proper nouns (not personal pronouns) can be used after person markers, although this is optional (section 4.5.1).

(133) *Kat mana si Mary mo jumpa John?*
 PREP where PERSON want meet
 ‘Where does Mary want to meet John at?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:02:44.4-00:02:47.3)

Also, in general, the noun phrase can precede verb phrase as subject, or follow verb as object (section 5.2). In examples (134) and (135) precede and follow the verb. The noun phrases *orang*

‘person’ and *dia* ‘3.SG’ precede the verb phrase as subjects, while the noun phrases *pokok* ‘tree’ and *tu orang* ‘that person’ follow the verb phrase as objects predicated by the verb.

- (134) *Orang panjat pokok.*
Person climb tree
‘A person climbs a tree.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:00:15.6-00:00:17.6)
- (135) *Dia téngok tu orang*
3.SG see that person
‘He saw that person.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:02:29.7-00:02:30.9)

4.1.1 Nominal morphology

Nouns can occur on their own without affixes, as with *anjing* ‘dog’ in (136) and *pokok* ‘tree’ in (137).

- (136) *Anjing tu gonggong.*
dog that bark
‘That dog barked.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:01:48.9-00:01:51.8)
- (137) *Empat pokok*
four tree
‘Four trees.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:25.7-00:00:26.8)

Nouns may also be reduplicated. In example (138), *adék* is singular for ‘sibling’ while the reduplicated form *adék-beradék* denotes the plural form ‘siblings’. In example (139), *pék* represents father’s elder brother, but this stem is never used on its own, thus the form *pék-pék* is preferred.

- (138) *Pangkat adék-beradék*
rank sibling -PL
‘Cousins.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-132, 00:03:41.4-00:03:43.8)

(139) *Pék-pék*⁶³

‘father’s elder brother.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-132, 00:07:10.4-00:07:11.4)

Noun reduplication is not as productive as verb reduplication is. Noun reduplication appears largely limited to reduplication of kinship terms as example (139) demonstrates (see section 3.2 for more details).

In addition to reduplicated forms, a noun may be created using the *-an* nominalizer suffix.

4.1.1.1 Derivational noun morphology

Derivational morphology that create new nouns is non-productive. In fact, the nominalizer suffix *-an* is only found with adjectives *manyak* ‘many’ and *kasair* ‘coarse’ in the corpus used. This creates nouns *manyak-an* ‘many’ and *kasair-an* ‘coarse manner’.

(140) *Manyak -an chakapbikin* *manyak*

many -NMZ speak make/do many

‘Many say *bikin* a lot.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:00:36.8-00:00:40.0)

(141) *Manyak -an chakaptua.hia*

Many -NMZ speak big.brother

‘Many say *tua.hia*.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-132, 00:01:38.6-00:01:38.5)

(142) *Kasair -an*

coarse-NMZ

‘Coarse manner.’

Aside from nominalizer *-an*, there is virtually no other derivational morphology concerning nouns in SBM.

⁶³ Earlier, *pék-pék* was represented as *peʔ-peʔ* since the phonological form was being cited. Here, the orthographic form is used.

4.1.1.2 Nominal compounds

Nominal compounds in BM are noun phrases comprising a noun modified by either noun, adjective, verb, or even classifier. The following table lists examples of nominal compounds found in the language, as well as their patterns of modification. For the purpose of illustration, nominal compounds that use *ati* ‘heart’, *mata* ‘eye’ and *mulot* ‘mouth’ are listed. Note that none of these patterns are productive ones.

<i>Nominal compound</i>	<i>Meaning of individual lexemes</i>	<i>Meaning of compound</i>	<i>Parts of speech</i>
<i>ati baik</i>	heart good	kind	N + A
<i>ati busok</i>	heart smelly	cruel	N + A
<i>ati it gor it chap</i>	heart one five one ten	nervous	N + NUM + NUM + NUM + NUM
<i>jantong ati</i>	heart	beloved	N + N
<i>mata ari</i>	eye day	sun	N + N
<i>mata ayé</i>	eye water	sweetheart	N + N
<i>mata beliak</i>	eye glare	protruding eyes	N + V
<i>mata gelap</i>	eye dark	detective	N + A
<i>mata ikan</i>	eye fish	wart	N + N
<i>mata juling</i>	eye squinty	cockeye	N + A
<i>mata kucing</i>	eye cat	longan fruit	N + N

<i>mata lembu</i>	eye cow	fried egg with yolk intact	N + N
<i>mata mata</i>	eye eye	policeman	N + N
<i>mata piso</i>	eye knife	blade of knife	N + N
<i>bijik mata</i>	CLF.small.round eye	favorite child, apple of one's eye	CLF + N
<i>tanda mata</i>	sign eye	heirloom	N + N
<i>mulot asin</i>	mouth salty	ability to make accurate predictions	N + A
<i>mulot béngok</i>	mouth twisted	twisted mouth	N + A
<i>mulot berat</i>	mouth heavy	inability to express oneself	N + A
<i>mulot bocho</i>	mouth leaky	inability to keep a secret	N + A
<i>mulot busok</i>	mouth smelly	bad breath	N + A
<i>mulot gatair</i>	mouth itchy	uncontrollable mouth	N + A
<i>mulot dunya</i>	mouth world	public opinions	N + N
<i>mulot jahat</i>	mouth evil	caustic mouth	N + A
<i>mulot kering</i>	mouth dry	inability to say more	N + A
<i>mulot manis</i>	mouth sweet	ability to speak sweetly	N + A
<i>mulot pantat ayam</i>	mouth buttocks chicken	deceitful mouth	N + N + N
<i>mulot ringan</i>	mouth slim	polite character	N + A

Table 26: List of some nominal compounds in Baba Malay

In these nominal compounds, the modifier occurs after the head, except for *bijik mata* ‘favorite child, apple of one’s eye’, where a classifier precedes the noun. The head+modifier form of these nominal compounds is interesting because modifiers do not always occur after nouns in BM (See sections 4.5.2 and 5.1.5 for examples). With most modifiers following nouns, except for classifiers, the word order in BM nominal compounds is similar to word order in the lexifier language, Malay. Modifiers occur after noun heads, except for classifiers. For example, two blue cars would be expressed in Malay as *dua buah keréta biru* ‘two CLF.big.object cars blue’. Note that modifiers occur before noun heads in the substrate language Hokkien.

4.1.2 Noun classifiers and partitives

Noun classifiers and partitives (mass classifiers) are common in BM, although they are non-obligatory. Noun classifiers are presented in the following table.

animals and young children	<i>ékor</i>
fruit	<i>buah</i>
bloom	<i>kuntum</i>
small round items	<i>bijik</i>
durian flesh	<i>uluair (ular, uluair: snake)</i>
long thin items	<i>batang</i>

Table 27: Noun classifiers in Baba Malay

These classifiers are used in front of nouns, together with numerals. Some examples of these are as follows:

- (143) *Dua ékor ikan*
 two CLF.animal fish
 two fish
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:01:57.9-00:01:59.1)

- (144) *Ini buah pear*
 this CLF.fruit
 ‘These are pears.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:27.9-00:00:28.9)
- (145) *Se-kuntum bunga*
 one-CLF.bloom flower
 ‘One flower.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-160, 00:02:04.7-00:02:05.7)
- (146) *Dua bijik intan*
 Two CLF.small.round diamond.chip
 ‘Two diamond chips.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:07:12.9-00:07:16.2)
- (147) *Dua uluair durian*
 two CLF.durian.flesh/snake
 ‘Two pieces of durian.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:04:39.5-00:04:41.3)

Besides noun classifiers, BM speakers also use measure words when speaking about mass nouns in partitive phrases.

bundle	<i>bungkus</i>
bottle	<i>botol</i>
slice	<i>kepéng</i>
cup	<i>changke</i>
bowl	<i>mangkok</i>
spoon	<i>séndok</i>
sheet	<i>hélai, lay, éla</i>
gallon	<i>gantang</i>
quarter gallon	<i>chupak</i>

Table 28: Mass classifiers in Baba Malay

Some examples of these partitive phrases are as follows.

- (148) *Satu bungkus nasik*
one **bundle** cooked.rice
'One bundle of rice.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-149, 00:27:14.4-00:27:15.7)
- (149) *Potong-kan satu kepéng kék*
cut-TR one **piece** cake
'One piece of cake.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-149, 00:01:55.5-00:01:59.3)
- (150) *Tuang-kan satu changke kopi*
pour-TR one **cup** coffee
'One cup of coffee'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-156, 00:16:08.0-00:16:10.1)
- (151) *Satu mangkok nasik*
One **bowl** cooked.rice
'One bowl of rice.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:26:58.0-00:27:01.4)
- (152) *Kasi gua dua séndok nasik*
Give 1.SG two **spoon** cooked.rice
'Give me two spoons of rice.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:24:47.0-00:24:50.0)
- (153) *Satu lay kain*
One **sheet** cloth
'One sheet of cloth.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-149, 00:08:02.5-00:08:06.0)

4.1.3 Pronouns

Pronouns here refer to a closed set of grammatical items that may substitute a noun phrase. The set of pronouns in BM include personal pronouns, reflexives, reciprocals and interrogative pronouns.

4.1.3.1 Personal pronouns

The following is the system of personal pronouns used in BM.

1.SG (refined)	<i>saya</i>
(coarse)	<i>gua</i>
1.PL	<i>kita</i>
2.SG	<i>lu</i>
2.PL	<i>lu-orang/ lu</i>
3.SG	<i>dia</i>
3.PL	<i>dia-orang/ dia</i>

Table 29: Personal pronouns in Baba Malay

The personal pronominal system is a mixed system derived from Hokkien and Malay. The words for the first person (coarse) *gua* and second person *lu* are derived from Hokkien, while the rest of the terms originate from Malay.

Some aspects are noteworthy. While terms are derived from both the lexifier language, Malay, and the substrate language, Hokkien, calquing occurs whereby terms for second person and third person plurals are directly translated from Hokkien terms. Second person plural is *lin-lâng* ‘2.SG-people’ and third person plural is *i-lâng* ‘3.SG-people’ in Hokkien. In the BM versions of these words, *nang* is replaced by the Malay word for people, *orang*. It is curious that the first person plural is not a calque of *goá-lâng* ‘1.SG-people (Hokkien)’.

Next, while speakers of other languages differentiate between formal and informal registers, BM speakers differentiate have two registers depending on whether they want to sound *alus* ‘refined’, or *kasair*⁶⁴ ‘coarse’ (see section 3.3.5). Within the pronominal system, the first person singular is the only concept that is ostensibly marked as being *alus* or *kasair*. The word that BM derives from Malay, *saya*, is used as an *alus* term, while the word that BM derives from Hokkien, *gua*, is used as a *kasair* term. It is interesting to note that that older females appear to use more *saya* more often than males or younger females do.⁶⁵ That *gua* is used more often than *saya* among the Singapore BM-speaking community may be indicative that the language is undergoing changes,

⁶⁴ Also occurring as *halus* (after initial h- deletion) and *kasar* (coarse form).

⁶⁵ In fact, while females have also been observed to use *gua* for the first person singular, the researcher had been reprimanded several times by an older male speaker for using *gua* during the course of fieldwork.

whereby the younger speakers no longer are familiar with what is *alus* ‘refined’ and what is *kasair* ‘coarse’. In addition to these two forms of first personal singular pronouns, Chia (1983) notes that it was common for pre-war⁶⁶ Peranakans to refer to themselves in the third person as a show of humility. This is no longer common nowadays. Other personal pronouns are not outwardly *alus* or *kasair*, although when addressing an interlocutor, it is more *alus* to use the interlocutor’s name (and title if appropriate) rather than the second person singular *lu*. For example, when enquiring if one has eaten, it is more *alus* to say “Uncle XX sudah makan? (Has Uncle XX eaten?)”, rather than “lu sudah makan? (Have you eaten?)”.

Finally, Pakir (1986:146) notes a distinction between first person plural inclusive, *kita-orang* and first person plural exclusive *gua-orang*, and states these terms are analogous to Hokkien *lan-lang* ‘first person plural inclusive’ and *gua-lang* ‘first person plural exclusive’. These terms are not found in the current dataset used. These are also not reported by Lee (1999). Other terms that are undergoing change include second person plural *lu-orang* and third person plural *dia-orang*. Some speakers use *lu* and *dia* instead for these concepts. This variation is also observed by Lee (1999:20), but not in the earlier work of Pakir (1986). This may indicate language shift (or decreolization since the language is shifting away from the substrate language and towards the lexifier language).

4.1.3.2 Reflexives

Reuland (1999) states that the reflexivity of predicates is licensed by either inherent properties of the verb or by the addition of a form such as *self*. In BM, reflexives in BM are formed using the reflexive pronoun *sendiri* ‘self’, as also noted by Lee (1999). As a reflexive marker, *sendiri* denotes the same referent with the one of the subject noun phrase. In the more typical instances, the subject and the object could have the same referent, such as with examples (154), (155) and (156). The word *sendiri* replaces the pronoun or noun referent in the object noun phrase position.

⁶⁶ War here refers to the 1942-1945 period of the Second World War, when the Malay Archipelago was invaded by the Japanese.

- (154) *Gua boléh mandi sendiri.*
 1.SG can bathe **self**
 ‘I can bathe myself.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-150, 00:12:45.2-00:12:46.4)
- (155) *Budak-budak sumua pi mandi sendiri.*
 Child-child all go bathe **self**
 ‘All the children go bathe themselves.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-154, 00:12:20.9-00:12:22.6)
- (156) *Kita mandi sendiri*
 1.PL bathe **self**
 ‘We bathe ourselves.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-154, 00:09:38.5-00:09:39.5)

In these instances where the reflexive replaces the object in a sentence, *sendiri* can also be used immediately after the subject as an emphatic reflexive, to reinforce the notion that the predicate concerns the subject himself, herself or itself.

- (157) *Lu sendiri suap sendiri.*
 2.SG **self** feed **self**
 ‘You yourself feed yourself.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-150, 00:15:29.5-00:15:32.6)
- (158) *Dia sendiri mandi sendiri.*
 3.SG **self** bathe **self**
 ‘He himself bathes himself.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-150, 00:17:13.4-00:17:16.0)

Similarly, *sendiri* can be used as an emphatic reflexive in intransitive constructions where there is no object, reinforcing that the subject carried out the action himself or herself, and no one else.

- (159) *Gua sendiri makan, lu jangan suap gua.*
 1.SG **self** eat, 2.SG do.not feed 1.SG
 ‘I eat (by) myself, you do not (have to) feed me.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-154, 00:13:08.5-00:13:14.7)
- (160) *Dia sendiri jatoh.*
 3.SG **self** fell
 ‘He himself fell.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-150, 00:27:33.2-00:27:34.9)

- (161) *Kita sendiri tau mandi.*
 1.PL self know bathe.
 'We ourselves know how to bathe.'
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-150, 00:14:35.9-00:14:37.6)

Finally, while Lee (1999:21) suggests that reflexives can also be formed with the expression *mia diri* 'POSS self', *mia diri* 'POSS self' is more accurately analysed as being part of the object phrase. For instance, in example (162), *kita mia diri* refers to 'our selves or bodies', where *diri* is not an anaphora of *kita*. In fact, this expression was judged to be strange by some but not by others due to the subject matter.

- (162) *Kita boleh mandi kita mia diri.*
 1.PL can bathe 1.PL POSS self
 'We can bathe our selves/bodies.'
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-150, 00:14:50.4-00:14:53.5)

4.1.3.3 Reciprocals

Reciprocals in BM are indicated by the expressions *satu sama satu* 'one with one', or *satu sama lain* 'one with all' depending on the number of people involved in the predicated activity. When there are only two people involved in a reciprocal relationship, the expression *satu sama satu* 'one with one' is used as an object noun phrase, as with example (163). When more than two people are involved, the expression *satu sama lain* 'one with other' is used in the same position, as with examples (164), (165) and (166).

- (163) *Mary sama Lucy sayang satu sama satu.*
 with love one with one
 'Mary and Lucy love each another.'
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-154, 00:00:16.0-00:00:20.2)

- (164) *Ini budak.budak semua sayang satu sama lain*
 this children all love one with other
 'These children love one another.'
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-154, 00:03:12.2-00:03:18.0)

- (165) *Dia adék.beradék sayang satu sama lain*
 3.SG siblings love one with other
 'Them siblings love one another.'
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-161, 00:00:28.0-00:00:32.7)

- (166) *Itu empat budak sayang satu sama lain*
 that four children love one with other
 ‘Those four children love one another.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-150, 00:05:03.6-00:05:15.8)

There is also some individual variation in expressing reciprocity. The following are examples of how one consultant prefers to express reciprocity, even though she is aware of the *satu sama satu/ satu sama lain* expressions. In these examples, the expression *sama sama* ‘same same’ is used either before or after the predicate (note that *sama* can mean both ‘with’ and ‘same’).

- (167) *Mark sama Lucy sama sama sayang*
 with same same love
 ‘Mark and Lucy love each other.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-150, 00:00:14.4-00:00:24.0)

- (168) *Dia empat budak sayang sama sama*
 3.SG four child love same same
 ‘Them four children love one another.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-150, 00:03:34.8-00:03:42.5)

4.1.3.4 Interrogative pronouns

In BM, there are five interrogative pronouns. These are *apa* ‘what’, *siapa* ‘who’, *siapa punya* ‘whose’, *berapa* ‘how many’, and *mana* for ‘which’. Interrogative pronouns are interrogatives that can substitute noun phrases. Not all question words are interrogative pronouns as not all can be substitutes for noun phrases. Interrogative adverbs include *bila* ‘when’, *mana*⁶⁷ ‘where’, *apa pasal/ pasair/ sair* ‘what reason (why)’, *apa macham/ apacham/ amcham* ‘what like (how)’ and the rhetorical *mana ada* ‘where EXIST’ (See section 4.4.2 for more details).

⁶⁷ Note that *mana* means both ‘where’ and ‘which’ in BM.

what	<i>apa</i>
who	<i>siapa</i>
whose	<i>siapa punya / mia / nia</i>
how many	<i>berapa</i>
which	<i>mana</i>

Table 30: Interrogative pronouns in Baba Malay

The following instances demonstrate how interrogative pronouns may be used in BM. More details can be found in 5.6.10.1.

- (169) *Ini apa?*
This **what**
‘This (is) what?’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:23:11.1-00:23:12.5)
- (170) *Siapa itu?*
who that
‘Who (is) that?’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:23:11.1-00:23:12.5)
- (171) *Siapa punya bubor itu?*
who **POSS** porridge that
‘Whose porridge (is) that?’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:04:10.1-00:04:12.1)
- (172) *Umor lu berapa taon?*
age 2.SG **how.many** year
‘How many years old are you?’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:14:51.4-00:14:53.7)
- (173) *Mana baik, mana satu tak baik?*
Which good **which** one NEG good.
‘Which (is) good, which one (is) not good?’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:23.2-00:01:26.3)

4.2 Verbs

Verbs in BM can be identified by the following distribution patterns. They occur after noun subjects. Transitive verbs precede noun complements. Verbs also may occur after aspect markers such as progressive *ada*, and with other verbs in serial verb constructions. The following are respective examples of these distribution patterns.

- (174) *Budak itu ketawa*
child that laughs
'That child laughs.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:09:16.0-00:09:17.8)
- (175) *Anjing ambek bakol*
dog take basket
'The dog takes the basket.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:01:19.6-00:01:21.9)
- (176) *Dia ada bikin kuéh*
3.SG PROG make cake
'She is making cake.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:26:46.5-00:26:47.6)
- (177) *Pi tidor*
go sleep
'Go sleep.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:03:24.1-00:03:25.5)

4.2.1 Verbal morphology

Verbs can occur on their own without affixes, such as with both *dapat* 'get' and *makan* 'eat' in example (178) and *senyum* 'smile' in example (179).

- (178) *Dapat makan*
get eat
'Gets to eat.'
(Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:00:18.2-00:00:19.3)

(179) *Budak*⁶⁸*tu senyum.*

child that **smile**

‘That child smiles.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:35.9-00:00:37.7)

Verbs may also occur in reduplicated forms. In the following instances, these verbs express tentativeness, meaning that the agents are not walking anywhere or looking at anything in particular in examples (180) and (181) (see section 3.2).

(180) *Jalan jalan*

walk walk

‘Take a walk’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:13.1-00:02:14.6)

(181) *Téngok téngok*

see see

‘Take a look’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:14.6-00:02:16.2)

Other than in reduplicated forms, verbs also can take on affixes. In general, verbs undergo more affixations than nouns. The different types of affixation are discussed in following sections.

4.2.1.1 Derivational verb morphology

Inflectional affixes neither change lexical meanings nor part of speech, while derivational affixes change lexical meaning, and may also change the part of speech. In BM, there are no inflectional affixes but a few derivational ones. Two derivational suffixes that are productive are the transitive suffix *-kan* and prefix *-ter* that is attached to verbs to emphasize that something was done accidentally or with movement.

Transitivity “involves a number of components” regarding the “effectiveness with which an action takes place”, and one of the many indicators of transitivity is valency (Hopper and Thompson 1980:251). Valency refers to the number of arguments a verbal predicate takes, and in

⁶⁸ For both male and female. Gender can be differentiated by adding the words *perompuan* ‘female’ or *jantan* ‘male’. Hence *budak jantan* would indicate ‘boy’, while *budak perompuan* would indicate ‘girl’.

BM, transitivity is increased by the suffix *-kan* which causes the verbal predicate to take two arguments instead of one. The productive transitive marker *-kan* is attached to verbs, regardless of whether they are inherently transitive as with examples (182) and (183), or intransitive as with examples (184) and (185).

(182) *Amék-kan gua mia aloji*
Take-TR 1.SG POSS small.clock
 ‘Take my small clock.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:10:09.4-00:10:11.7)

(183) *Bukak -kan itu kepok.*
open-TR that box
 ‘Open that box.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:02:07.9-00:02:09.4)

(184) *Auntie Jane selalu ketawa-kan gua*
 always **laugh-TR** 1.SG
 ‘Auntie Jane always laughs (at) me.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:18:03.5-00:18:06.4)

(185) *Tak tau apa sair dia pekék-kan si John*
 NEG know what reason 3.SG **shout-TR** PERSON
 ‘Don’t know why she shouted (at) John.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:11:50.7-00:11:53.2)

Other than on verbs, the transitive suffix *-kan* can be attached to adjectives and adverbs to convert them into transitive verbs. In example (186), *panas* ‘hot’ is an adjective, while in example (187), *panas-kan* is a transitive verb meaning ‘to heat something up’. Similarly, in example (188), *kechik* ‘small’ is an adjective, while in example (189), it is a transitive verb meaning ‘to make something small.’ Example (189) also shows how *-kan* can be suffixed to an adverb (*kurang* ‘less’) to make it a transitive verb.

(186) *Ala panas sair, pakay kopiah dia.*
 EXCLAM **hot** CONF wear hat 3.SG
 ‘(Goodness it is) hot indeed, wearing a hat he is.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:02:26.7-00:02:29.2)

- (187) *Panas-kan itu kuah.*
hot-TR that gravy
 ‘Heat (up) that gravy.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:44:57.0-00:44:59.9)
- (188) *Tidor ranjang kechik*
 sleep bed **small**
 ‘Sleep (on the) small bed.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:02:01.3-00:02:03.3)
- (189) *Kalo mo kechik-kan badan kurang-kan roti*
 if want **small-TR** body **less-TR** bread
 ‘If (you) want to make your body smaller, lessen (your intake of) bread.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:12:27.5-00:12:32.0)

The other derivational affix in BM is *ter-*. It has two main uses in BM, the first of which is to indicate that an action occurred accidentally, as shown in examples (190), (191), and (192), while the second emphasizes that there is movement to the action that is taking place, as shown in examples (193), (194), and (195). It is possible that the function of *ter-* meaning ‘movement’ extended from its use as an accidental marker, since the accidental interpretation is also associated with *ter-* in Bahasa Melayu and Bahasa Indonesia. Some of the examples could possibly be interpreted as expressing either ‘accidental’ or ‘movement.’ The glosses in the examples given below are hence based on context.

- (190) *Itu tortoise, dia sendiri ter-balék*
 that 3.SG self **ACD-return**
 ‘That tortoise accidentally flipped itself (upside down).’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:10:54.4-00:10:56.6)
- (191) *Dia ter-gui*
 3.SG **ACD-kneel**
 ‘He accidentally fell on his knees.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:07:44.9-00:07:46.1)
- (192) *Dia jalan ter-peléchok*
 3.SG walk **ACD-twist.foot**
 ‘He walked (and) accidentally twisted his foot.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:07:43.1-00:07:44.9)

- (193) *Mata dia ter-kelék.kelék*
 eye 3.SG MVT-blink
 ‘His eyes blinked.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:00:23.9-00:00:24.9)
- (194) *Ter-lompat lompat*
 MVT-hop hop
 ‘Hopping (here and there).’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:00:38.6-00:00:39.8)
- (195) *Ter-kejar ni mia apa?*
 MVT-chase this REL what
 ‘(It) chased what is this?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:08.7-00:02:10.0)

While *ter-* is necessary in indicating that something is accidentally taking place, it is optional in indicating movement. This is demonstrated by example (196), which can be compared to example (194).

- (196) *Dia lompat sini, dia lompat sana*
 3.SG hop here 3.SG hop there
 ‘He hops here and there.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:08.7-00:02:10.0)

The transitive suffix *-kan* occurs in Malay (both spoken in Singapore/ Malaysia and Indonesia), where it is said to license arguments that are not syntactically licensed by the base verb (Cole and Son 2004), in effect, changing the valency of the verb. The prefix *ter-* also occurs in both varieties of standard Malay, but its meaning differs slightly than in BM. Again, in Bahasa Melayu and Bahasa Indonesia, *ter-* has been analysed as an accidental marker (Sneddon 1996, Wee 1995), as well as an active marker, a superlative marker, a passive marker, and an abilitative marker (Chung 2011). The prefix *ter-* in BM which has the same ‘accidental’ meaning, has developed its own function, indicating ‘movement’. However, it cannot be used as an active marker, a superlative marker, or as an abilitative marker.

4.2.1.2 Non-productive verb morphology

There are two more suffix forms in BM that are unproductive. These are *ber-* and *me-*. The prefix *ber-* in Malay is said technically to correlate with the notion of middle voice (Benjamin

2009, Windstedt 1927, Windstedt 1945), wherein the agent of the verb performs an action that involves the agent herself or himself. In BM, some verbs retain reflexes of the *ber-* prefix but this prefix has no lexical meaning of its own in BM. The following examples feature words that begin with the prefix form *ber-*. BM speakers however do not analyze any of these words as having two separate components, as a Malay speaker would. For example, a Malay speaker would recognize that the *bergerak* in (197) comprises the *ber-* middle voice prefix and *gerak*, meaning ‘movement’.

- (197) *Tu bangkuang tak bergerak*
 that turnip NEG move
 ‘That turnip does not move.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:01:00.5-00:01:03.3)
- (198) *Itu orang tak tau berenang*
 that person NEG know swim
 ‘He doesn’t know (how to) swim.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:16:31.0-00:16:33.2)
- (199) *Dia-orang bertengkar*
 3.PL argue
 ‘They argued.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:11:47.4-00:11:50.7)
- (200) *Bergetair lah*
 Tremble EMP
 ‘Tremble.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:05:00.2-00:05:01.7)

The presence of *ber-* is infrequent, and not used to indicate middle voice. Instead, where speakers of standard Malay would use *ber-* to form middle voice forms of *berchakap* ‘speak’, *berfikir* ‘think’⁶⁹, and *berlari* ‘run’ respectively in examples (201), (202), and (203), BM speakers prefer to produce forms without *ber-*.

- (201) *Kita chakap Peranakan*
 1.PL speak
 ‘We speak Peranakan.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:05:00.2-00:05:01.7)

⁶⁹ Malay *fikir* ‘think’ corresponds to BM *pikay*.

- (202) *Goldilocks tak pikay siapa punya bubor itu*
 NEG think who REL porridge that
 ‘Goldilocks did not think whose porridge that (was).’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:04:08.7-00:04:12.1)
- (203) *Terus dia lari*
 straight 3.SG run
 ‘Straight he ran.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:03:01.5-00:03:02.8)

In addition to the prefix form *ber-*, the prefix form *me-* is also found, occurring unproductively in BM. The prefix *me-* is commonly analyzed as an active voice marker in the standard Malays of Singapore and Malaysia, as well as Indonesia (Chung 1976, Nomoto and Shoho 2007, Son and Cole 2008). Again, in the examples below, BM speakers do not analyze these words as having two separate components, namely a *me-* active prefix and a root. They state that they do not know any component words such as *nyalap* ‘howl’ or *ngantok* ‘yawn’.

- (204) *Itu kuah sudah menidi.*
 that gravy already boil
 ‘That gravy (is) already boiled.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:45:15.2-00:45:17.3)
- (205) *Tu anjing menyalap.*
 that dog howl
 ‘That dog howled.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:01:40.2-00:01:44.6)
- (206) *Mengantok.*
 yawn
 ‘Yawn.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:29:44.0-00:29:45.5)
- (207) *Sini dua ekor memisék⁷⁰*
 here two CLF.animal whisper
 ‘Here (these) two are whispering.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043,00:13:10.2-00:13:13.5)

⁷⁰ Note that the Malay version of *me-bisek* is *membisek* ‘whisper’, the coda of the suffix agreeing in place of articulation with the onset of the first syllable on the root form *bisek*. This does not occur in BM.

In other instances, speakers produce two versions of the verb, one with *me-* and one without *me-*. Out of the lexicon collection of about 1100 words, there are only two words that commonly show up with differing forms. They are *nyanyi* and *menyanyi* ‘sing’ in examples (208) and (209), as well as *nangis* and *menangis* ‘cry’ in examples (210) and (211). While the form of *nyanyi* in (208) might be explained by the fact that the sentence is an imperative, as compared to (209), examples (210) and (211) show that *me-* does not appear to be particularly indicative of the active form as it would have been in Malay, otherwise both verb form in (210) would have to take on the *me-* prefix. Examples (5) and (6) are replicated as (210) and (211).

(208) *Nyanyi lah, lagu terang bulan ka,*
sing EMP song bright moon or
 ‘Sing the Bright Moon song or (something else),’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:08:26.2-00:08:28.9)

(209) *Kita sama sama menyanyi.*
 1.PL same same **sing**
 ‘We sing together.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:08:26.2-00:08:28.9)

(210) *Téngok bapak dia mati, dia tak nangis.*
 see father 3.SG die 3.SG NEG **cry**.
 ‘See, his father passed away, (and) he did not cry.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-021, 00:32:44.0-00:32:46.3)

(211) *Dia kuluar dari rumah menangis.*
 3.SG go.out from house **cry**
 ‘She left the house crying.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:04:18.3-00:04:20.9)

Considering that *me-* is not lexically meaningful, it is better to treat both versions of each word above as different variations of the same word, but denoting the same notion. Again, it is much more common for BM speakers to not use the *me-* form at all, as demonstrated in the following examples. Malay speakers would otherwise replace *bacha* ‘read’ with *membacha* in (212), *dengair* ‘listen’ with *mendengar*⁷¹ in (213), and *jual* ‘sell’ with *menjual* in (214) to indicate that these verbs are active.

⁷¹ There is no *-air* form in Malay.

- (212) *Mak tu ada bacha lagik.*
 mother that PROG read more
 ‘That mother is reading more.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:08:53.3-00:08:57.0)
- (213) *Gua tak dengair apa lu chakap.*
 1.SG NEG hear what 2.SG speak
 ‘I do not hear what you speak.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:09:17.6-00:09:20.7)
- (214) *Orang jual batu*
 person sell stone
 ‘The person sells stone.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:03:06.1-00:03:07.7)

Hence, both *ber-* (middle voice marker in Malay) and *me-* (active marker in Malay) prefix forms are neither productive nor lexically meaningful in BM.

4.2.1.3 Verbal compounds

Verbal compounds in BM are verb phrases comprising a verb and a modifier. Modifiers include noun, adjective, verb, as well as preposition phrase. The following table lists examples of verbal compounds found in in the language, as well as their patterns of modification. For the purpose of demonstration, the verbs *buang* ‘throw’, *buat* ‘do, make’, and *naik* ‘ascend’ are used (note that there are two words for ‘do, make’ in BM, the other word being *bikin*). None of these verbal patterns are productive ones

<i>Verbal compound</i>	<i>Meaning of individual lexemes</i>	<i>Meaning of compound</i>	<i>Parts of speech</i>
<i>buang ayé</i>	throw water	urinate	V + N
<i>buang ayé besair</i>	throw water big	defecate	V + N + A

<i>buang ayé kechik</i>	throw water small	urinate	V + N + A
<i>buang buang</i>	throw throw	exorcise	V + V
<i>buang mata</i>	throw eye	keep an eye on someone or something	V + N
<i>buang mulut</i>	throw mouth	convey by speaking (as opposed to writing)	V + N
<i>buang se-belah</i>	throw one-side	aside from	V + NUM + PREP
<i>buang segan</i>	throw shy	stretch upon waking up	V + A
<i>buang terbiat</i>	throw attitude	throw tantrum	V + N
<i>buat bodoh</i>	do/ make stupid	feign ignorance	V + N
<i>buat mahal</i>	do/ make expensive	play hard to get	V + A
<i>buat malu</i>	do/ make embarrassed	cause embarrassment	V + A
<i>buat mungka</i>	do/ make face	be sour-faced	V + N
<i>buat suay</i>	do/ make unlucky	cause misfortune	V + A
<i>buat susah</i>	do/ make difficult	cause difficulty	V + A
<i>buat tak dengair</i>	do/ make not hear	pretend to not hear	V + NEG + V
<i>naik baik</i>	ascend good	change for good	V + A
<i>naik chuan</i>	ascend breathless	be breathless	V + A
<i>naik darah</i>	ascend blood	be angry	V + N

<i>naik geléték</i>	ascend tickle	be up to mischief	V + V
<i>naik geram</i>	ascend furious	be furious	V + A
<i>naik gila</i>	ascend mad	be mad	V + A
<i>naik lemak</i>	ascend cooked.in.coconut.milk	be up to mischief	V + A
<i>naik pangkat</i>	ascend rank	be promoted	V + N
<i>naik sedap</i>	ascend delicious	be satisfied	V + A
<i>naik seram</i>	ascend frightening	be frightened	V + A

Table 31: List of some verbal compounds in BM

In general, the shape of verbal compounds appear to follow that of a verb phrase comprising verb and complement.

4.3 Adjectives

Adjectives in BM occur before or after nouns they modify. When used before a noun, they usually occur with relative marker *punya* (shortened versions being *mia* and *nia*), as in example (216). Adjectives may also occur after the verb *ada* when it functions as a copula (see section 5.2 for other usages of *ada*), as in example (217). It is however much more common for the copula to be omitted, as with example (218).

- (215) *Beruang kechik ada mangkok kechik*
bear small have bowl small
‘Small bear has (a) small bowl.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:01:20.9-00:01:22.4)

- (216) *Cherita-kan gua betol mia cherita*
story-TR 1.SG real REL story
‘Tell me a real story.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:03:20.8-00:03:26.7)

(217) *Ada baik tak a?*
 COP **good** NEG COP
 ‘Are (you) good or not?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:09.9-00:00:11.2)

(218) *Taukay kebun ni kiam sair*
 boss garden this **miserly** CONF
 ‘This boss (of the) garden is miserly indeed.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:01:03.7-00:01:05.8)

4.3.1 Adjectival morphology

Adjectives can occur on their own, as with *mahal* ‘expensive’ and *chanték* ‘beautiful’ in examples (219) and (220).

(219) *Tapi mahal lah*
 but **expensive** EMP
 ‘But (it is) expensive.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:30.8-00:02:32.9)

(220) *Chanték sekali*
beautiful very
 ‘Very beautiful.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:35.7-00:02:37.2)

Adjectives can occur in reduplicated forms, to indicate tentativeness (see section 5.2.5.7).

(221) *Tawair-tawair*
Tasteless-tasteless
 ‘Kind of tasteless.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:11:12.7-00:11:15.0)

In other instances, when reduplicated, adjectives function as adverbs (see section 4.4).

Adjectives are also not formed with affixes. When affixed with nominalizer *-an*, nouns are formed instead (see section 4.1.1.1).

4.4 Adverbs

Adverbs modify adjectives or verbs. They may thus precede or follow adjectives, as with examples (222) and (223). They may also precede or follow the verb phrases that they modify, as with examples (224) and (225).

- (222) *Kambing dia kuat sekali.*
 goat 3.SG strong **very**
 ‘The goat, it (is) very strong.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:01:52.8-00:01:54.6)
- (223) *Ada satu bangkuang dia téngok betol besair.*
 EXIST one turnip 3.SG see **really** big
 ‘There was one turnip he saw (that was) really big.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:00:27.0-00:00:30.5)
- (224) *Dia [peték tu buah pear] lagik*
 3.SG pluck that CLF.fruit **again**
 ‘He plucked those bears again.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:01:50.5-00:01:53.0)
- (225) *Si tua tu pelan-pelan [angkat satu-satu.]*
 PERSON old that **slow-slow** lift one-one
 ‘That old person slowly lifted one by one.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:04:10.4-00:04:11.5)

4.4.1 Adverbial morphology

Adverbs may occur on their own, as with *selalu* ‘always’ and *macham* ‘like that’ in examples (226) and (227). Note that the short form of *pegi* ‘go’ is *pi*.

- (226) *Dia selalu pi sana.*
 3.SG **always** go there
 ‘She always goes there.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:39.8-00:02:41.2)
- (227) *Dia-orang sudah biasa macham.*
 3.PL already used.to.it **like.that**
 ‘They are already used to it more or less.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:04:58.9-00:05:01.1)

There is no affixation process that creates adverbs. However, reduplication may result in the formation of adverbs, the base form for reduplication being adjectives in these instances.

- (228) *Pelan-pelan, satu-satu, budak masok-kan*
slow-slow one-one child enter-TR
 ‘Slowly, one-by-one, the children put (the pears) in.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:04:58.9-00:05:01.1)
- (229) *Apa sair gua chakap lu dudok diam-diam?*
 what reason 1.SG speak 2.SG sit **quiet-quiet**
 ‘Why are you sitting quietly (while) I speak?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:19:50.6-00:19:52.8)
- (230) *Sandah baik-baik*
 lean **good-good**
 ‘Lean properly.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:06:23.4-00:06:24.7)
- (231) *Misti angkat baik-baik*
 Must carry **good-good**
 ‘Must carry properly.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:05.5-00:01:07.2)

However, these adverbs may also occur in a non-reduplicated form in casual speech, as with examples (232) and (233). Essentially, adjectives may function as adverbs. Note that example (232) features both the reduplicated form *lekas lekas* ‘quickly’, and the non-reduplicated form *lekas*, which also functions as an adverb that modifies the verb, instead of as an adjective.

- (232) *Lekas-lekas datang lekas datang.*
quick-quick come **quick** come
 ‘Quickly come quickly come.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:04:58.9-00:05:01.1)
- (233) *Ketawa ketawa lekair senyum*
 laugh laugh **quick** smile
 ‘Laugh, laugh and quickly smile.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:06:16.5-00:06:18.0)

Besides not having affixes that form adverbs, BM also appears to not have adverbial compounds.

4.4.2 Interrogative adverbs

Interrogative adverbs include *bila* ‘when’, *mana* ‘where’, *apa pasal*, literally ‘what reason’, and *apa macham*, ‘what like’. *Apa pasal* and *apa macham* function as ‘why’ and ‘how’ respectively. *Mana ada*, literally meaning ‘where EXIST’ is used as a rhetorical question, where the implied answer is negative. While interrogative pronouns can substitute for the noun phrase, interrogative adverbs are able to substitute the adverbial clause. Interrogative adverbs usually concern time, location, or the manner in which something was done, except for *mana ada*. Note that there are two other forms for *pasal* ‘reason’, the first being the refined form *pasair* (see section 3.3.5), the second being *sair*, a shortened version of *pasair*. *Sair* is the most commonly used form in casual speech. The shortened version of *macham* in *apa macham* is *cham*.

when	<i>bila</i>
where	<i>mana</i>
why ‘what reason’	<i>apa pasal/ pasair/ sair</i>
how ‘what like’	<i>apa macham/ cham</i>
rhetorical ‘where EXIST’	<i>mana ada</i>

Table 32: Interrogative adverbs in BM

The following examples demonstrate how these interrogative adverbs are used in BM. More details can be found in section 5.6.10.1

- (234) ***Bila*** *gua panggay John datang?*
when 1.SG ask come
 ‘When did I ask to come?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-093, 00:07:47.7-00:07:49.6)
- (235) ***Mana pi si Mary?***
where go PERSON
 ‘Where goes Mary?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:08:42.2-00:08:43.5)

- (236) *Apa sair gua chakaplu tak jawab gua?*
what reason 1.SG speak 2.SG NEG answer 1.SG
 ‘Why are you not answering when I speak?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:20:03.9-00:20:06.9)
- (237) *Apa macham mo masak iték tim?*
what like want cook duck double-boil
 ‘How (do I) want to cook double-boiled duck (soup)?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:04:30.9-00:04:33.1)
- (238) *Mana ada piso?*
where EXIST knife
 ‘Where is the knife (implying there is no knife)?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:21:02.8-00:21:05.6)

4.5 Determiners

Determiners in BM include person markers, demonstratives, numerals and other quantifiers.

4.5.1 Person marker

Si is used as a person marker in BM. It can be used directly before a person’s name, or with attributes that describe a person (usually with personal pronouns, nouns and adjectives). The usage of *si* is not mandatory. For example, *si* does not precede *John* in example (239).

- (239) *Apa sair si Mary bunoh John?*
 what reason **PERSON** **Mary** kill John?
 ‘Why did Mary kill John?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:01:06.7-00:01:09.1)
- (240) *Si sa ko bongsu.*
PERSON **third paternal.aunt** **youngest.child**
 ‘Youngest third paternal aunt.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:09:01.2-00:09:02.9)
- (241) *Téngok si tua tu peték dia mia pear.*
 look **PERSON** **old that** pluck 3.SG POSS
 ‘Look at the old man plucking his pear.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:36.9-00:05:38.7)

- (242) *Bila gua boléh jumpa si ano ni?*
 when 1.SG can meet PERSON anonymous this
 ‘When can I meet this anonymous person?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:12:11.7-00:12:15.2)

Si in BM can be attributed to Malay, where it is also used as an optional person marker. There is no similar person marker in Hokkien, except for *eng* which is used with kinship terms (see Appendix A).

4.5.2 Demonstratives

Four demonstratives in BM are *ini* ‘this’ and *itu* ‘that’, as well as *sini* ‘here’ and *sana* ‘there’. The shortened forms of *ini* ‘this’ and *itu* ‘that’ are *ni* and *tu* respectively. *Ini* and *itu* can be used as demonstrative pronouns as with examples (243) and (244), where these words stand in for nouns in a deictic way – these words can only be understood in context.

- (243) *Ini buah pear.*
 this CLF.fruit
 ‘This is a pear.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:27.9-00:00:28.9)
- (244) *Lepas itu kita boléh buat soup bangkuang*
 after that 1.PL can make turnip
 ‘After that we can make turnip soup’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:00:38.8-00:00:46.6)

Besides demonstrative pronouns, *ini* and *itu* can also be used as demonstrative determiners that occur together with nouns. As determiners, *ini* and *itu* can precede or follow nouns, as with examples (245) and (246), and examples (247) and (248) respectively.

- (245) *Ni tiga ékor babi tinggal sama mak babi*
 this three CLF.animal pig live with mother pig
 ‘These three pigs live with mother pig.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:00:23.0-00:00:25.0)
- (246) *Itu bangkuang pun tak bergerak.*
 that turnip also NEG move
 ‘That turnip also does not move.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:01:20.0-00:01:22.2)

- (247) *Budak ini sangat kechik lah.*
 child **this** very small EMP
 ‘This child is very small.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:23.5-00:03:25.3)
- (248) *Sekarang dia tolak bicycle itu.*
 now 3.SG push **that**
 ‘Now he is pushing that bicycle.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:04:34.6-00:04:36.3)

Sini ‘here’ and *sana* ‘there’ are also to be interpreted deictically, based on context. (Note that the shortened version of *tak ada* ‘NEG have’ is *tak -a*.) Examples of their usage are as follows.

- (249) *Kita sini tak-a pokok.*
 1.PL **here** NEG-have tree
 ‘We do not have the tree here.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:35.5-00:00:36.9)
- (250) *Lama tak jalan sini.*
 Long.time NEG walk **here**
 ‘(I have) not walked here (for) a long time.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:06.5-00:00:08.3)
- (251) *Pusing sini pusing sana.*
 turn **here** turn **there**
 ‘Turns here and turns there.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:22:09.8-00:22:11.6)
- (252) *Dia tau sana ada orang.*
 3.SG know **there** EXIST people
 ‘He knows there is someone (there).’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:23.1-00:02:24.8)

4.5.3 Numerals

Two sets of cardinal numerals are used. One set is derived from Malay, while the other from Hokkien. In general, the Malay terms for numbers are used. Hokkien numerical terms are used specifically with lunar calendar dates and kinship terms.

The following table shows the set of Malay numbers used in BM. Note that *se-* is short for *satu* ‘one’, and is the only numeral that has its own shorter prefix version.

1	<i>satu</i>	11	<i>se-belas</i>
2	<i>dua</i>	12	<i>dua-belas</i>
3	<i>tiga</i>	13	<i>tiga-belas</i>
4	<i>empat</i>	20	<i>dua-puluh</i>
5	<i>lima</i>	30	<i>tiga-puluh</i>
6	<i>enam</i>	40	<i>empat-puluh</i>
7	<i>tujuh</i>	41	<i>empat-puluh satu</i>
8	<i>lapan</i>	100	<i>se-ratus</i>
9	<i>semilan</i>	200	<i>dua-ratus</i>
10	<i>se-puluh</i>	1000	<i>se-ribu</i>

Table 33: Numerals derived from Malay

Examples (253) and (254) show how these Malay-derived general numerals are used.

- (253) *Ada satu orang tua lah*
 EXIST **one** person old EMP
 ‘There is an old person.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:00:21.5-00:00:24.0)

- (254) *Ada tiga ekor beruang*
 EXIST **three** CLF.animal bear
 ‘There are three bears.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:00:25.2-00:00:27.3)

On the other hand, the Hokkien-derived numbers are only used for particular domains that are discussed using Hokkien terms. These include kinship and lunar calendar dates. The lunar system is used in particular for keeping track of Chinese festivities, which are traditionally important to the Peranakans. Note that there are fifteen days to the lunar month, thus Hokkien numbers used in BM do not usually exceed fifteen.

1	<i>it</i>	9	<i>gau</i>
2	<i>ji</i>	10	<i>chap</i>
3	<i>sa</i>	11	<i>chap-it</i>
4	<i>si</i>	12	<i>chap-ji</i>
5	<i>gor</i>	13	<i>chap-sa</i>
6	<i>lak</i>	14	<i>chap-si</i>
7	<i>chit</i>	15	<i>chap-gor</i>
8	<i>puay</i>		

Table 34: Numerals derived from Hokkien

Examples (255) and (256) show how these Hokkien-derived numeral are used.

- (255) *Ji pék.*
two father's.elder.brother
 'Second uncle (who is older than one's father).'
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-132, 00:07:50.6-00:07:52.0)

- (256) *Bila chuay it chap-go gua pi sohio.*
 when beginning.lunar.month **one fifteen** 1.SG go burn.incense
 'I will go burn incense on the first and the fifteenth of the lunar month.'
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-171, 00:09:53.9-00:09:56.8)

4.5.3.1 Other numerical expressions

Odd-numbered items are referred to as *ganjil*, while even-numbered items are referred to as being in pairs or *pasang*.

- (257) *Bila kita tangkap gambair, kita tak suka ganjil, mesti ber-pasang.*⁷²
 when 1.PL capture picture 1.PL NEG like **odd.number** must Poss-pair
 ‘when we take photographs, we do not like odd numbers, (we) must have pairs.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:12:02.8-00:12:09.8)

While *gangil* cannot be used to count, *pasang* can be used for counting items.

- (258) *Se-pasang gelang tangan*
 one-pair bracelet hand
 ‘One pair of hand bracelets.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:09:56.6-00:10:06.7)

- (259) *Se-pasang kasot manék*
 one-pair shoes bead
 ‘One pair of beaded shoes.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:08:24.0-00:08:31.3)

Pasang can also be used to refer to two items that may not form natural pairs such as shoes or bracelets.⁷³

- (260) *Se-pasang suluair*
 one-pair pants
 ‘one pair of pants (technically two pants).’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:08:54.3-00:08:56.4)

Other numerical expressions include *satu satu* (one by one), *dua dua* (two by two), and so on and so forth.

- (261) *Satu-satu, budak masok-kan*
 one-one child put.in-TR
 ‘One by one, children put in (the pears).’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:51.4-00:05:42.3)

4.5.4 Quantifiers

Quantifiers in BM include *sumua* ‘all’, *manyak* ‘many’, *berapa* ‘some’, *sikit* ‘little’. These can be used on their own as nouns, or be used to modify other nouns and adjectives (see section 5.3.6

⁷² Only instance where *ber-* appears to indicate some notion of possession.

⁷³ The Peranakan bride wears one on each hand.

for modification of adjectival phrases). The following examples demonstrate how they are used in BM. Example (262) shows how the quantifier can be used as a noun on its own. Examples (263), (264) and (265) show how these typically modify nouns, and example (266) demonstrates how quantifiers can modify adjectives too.

(262) *Sumua kata sedap.*
all say delicious
 All say (it's) delicious.
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:10:04.8-00:10:05.8)

(263) *Kita sumua ada chakap sama tu dua budak.*
 1.PL **all** PROG talk with that two child
 'All of us are talking to the two boys.'
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-071, 00:16:41.5-00:16:49.8)

(264) *Makan ikan banyak.*
 eat fish **many**
 'Eat a lot of fish.'
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:03:44.4-00:03:46.6)

(265) *Tu dua budak chakapsama kita berapa orang*
 that two child talk with 1.SG **some** people
 'Those two children talked to some of us.'
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-071, 00:18:06.8-00:18:09.3)

(266) *Asien-asin sikit.*
 somewhat.salty **little**
 'A little salty.'
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:11:18.0-00:11:20.8)

4.6 Prepositions

Prepositions here refer to a closed set of lexemes that occur before noun phrase complement. They form adverbial phrases with these noun phrase complements. Prepositions in BM include general prepositions *dekat* and *di*, *dalam* 'inside', *dari* 'from', *sampay* 'until' and *sama* 'with'. *Dekat* (short version *kat*) literally means 'near', as shown by its use in example (267). It has been grammaticalized and is used as a general preposition. It is used to denote the concepts of *at*, as

demonstrated by examples (268) and (269), *in*, as with example (270), *to*, as with example (271), and *on*, as with example (272).

general prepositions	<i>dekat</i>
at, in, to, on	<i>di</i>
inside	<i>dalam (dekat dalam, di dalam)</i>
front	<i>depan (dekat depan, di depan)</i>
beside	<i>se-belah (dekat se-belah, di se-belah)</i>
behind	<i>belakang (dekat belakang, di belakang)</i>
top	<i>atas (dekat atas, di atas)</i>
bottom	<i>bawah (dekat bawah, di bawah)</i>
from	<i>dari</i>
until	<i>sampay</i>
with	<i>sama</i>

Table 35: List of prepositions in Baba Malay

- (267) *Kupukupu trebang di dekat itu mia anjing.*
butterfly fly **PREP** near that REL dog
‘The butterfly flew near that dog.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:01:36.8-00:01:41.7)
- (268) *Dia tarok kat depan*
3.SG put **PREP** front
‘He put (it) in front.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:05.0-00:03:06.4)
- (269) *Dekat dia nia kebun*
PREP 3.SG POSS garden
‘At his garden.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:00:22.0-00:00:23.4)

- (270) *Se-kali dia angkat tulang itu dekat mulot dia.*
 one-time 3.SG lift bone that **PREP** mouth 3.SG
 ‘Once it lifted that bone in its mouth.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:46.5-00:02:48.6)
- (271) *Dia salin-kan dekat bakol*
 3.SG transfer-TR **PREP** basket
 ‘He transferred (the pears) to the basket.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:50.8-00:00:55.4)
- (272) *Satu couple ada tarok dia mia kain dekat rumput*
 one PROG put 3.SG POSS cloth **PREP** grass
 ‘One couple is putting their cloth on the grass.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:05:36.3-00:05:41.3)

Di is another general preposition whose use is derived from Malay where it is used to indicate ‘at’. It is used in sentences where *dekat* is already used to denote ‘near’. Example (267) is replicated here as example (273). *Di* is much less frequently used than *dekat*, and may be a later development in BM, as speakers begin to borrow more from Malay.

- (273) *Kupukupu terbang di dekat itu mia anjing.*
 butterfly fly **PREP** near that REL dog
 ‘The butterfly flew near that dog.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:01:36.8-00:01:41.7)
- (274) *Kerekot di dekat panjang*
 bent.and.curled.up **PREP** near bed
 ‘Bent and curled up on the bed.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:04:26.2-00:04:28.5)
- (275) *Datang di tepi sunggay*
 come **PREP** side river
 ‘Came to the side of the river.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:14.1-00:02:17.3)
- (276) *Tulang di mulot, buang di sunggay*
 bone **PREP** mouth throw **PREP** river
 ‘Bone in the mouth, thrown into the river.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:14.1-00:02:17.3)

Another preposition in BM is *dalam* ‘inside’. It can also be used together with the general preposition *dekat*, as demonstrated by examples (278) and (279).

- (277) *Dia-orang* *punggot* *itu* *buah* *pear* *masuk* ***dalam*** *bakol*
 3.PL pick.up that CLF.fruit enter **inside** basket
 ‘They picked up those pears (and) put (them) inside the basket.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:03:56.1-00:03:58.5)
- (278) *Dia* *tarok* *bunga* ***kat*** ***dalam*** *kepok*.
 3.SG put flower **PREP inside** box
 ‘He put flowers inside the box.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:04:07.9-00:04:09.9)
- (279) *Dia* *mia* *kodok* ***kat*** ***dalam*** *sana*
 3.SG POSS frog **PREP inside** there
 ‘His frog (is) inside there.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:09:50.3-00:09:52.8)

Depan is used to indicate ‘front’. Again, other prepositions may occur before *depan*, as with examples (282) and (283).

- (280) ***Depan*** *carpark* *ada* *ini* *kebun* *bunga*
front EXIST this garden flower
 ‘In front of the carpark there is this flower garden.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:07:32.8-00:07:38.7)
- (281) *Kebun* *bunga* ***depan*** *carpark*
 garden flower **front** carpark
 ‘The flower garden is in front of the carpark.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:07:51.2-00:07:54.6)
- (282) ***Di*** ***depan*** *carpark* *nanti* *jumpa* *satu* *kebun* *bunga*
PREP front carpark later meet one garden flower
 ‘In front of the carpark (you) will see one flower garden.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:08:12.0-00:08:17.2)
- (283) *Dia* *tarok* ***kat*** ***depan***
 3.SG put Prep front
 ‘He puts (it) in front.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:04.4-00:03:06.4)

Se-belah, literally ‘one-side’ is used to indicate ‘besides.’ It can be used with a preceding general preposition, as demonstrated by example (286).

- (284) *Se-belah park satu keday kopi*
one-side one shop coffee
 ‘Beside the park (is) one coffee shop.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-103, 00:11:30.9-00:11:34.1)
- (285) *Lu mesti seberang sempang se-belah supermarket.*
 2.SG must cross traffic.junction **one-side**
 ‘You must cross the junction beside the supermarket.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-103, 00:14:22.6-00:14:27.1)
- (286) *Di se-belah, dia ketok apa?*
PREP one-side 3.SG knock what
 ‘Beside, they are knocking what?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:14:31.8-00:14:33.5)

Belakang is used to denote ‘behind’, as demonstrated by examples (287) to (289).

- (287) *Tempat buang ayé belakang restaurant sair*
 place throw water **behind** CONF
 ‘The toilet is behind the restaurant indeed.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-103, 00:13:29.4-00:13:33.5)
- (288) *Belakang restaurant tu ada chiwan*
behind that EXIST toilet
 ‘Behind that restaurant there is a toilet.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:05:29.4-00:05:34.2)
- (289) *Anjing utan kejair belakang dia*
 dog jungle chase **behind** 3.SG
 ‘The wolf chased behind him.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:04:22.3-00:04:23.9)

Atas is used to indicate ‘top’, as shown in examples (290) to (291).

- (290) *Dia naik tangga atas pokok.*
 3.SG ascend ladder **top** tree
 ‘He climbs a ladder to the top of the tree.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:08.6-00:01:12.1)
- (291) *Rumah-rumah atas bukit*
 house-PL **top** mountain
 ‘The houses are on top of the mountain.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:58.2-00:04:00.6)

- (292) *Kodok dekat atas itu yacht senang*
 frog PREP top that yacht relax
 ‘The frog is on top of that yacht relaxing.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:53.5-00:00:56.4)

Bawah is used to denote ‘bottom’.

- (293) *Dia dudok bawah pokok.*
 3.SG sit bottom tree
 ‘He sat at the bottom of the tree.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:34.9-00:01:37.1)

The notion ‘from’ is expressed by the preposition *dari*.

- (294) *Si sa ko dari mana*
 PERSON three paternal.aunt from where
 ‘Third paternal aunt from where?’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:08:49.3-00:08:53.2)
- (295) *Dia jatoh dari ranjang*
 3.SG fall from bed
 ‘She fell from the bed.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:08:29.1-00:08:31.3)
- (296) *Dari tempat letak keréta, lu mesti seberang satu sempang*
 from place park car 2.SG must cross one traffic.junction
 ‘From the car park, you must cross one junction.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-103, 00:11:02.0-00:11:05.1)

Sampay undergoes grammaticalization. Its original meaning that is also retained in BM is ‘reach’ while the preposition it forms is ‘until’, a related notion. *Sampay* can be used to indicate ‘until’ with regard to place, as with example (298), time, as with example (299), and state, as with example (300). Example (297) shows how *sampay* is used to mean ‘reach’.

- (297) *Apa macham gua boléh sampay sana?*
 what like 1.SG can reach there
 ‘How can I reach there?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-103, 00:18:27.4-00:18:31.7)
- (298) *Gua sudah jalan sampay sini*
 1.SG already walk until here
 ‘I already walked until here.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:16:04.9-00:16:07.3)

- (299) *Dudok sampay pukol lima*
 sit until strike five
 ‘Sit until five o’clock.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:03:03.3-00:03:05.6)
- (300) *Dia dudok bicycle chepat pulak sampay dia nia buah pear jatoh*
 3.SG sit fast instead until 3.SG POSS CLF.fruit fall
 ‘He rode the bicycle fast instead, until his pears fell.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:03:11.6-00:03:13.0)

Besides being used literally to mean ‘same’, and as a conjunction (see section 4.7), *sama* is also used as a preposition, denoting *with*. General use of *sama* is demonstrated in examples (301) to (304).

- (301) *Orang itu pukol kucing sama kayu.*
 person that hit cat with stick
 ‘That person hit the cat with the stick.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:40:58.1-00:41:00.2)
- (302) *Ini kawan, dia datang sama apa?*
 this friend 3.SG come with what
 ‘This friend, he came with what?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:34.9-00:01:37.1)
- (303) *Pi masak sama dia*
 go cook with 3.SG
 ‘Go cook with it (fish).’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:12:43.3-00:12:45.2)
- (304) *Dia chakapsama dia*
 3.SG speak with 3.SG
 ‘He is talking to him.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:44.4-00:02:46.0)

Sama is also used in comparatives, as demonstrated by example (305) (refer to section 5.3.1 for more details and examples).

- (305) *Dia tinggi sama ngko dia.*
 3.SG tall with older.brother 3.SG
 ‘He is as tall as his brother.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:10:31.5-00:10:33.6)

Additionally, *sama* is used to connect verb with its object noun phrase, as compared to the obliques in examples (302) to (305). This is demonstrated in examples (307) to (309). Example (306) shows that this is optional for transitive verbs, when compared to example (307). Inserting *sama* is also a way of making an intransitive verb transitive, as example (309) demonstrates.

- (306) *Dia sumua halo tu mia anjing*
 3.SG all chase that REL dog
 ‘They all chased that dog.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:04:32.1-00:04:34.1)
- (307) *Halo sama budak ini*
 chase.away with boy this
 ‘Chase away this boy.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:06:02.5-00:06:04.6)
- (308) *Rindu sama gua*
 miss with 1.SG
 ‘Miss me.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:29:51.9-00:29:54.3)
- (309) *Senyum sama gua*
 smile with 1.SG
 ‘Smile at me.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:18:07.3-00:18:08.7)

4.7 Conjunctions

The following table features conjunctions that are commonly used in BM.

and	<i>sama</i>
finish	<i>habi/abi</i>
or	<i>ka</i>
but	<i>tapi</i>
because	<i>pasal/ pasair</i>
if	<i>kalu</i>
when	<i>bila</i>
although (literally ‘really also’)	<i>sunggu pun</i>

Table 36: List of conjunctions in Baba Malay

The following example sentences show how they are used. Examples (310) to (312) demonstrate the using of coordinating conjunctions. These are used to conjoin similar phrases, or clauses at the same level. In example (313), the subject noun is unexpressed in the second clause (see section 5.6.2), and both components being conjoined are clauses. Constituents involved in conjunction are presented in parentheses.

- (310) [*Ini kupukupu*] **sama** [*anjing ni*].
 this butterfly **and** dog this
 ‘This butterfly and this dog’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:01:48.0-00:01:50.4)
- (311) [*Mak pegi pasair*], **habis** *tu* [*dia pegi kopitiam*].
 mother go market **finish** that 3.SG go coffee.shop
 ‘Mother went to the market. After that she went to the coffeeshop.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-110, 00:01:23.4-00:01:29.3)
- (312) *Gua rasa ini* [*anak dia*] **ka** [*chuchu dia*].
 1.SG think this child 3.SG **or** grandchild 3.SG
 ‘I think this is his child or grandchild.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:13.0-00:02:15.0)
- (313) [*Gua mia adék bikin kék*] **tapi** [*tak sedap*].
 1.SG POSS sibling make cake **but** NEG delicious
 ‘My sister baked a cake but it was not delicious.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-117, 00:14:38.9-00:14:44.6)

Examples (314) to (316) show the use of subordinating conjunctions. These ones link subordinate clauses to main clauses. The subordinate clauses in the following examples are presented in parentheses. Example (315) is specifically a conditional expression.

- (314) *Tiga minggu macham, dia balék [pasair dia tinggal Singapore]*
 three week like.that 3.SG return **because** 3.SG live Singapore
 ‘For about three weeks, she returns because she lives in Singapore.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:04:03.8-00:04:09.2)
- (315) [*Kalo gua tau lu mo datang*] *gua tentu jumpa lu dekat airport*
if 1.SG know 2.SG want come 1.SG definite meet 2.SG PREP
 ‘If I knew you were coming, I am definite (I would) meet you at the airport.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:13:32.1-00:13:37.9)
- (316) [*Bila bahru gua balék rumah*], *bahru gua dapat tau pasair tu accident*
when just 1.SG return home just 1.SG get know matter that
 ‘When I just returned home, I just got to know (about) that accident.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:02:22.8-00:02:26.3)
- (317) [*Sunggu.pun dia tak standard*], *dia dapat ini kreja.*
although 3.SG Neg 3.SG get this work
 ‘Although he (is of) no standard, he got this work.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:20:53.4-00:20:57.2)

More details on the different types of conjunctions and their syntactic patterns can be found in section 5.6.8. Conditionals are discussed in section 5.6.7.

4.8 Discourse elements

BM discourse is characterized by its heavy use of interjections and particles.

4.8.1 Interjections

Interjections belong to a non-productive word class, whose main function is emotive (Crystal 1997). In this grammar, it refers to both single words and short two word utterances that can be used on their own or right at the beginning of sentences. Interjections are characteristic of daily conversations in BM, with participants engaging in *mincharok*, or using curse words. While there is usually little illocutionary force in *mincharok*, and hence no intent on the speaker’s part to curse her or his interlocutor or to wish bad things upon them, it is considered *kasar* ‘coarse’ to

engage in *mincharok*. It is also less appropriate for younger speakers to use *mincharok* with older speakers, although it is normal for older speakers to use it with younger speakers regardless of occasion. Note that not all interjections are forms of *mincharok* ‘curse’. A list of common interjections is given in the following table, some of which are *mincharok*.

<i>aiyo</i>	exclamation of irritation
<i>ala</i>	exclamation of regret
<i>ayi</i>	exclamation of surprise
<i>alamak</i>	exclamation of dismay
<i>éh</i>	exclamation in a jibing manner
<i>amboey</i>	exclamation of surprise
<i>adoey</i>	exclamation of pain
<i>mati</i>	die
<i>mampus</i>	dead
<i>chilaka</i>	cursed one
<i>kus semangat</i>	cry to a dead spirit

Table 37: List of common interjections in Baba Malay

The following are some examples of how these interjections are used. While some cannot be directly translated into English, close translations are provided.

- (318) *éh, ho mia lah lu.*
EXCLAM.jibe good life EMP 2.SG
 ‘Hey, you have a good life.’ (said in a jibing manner)
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:26:20.3-00:26:22.9)
- (319) *Ayi mampus kambing dia kuat sekali.*
EXCLAM.surprise **dead** goat 3.SG strong very
 ‘Surprising, goodness, the goat it (is) very strong.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:52.8-00:01:54.6)

- (320) *Amboey.* *dia* *kasi-kan* *kopiah.*
EXCLAM.surprise 3.SG give-TR hat.
 ‘Surprising. He gave (him) the hat.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:05:06.6-00:05:10.2)

- (321) *Alamak,* *chilaka,* *kepala gua* *sakit.*
EXCLAM.dismay **cursed.one** head 1.SG sick
 ‘Goodness, cursed one, my head hurts.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:05:45.8-00:05:47.8)

4.8.2 Particles

While interjections are used on their own, particles are used at the end of sentences. These discourse particles do not change the meaning of the utterances they are attached to. The two most commonly used particles are *lah*, and *sair*.

Lah, usually accompanied with a falling pitch is used emphatically. Emphatic *lah* can be attributed to Hokkien influence, just as *lah* in Colloquial Singapore English (a creole with a Hokkien substrate) is said to have been derived from Hokkien (Platt and Ho 1989). In the following examples, the insertion of *lah* at the end of the utterances emphasizes whatever had been said in the utterance. Semantically, the insertion of this pragmatic particle does not change the meanings in these utterances. The usage of *lah* is shown in examples (322) to (327).

- (322) *Abi* *dia* *naik* *geram* **lah**
 finish 3.SG ascend furious **EMP**
 ‘After that he became furious.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:00.2-00:03:02.0)

- (323) *Budak ni* *ada* *jaga* *dia* *mia* *kambing* **lah**
 child this PROG guard 3.SG POSS sheep **EMP**
 ‘This child is guarding his sheep.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:04:00.6-00:04:03.4)

- (324) *Dia-orang* *sumua* *marah* *dia* **lah.**
 3.PL all angry 3.SG **EMP**
 ‘They all (were) angry (at) him.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:04:37.9-00:04:40.0)

- (325) *Chakiak orang putéh lah*
 clogs person white EMP
 ‘Western-style clogs.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:07:15.6-00:07:17.5)
- (326) *Tak mo lah*
 NEG want EMP
 ‘(I) don’t want to.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:06:17.9-00:06:19.1)
- (327) *Sini masak nia sedap lah*
 here cook REL delicious EMP
 ‘(The food) that is cooked here (is) delicious.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:06:45.8-00:06:48.1)

The other particle that is commonly used is *sair*, which has been analysed as being derived from *sekali* ‘very’ (Gwee 1998, Gwee 2006). This particle is usually accompanied with a rising pitch. Gwee translates *sair* as ‘indeed’, and in line with this, *sair* appears to have the effect of confirming the utterance itself, that what is being said is indeed the case. It is thus glossed as a confirmative particle in this grammar. As a discourse particle, *sair* does not significantly affect the meanings of the utterances to which it is attached. The usage of *sair* is demonstrated by examples (328) to (333). Note that the *sair* particle is not the same lexical item as *sair* in *apa sair*, which is the short for the refined form *apa pasair* ‘what reason’ (coarse form: *apa pasal*). Example (218) is replicated below as example (329).

- (328) *Panas sair*
 hot CONF
 ‘hot indeed.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:02:25.1-00:02:26.3)
- (329) *Taukay kebun ni kiam sair*
 boss garden this miserly CONF
 ‘This boss (of the) garden is miserly indeed.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:01:03.7-00:01:05.8)
- (330) *Dia chakap Peranakan lanchang sair*
 3.SG speak Peranakan fluent CONF
 ‘He speaks Peranakan fluently indeed.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:19:18.8-00:19:22.4)

- (331) *Tapi bukan anak sair*
but NEG child CONF
'but (it is) not the children indeed.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:09:54.0-00:09:56.2)
- (332) *Jangan bising sair*
do.not noisy CONF
'do not (be) noisy indeed.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:15:21.3-00:15:23.9)
- (333) *Suay sair*
unlucky CONF
'Unlucky indeed.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:15:55.6-00:15:56.5)

5. Syntax

Being more of an isolating language, BM uses word order predominantly in expressing syntactic relations. This chapter concerns the syntax of noun phrases, verb phrases, adjectival phrases, adverbial phrases, as well as more complex clauses. From the examples in this chapter, it will become clear that BM is neither predominantly head-initial nor head-final.

5.1 Noun phrases

A noun phrase (NP) comprises pronoun or a noun phrase and optional modifiers that both precede and follow the noun head. Modifiers include demonstratives, person marker, numerals, quantifiers, and adjectives. Modifiers may also come in the forms of genitive clauses as well as relative clauses. The following are examples of these.

(334) Pronoun:

Dia tolak.

3.SG push

‘He pushed.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:04:27.7-00:04:28.7)

(335) Noun without modifier:

Anjing sudah bangun

dog alreadywake.up

‘The dog has already woken up.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:17.1-00:00:18.7)

(336) Noun with preceding demonstrative:

Dia-orang tolong itu budak

3.PL help **that** **child**

‘They helped that child.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:04:09.1-00:04:11.1)

(337) Noun with following demonstrative:

Budak ini sangat kechik.

child this very small

‘This (is) very small.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:23.5-00:03:25.3)

- (338) Noun with preceding person marker

Si tua

PERSON old

‘That old person.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:04:10.4-00:04:10.8)

- (339) Noun with preceding numeral:

Satu bakol sudah ilang.

one basket alreadylost

‘One basket has gone missing.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:47.1-00:05:48.9)

- (340) Noun with preceding demonstrative, numeral and classifier:

Ayi dia jumpa tiga ékor budak

EXCLAM.surprise 3.SG meet **three CLF.animal child**

‘Surprising he met three children.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:52.6-00:05:54.5)

- (341) Noun with preceding quantifier

Sumua orang sekarang tak tau chakap

all person now NEG know speak

‘Everyone now does not know (how to) speak.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:07:37.8-00:07:40.9)

- (342) Noun with following quantifier

Angkat changkay changkay sumua pegi belakang.

carry **cup cup** all go behind

‘Carry all the cups (and) go to the back.’

(Jane Quek oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:06:47.4-00:06:50.1)

- (343) Noun with preceding adjective

Cherita-kan gua betul mia cherita.

story-TR 1.SG **true REL story**

‘Tell me a true story.’

(Jane Quek oai:scholarspace.manoa.hawaii.edu: NL1-100, 00:02:56.4-00:03:00.5)

- (344) Noun with following adjective

Babi kechik chakap.

pig small speak

‘The small pig spoke.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:03:44.7-00:03:45.9)

relationship, is by using the possessed-possessor word order. The following demonstrates both structures.

- (349) *William nia bapak*
 POSS father
‘William’s father.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:45:06.9-00:45:09.7)
- (350) *Asam gugol mia kulit*
tamarind dried.fruit **POSS** skin
‘The dried tamarind’s skin’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:11:51.6-00:11:54.0)
- (351) *Kita mia orang*
1.PL **POSS** people
‘Our people’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:05:49.3-00:05:50.4)
- (352) *Dia mia nama*
3.SG **POSS** name
‘Its name’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:00:37.9-00:00:41.4)
- (353) *Dia mia kawan*
3.SG **POSS** friend
‘His friend’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:05:58.2-00:05:59.0)

While expressions using *punya* as a possessive marker can be used to mark genitive relations between a proper noun, common noun or a pronoun and the possessed noun, only pronouns can function as possessor in the possessed-possessor structures. There are no instances of non-pronouns functioning as the possessor in utterances of the following sort. Note that third person pronoun *dia* occasionally occur as *nia*.

- (354) *Badan lu*
body 2.SG
‘Your body.’
(Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:11:33.1-00:11:36.1)
- (355) *Kambing dia*
sheep 3.SG
‘His sheep’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:05:47.9-00:05:48.0)

- (356) *Bini dia*
 wife 3.SG
 ‘His wife.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:01:32.0-00:01:33.3)
- (357) *Mata dia*
 eye 3.SG
 ‘His eyes’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:00:23.9-00:00:24.5)
- (358) *Rumah gua*
 house 1.SG
 ‘my house.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:02:41.9-00:02:43.0)

Considering that *punya* constructions have more functionality than possessed-possessor constructions (which can only express genitive relations where pronouns are concerned), BM’s more dominant word order appears to be GEN N, rather than N GEN.

5.1.2 Noun phrases with determiners

Determiners in this grammar refers both demonstratives, person marker, numerals, classifiers and quantifiers, these items co-occurring with nouns to express semantic contrasts, for example, distance and quantity. These determiners have different distributions. In general, demonstratives as well as quantifiers can precede or follow nouns, while numerals and classifiers precede nouns.

In addition, it is important to note that there are no determiners that express a definite-indefinite contrast. Context is largely used to determine if a noun is definite or indefinite. Example (108) is replicated here as (359), and example (335) as (360).

- (359) First mention of a noun
Orang panjat pokok.
 person climb tree
 ‘A person climbs a tree.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:00:15.6-00:00:17.6)

(360) Subsequent mention of a noun

Anjing sudah mangun
dog alreadywake.up
'The dog has already woken up.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:17.1-00:00:18.7)

Both *orang* 'person' and *anjing* 'dog' in the above examples do not co-occur with any determiners, yet *orang* is interpreted to be indefinite, and *anjing* definite. This is due to the contexts in which they occur. Where (359) is specifically concerned, *orang* is the first utterance in a narrative and there is no precedent. The word is hence interpreted to be indefinite. In the case of (360), *anjing* has been mentioned prior to this occurrence, and it is therefore understood by the listener as being definite.

5.1.2.1 Noun phrases with demonstratives

While there is no definite-indefinite contrast in BM, definite determiners can be contrasted based mainly on distance or space. The demonstratives *ini* 'this' and *itu* 'that' are used based on deictic notions. *Ini* is proximal, co-occurring with nouns that are typically close to the speaker or the action taking place, while *itu* is distal, and it co-occurs with nouns that are typically further away from the speaker or the action taking place. There is no preferred word order for noun phrases comprising demonstrative and noun. Note that the shortened versions of *ini* and *itu* are *ni* and *tu* respectively.

(361) ***Ini*** *kebun*
this garden
'this garden.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:32.0-00:00:32.8)

(362) ***Ini*** *rumah-rumah*
this house-PL
'these houses.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:52.6-00:03:53.2)

- (363) *Ni orang*
this person
 ‘This person.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:29.0-00:00:30.0)
- (364) *Tu budak perempuan*
that child female
 ‘That girl.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:03:28.0-00:03:29.2)
- (365) *Itu kopiah*
that hat
 ‘That hat.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:04:37.0-00:04:38.5)

While the above examples show noun phrases where demonstratives precede nouns, the following are examples of nouns preceding demonstratives. Example (366) replicates(130).

- (366) *Budak ini*
 child **this**
 ‘this child.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:23.5-00:03:24.0)
- (367) *Anjing hutan ini*
 dog jungle **this**
 ‘this wolf’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:04:24.7-00:04:26.6)
- (368) *Anjing tu*
 Dog **that**
 ‘that dog’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:01:41.7-00:01:42.3)
- (369) *Kuching belanda tu*
 cat Holland **that**
 ‘That rabbit.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:01:18.0-00:01:19.4)
- (370) *Si tua tu*
 PERSON old **that**
 ‘That old person.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:04:10.4-00:04:11.0)

In addition, *demonstratives* may occur together with *punya*, as demonstrated in the following examples. In these instances, *punya*'s function appears to be similar to when it is a relative clause marker (see section 5.6.3), a relative clause being a subordinate clause that modifies the head noun. The subordinate clauses contain deictic information in these instances. For these constructions, the demonstrative always occurs before noun.

(371) *Ini mia budak*
this REL child
 'This child.'
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:49.0-00:03:50.0)

(372) *Tu mia anjing*
that REL dog
 'That dog.'
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:32.3-00:04:34.1)

Although demonstrative precedes noun in the *punya* construction, this is not the most basic type of noun phrase that features a demonstrative. Both DEM-NP and NP-DEM orders are equally common in BM and there is no preferred word order.

5.1.2.2 Noun phrases with person marker

Noun phrases featuring the person marker *si* always have the order *si*-NP. Example (240) is replicated as (374).

(373) *Si Mary*
PERSON
 'Mary.'
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:01:07.1-00:01:08.2)

(374) *Si sa ko bongsu.*
PERSON third paternal.aunt youngest.child
 'Youngest third paternal aunt.'
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:09:01.2-00:09:02.9)

(375) *Si tua tu*
PERSON old that
 'That old person.'
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:04:10.4-00:04:11.0)

- (376) *Si ano ni?*
PERSON anonymous this
 ‘This anonymous person?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:12:13.5-00:12:15.2)

5.1.2.3 Noun phrases with numerals, noun classifiers and partitives

Similar to noun phrases comprising person marker, noun phrases are head-final when they comprise numerals or numerals and classifiers. The following are examples of noun phrases comprising numerals, and numerals with classifiers. It is interesting to note that the classifier used for animals can also be used on young children. The relevant examples are (382) and (386).

- (377) *Satu anjing*
one dog
 ‘One dog.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:57.0-00:02:57.9)

- (378) *Satu tulang*
one bone
 ‘One bone.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:55.0-00:02:55.8)

- (379) *Empat ratus taon*
four hundred year
 ‘Four hundred years.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:13:45.4-00:13:47.0)

- (380) *Dua bulan*
two month
 ‘Two months.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:03:33.6-00:03:35.0)

- (381) *Ni empat minggu*
 this **four** week
 ‘These four weeks.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:16:55.0-00:16:56.5)

- (382) *Se-kor kambing*
one-CLF.animal goat
 ‘One goat.’
 (Victor oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:32.4-00:01:34.7)

- (383) *Dua bijik pear*
two CLF.small.round
 ‘Two pears.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:04.5-00:05:05.4)
- (384) *Satu bijik timun*
one CLF.small.round cucumber
 ‘One cucumber.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-149, 00:03:17.8-00:03:19.4)
- (385) *Se-kuntum bunga*
one-CLF.bloom flower
 ‘One flower’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-160, 00:01:59.2-00:02:00.7)
- (386) *Ni tiga ékor budak*
 this **three CLF.animal** children
 ‘These three children.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:53.5-00:05:54.5)

Partitive phrases are also considered here. These are semantically different than noun phrases featuring noun classifiers, as they are used to refer to a part or a quantity of a mass noun. However, they are syntactically alike noun classifiers with modifier preceding head.

- (387) *Satu botol susu*
one bottle milk
 ‘One bottle of milk.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:21:05.5-00:21:08.5)
- (388) *Satu changkay ayé*
one cup water
 ‘One cup of water.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:21:41.5-00:21:42.6)
- (389) *Satu kepéng ayam*
one piece chicken
 ‘One piece of chicken.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:21:16.8-00:21:17.9)
- (390) *Satu éla mia kain*
one sheet REL cloth
 ‘One sheet of cloth.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:23:27.5-00:23:28.5)

- (391) *Satu baldi mia ayé*
one pail REL water
 ‘One pail of water.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-148, 00:21:42.7-00:21:44.7)

Hence, noun phrases can comprise Num-NP, Num-Clf-NP, and Num-Part-NP sequences. All of these sequences alike in that they are head-final.

5.1.2.4 Noun phrases with quantifiers

Noun phrases may also have quantifiers as modifiers. In these instances, the modifier may precede or follow the head noun.

- (392) *Sumua adék- beradék*
all sibling-PL
 ‘All siblings.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:07:37.9-00:07:38.5)
- (393) *Orang orang sumua*
 people people **all**
 ‘All people.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:01:23.6-00:01:24.6)
- (394) *Manyak orang*
many people
 ‘Many people.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:24.0-00:00:25.5)
- (395) *Dagin putéh manyak*
 meat white **many**
 ‘A lot of white meat.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:03:52.5-00:03:53.6)
- (396) *Berapa orang*
some people
 ‘Some people.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-071, 00:18:08.3-00:18:09.3)
- (397) *Umor sikit*
 age **little**
 ‘Little age.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:06:23.0-00:06:24.1)

Other than noun phrases that are made up of nouns and quantifiers in either phrase-initial or phrase-final positions (Quan-NP, NP-Quan), it is also possible for quantifiers to take the position of the noun itself.

(398) *Manyak tak sama*
many NEG same
'Many are not the same.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:02:13.0-00:02:14.1)

(399) *Sumuabising*
all noisy
'All are noisy.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:01:33.9-00:01:34.5)

(400) *Dia minum sikit*
3.SG drink **little**
'He drinks a little.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:22:41.0-00:44:42.7)

The only quantifier that has not been observed to replace the noun itself is *berapa* 'some'. Note that *berapa* is also an interrogative meaning 'how many'. Only context can help differentiate if it is meant as a quantifier or as an interrogative.

5.1.3 Noun phrases with adjectival modifiers

Adjectival modifiers in BM can follow or precede the head in a noun phrase. The more basic noun phrase structure comprises of noun followed by adjective.

(401) *Dagin mérah*
meat **red**
'Red meat.'
(Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:03:50.7-00:03:51.7)

(402) *Kepok kosong*
box **empty**
'Empty box.'
(Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:12:56.5-00:12:57.7)

- (403) *Beruang* ***besair***
 bear **big**
 ‘Big bear.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:04:24.7-00:04:27.2)
- (404) *Kayu* ***panjang***
 wood **long**
 ‘Long wood.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:01:20.3-00:01:22.4)
- (405) *Pintu* ***chanték***
 door **beautiful**
 ‘Beautiful door.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:01:55.7-00:01:56.8)

Adjectives may also precede nouns in an adjectival noun phrase when relative marker *punya* is used, as with examples (406) to (410). In these, the adjectives and *punya* function as a subordinate clause that gives information about the head noun, so that an example such as (406) *betol mia cherita*, may also be interpreted as ‘a story that is real’. Example (16) is replicated here as (406).

- (406) ***betol*** ***mia*** *cherita*
 real REL story
 ‘Real story.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:03:24.8-00:03:26.0)
- (407) ***sekarang*** ***mia*** *orang*
 now REL people
 ‘Modern people’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:06:52.9-00:06:54.2)
- (408) ***dulu*** ***mia*** *lauk*
 long.ago REL cook.food
 ‘Old-fashioned dishes.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:09:47.6-00:09:48.5)
- (409) ***Bising*** ***mia*** *pasair*
 noisy REL matter
 ‘Noisy matters.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:01:34.1-00:01:35.5)

- (410) ***Betol mia nasik lemak***
real REL rice cooked.in.coconut.milk
 ‘Real rice cooked in coconut milk (a dish).’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:08:10.0-00:08:11.5)

In effect, while both NP-A and A-NP orders are possible in BM, the basic form of the adjectival noun phrase is NP-A, since it does not require the use of an additional relative clause marker, even though the A-*punya*-NP structure is also commonly used.

5.1.4 Negation of noun phrases

Noun phrases can be negated in the following way with *bukan* preceding the noun phrase. This is akin to the use of ‘not’ in English, rather than ‘no’ in these instances.

- (411) ***Bukan anak***
NEG children
 ‘Not children.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:09:54.5-00:09:56.0)

- (412) ***Bukan bawang***
NEG onion
 ‘Not onions.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:10:59.7-00:11:00.0)

- (413) ***Bukan saya***
NEG 1.SG
 ‘Not me.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-018, 00:03:56.7-00:03:57.3)

- (414) ***Bukan Singapore***
NEG
 ‘Not Singapore.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:19.0-00:02:22.0)

- (415) ***Bukan [chakiak kita china mia]***
NEG clogs 1.PL Chinese POSS
 ‘Not our Chinese(-style) clogs.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:07:17.5-00:17:19.3)

Note that the noun phrase is not negated when expressing ‘no’ semantically. Instead, a verb phrase containing the existential marker, *ada*, is negated with a separate negative marker *tak*.

Tak-a is the shortened version of *tak ada*. (Refer to section 5.2.8 for more details and examples of the *tak* verbal phrase negation strategy.)

(416) *Tak-a pokok*

NEG-EXIST tree

‘There are no trees.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:36.0-00:00:36.9)

In general, where noun phrases are concerned, *bukan* is used in a *bukan*-NP sequence. Note that *bukan* can also be used to negate statements (see section 5.6.10.2).

5.1.5 Order of elements in noun phrases

It is difficult to state if the noun phrase is strictly head-final or head-initial. Demonstratives as well as quantifiers may precede or follow main noun phrase. Adjectival modifiers are observed to precede and follow nouns, although the default order appears to be noun followed by adjective, since a relative marker, *punya*, is required to link adjective to noun when the adjective precedes the noun (see section 5.1.3). The same goes for genitive constructions. However, there is also support for the preference of a head-final noun phrase, considering that numerals, noun classifiers and partitives, the person marker *si*, as well as the negative marker strictly precede the main noun phrase. Overall, it is tenuous to state that there is a general preference for modifier to precede head, or head to precede modifier. More noun phrase word order is discussed in the section on relative clauses (see section 5.6.3).

5.2 Verb phrases

A verb phrase (VP) comprises verb, verb and complement(s), or sequences of more than one verb. Auxiliaries may precede the main verb phrase when they are used to express passivization, modality, or tense and aspect. Verb phrases may also be made up of more than one main verb in the case of serial verb construction. Modifiers include negation markers and other adverbs. The following are examples of these.

(417) Intransitive verb

Berenang.

swim

‘Swim.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:16:20.1-00:16:22.1)

(418) Transitive verb with direct object complement

Bukak itu pintu

open that door

‘Open that door.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:20:35.9-00:30:39.1)

(419) Transitive verb with direct and indirect object complements

Dia kasi dia dua bijak pear

3.SG give 3.SG two CLF.small.round

‘He gave him two pears.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:04.0-00:05:05.4)

(420) Causative verb with clause complement

Yauguai, kasi gua terperanjat sekali.

demon cause 1.SG be.shocked very

‘Demon, you let me be very shocked.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:040:26.2-00:40:28.2)

(421) Auxiliary verb and verb

Sumua boléh chakap Peranakan.

all can speak

‘All can speak Peranakan.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:06:12.2-00:16:14.9)

(422) Adverb and verb

lekair senyum

quick smile

‘Quickly smile.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:06:17.5-00:06:18.6)

(423) Verb and adverb

Lu mesti mo chobak dulu

2.SG must want taste first

‘You must taste first.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:21:46.0-00:21:48.3)

(424) Serial verb construction

Dudok téngok

sit look

‘Sitting and looking.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:50.0-00:00:51.0)

(425) Negation marker and verb

Tak tau

Neg know

‘(do) not know.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:08:16.4-00:08:17.2)

The examples show that verb phrases are head-initial and head-final. The following sections provide more details.

5.2.1 Copula constructions

Copulas are technically verbs that link the subject and a complement. In BM, there are two types of copula constructions. The copula verb *ada* can exist. It can be used as a regular copula, and it may be used in tag interrogatives, as with examples (426) and (427) (see section 5.6.10.2). However, in most instances, it is omitted.

(426) *Ada baik tak-a*

COP good NEG-COP

‘Are (you) well or not?’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:09.2-00:00:11.2)

(427) *Rumah gua ada chanték tak*

house 1.SG **COP** beautiful NEG

‘My house is beautiful or not?’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:02:41.9-00:02:43.7)

The following examples show typical copula constructions where the copula verb is omitted. In the instances below, these copula constructions, by virtue of their word order (subject before complement), associate subjects with complements ranging from adjective phrases to preposition phrase. Example (291) has been replicated below as (432).

- (428) *Gua lapair*
 1.SG hungry
 ‘I (am) hungry.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:14:16.4-00:14:17.5)
- (429) *Itu barang mentah*
 that thing unripe
 ‘That thing (is) unripe.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:22:15.6-00:22:17.1)
- (430) *Kemantin chanték*
 bride beautiful
 ‘The bride (is) beautiful.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:24:57.7-00:24:58.8)
- (431) [*Dia mia barang*]mahal
 3.SG POSS thing expensive
 ‘Its things (are) expensive.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:32.9-00:02:33.8)
- (432) *Rumah rumah atas bukit*
 house house top mountain
 ‘The houses (are) on top of the mountain.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:58.2-00:04:00.6)

5.2.2 Modality

Verb phrases may also comprise an auxiliary verb preceding the predicate. There are several functions of auxiliaries, including modality, passivization, and aspect. This section is concerned with the structures that expresses modality, modality being associated with the semantic expression of beliefs, attitudes, obligations and ability. While there are differing opinions as to what deontic modality means (Traugott 1989), it usually concerns will, permission and obligation (Lyons 1977, Traugott 1989). The epistemic modality on the other hand expresses one’s belief state or attitude towards a certain proposition (Traugott 1989), while dynamic modality concerns one’s capacity to do something (Nuyts 2006). As with other languages such as English, the same auxiliary may have overlapping functions.

The following sentences show the auxiliary verb *boléh* ‘can’ to express all three modalities. In examples (433) and (434), the speakers are discussing the abilities of the agents to perform a particular task. Hence the dynamic modality is expressed by placing *boléh* before the main verb phrase. Example (421) is replicated here as example (433)

(433) *Sumua boléh* [*chakap Peranakan*].
 all can speak
 ‘All can speak Peranakan.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:06:12.2-00:16:14.9)

(434) *Kita boléh* [*masak*]
 1.PL can cook
 ‘We can cook.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:10:21.0-00:10:22.0)

In the next example, epistemic modality is expressed. The speaker is saying that something should be eaten with rice, and thus expressing a belief about a particular subject.

(435) *Ini sumua boléh* [*makan sama nasik*].
 this all can eat with rice
 ‘All this (you) can eat with rice.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:08:46.5-00:08:48.1)

Boléh can also be used to express deontic modality. In examples (436) and (437), the speakers are not questioning the ability of the agents to perform particular tasks, but asking them if they are willing to perform them.

(436) *Lu boléh* [*chakap Peranakan tak?*]
 2.SG can speak NEG
 ‘Can you speak Peranakan?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:51:13.1-00:51:17.7)

(437) *Mak, boléh* [*masakini tak?*]
 mother can cook this NEG
 ‘Mother, can you cook this?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:09:28.4-00:09:29.8)

Similarly, *mesti* has two uses. *Mesti* can be used to express the epistemic modality. In the following examples, the speakers express certain beliefs that they have about the world and what they have observed.

(438) *Kuching itu mesti makan ikan*
 cat that **must** eat fish
 ‘That cat must eat fish.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:43:41.4-00:43:43.6)

(439) *Dia mesti ter- langgair*
 3.SG **must** ACC- crash
 ‘He must have crashed.’
 (Peter, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:32.5-00:03:34.1)

Mesti can also be used to express the deontic modality. In the following two examples, the speaker talks about his and his group’s obligation to carry out particular tasks.

(440) *Kita orang mia Peranakan, kita mesti [belajair].*
 1.PL people POSS 1.PL **must** learn
 ‘Our people’s Peranakan, we must learn.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:51:13.1-00:51:17.7)

(441) *Gua mesti [panggay engko]*
 1.SG **must** call older.brother
 ‘I must call (a male acquaintance) older brother.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:05:25.7-00:05:27.3)

In BM, the notion of *mesti*, whether as an epistemic or deontic auxiliary, can be reinforced and emphasized by the form *mesti mo*, literally translated as ‘must want’. This appears to have been transferred directly from Hokkien *beh ai* (see Pakir 1986). In examples (442) and (443), the speaker is saying that she believes this is what has to happen when a person grows old, thus expressing beliefs. In examples (444) and (445), the speakers are expressing that the interlocutors have obligations to carry out particular tasks. While all speakers translate *mesti mo* as ‘want’, it is as though by adding volitional *mo* to *mesti*, the speaker believes strongly that the predicated proposition must be what is desired, whether it is part of their belief system (epistemic) or something that they want instilled in someone else’s (deontic).

(442) *Makan mesti mo [orang suap].*
 eat **must want** person feed
 ‘(when this person) eats, there must be a person to feed (him).’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:05:01.1-00:05:02.4)

- (443) *Kenching orang mesti mo [jaga apa].*
 urinate person **must want** take.care what
 ‘(when this person) urinates, a person must take care of whatever it is.
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:05:02.4-00:05:04.3)
- (444) *Mesti mo [ingat-kan gua].*
must want remember-TR 1.SG
 ‘(you) must remember me.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:06:26.0-00:06:27.4)
- (445) *Kalu sudah tua, mesti mo [jaga badan].*
 if alreadyold **must want** take.care body
 ‘If (you are) already old, (you) must take care of (your)body.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:04:49.6-00:04:53.6)

On its own, *mo* ‘want’ is also an auxiliary verb that expresses the future tense (see section 5.2.5.1), and or the deontic modality.

- (446) *Dia mo [piléh].*
 3.SG **want** choose
 ‘He wants (to) choose.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:19.3-00:01:20.0)
- (447) *Gua mo pegi itu kebun bunga.*
 1.SG **want** go that garden flower
 ‘I want (to) go (to) that flower garden.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:12:32.0-00:12:33.7)

In general, *boléh* ‘want’, *mesti* ‘must’ and *mesti mo* ‘must want’ are auxiliaries that express modalities when they precede the main verb phrase (Aux VP).

5.2.3 Passivization

Passivization is a process that promotes the object NP into the subject position. In BM, there are two patterns of passivization, the first of which uses the passive verb *kena*, and the other uses the ditransitive and causative verb *kasi* ‘give’ (see section 5.2.4). The form of passivization that is more commonly used is the *kena* passivization. When the passive verb *kena* is used, the logical subject is not expressed in the utterance. Examples (448) and (450) are the active counterparts of the passive sentences in examples (449) and (451).

- (448) *Orang itu pukol kucing.*
 person that hit cat
 ‘That person hit the cat.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:42:26.0-00:42:27.8)
- (449) *Itu kucing kena pukol.*
 that cat PASS hit
 ‘That cat was hit.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:42:37.0-00:42:38.8)
- (450) *Itu kucing makan ikan.*
 that cat eat fish
 ‘That cat eats fish.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:37:23.2-00:37:25.5)
- (451) *Itu ikan sudah kena makan*
 That fish already PASS eat
 ‘That fish has been eaten.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:38:17.8-00:38:19.9)

It is important to note that *kena* is an adversative passive, and semantically incompatible with sentences that have non-adversative connotations. Note that besides being a passive marker, *kena* is also used as a non-volitional verb meaning ‘subjected to’. While *kena* functions as a passive marker in Malay, its non-volitional function is said to have been derived from Hokkien passive and non-volitional marker *tio?* (see Lim 1988, section 7.4.1). The following are examples that show non-volitional *kena* is used.

- (452) *Tapi lu ada kena kaki chaukah, alamak*
 but 2.SG PFV **subjected.to** friend bad.sport EXCLAM.dismay
 ‘But if you have been subjected to friends who are bad sports, oh no.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:15:12.8-00:15:15.9)
- (453) *Mak kena masak*
 mother **subjected.to** cook
 ‘Mother was subjected to cook.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:06:55.0-00:06:58.0)

As demonstrated by example (452), a noun can follow *kena* as compared to instances of *kena* passivization where the logical subject is not expressed. It is thus important to differentiate

examples of non-volitional *kena*, which is the auxiliary verb that precedes the main VP in (452) and (453), from true instances of passivization, where the object NP is promoted to subject position.

Whereas the oblique is not kept in passive *kena* sentences, it is maintained in *kasi* passives. *Kasi* literally means ‘give’, except that there is no volition in these passives on the part of the patients, or the subjects of the passives that undergo the event. These passives are also adversative passives, and are most likely derived from Hokkien (see Pakir 1986, Lim 1988, Shih 2009, and section 7.4.3). In *kasi* passives (but not *kena* passives), the original object or semantic patient of the active sentence is promoted to subject in the passive sentence, and the logical subject becomes the subject of embedded clause. Examples (25) and (26) are replicated as (454) and (455).

(454) *Mary bunoh dia*
 kill 3.SG
 ‘Mary killed him.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:39:17.1-00:39:18.8)

(455) *Dia kasi [Mary bunoh].*
 3.SG PASS kill
 ‘He was killed by Mary.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:40:01.8-00:40:14.9)

(456) *Orang kasi [embok-embok chakap]*
 person PASS traditional.Peranakan.elders talk
 People were talked about by the traditional Peranakan elders.
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:07:59.1-00:08:01.0)

(457) *Siapa kasi [si Mary bunoh]?*
 who PASS PERSON kill
 ‘Who was killed by Mary?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:39:39.8-00:39:43.2)

5.2.4 Ditransitive, causative and benefactive constructions

Ditransitive, as well as causative and benefactive constructions are related in BM. The typical ditransitive construction features the verb *kasi*, which literally means ‘give’, is also used for causative and benefactive constructions as well.

5.2.4.1 Ditransitive constructions

The typical ditransitive verb, *kasi* ‘give’ takes two arguments. In its complete form, the ditransitive verb phrase comprises the *kasi* verb, and a sequence comprising an indirect object (the semantic beneficiary), and a direct object NP (the semantic theme). Example (419) is replicated below as (458).

(458) *Dia kasi [dia] [dua bijik pear].*
3.SG give 3.SG two CLF.fruit
‘He gave him two pears.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:04.0-00:05:05.4)

(459) *Dia kasi [dia] [tiga].*
3.SG give 3.SG three
‘He gave him three.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:20.9-00:05:22.6)

(460) *Mak sudah kasi [gua] [duit].*
mother already give 1.SG money
‘Mother already gave me money.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-094, 00:01:24.0-00:01:25.4)

(461) *Mak.ko ada gula-gula mo kasi [lu] [dua].*
Eldest.paternal.uncle’s.wife have sweet-sweet want give 2.SG two
‘Eldest paternal uncle’s wife has sweets that she wants to give you two.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-094, 00:06:31.0-00:06:48.1)

(462) *Dia kasi [gua] [itu lauk].*
3.SG give 1.SG that cooked.food
‘He gave me that cooked food.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:17:42.3-00:17:45.0)

Other ditransitive verbs observed in the corpus include *ajair* ‘teach’, *pinjam* ‘lend’, and *tunjok* ‘show’. Note that Hokkien uses the same lexeme for both notions of lend and borrow, while

Malay uses *pinjam* for borrow and *memberi pinjam*, literally ‘give borrow’ for the concept of lend. In BM, lend is ditransitive, while borrow is simply transitive, as demonstrated by examples (463) and (464). Example (465) shows how *tunjok* ‘show’ is used in ditransitively.

(463) *Lu pinjam [gua] [lu mia bég].*
 2.SG **lend** 1.SG 2.SG REL bag
 ‘You lend me your bag.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-089, 00:28:07.5-00:28:10.0)

(464) *Gua pinjam [lu mia bég kechik].*
 1.SG **borrow** 2.SG POSS bag small
 ‘I borrow your small bag.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-089, 00:28:16.1-00:28:18.5)

(465) *Gua tunjok [lu] [keday].*
 1.SG **show** 2.SG shop
 ‘I showed you the shop.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:38:38.1-00:38:39.7)

In general, in ditransitive instances where the verb takes an additional argument, the ditransitive verbs precede the beneficiary, an indirect object NP, and the theme, a direct object NP, in that order. Also, it is possible for either indirect object to not be expressed. Example (464) is replicated as (466).

(466) *Gua pinjam [lu mia bég kechik].*
 1.SG **borrow** 2.SG POSS bag small
 ‘I borrow your small bag.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-089, 00:28:16.1-00:28:18.5)

(467) *Gua mo kasi [angpau].*
 1.SG want **give** red.packet.of.monetary.gift
 ‘I want to give a red packet.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:44:51.2-00:44:52.4)

There are no instances in the corpus where direct object is not expressed.

5.2.4.2 Causative and benefactive constructions

Causatives in BM can also be formed with the verb *kasi*. These causatives have been identified in previous literature as being highly similar to a Hokkien construction (Lim 1981, Lim 1988,

Pakir 1986, Shih 2009), see section 7.4.1 for more discussion. In effect, the causative construction appears similar to the ditransitive *kasi* ‘give’ construction, except that the indirect object NP does not necessarily benefit positively, and that the direct object theme is a verb phrase representing the caused event that affects the indirect object NP. In effect, the indirect object NP and the verb phrase constitute an embedded clause, similar to the *kasi* passive pattern (see section 5.2.3). The following sentences feature the causative verb *kasi*. Example (420) is replicated here as example (468).

(468) *Yauguai, kasi [gua terperanjat sekali].*
 demon cause 1.SG be.shocked very
 ‘Demon, you made me be very shocked.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:040:26.2-00:40:28.2)

(469) *Dia kasi [gua marah sekali].*
 3.SG cause 1.SG angry very
 ‘She made me very angry.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:39:42.7-00:39:45.4)

Other than *kasi* ‘give’, *paksa* ‘force’ can also be used in causative constructions, such as with example (470).

(470) *John paksa [Mary lupa-kan si Peter].*
 force forget-TR PERSON
 ‘John forced Mary (to) forget Peter.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-057, 00:07:04.2-00:07:10.4)

There are instances where the indirect object in fact benefits from the particular event predicated by the embedded verb phrase. These are known as benefactive constructions. Note that *paksa* ‘force’ cannot be used in benefactive constructions. Only *kasi* ‘give’ can be used in these instances, as with examples (471) to (473). This phenomenon is also recognised by Lim (1988)

(471) *Mak kasi [gua pinjam satu ratus].*
 mother let 1.SG borrow one hundred
 ‘Mother let me borrow one hundred.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-094, 00:00:40.5-00:00:46.0)

- (472) *Saya ada satu lagu mo kasi [lu dengar].*
 1.SG have one song want **let** you hear
 ‘I have a song that I want to let you hear.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-094, 00:01:05.4-00:01:12.7)
- (473) *Jangan kasi [dia dengair ini perkara].*
 do.not **let** 3.SG hear this matter
 ‘Do not let him hear about this matter.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-094, 00:05:39.4-00:05:48.2)

5.2.5 Tense and aspect

BM utilizes an optional system that combines tense and aspect, although the aspectual system is more complex than its tense system. A tense system is one that relates to the time of speaking, while an aspectual system is one that focuses on “different ways of viewing the internal temporal constituency of a situation” (Comrie 1976:3). The following figure represents the tense and aspectual system of BM. The only tense that is available in BM is the future. These are expressed with the adverbs *belum* ‘not yet’ and *nanti* ‘later’, and auxiliary verb *mo* ‘want’, which indicate that a current event has not taken place as of current time but will take place later. The aspectual system comprises adverb *sudah* ‘already’, which has a perfective meaning, and the auxiliary verb *ada* ‘have’, which is used to indicate several aspects, including the perfective, the progressive, experiential perfect and habitual, the adverb *baru*, which has a recent perfect meaning, and the adverb *pernah* ‘ever’ which is used to indicate the experiential perfect aspect. Tentative aspect is signaled by the use of reduplication (see section 3.2.1). This tense and aspect system is optional, especially when there is enough context provided to tell the interlocutor if the event has been completed, or if it is still going on or will happen in the future. More details on how sentences can be modified with adverbial phrases that provide information of this sort can be found in section 5.6.6.1. Note also that most aspect markers are glossed in accordance to their literal meanings, because they contribute to the understanding of the particular aspect they are used for, except for *ada*, which has been glossed according to the list of linguistic abbreviations provided in section 2.2.3.

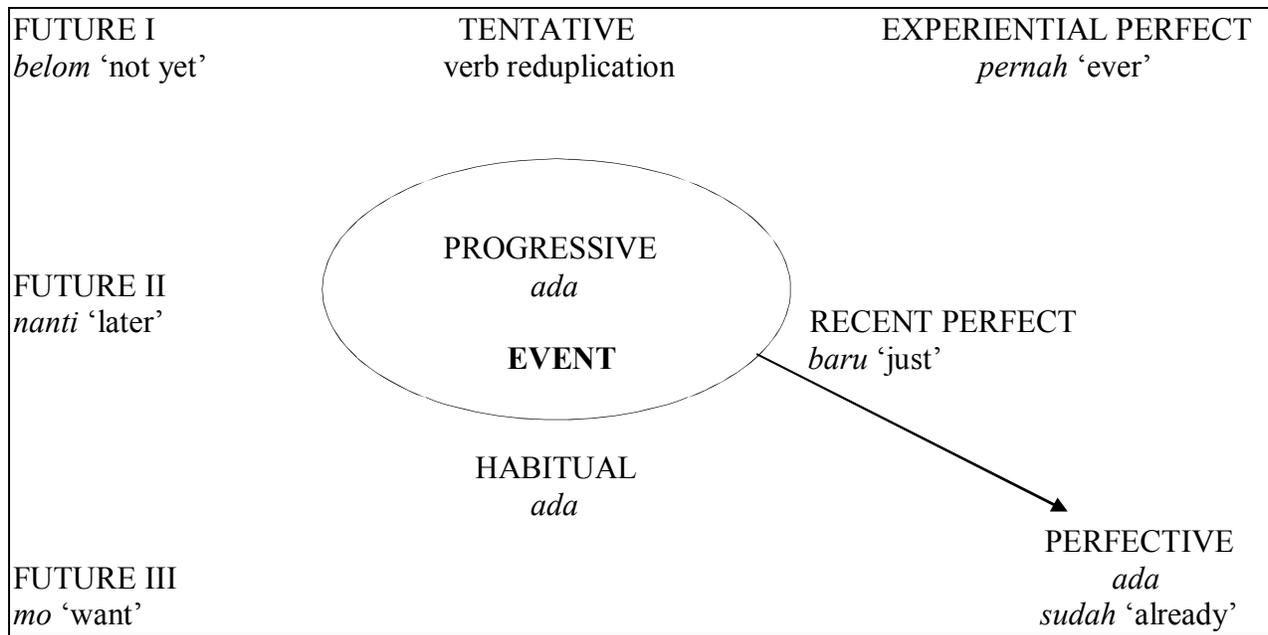


Figure 31: Tense and aspect system of Baba Malay.

5.2.5.1 Future tense

Future is indicated in three ways, through the use of the adverbs *belum* 'not yet' and *nanti* 'later', as well as by attaching the auxiliary verb *mo* 'want'. The following show how these adverbs and auxiliary are used to express the future tense. This is done by attaching these before the main verb phrase. Example (447) is replicated below as example (480). *Mo* in examples (480) and (481) express that something will happen in the future in addition to deontic modality (see section 5.2.2).

- (474) *Saya **belum** kasi tau mak.*
 1.SG **not.yet**let know mother
 'I have not let mother know.'
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-094, 00:02:03.4-00:02:05.1)

- (475) *Dia **belum** habis.*
 3.SG **not.yet**finish
 'He has not finished.'
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:15:29.2-00:15:30.3)

- (476) *Budak-budak **belum** pegi.*
 child-PL **not.yet** go
 ‘The children have not gone.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-100, 00:11:33.1-00:11:37.1)
- (477) *John **nanti** belajardua jam kat library*
 later study two hour PREP
 ‘John will study (for) two hours at the library.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:28:40.5-00:28:50.1)
- (478) *Gua **nanti** tidor sampay bésok*
 1.SG **later** sleep until tomorrow
 ‘I will sleep until tomorrow.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-058, 00:05:51.3-00:05:53.7)
- (479) *Gua **nanti** tak makan sampay dia masak ayam*
 1.SG **later** NEG eat until 3.SG cook chicken
 ‘I will not eat until she cooks chicken.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-058, 00:06:36.8-00:06:41.5)
- (480) *Gua **mo** pegi itu kebun bunga.*
 1.SG **want** go that garden flower
 ‘I want (to) go (to) that flower garden’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:12:32.0-00:12:33.7)
- (481) *Gua **mo** pi belakang*
 1.SG want go behind
 ‘I want to go to the back.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:17:37.6-00:17:39.6)

Note that both *belum* and *nanti* are also adverbs that can modify the entire clause when placed in front of the clause.

- (482) ***Belom** [dia pulang], gua sudah dapat tau.*
before 3.SG return.home 1.SG alreadyget know
 ‘Before she returned home, I already got to know.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-096, 00:03:17.4-00:03:40.4)
- (483) ***Nanti** [Mary tinggal dekat Singapore sampay January].*
later stay PREP until
 ‘Later Mary (will) stay in Singapore until January.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:18:24.3-00:18:27.8)

5.2.5.2 Perfective aspect

Where aspect is concerned, BM differentiates between the perfective and the imperfective progressive. The perfective “indicates the view of a situation as a single whole, without distinction of the various separate phases that make up that situation” (Comrie 1976:16). The perfective is thus often associated with completed action (Comrie 1976). In BM, there are two ways to indicate the perfective. These are the use of the auxiliary verb *ada* and the adverb *sudah*. While *ada* is also used to mean ‘have’, and as an existential marker, a copula, a progressive marker and a habitual marker, *sudah* literally means ‘already’. A shortened version of *sudah* in SBM is *sua*. The following are examples of how *sudah* and *ada* may be used to express the perfective aspect.

- (484) *Gua sudah kata lu, betul?*
1.SG **already** tell 2.SG correct
‘I told you, right?’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:38.0-00:03:39.6)
- (485) *Lu sudah jatoh lu mia kopiah*
2.SG **already** dropped 2.SG POSS hat
‘You dropped your hat.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:04:55.5-00:04:57.5)
- (486) *Gua sudah tau.*
1.SG **already** know
‘I knew.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:19:80.1-00:19:19.4)
- (487) *Dia ada beli apple, bukan?*
3.SG **PFV** bought no
‘She bought an apple, no?’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:40:03.5-00:40:06.9)
- (488) *Ada tukar itu burung*
PFV change that bird
‘(It) changed (into) that bird.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:04:15.0-00:04:18.0)

- (489) *Gua ada tutop.*
 1.SG **PFV** CLOSE
 ‘I closed (the door).’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:22:15.8-00:22:16.7)

5.2.5.3 Recent perfect aspect

Baru is used as an adverb to indicate the recent perfect, *baru* also being an adjective indicating ‘new’. It refers to a situation recently completed, and usually still affecting the (current) moment of speech. This is demonstrated by examples (490) and (491).

- (490) *Dia baru datang.*
 3.SG **just** come
 ‘He just came.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:30:58.1-00:30:59.2)

- (491) *Dia baru dapat tau.*
 3.SG **just** receive know
 ‘He just got to know.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-096, 00:00:24.2-00:00:25.8)

Note that in a separate construction, *baru* may also modify the clause that follows it, *baru* in this instance meaning ‘just then’, or just at that moment. This is shown in examples (492) and (493).

- (492) *Lagik satu minggu, baru dia rasa baik*
 more one week **just.then** 3.SG feel good
 ‘One more week, just then he felt better.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:29:33.9-00:29:36.4)

- (493) *Baru masok buah paya*
just.then enter CLF.FRUIT papaya
 ‘Just then (you) put in papaya.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:12:45.7-00:12:48.0)

5.2.5.4 Progressive aspect

The progressive aspect is imperfective, which means that it “makes explicit reference to the internal temporal structure of a situation, viewing a situation from within” (Comrie 1976: 24). As

a specific type of imperfective, the progressive refers to a temporary, continuous state (Comrie 1976). The progressive aspect in BM is also expressed by using *ada* as an auxiliary verb.

- (494) *Itu perempuan ada baca magic book*
 that woman **PROG** read
 ‘That woman is reading a magic book.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:08:46.5-00:08:49.1)
- (495) *Mak tu ada baca lagik*
 mother that **PROG** read more
 ‘That mother is reading more.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:08:54.3-00:08:57.0)
- (496) *Lu ada bikin apa?*
 2.SG **PROG** make what
 ‘You are making what?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:12:50.2-00:12:51.8)

5.2.5.5 Habitual aspect

Another imperfective aspect is the habitual aspect. This “describes a situation which is characteristic of an extended period of time” (Comrie 1976:27-28). Again, *ada* can be used as an auxiliary verb that indicates the habitual aspect.

- (497) *Ada pegi.*
HAB go
 ‘(I) usually go.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:40.0-00:00:41.0)
- (498) *Nampak gua dia ada senyum.*
 see 1.SG 3.SG **HAB** smile
 ‘(when she) sees me she usually smiles.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:19:32.2-00:19:33.9)
- (499) *Tak senang, gua ada masak*
 Neg free 1.SG **HAB** cook
 ‘Not free, I usually cook.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:02:07.6-00:02:09.7)

5.2.5.6 Experiential perfect aspect

Different than the previous aspects discuss, the perfect describes not the situation itself, but “relates some state to a preceding situation” (Comrie 1976:52). The experiential perfect specifically indicates that a situation has happened “at least once during some time in the past leading up to the present” (Comrie 1976:58). In a sense, this is a combination of both time and event, and hence relates to both tense and aspect. The experiential perfect in BM is indicated by auxiliary verb *pernah* ‘ever’. Both *pernah* ‘ever’ and *tak pernah* ‘never’ are commonly used in BM.

- (500) *Dia pernahjumpa gua.*
3.SG ever meet 1.SG
‘He has met me before.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:03:37.0-00:03:38.2)
- (501) *John rasa dia pernahjumpa Mary.*
think 3.SG ever meet
John thinks that he has met Mary before.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:42:23.6-00:42:25.7)
- (502) *Dia tak pernahbikin ini sumua.*
3.SG NEG ever do this all
‘She never did this all.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:42:23.6-00:42:25.7)

In summary, other than perfective, progressive, and habitual marker *ada* which is an auxiliary verb, the rest of the tense and aspect markers are adverbs. These include *belum* ‘not yet’, *nanti* ‘later’, *sudah* ‘already’ and *pernah* ‘ever’. All tense and aspect markers regardless of whether they are auxiliary verb or adverbs occur before the main verb phrase.

5.2.5.7 Tentative aspect

Tentative aspect can be expressed in BM by verb reduplication (see section 3.2.1). Examples are as follows. The tentative aspect is being expressed in these instances since there is no particular goal or purpose expressed, and that the action predicated is meant to take place for a short duration of time. Smith (1991) claims that it is a type of perfective aspect, since it represents a closed situation, of short duration and little importance. Verbal reduplication that

expresses tentative aspect is commonly found in Sinitic languages, such as Hokkien and Mandarin (Tsao 2004).

- (503) *Jalan jalan*
walk walk
'Take a walk.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:13.1-00:02:14.6)

- (504) *Téngok téngok*
look look
'Take a look.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:14.6-00:02:16.2)

5.2.6 Serial verb constructions

The notion of serial verbs here follows that of Sebba (1987) and Aikhenvald (2005). According to Sebba's criteria, both verbs must be lexical verbs that have to be interpreted as having the same categories of tense-aspect-mood. There should also be neither clause boundary nor conjunction between the two verbs (Sebba 1987). In addition, Aikhenvald (2005) states that each component in the serial verb construction must be able to occur on its own. This is not the case with periphrastic constructions. Serial verb constructions also have to be differentiated from compound verbs such as *buang buang*, literally 'throw throw', meaning 'exorcise', or *naik geleték*, literally 'ascend tickle', meaning 'to be up to mischief.' The following are some examples of serial verb construction in BM. Note that serial verb constructions comprising three verbs are also possible, as in example (507).

- (505) *Dia [turun] [masok]*
3.SG descend enter
'He descended and entered.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:05:37.3-00:05:39.8)

- (506) *Kodok [turun] [kuluair]*
frog descend go.out
'The frog gets down and out.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:08:31.4-00:08:33.3)

- (507) *Gua mo [kuluair] [pi] [jalan]*
 1.SG want **go.out** **go** **walk**
 ‘I want to go out, and go and walk.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:53:39.7-00:53:41.8)
- (508) *Mama gua mo [pi] [buang-ayé]*
 grandmother 1.SG want **go** **throw –water**
 ‘Grandmother I want to go urinate.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:48:25.8-00:48:27.5)
- (509) *[Dudok] [tengok] TV*
sit **watch**
 ‘Sit and watch the television.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:03:17.1-00:03:19.0)
- (510) *Kodok tu [naik] [panjat].*
 frog that **ascend climb**
 ‘The frog is ascending and climbing.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:07:37.3-00:07:41.3)
- (511) *Kawan -kawan dia [datang] [tolong]*
 friend -friend 3.SG **come** **help**
 ‘His friends came and helped’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:58.4-00:04:00.9)
- (512) *Kawan dia [datang] [angkat satu kambing].*
 friend 3.SG **come** **hold** **one** **goat**
 ‘His friend comes holding one goat.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:41.8-00:01:44.8)

All instances of serial verb construction in BM appear to involve a sequence of dynamic verbs and not stative ones. Speakers also state that it is unnatural for verb sequences to be broken up by conjunctions. Thus serial verb constructions in BM follow an uninterrupted VP-VP sequence.

5.2.7 Verb phrases with adverbial modifiers

Verb phrases may be also be modified by adverbs. This has been earlier demonstrated in section 5.2.5, where adverbs precede verbs to express tense and aspect. The following examples show other examples that may precede or follow the verb phrase.

- (513) *Itu bangkuang pun tak bergerak.*
that turnip **also** NEG move
'That turnip also does not move.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:01:20.0-00:01:22.2)
- (514) *Aunty Jane selalu ketawa-kan gua.*
always laugh-TR 1.SG
'Aunty Jane always laughs at me.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:18:03.5-00:18:06.4)
- (515) *Dia langsung tidur.*
3.SG **straightaway** sleep
'She slept straightaway.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:06:54.0-00:06:55.8)
- (516) *Gua terus lari*
1.SG **straight** run
'I ran straight.'
(Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:00:56.7-00:00:57.5)
- (517) *Dia pelan-pelan makan kurang.*
3.SG **slow-slow** eat less
'They slowly eat less.'
(Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:03:42.6-00:03:44.4)
- (518) *Lu dudok diam-diam*
2.SG sit **quiet-quiet**
'You sit quietly.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:19:50.0-00:19:52.8)
- (519) *Ingat baik-baik*
remember good-good
'Remember well.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:21:20.1-00:21:22.7)

- (520) *Boléh lu tolong gua masak lagik*
 can 2.SG help 1.SG cook **again**
 ‘Can you help me cook again?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:13:30.1-00:13:34.1)
- (521) *Dia lalu pulak.*
 3.SG pass **instead**
 ‘He passed by instead.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-028, 00:02:14.0-00:02:15.1)
- (522) *Masak sahja*
 cook only
 ‘(I) cook only.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:05:55.9-00:05:57.1)

In fact, functional adverbs have fixed positions. Some may always precede the verb phrase while others always follow the verb phrase. The following table is a list of commonly used functional adverbs that modify verb phrases, and their positions in relation to the verb phrase.

<i>pun/kun</i>	also	_ VP
<i>selalu</i>	always	_ VP
<i>baru</i>	just	_ VP
<i>balék</i>	again, back	VP _
<i>dulu</i>	first (before doing something else)	VP _
<i>lagik</i>	still, more, again	VP _
<i>pulak</i>	instead	VP _
<i>sahja</i>	only	VP _
<i>jugak</i>	also	VP _

Table 38: List of commonly used adverbs that modify verb phrases and their distribution

Considering the above, there is no preferred position for the adverb that modifies the verb phrase. A verb phrase can comprise Adv VP, or VP Adv. Both are equally common.

5.2.8 Negation of verb phrases

There are a couple of ways in which a verb can be negated, the most general negative marker being *tak*. It precedes the main verb phrase that it negates.

- (523) *Tak jadi.*
 NEG happen
 ‘(It) did not happen.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:55.0-00:02:56.3)
- (524) *Gua tak reti.*
 1.SG NEG understand
 ‘I do not understand.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:16:31.2-00:16:32.0)
- (525) *Dia tak pi.*
 3.SG NEG go
 ‘She does not go.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:04:50.5-00:04:51.2)
- (526) *Itu kambing, gua tak tau perempuan ka jantan.*
 That goat 1.SG NEG know female or male
 ‘That goat, I don’t know (if it is) female or male.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:47.3-00:01:49.8)
- (527) *Gua rasa pear tu tak boleh makan*
 1.SG think that NEG can eat
 ‘I think that pear cannot be eaten.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:24.3-00:05:25.7)

Commonly used negative forms of verbs have been developed by contracting negative marker and verb. These are *tak-a*, from *tak ada* ‘NEG have’, and *toksa*, from *tok usa* ‘NEG need’.

- (528) *Tak-a*
 NEG-PFV
 ‘(Did) not.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:22:14.8-00:22:15.8)

- (529) *Kita sini tak-a pokok.*
 1.PL here NEG-have tree
 ‘Here we do not have trees.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:35.5-00:00:36.9)
- (530) *Toksa*
 NEG.need
 ‘No need.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:15:25.0-00:15:25.5)
- (531) *Toksa tanya*
 NEG.need ask
 ‘(Do) not need to ask.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:47:31.6-00:47:33.4)
- (532) *Toksa marah*
 NEG.need angry
 ‘(Do) not need to be angry.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:40:02.6-00:40:01.0)

Other than *tak*, *jangan* is also used in the negation of verb phrases. More specifically, *jangan* is used in imperatives or commands. More details on non-negative imperatives can be found in section 5.6.11.

- (533) *Jangan raba.*
 do.not touch
 ‘Do not touch.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:21:17.3-0021:18.7)
- (534) *Jangan rindu gua*
 do.not miss 1.SG
 ‘Do not miss me.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:29:30.0-00:29:32.4)
- (535) *Jangan nangis*
 do.not cry
 ‘Do not cry.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:29:32.4-00:29:33.7)

(538) Comparison of equality

Dia tinggi sama ngko
3.SG tall same older.brother

‘He (is) tall like older brother.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:10:16.2-00:10:17.5)

(539) Comparison of similarity

Kerair begi batu
hard like rock

‘Hard like a rock.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:25.7-00:05:25.9)

(540) Superlative (relative clause)

Ini apple yang manis sekali
this REL sweet very

‘This (is the) apple that (is) very sweet.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:35:46.0-00:35:47.8)

(541) Excessive degree

Terlalu sejok
too cold

‘Too cold.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:04:42.3-00:04:43.6)

(542) Adverb and Adjective

Budak ini sangat kechik lah.
child this very small EMP

‘This child (is) very small.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:23.5-00:03:25.3)

(543) Adjective and Adverb

Kambing dia kuat sekali.
goat 3.SG strong very

‘His goat (is) very strong.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:53.3-00:01:54.6)

(544) Negation of Verb Phrase with Adjective

Barang tak hak tak beli lah.
thing NEG suitable NEG buy EMP

‘The thing (is) not suitable, (we) do not buy.’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:18.5-00:02:20.2)

Note in example (544) that while it appears as though the negative marker *tak* precedes the adjective directly, copulas are usually unexpressed in BM, and that *tak* negates the verb phrase that comprises *hak* as its complement. Verb phrase negation is covered in section 5.2.8.

5.3.1 Comparatives

Comparatives in BM are expressed by the structure *lagik-AP-lagik*, *lagik-AP* and *AP-lagik*, as shown in the following examples. This form of comparative has also been observed by Lee (1999). *Lagik* has the literal meaning of ‘more’ in these instances.⁷⁵ The noun phrase that follows the comparative expression is an oblique. It is optional as shown by examples (549) and (550). Example (537) is replicated as example (545).

- (545) *Ini apple lagik manis lagik itu apple.*
 this more sweet more that
 ‘This apple (is) sweet than that one.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:35:27.2-00:35:29.8)
- (546) *Ini budak pandaylagik itu budak.*
 this child clever more that child
 ‘This child (is) more clever than that child.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:34:17.0-00:34:22.7)
- (547) *Rumah batu lagik bagus lagik rumah yang lain.*
 house rock more good more house Rel other
 ‘The rock house (is) better than the house that (is) the other.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:03:16.9-00:03:27.0)
- (548) *Rumah kayu lagik bagus lagik rumah rumput kering.*
 house wood more good more house grass dry
 ‘The wooden house (is) better than the hay house.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:02:46.1-00:02:50.2)
- (549) *Lagik senang.*
 more easy
 ‘(It is) easier.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:12:50.4-00:12:50.9)

⁷⁵ *Lagik* also means ‘still, later, and again.’

- (550) *Ini lagik shiok⁷⁶ kan?⁷⁷*
 this **more** satisfied.feeling no
 ‘It feels better, no?’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:07:24.0-00:07:25.7)

Expressions of comparatives always follow the noun phrase that it modifies. Again, the copula is non-obligatory.

5.3.2 Comparison of equality

It is also possible in BM to express equality using the structure in the following examples, where *sama* is used, linking adjective with the noun phrase the subject is being compared to.

Technically, *sama* NP appears to function as an adverb that modifies the adjective by following it. Example (538) is replicated below as example (551).

- (551) *Tapi dia tinggi sama ngko*
 But 3.SG tall **same** older.brother
 ‘But he (is) tall like older brother.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:10:16.2-00:10:17.5)

- (552) *Derek tinggi sama ngko dia*
 tall **same** older.brother 3.SG
 ‘Derek (is) tall like his older brother.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:10:31.5-00:10:33.6)

- (553) *Panday sama anak dia*
 clever **same** son 3.SG
 ‘Clever like her son.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:08:10.1-00:08:12.1)

Again, the main noun phrase precedes the adjectival phrase, except for example (553) where the subject is dropped. Subjects are optional in BM (see section 5.6.2).

⁷⁶ This is the only word where ʃ is found, hence *shiok* [ʃjoʔ] is not part of the stable consonant inventory. The word itself may have been derived from Punjabi *shauk* [ʃjawʔ], which is an exclamation that is akin to ‘great’.

⁷⁷ *Kan* in this instance is a shortened version of *bukan* ‘no’.

5.3.3 Comparison of similarity

Another type of comparison in BM is that of similarity. These expressions are akin to stating that the subject in question is similar to something else, but not completely alike. For these expressions, *begitu* or *macham* are used, both meaning ‘like’. The short form of *begitu* is *begi*. Similar to comparatives of equality, *begi* NP and *macham* NP appear to be adverbs that modify the adjectives or verbs they follow, except for examples (557) and (559) where they modify the noun phrases that they follow. Note that many of these expressions featuring *begitu* or *macham* are idiomatic, as with examples (555), (556), (558), and (559). Example (539) is replicated here as example (554).

- (554) *Kerair begi batu*
hard like rock
‘Hard like a rock.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:25.7-00:05:25.9)
- (555) *Dia-orang seronoh begi kain lipat*
3-PL proper like cloth fold
‘They (are) proper like a folded cloth.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:03:03.8-00:03:13.0)
- (556) *Dia chakap begi lidah tak-a tulang.*
3.SG speak like tongue NEG-have bone
‘He speaks like a tongue without bone (uncontrollable).’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-104, 00:13:51.4-00:13:53.8)
- (557) *Itu macham Chettiar Melaka*
that like
‘That (is) like the Chettiars in Malacca (known for being moneylenders)’
(Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:02:21.7-00:02:23.4)
- (558) *Orang ni pandaysekali macham gauchaytian*
person this clever very like monkey.god
‘This person (is) very clever like the monkey god.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-104, 00:22:36.7-00:22:40.8)
- (559) *Dia macham kain lipat*
3.SG like cloth fold
‘He (is) like a folded cloth (very proper).’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:03:28.7-00:03:21.3)

5.3.4 Superlatives

There is no actual grammatical construction in BM that is strictly a superlative. In fact, the interpretation of the notion of a superlative usually depends on the given context, given that superlatives take the form of relative clauses (see section 5.6.3). For example, the speaker produces (560) and (561) after giving examples about what constitutes comparatives, the earlier example (545) being the comparative counterpart of these sentences. Similarly, example (562) is produced in this speaker's retelling of Grimm's *The Three Little Pigs*, and after producing this utterance, he says in English, "the youngest". Example (540) is replicated here as (560).

(560) *Ini apple yang manis sekali*
this REL sweet very
'This is the apple that is very sweet (the sweetest apple)'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:35:46.0-00:35:47.8)

(561) *Ni yang manis sekali.*
this REL sweet very
'This is that which is very sweet (this which is the sweetest).'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:35:43.0-00:35:43.7)

(562) *Si babi yang kecil*
PERSON pig REL small
'The pig that is small (the pig that is the smallest)'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:01:12.0-00:01:17.0)

These sentences that are to be interpreted as superlatives are essentially relative clauses that post-modify the noun phrase head. Lee (1999) also interprets similar constructions to be superlatives, although she does not identify these as relative clauses. More information on relative clauses can be found in section 5.6.3.

5.3.5 Excessive degree

Excessive degree is expressed in BM by pre-modifying the adjective concerned with the adverb *terlalu* 'too'. In Malay, *ter-* forms the superlative when attached to an adjective. However in this instance, *ter-* is attached to the verb *lalu* which means to 'pass by something or someone', and the resulting adverb expresses the notion of 'too', or to an excessive degree. The prefix *ter-* is otherwise not productive in this manner, and it is not used for purposes other than to express that

a verb is accidental or involves movement (see section 4.2.1.1). Examples of utterances expressing excessive degree are shown below.

- (563) *Itu kuéh kuéh terlalu manis.*
that cake cake **too** sweet
'Those cakes (are) too sweet.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:05:50.4-00:05:54.5)
- (564) *Tu kuéh rasa terlalu manis*
that cake feel **too** sweet
'That cake (I) feel (is) too sweet.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:06:27.6-00:06:30.3)
- (565) *Itu bubor terlalu panas*
that porridge **too** hot
'That porridge (is) too hot.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:04:27.2-00:04:30.6)
- (566) *Itu kerosi terlalu keras.*
that chair **too** hard
'That chair (is) too hard.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:05:19.8-00:05:21.4)
- (567) *Kerosi tu terlalu lembéh.*
chair that **too** soft
'That chair (is) too soft.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:05:29.2-00:05:31.8)
- (568) *Itu pun terlalu tinggi.*
that also **too** tall
'That (is) also too tall.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:06:34.1-00:06:35.7)

Hence, where expressions of excessive degree are concerned, the adjectival phrase is head-final, with preceding modifier *terlalu* 'too'.

5.3.6 Adjectival phrases with adverbial modifiers

In general, adverbs may precede or follow the main adjectival phrase. Example (542) is shown below as example (569).

- (569) *Sunggu* *chukop*
really enough
‘Just nice.’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:05:50.8-00:05:52.3)
- (570) *Budak ini* *sangat kechik lah.*
child this **very** small EMP
‘This child (is) very small.’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:03:23.5-00:03:25.3)
- (571) *Manyak* *panas, dia tak boléh tahan.*
Many hot 3.Sg Neg can withstand
‘(It is) very hot, she cannot stand (it).’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:04:28.6-00:04:30.6)
- (572) *Mahal* *sikit*
expensive **little**
‘A little expensive.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:48.7-00:02:50.0)
- (573) *Chanték* *sekali*
beautiful **very**
‘Very beautiful.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:35.7-00:02:37.2)
- (574) *Dulu* *dia jahat sikit.*
long.ago 3.SG evil **little**
‘Long ago he (was) a little evil.’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:35.7-00:02:37.2)

There is no preference for either head-initial or head-final adjectival phrase. Both are observed to be equally common in BM.

5.3.7 Order of elements in adjectival phrases

The different types of adjectival phrases covered in this section are general comparatives, comparatives of equality, comparatives of similarity, superlatives (in the form of relative clauses), and expressions of excessive degree, as well as simple adjectival phrases that are either pre-modified or post-modified by adverbs. In many instances (such as with the different comparatives), the modifiers mostly follow the head adjective. However, it is also common to

have modifiers preceding the adjectives, as with expressions of excessive degree and general modifications of the simple adjectival phrase. Note that adverbs may also precede or follow basic adjectival phrases. In general, similar to the noun phrase it is inaccurate to state that the adjectival phrase has a preference for being head-initial or head-final.

5.4 Adverbial phrases

Examples in earlier sections demonstrated how adverbs may be used to modify verbs and adjectives (see sections 5.2.7, 5.3.3, 5.3.5, and 5.3.6 for examples). This section shows how adverbial phrases may provide more information on matters of time, location, manner, intensity, and degree. These adverbial phrases are not strictly limited to phrases containing adverbs. Rather, they may also be noun phrases, prepositional phrases, or adjectival phrases that function as modifiers of verb phrases and other adjectival phrases. These adverbial phrases also have to be differentiated from adverbial clauses that modify other clauses (see section 5.6.6).

(575) Temporal

Mary nanti tinggair Melaka dua minggu.
later live two week

‘Mary will stay in Malacca for two weeks.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:28:06.4-00:28:12.4)

(576) Location

Mary beli apple kat pasair.
buy PREP market

‘Mary bought the apple at the market.’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:33:27.0-00:33:32.4)

(577) Manner

Kuching belanda lari chepat sekali
rabbit run fast very

‘The rabbit ran very fast.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:00:51.0-00:00:55.1)

5.4.1 Temporal adverbial phrases

Adverbial phrases may be used to indicate information regarding time. In BM, it is common to have noun phrase or preposition phrase functioning as an adverbial phrase that modifies the verb phrase. Example (575) is replicated below as example (578).

- (578) *Mary nanti tinggair Melaka dua minggu.*
later live two week
'Mary will stay in Malacca for two weeks.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:28:06.4-00:28:12.4)

- (579) *John nanti belajar dua jam kat library.*
later study two hour PREP
'John will study for two hours at the library.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:28:40.5-00:28:50.1)

- (580) *Kukus tu sayur dalam se-puluh minit*
steam that vegetable inside one-ten minute
'Steam those vegetables for ten minutes.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:27:29.9-00:27:36.5)

Adverbial phrases regarding time usually follow the verb phrases that they modify instead of preceding them. A sentence may also have more than one adverbial phrase, as with example (579), which has two adverbial phrases, the first one following the verb phrase expressing time, and the second one expressing location.

5.4.2 Location adverbial phrases

Adverbial phrases in BM that usually comprise prepositional phrases are used to give information regarding location. Examples of these have also been presented in the earlier section on prepositions (see 4.6). Example (293) is replicated here as example (581), (267) as (582), and (579) as (583).

- (581) *Dia dudok bawah pokok.*
3.SG sit bottom tree
'He sat at the bottom of the tree.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:34.9-00:01:37.1)

- (582) *Kupukupu trebang di dekat itu mia anjing.*
 butterfly fly PREP near that REL dog
 ‘The butterfly flew near that dog.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:01:36.8-00:01:41.7)
- (583) *John nanti belajardua jam kat library.*
 later study two hour PREP
 ‘John will study for two hours at the library.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:28:40.5-00:28:50.1)
- (584) *Ada gui kat sana.*
 PROG kneel PREP there
 ‘(They) are kneeling there.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:05:43.6-00:05:45.8)

Again, these adverbial phrases that modify the verb phrase usually following rather than precede it. A verb phrase may also be modified by more than one adverbial phrases, as with example (583). Consultants state that there is no preference for any specific order where these adverbial phrases are concerned. Temporal adverbial phrase may precede location adverbial phrase, or vice versa.

5.4.3 Manner adverbial phrases

Adverbial phrases may also be used to indicate the manner in which something is done. This happens when an adverbial phrase is used to modify a verb phrase. Example (577) is replicated as example (585) here. Again, adverbial phrases can comprise adjectives used as adverbs, as with (585) and (586). Adverbial phrases expressing manner may also comprise preposition phrases featuring *sama* ‘with’⁷⁸. This is shown in examples (587) and (588).

- (585) *Kuching.belanda lari chepat sekali*
 rabbit run fast very
 ‘The rabbit ran very fast.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:00:51.0-00:00:55.1)
- (586) *Labi.labi main pelan sair.*
 tortoise play slow CONF
 ‘The tortoise played slowly indeed.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:00:58.4-00:01:00.6)

⁷⁸ *Sama* is also used as general conjunction, ‘and’.

- (587) *Lu boléh bukak itu pintu sama ini konchi.*
 2.SG can open that door **with this key**
 ‘You can open that door with this key.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-123, 00:05:30.3-00:05:33.8)
- (588) *Dia pinjak itu kachua sama dia mia kasot.*
 3.SG step.onthat cockroach **with 3.SG POSS shoe**
 ‘He stepped on that cockroach with his shoe.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-123, 00:06:06.1-00:06:09.8)

5.4.4 Order of elements in adjectival phrases

While single-word adverbs may precede or follow the verb phrase (see section 5.2.7 for examples), more complex adverb phrases like the ones in these sections usually follow the verb phrase instead of preceding it. These adverbial phrases expressing time, location, and manner are mostly head-initial, with the modifiers (mainly complements) following the heads, rather than preceding them.

5.5 Summary of word order at the phrase level

The following table sums up the different types of word order at the phrase level, and the default order of individual phrases, if any.

NP	DEM N / N DEM A <i>punya</i> N / N A (basic) GEN <i>punya</i> N (basic) / N GEN (only when possessor is a pronoun) PERSON N NUM N CLF N NEG N	No dominant tendency
VP	AUX V ADV V (strictly so for tense and aspect) V ADV NEG V	Modifier-head tendency
AP	ADV A A ADV	No dominant tendency
AdvP	ADV COMP	Head-modifier tendency
Prepositions		

Table 39: List of phrases, their word orders and general tendencies

The observed word orders in BM are interesting because the language does not appear to fall within the common parameters that Greenberg (1963) had postulated for SVO languages, BM being very clearly SVO (see section 5.6.2). For example, in languages with prepositions, the genitive is said to almost always follow the governing noun. This is untrue of BM, since the

default order is GEN N (N GEN can only be used when N is a pronoun). These word order properties of BM make it typologically interesting.

5.6 Clauses

While the preceding sections focused on phrase-internal structure, the following section focuses on order at the higher clause level. Essentially, these sections are concerned with what phrases constitute a clause (or what some may refer to as a sentence), and what more complex clauses (such as relative clauses, complement clauses, among others) comprise.

5.6.1 Word order at the clause level

Clauses are generally predicate-final, the predicate being a verb in most instances. Example (23) is shown below as (589).

- (589) *Budak tu senyum*
child that **smile**
'That child smiles.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:35.9-00:00:37.7)

- (590) *Budak pakay baju*
Child **wear clothes**
'The child puts on clothes.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:59.4-00:01:01.8)

Sentences featuring adjectives like the ones that follow, can be interpreted to have an unexpressed copula predicate. For the purpose of comparison, example (593) features an expressed copula, copulas being optional in BM.

- (591) *Ini orang kiamsiap sair.*
this person **miserly** CONF
'This person (is) miserly indeed.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-105, 00:02:06.1-00:02:07.5)

- (592) *Ini orang tuakang sekali.*
this person **generous** **very**
'This person (is) very generous.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-105, 00:01:29.8-00:01:32.5)

- (593) *Ini piso ada tajam.*
 This knife COP sharp
 ‘This knife is sharp.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-105, 00:03:58.5-00:03:59.9)

Similarly, sentence-final predicates can include noun phrases. Again the copula is not expressed.

Example (144) is replicated here as (594).

- (594) *Ini buah pear*
 this CLF.fruit
 ‘These (are) pears.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:27.9-00:00:28.9)

- (595) *Peter punya bapak sama William punya bapak adék-beradék*
 POSS father and POSS father sibling-PL
 ‘Peter’s father and William’s father (are) siblings.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:45:01.7-00:45:06.9)

In general, the order of basic clauses in BM appear to be NP-VP, when subjects are expressed. Again, it is possible for subjects to not be expressed (see 5.6.2). However, as section 5.6.9 shows, this general order may be disrupted when topicalization occurs, as it commonly does in BM.

5.6.2 Grammatical relations and alignment

BM is a SVO (subject verb object) language. The notion of subject here follows from that of Comrie, who states that “the prototype of subject represents the intersection of agent and topic (1989:107). In BM, the subject NP is not marked differently from the object NP. Instead, in most instances, the subject of a clause can be determined by word order. The syntactic subject precedes the main verb phrase. This is demonstrated by examples (596) to (599). Example (197) is shown below as example (598).

- (596) *Beruang kechik ada mangkok kechik*
 bear small Poss bowl small
 ‘The small bear has a small bowl.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:01:20.9-00:01:24.5)

- (597) *Beruang kechik dudok kerosi kechik.*
bear small sit chair small
 ‘The small bear sits on a small chair.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:01:44.0-00:01:46.8)
- (598) *Tu bangkuang tak bergerak*
that turnip NEG move
 ‘That turnip does not move.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:01:00.5-00:01:03.3)
- (599) *Ada satu tukang kebun, sama bini dia, tanam bangkuang.*
EXIST one labourer garden and wife 3.SG plant turnip
 ‘There was a gardener, and his wife, (who) planted turnips.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-034, 00:00:15.9-00:00:22.0)

While both (598) and (599) show that it is difficult to always correlate subject with the most agent-like NP in a clause, and hence provides support for the notion of a prototypical subject instead of a definite one, (599) also demonstrates that the subject is not always overtly expressed in BM. BM is a null-subject language.

In terms of grammatical relations, it is also possible for object to undergo passivization with the use of passive marker *kena*, so that the object is promoted to subject. The logical subject is not expressed. This has also been discussed in section 5.2.3.

- (600) *Orang itu pukol kucing.*
person that hit cat
 ‘That person hit the cat.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:42:26.0-00:42:27.8)
- (601) *Itu kucing kena pukol.*
that cat PASS hit
 ‘That cat was hit.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:42:37.0-00:42:38.8)

Where alignment is concerned, BM is a nominative-accusative language. The subject of the intransitive verb is treated equivalently to the agent of the transitive verb. This is demonstrated by the following examples.

(602) *Kuching tu lompat.*
cat that jump
'That cat jumped.'

(Jane Quek oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:10:58.0-00:10:59.8)

(603) *Anjing ambek bakol.*
dog take basket
'The dog takes the basket.'

(Jane Quek oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:01:19.6-00:01:21.9)

The subject of the intransitive verb, as with *kuching tu* 'that cat' in (602), and the agent of the transitive verb, as with *anjing* 'dog' in (603), always precede the verb, whereas the object of the transitive verb always follows the verb.

It is important to note that there are no overt coding devices that reflect the nominative-accusative alignment in BM. There are also no other coding devices found on nouns or verbs to indicate alignment, agree, or cross-referencing. Relations are expressed mainly through word order.

5.6.3 Relative clauses

Relative clauses here refer to subordinate clauses that modify a noun phrase. Comrie (1989:143) states that relative clauses "consists necessarily of a head and a restricting clause" and that the prototypical relative clause is restrictive rather than non-restrictive. However, there is no basis for excluding non-restrictive relative clauses, so a more general view of relative clauses is adopted. More pertinently, Comrie (1989) argues for an accessibility hierarchy for relativization: Subject \supset Direct Object \supset Indirect Object \supset Possessor. That is, if a language can have relative clauses modifying noun phrases at a given position lower on the hierarchy, then it can relativize on all positions higher (to the left) on the hierarchy. The hierarchy of accessibility he identifies assumes that subjects are more easily relativized than direct objects, which are more easily relativized than indirect objects. Indirect objects are then in turn more easily relativized than possessors. In addition to the accessibility hierarchy, Comrie also highlights different strategies of relativization. The strategy that is most relevant to BM is the gap-type strategy. The

gap-type strategy is one that does not provide any overt indication of the role of the head within the relative clause.

In BM, there are two relative clause markers. One of the relative clause markers is *yang* (or sometimes *nang* in BM), which post-modifies the noun phrase head, as it does in Malay, the language from which it is derived. The other relative clause is *punya* (shortened forms: *mia*, *nia*), used in its basic form to indicate possession (see section 5.1.1). That *punya* is also used for relative clauses may be attributed to influence from substrate Hokkien relative clause marker *e*, which is also incidentally used to indicate possession (see section 7.4.4.2). Note that this does not imply that the possessive function of *punya* (described in section 5.1.1) comes from Hokkien, since *punya* is also to indicate possession in other varieties of Malay, including Jakarta Malay, Manado Malay and Moluccan Malay. That *punya* derives its relativizer function from Hokkien *ê* has been identified previously by Lee (2012) with regard to SBM and Lim (1988) with regard to MBM. Whereas *yang* relative clauses post-modifies noun heads, *punya* relative clauses pre-modify noun heads. The following are examples of both.

Both *yang* and *punya* can relativize subject, as demonstrated by examples (604) to (607), albeit *yang* relativizes postnominally while *punya* relativizes preminally. Example (21) is replicated here as (607).

(604) Subject

Ini budak [nang ter- teriak wolf].

this child REL ACD- call.out

‘This boy that accidentally cried wolf.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:34.9-00:03:37.3)

(605) *Ini sumua dia mia kawan [nang jaga kambing].*

this all 3.SG POSS friend REL guard sheep

‘These are all his friends who guard sheep.’

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:34.9-00:03:37.3)

(606) [*téngok saya punya*] *kuda punya kaki sudah patah*

look 1.SG REL horse Poss leg already snap

‘The horse that was looking at me had one of its legs already snapped.’

(Lee 2012)

- (607) [*Anak perempuan nia*] *satu*
 child female REL ONE
 ‘The one that is a girl.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:44.8-00:00:47.0)

Both *yang* and *punya* can also be used to relativize direct object, as shown in examples (608) to (611). Again, whereas Malay-derived *yang* modifies the main noun phrase postnominally, *punya*, which has derived its characteristics from Hokkien (see section 7.4.4.2, Lee 2012), modifies the main noun phrase preminally.

- (608) Direct object

Itu kerosi kechik [yang dia pechah]
 this chair small REL 3.SG break
 ‘That chair that she broke.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:06:03.1-00:06:05.6)

- (609) *Ini [yang dia kejar]*.
 this REL 3.SG chase
 ‘This that he chased.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:02:06.0-00:02:07.5)

- (610) *Gua nampak [orang tarék punya] chia*
 1.SG see person pull REL car
 ‘I saw the car (rickshaw) that the man pulled.’
 (Lee 2012)

- (611) [*Satu orang masak mia*].
 one person cook REL
 ‘(The one) that one person cooked.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:11:08.5-00:11:09.6)

Yang and *punya* may also relativize indirect object. Again, *yang* is used as a postnominal relativizer while *punya* is used as a prenominal relativizer.

- (612) Indirect object

Itu perempuan [yang mak kasi lauk], sudah balék
 that female REL mother give cook.food alreadyreturn
 ‘That female that mother gave cooked food to, has returned.’
 (Lee 2012)

- (613) *Saya nampak* [*saya kasi kuéh punya*] *orang*
 1.SG see 1.SG give cake REL person
 ‘I saw the man whom I gave the cake to.’
 (Lee 2012)

The only position that *yang* can relativize, but not *punya*, is the possessor. This may be due to the fact that *punya* is already used to indicate possession (see section 5.1.1), hence making it confusing if it is also used to relative possessor. The sentence in (616) is constructed and presented to speakers, who judged it to be ungrammatical. They state that it is not possible to construct a sentence using *punya* to express ‘that person whose friends hit me’ (Lee 2012).

- (614) Possessor
Saya kenal satu anjing [yang ada lima anak kechik].
 1.SG know.person one dog REL POSS five child small
 ‘I know a dog that has five small children.’
 (Lee 2012)

- (615) [*Yang ada umur sikit*]
 REL have age little
 ‘(Those that) have a little age /Those that are a little old.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:06:22.4-00:06:24.1)

- (616) **itu (orang) punya kawan-kawan pukol gua punya orang*
 that (person) possess friend-friend hit 1.SG REL person
 ‘that person whose friends hit me.’
 (Lee 2012)

Hence, the two relative clause markers, *yang* and *punya* have different patterns of use. While *yang* modifies a head noun phrase postnominally, *punya* modifies a head noun phrase prenominally. Speakers can use *yang* to relativize subject, direct object, indirect object, and possessor, while they can use *punya* to relative subject, direct object, and indirect object, but not possessor. Both relative clause structures use the gap-type strategy, where there is no overt indication of role of the head within the relative clause (Comrie 1989). More discussion on how *punya* derives its functions from a Hokkien relative clause marker is provided in section 7.4.4.2.

5.6.4 Complement clauses

In BM, complement clauses can be predicated by verb or adjective. In a sense, the complement clause is required to realise the meaning of the verb or adjective. The zero strategy is used for complement clauses, the main clause and the subordinate complement clause being juxtaposed against each other. Examples (501) and (527) are replicated here as (617) and (618) respectively.

- (617) *John rasa [dia pernah jumpa Mary].*
 think 3.SG ever meet
John thinks that he has met Mary before.
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:42:23.6-00:42:25.7)
- (618) *Gua rasa [pear tu tak boléh makan].*
1.SG **think** that NEG can eat
'I think that pear cannot be eaten.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:05:24.3-00:05:25.7)
- (619) *Gua rasa [ni orang ada peték buah pear].*
1.SG **think** this person PROG pluck CLF.fruit
'I think this person is plucking pears.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:00:28.9-00:00:31.2)
- (620) *Gua tau [lu mo datang].*
1.SG **know** 2.SG want come
I know you want to come.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:13:49.3-00:13:50.8)
- (621) *Mak dia harap [dia balék siang].*
mother 3.SG **hope** 3.SG return early
'His mother hopes he returns early.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-090, 00:05:46.0-00:05:58.6)
- (622) *Gua tak pikay [gua pernah jumpa lu].*
1.SG NEG **think** 1.SG ever meet 2.SG
'I do not think I have ever met you.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-096, 00:52:09.0-00:52:11.3)

- (623) *Tak* **sangka** [*boléh jumpa orang Peranakan*].
 NEG **expect** can meet person
 ‘(I) did not expect (I) could meet Peranakans.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:04:31.4-00:04:33.7)
- (624) *Gua* **tentu** [*jumpu lu dekat airport*].
 1.SG **definite** jumpa 2.SG PREP
 ‘I (am) definite (I would) meet you at the airport.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:13:34.1-00:13:37.9)

In instances where the subject is not expressed in the subordinate clause, the unexpressed subject always shares the same referent as the subject of the main clause, as with examples (623) and (624). These are instances of subject control, whereas the rest of the examples, (617) to (621), are instances of object control. Object control occurs when the object of the main verb is also the subject of the verb in the subordinate clause. It is also possible for the subject to be overtly expressed in the subordinate clause in cases of subject control (where subject of main clause has the same referent as subject of subordinate clause). This is demonstrated by example (622). However, this is less frequently observed, possibly due to the efficiency of dropping subjects and the preference for null-subject in BM.

5.6.5 Direct and indirect speech

In BM, the zero strategy is used for both direct and indirect speech. There is very little difference between direct and indirect speech, except that the pronoun used within the subordinate clause changes accordingly. In examples (625) to (627), the speakers are using direct speech, either reporting what they said themselves, or what others have said in its original form. In instances such as these, the subordinate clause usually involve first person or second person pronouns.

- (625) *Gua* **kata**, [*gua sudah lama kun tak makan*].
 1.SG **say**, 1.SG already long.time also NEG eat
 ‘I said, “I also have not eaten (this) for a long time.”’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:00:52.3-00:00:55.7)

- (626) *Mama mama sumua chakap*, [*amek gua mia aloji kulucair*].
 grandmother grandmother all **speak** take 1.SG POSS small.clock out
 ‘Grandmothers all said, “Take my small clock out.”
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:10:19.5-00:10:21.4)
- (627) *Tony gua chakap*, [*oh lu Catholic ada chutsi eh*]?⁷⁹
 1.SG **speak** 2.SG have rebirth Q
 ‘My Tony said, “Oh you Catholics have the concept of rebirth?”
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:06:05.7-00:06:10.8)

Where indirect speech is concerned, the zero strategy is also used, and the only difference between direct and indirect speech is that third person pronouns are usually used in indirect speech. Examples (628) and (629) show instances of indirect speech.

- (628) *Bill chakap*, [*dia sudah pi France*].
speak 3.SG already go
 ‘Bill said, he has already been to France.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:09:59.6-00:10:02.8)
- (629) *Dia kata*, [*lama dia tak makan*].
 3.SG **say** long.time 3.SG NEG eat
 ‘She said she has not eaten (this) for a long time.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:00:31.2-00:00:33.5)

5.6.6 Adverbial clauses

The adverbial clause functions as a modifier of the larger component clause. The adverbial clause is an adjunct and optional, the main clause being grammatical on its own in the absence of the adverbial clause. However, it is often valuable, offering additional information on time, location, and to some extent, manner. Examples of the different types of adverbial clauses are shown here.

- (630) Temporal
 [*Dulu minggu*], *Mary jatoh sakit*.
 before week fall sick
 ‘Last week Mary fell sick.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:04:39.2-00:04:43.8)

⁷⁹ Intended as a joke.

(631) Location

[*Se- lepas restaurant itu*], *lu nanti jumpa satu cake shop*.⁸⁰
one- after that 2.SG later see one
'Immediately after that restaurant, you will see a cake shop.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:13:51.3-00:13:57.8)

(632) Manner

Dia menangis sedéh sedéh [bila dia kuluar rumah].
3.SG cry sad sad when 3.SG go.out house
'She cried sadly when she left the house.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:04:52.8-00:04:45.7)

5.6.6.1 Temporal adverbial clauses

One of the most important functions of adverbial clauses is to indicate information regarding time, considering that BM has a more elaborate system of aspect than of tense. Tense is hence often implied through the use of adverbial clauses. The following examples show how past, present and future are implied through the use of adverbial clauses. These clauses usually precede the main clause, although they may also follow the main clause. The events in examples (633) to (635) are to be interpreted as having happened in the past, the events in examples (636) to (638) are current, and those in examples (639) to (641) will take place in the future.

(633) [*Empat puluh taon lepas*], *Tan nia chepuat tinggair dekat Katong*.
four ten years after REL family live PREP
'Forty years ago, the Tan family lived in Katong.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:00:38.5-00:00:54.5)

(634) [*Dua bulan lepas*], *Mary beli rumah dekat Katong*.
four month after buy house PREP
'Four months ago, Mary bought a house in Katong.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:03:33.6-00:03:35.4)

(635) *Ini cherita jadi dia [lima ratus taon lepas]*.
this story happen 3.SG five hundred year after
'This story happened five hundred years ago.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:01:42.4-00:01:46.5)

⁸⁰ It is possible for *se-* 'one' to be prefixed to *lepas* 'after', to denote immediacy, or that something took place one moment after another did. Similarly, it is possible for *se-* 'one' to be prefixed to *belom*, to denote that something took place immediately before something else.

- (636) [*Sekarang*] *tak tau chakap.*
 now NEG know speak
 ‘Now, (people) do not know how to speak.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:07:39.2-00:07:40.9)
- (637) [*Ini ari*], *Mary lu mo pegi mana?*
 this day 2.SG want go where
 ‘Today, where do you want to go, Mary?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:02:28.1-00:02:30.6)
- (638) [*Hari ini*], *Mary jumpa Jane dekat pasair*
 day this meet PREP market
 ‘Today, Mary is meeting Jane at the pasair.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:07:20.9-00:07:25.3)
- (639) [*Bésok*], *Mary pi pasair.*
 tomorrow go market
 ‘Tomorrow, Mary will go to the market.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:22:27.0-00:22:29.4)
- (640) [*Lain minggu*], *Mary pi Melaka.*
 another week go
 ‘Next week Mary will go to Malacca.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:25:03.5-00:25:06.3)
- (641) [*Lagik bulan*], *Mary pi Melaka.*
 More month go
 ‘A month later, Mary will go to Malacca.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:26:17.6-00:26:20.9)

It is interesting to note that the word *lepas* which is literally interpreted as an adverb meaning ‘after’, is used in conjunction with noun phrases that indicate a quantity of time, to denote the past and not the future. This is seen in examples (633) to (635). However, its counterpart *belom* ‘before’ is not used to indicate the future. Instead, the future is expressed when *lagik* ‘more’, precedes the relevant noun phrases, as with example (641). *Lain* ‘another’, as shown in example (640) has a more limited function of expressing next day, week, month or year. A list of words concerning time and date can be found in Appendix A.

5.6.6.2 Location adverbial clauses

Location can also be expressed by adverbial clauses. These also usually precede the main clause, but may also follow the main phrase. Example (631) is replicated here as example (642) and example (286) as (643).

- (642) [*Se-lepas restaurant itu*], *lu nanti jumpa satu cake shop*.
one-after that 2.SG later see one
'Immediately after that restaurant, you will see a cake shop.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:13:51.3-00:13:57.8)
- (643) [*Di se-belah*], *dia ketok apa?*
PREP one-side 3.SG knock what
'Beside, they are knocking what?'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:14:31.8-00:14:33.5)
- (644) [*Depan carpark*], *ada ini kebun bunga*.
front EXIST this garden flower
'In front of the carpark, there is a flower garden.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:07:32.8-00:07:38.7)
- (645) [*Di tepi library*] *ada satu kedai jual kuéh*.
PREP side EXIST one shop sell cake
'At the side of the library, there is a shop selling cake.'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:08:41.5-00:08:45.0)
- (646) *Anjing sudah mangun*, [*dekat dalam basket*].
dog alreadywake.up PREP inside
'The dog had already woken up, inside the basket.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:00:17.1-00:00:21.7)
- (647) *Itu budak main itu speedboat*, [*dekat itu laut sana*].
that child play that PREP that pond there
'That child is playing with that speedboat, at that pond over there.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-088, 00:05:22.1-00:05:27.0)

Note that in examples (646) and (647), the adverbial clause modify the entire clause. They are not adverbial phrases that modify only the verb phrase, as evidenced by the pause between both clauses.

5.6.6.3 Manner adverbial clauses

Adverbial clauses may also be used to modify other clauses to express a manner in which something is done. Example (482) is replicated as (650).

- (648) *Dia menangis sedéh sedéh [bila dia kuluir rumah].*
3.SG cry sad sad when 3.SG go.out house
'She cried sadly when she left the house.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:04:52.8-00:04:45.7)

- (649) [*Lepas makan pagi bahru*], *dia pi kreja.*
after eat morning just 3.SG go work
'After just eating breakfast, she goes to work.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:32:20.1-00:32:22.8)

- (650) [*Belom dia pulang*], *gua sudah dapat tau.*
before 3.SG return.home 1.SG alreadyget know
'Before she returned home, I already got to know (about something).'

Thus far, the examples discussed in these sections show that subordinate clauses may precede or follow the main clause, although there appears to be a stronger preference for subordinate clause to precede main clause. More adverbial clauses are discussed in the section on subordinating conjunctions (see section 5.6.8.2).

5.6.7 Conditionals

Kalu (alternative form: *kalo*) 'if' is used to connect clauses that semantically express conditionals or *if-then* constructions (Traugott et al. 1986), where in a simplified sense, a proposition is implied to be true if the conditions of another are fulfilled. *Kalu* can also be used for counterfactuals with imagined states, as with example (653). It is typically used at the beginning of the first clause that describes the antecedent. The second clause describes the consequence if the antecedent conditions are met. Although this is rare, *kalu* and its antecedent can also form the second clause, with the consequent clause fronted, as with example (654). Example (189) is replicated here as example (652).

- (651) [*Kalo gua tau lu mo datang*], *gua tentu jumpa lu dekat airport*
if 1.SG know 2.SG want come 1.SG definite meet 2.SG PREP
 ‘If I knew you were coming, I am definite (I would) meet you at the airport.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:13:32.1-00:13:37.9)
- (652) [*Kalu mo kechik-kan badan*], *kurang-kan roti*
If want small-TR body less -TR bread
 ‘If (you) want to make your body smaller, lessen (your intake of) bread.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:12:27.5-00:12:32.0)
- (653) [*Kalu chutsi*], *gua tak mo angkat pot*
if rebirth 1.SG NEG want carry
 ‘If there is rebirth I do not want to carry a pot (out into this world).’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:05:51.6-00:05:53.4)
- (654) *Gua bésok datang*, [*kalu Anne mo gua datang*].
 1.SG tomorrow come **if** want 1.SG come
 ‘I will come tomorrow if Anne wants me to come.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:14:56.9-00:15:03.4)

Aside from *kalu* ‘if’, it is also noted that *bila* ‘when’ can be used to expressed conditional relationships, as with (655), except that it is rarely used for conditionals (see section 5.6.8).

- (655) *Kita chakap Peranakan*, [*bila kita jumpa kita mia kawan*].
 1.PL speak Peranakan **when** 1.PL meet 1.PL REL friend
 ‘We speak Peranakan when we meet our friends.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:04:07.0-00:04:09.1)

5.6.8 Conjunctions

There are two types of conjunctions, as mentioned in section 4.7, these being coordinating conjunctions and subordinating conjunctions.

5.6.8.1 Coordinating conjunctions

Coordinating conjunctions are used to conjoin a phrase with another phrase, or a clause with another clause. The three coordinating conjunction markers in BM are *sama* ‘and’, *habis* ‘finish’, *ka* ‘or’, and *tapi* ‘but’.

In general, *sama* ‘and’ can be used to conjoin noun phrase to noun phrase or adjective phrase to adjective phrase. Examples (656) and (657) show how *sama* acts as a link between noun phrases, while (658) to (659) demonstrate how *sama* ‘and’ is used between adjective phrases.

- (656) [Mak] **sama** [ko] *pegi pasair*
 mother **and** paternal.aunt go market
 ‘Mother and paternal aunt went to market.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-110, 00:00:17.4-00:00:20.1)
- (657) *Mak masak [iték tim], [babi buah keluak] sama [sambal petay].*
 mother cook duck double.boil pig CLF.fruit Pangium.edule **and** chili.paste flat.bean
 ‘Mother cooked double boiled duck, pork cooked with Pangium edule, and chili with flat beans.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-110, 00:06:51.5-00:06:57.2)
- (658) *Dia kasi mak [marah] sama [sedéh].*
 3.SG cause mother angry **and** sad
 ‘She made mother angry and sad.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-110, 00:10:46.5-00:10:53.1)
- (659) *Ini ari [panas] sama [melekat]*
 this day hot **and** sticky
 ‘Today (is) hot and sticky.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-110, 00:13:24.5-00:13:26.3)

It is more rare to find *sama* ‘and’ being used to conjoin verb phrases. Speakers state that it is not possible to do two things at once, and prefer to use *habis* ‘finish’ to link two separate clauses together sequentially. Another version of the word *habis* is *abi*. That fact that it is not possible to express coordinated simultaneous action is also acknowledged by Lim (1988). He states that often, sequential events occur as two separate sentences in MBM. The following examples show how *habis* is used to link these separate clauses in SBM. Example (311) is replicated here as (660).

- (660) [Mak *pegi pasair*], **habis** *tu* [dia *pegi kopitiam*].
 mother go market finish that 3.SG go coffee.shop
 ‘Mother went to the market. After that she went to the coffeeshop.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-110, 00:01:23.4-00:01:29.3)
- (661) [Mak *bikin kuéh*], **habis** *tu* [dia *siram ayé pokok bunga*].
 mother make cake finish that 3.Sg flush water tree flower
 ‘Mother baked cake. After that she watered the plants.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-110, 00:03:41.6-00:03:57.2)

- (662) [*Jalan jalan*] **habi**, [*dudok kopi house*]
 walk walk **finish**, sit coffee
 ‘After taking a walk, (I) sit (in the) coffee house.’
 (Jane Quek oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:53.8-00:02:56.1)

Ka ‘or’ is used to connect any type of phrase to its equivalent phrase, or clause to another clause. Example (663) shows how it connects noun phrases, example (664), verb phrases, example (665), proper nouns, and example (666), preposition phrases. Example (312) is replicated as (663).

- (663) *Gua rasa ini [anak dia] ka [chuchu dia]*
 1.SG think this child 3.SG **or** grandchild 3.SG
 ‘I think this is his child or grandchild.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:13.0-00:02:15.0)

- (664) [*Mandi*] **ka** [*bikin apa*]?
 bathe **or** do what
 ‘Bathe or do what?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:07:54.8-00:07:56.5)

- (665) [*Terang bulan*] **ka** [*di tanjong katong*] *ka*.
 Bright moon **or** PREP cape uncertain or
 ‘Bright moon or at the uncertain cape or (both are names of songs).’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:08:27.8-00:08:31.0)

- (666) [*Dekat longkang*] **ka** [*dekat kopi tiam ka*],
 PREP drain **or** PREP coffee shop or,
 ‘At the drain or at the coffee shop.’

Tapi ‘but’ is used between two clauses, as with the following examples. Example (313) is replicated below as example (667).

- (667) [*Gua mia adék bikin kék*] **tapi** [*tak sedap*].
 1.SG REL sibling make cake **but** NEG delicious
 ‘My sister baked a cake but it was not delicious.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-117, 00:14:38.9-00:14:44.6)

- (668) [*Mary pi sekolah*] **tapi** [*dia tak suka bacha surat*].
 go school **but** 3.SG NEG like read letter
 ‘Mary goes to school but she doesn’t like to study.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-117, 00:16:26.0-00:16:29.5)

- (669) [*Bunga tu chanték*] **tapi** [*mak gua tak suka*].
 flower that beautiful **but** mother 1.SG NEG like
 ‘That flower is beautiful but my mother does not like (it).’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-117, 00:16:50.9-00:16:53.7)

5.6.8.2 Subordinating conjunctions

Subordinating conjunctions are used to conjoin subordinate clause to main clause. In these instances, the subordinate clause also usually are adverbial clauses, functioning as modifiers of the main clause. Two commonly used subordinating conjunctions are *pasal* (refined form: *pasair*) ‘because’, *bila* ‘when’, and *sunngu pun* (literally: really also) ‘although’.

Pasal, which literally means ‘matter, reason’, is grammaticalized, becoming ‘because’. *Pasal* ‘because’ hence is typically used between different clauses, where the first clause describes the outcome, and the second clause describes the reason behind the outcome. Example (670) to (672) show examples of how *pasal* is used, example (670) being a replication of (314). Note that the *pasal* construction may be the only subordinate clause that usually follows the main clause, rather than precede it (see section 5.6.6 for more examples of subordinate clauses preceding main clause).

- (670) *Tiga minggu macham dia balék [pasair dia tinggal Singapore]*
 three week like.that 3.SG return **because** 3.SG live Singapore
 ‘For about three weeks, she returns because she lives in Singapore.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:04:03.8-00:04:09.2)

- (671) [**Pasair** *orang tau*].
because person know
 ‘Because people know’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:09:41.1-00:09:43.0)

- (672) *Kita kena masak [pasair kita tau masak]*
 1.PL PASS cook **because** 1.PL know cook
 ‘We are made to cook because we know how to cook.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:09:49.5-00:09:54.0)

While most BM speakers use *pasal* for ‘because’, there also individual variation observed in the corpus, where the Malay form *kerana* ‘because’ is used.

- (673) *Gua angkat banyak tissue [kerana gua séisema]*
 1.SG carry many because 1.SG have.a.cold
 ‘I carry a lot of tissue because I have a cold.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:17:19.4-00:17:25.9)

Bila ‘when’, which is also used as an interrogative (see sections 4.1.3.4 and 5.6.10.1), can be used as a conjunction between a main clause and an adverbial clause. It is more common for the main clause to follow the adverbial clause, as with examples (674) and (675) rather than precede it, as with example (676). Recall that *bila* may also be used in conditionals, even though it is rarely used to do so (see section 5.6.7).

- (674) [*Bila bahru gua balék rumah*], *bahru gua dapat tau pasair tu accident*
when just 1.SG return home just 1.SG get know matter that
 ‘When I just returned home, I just got to know (about) that accident.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:02:22.8-00:02:26.3)

- (675) [*Bila gua senang*], *gua jalan jalan lah.*
when 1.SG free 1.SG walk walk EMP
 ‘When I (am) free, I take walks.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:01:50.1-00:01:52.2)

- (676) *Gua dapat tau pasair tu accident [bila gua jumpa dia].*
 1.SG get know matter that **when** 1.SG meet 3.SG
 ‘I got to know (about) that accident when I met her.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:00:30.5-00:00:36.0)

Sunggu pun (literally ‘really also’) functions as subordinating conjunction ‘although’. The subordinate clause it introduces usually precedes the main clause, although it is also acceptable for the subordinate clause to follow the main clause. Example (677) is a replication of (317).

- (677) [*Sunggu.pun dia tak standard*], *dia dapat ini kreja.*
although 3.SG NEG 3.SG get this work
 ‘Although he (is of) no standard, he got this work.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:20:53.4-00:20:57.2)

- (678) [*Sunggu.pun lu tak suka dia*], *lu mesti perama.*
 although 2.SG NEG like 3.SG 2.SG must polite
 ‘Although you do not like him, you must be polite.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:23:21.7-00:23:25.2)
- (679) *Dia boléh buang krejar*, [*sunggu.pun dia minum sikit.*]
 3.SG can throw work **although** 3.SG drinks little
 ‘He can lose his job although he drinks little.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-053, 00:22:38.1-00:22:42.7)

Similar to observations made in section 5.6.6, most of subordinate clauses usually precede the main clause, the only exception being sentences featuring *pasal* ‘because’.

5.6.9 Topicalization

Topicalization occurs when a particular constituent is fronted in a clause so that it is the most prominent component of the clause. Topicalization is commonly used in BM, and speakers topicalize the elements they believe to be most important. Lim (1988) goes as far as to suggest that BM has a basic sentence structure of topic followed by comment. He alludes to Li and Thompson (1976) who suggest that many Asian languages have Topic-Comment as their basic structure, instead of subject and predicate. BM, having derived its characteristics from substrate Hokkien, may have adopted its Topic-Comment structure. However, one hesitates to state that BM is more Topic-Comment than it is Subject-Predicate, simply because there are many instances of Subject-Predicate examples, such as with the recently discussed examples (677) to (679). Instead, it may be preferable to assert that topicalization often happens in BM.

The different constituents that can be fronted through topicalization include noun phrase object, adjectival phrase, adverbial phrase and also adverbial clause. Verb phrases usually do not have to be topicalized, since it is common for subjects to not be explicitly expressed, leaving verb phrases to already front the clause. Note that commas may be provided in some of the examples below to help in the parsing of these clauses, even though the speakers do not pause between the constituent being topicalized and the rest of the clause in some instances.

The following examples show topicalization of object noun phrase.

- (680) [*Kerosi*], *dudok, belom panas, sudah jalan*
 chair, sit not.yet hot already walk
 ‘That chair is not yet hot from sitting, (and you) already are going.’
 (Lilian, oai:scholarspace.manoa.hawaii.edu: NL1-079, 00:15:11.8-00:15:14.3)
- (681) [*Ikan kuning*], *tarok asam*
 fish yellow put tamarind
 ‘Yellow fish, put tamarind (on it).’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:07:29.4-00:07:32.2)
- (682) [*Teloh*], *goréng.*
 egg fry
 ‘The egg, fry (it).’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:07:44.8-00:07:47.5)

While noun phrase is topicalized in the preceding examples, the following ones show topicalization of adjectives. At first glance, it is not clear if what is being topicalized is the adjectival phrase or the verb phrase, considering there is very little use of the copula verb *ada* in BM. It is not possible to make presumptions about whether *ada* should also be fronted, since no such data exists. However, it is interesting to note again that there is no other instance of verb phrase fronting. In most instances, when verb phrases occur at the beginning of the clause, the clause carries no explicit subject, whereas in the examples that follow, the subject remains, albeit no longer in first position. Thus, these should be analysed as the topicalization of something other than the verb phrase, more likely, the adjectival phrase.

- (683) [*Samplang*], *dia.*
 promiscuous 3.SG
 ‘Promiscuous, he is.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:02:57.4-00:02:58.5)
- (684) [*Betol*], *ini mas*
 real this gold
 ‘Real, this gold is.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-100, 00:52:54.1-00:52:55.1)
- (685) [*Senyap*], *satu rumah*
 silent one house
 ‘Silent, the one house’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:03:00.6-00:03:03.5)

It is also possible for adverbial phrases and adverbial clauses to be fronted, although adverbial phrases modifying time, location and manner usually follow the verb phrase rather than follow it (see section 5.4). Instances of adverbial phrase topicalization are hence much more rare, as with example (686). On the other hand, it is very common for adverbial clauses to be topicalized, adverbial clauses having been observed to occur at both the beginning and end of sentences (see section 5.6.6). Examples (630) and (631) are replicated as examples (687) and (688).

(686) [*Betol lawa*], *dia pakay*
 real stylish 3.SG wear
 ‘Really stylish, he dresses.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-101, 00:06:22.8-00:06:27.3)

(687) [*Dulu minggu*], *Mary jatoh sakit.*
 before week fall sick
 ‘Last week Mary fell sick.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:04:39.2-00:04:43.8)

(688) [*Se-lepas restaurant itu*], *lu nanti jumpa satu cake shop.*
 one-after that 2.SG later see one
 ‘Immediately after that restaurant, you will see a cake shop.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:13:51.3-00:13:57.8)

5.6.10 Questions

The types of questions in BM can be divided into two kinds, content questions and tag questions.

5.6.10.1 Content questions

Following sections 4.1.3.4 and 4.4.2, the interrogative pronouns and adverbs of BM are summarized in the following table.

what	<i>apa</i>	Interrogative pronouns
who	<i>siapa</i>	
whose	<i>siapa punya/ mia/ nia</i>	
how many	<i>berapa</i>	
which	<i>mana</i>	
when	<i>bila</i>	Interrogative adverbs
where	<i>mana</i>	
why ‘what reason’	<i>apa pasal/ pasair/ sair</i>	
how ‘what like’	<i>apa macham/ cham</i>	
rhetorical ‘where EXIST’	<i>mana ada</i>	

Table 40: List of interrogative pronouns and adverbs in Baba Malay

The following are examples of how these interrogative pronouns and adverbs are used, and the content questions that they form. The interrogatives can either be fronted or remain *in situ*. Note that questions may occur with *ah* as an optional question particle.

Apa literally means ‘what’, and it can replace a noun phrase in a sentence. *Apa* can be fronted, as with example (689). Example (691) shows how questions with *apa* are usually answered.

Example (169) is replicated here as (690).

- (689) *Apa ini ah?*
what that Q
‘What is this?’
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:08:32.1-00:08:33.4)

- (690) *Ini apa?*
This **what**
‘This (is) what?’
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:23:11.1-00:23:12.5)

- (691) *Ini hospital*
 this
 ‘This (is) a hospital.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-103, 00:01:01.4-00:01:03.6)

Siapa ‘who’ is also an interrogative pronoun, and may replace noun phrases in utterances, as demonstrated by the following examples. Example (694) demonstrates how questions with *siapa* may be answered. Example (170) is replicated as (692).

- (692) ***Siapa itu?***
who that
 ‘Who (is) that?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:23:11.1-00:23:12.5)

- (693) *Mary ini siapa?*⁸¹
 this **who**
 ‘This Mary (is) who?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:23:11.1-00:23:12.5)

- (694) *Ini anak dia*
 this child 3.SG
 ‘This (is) his child.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:02:13.5-00:02:14.4)

The notion of ‘whose’ is expressed when *siapa* is used together with possessive *punya*. The short versions of *punya* are *mia* and *nia* (see section 5.1.1), although there are no instances of *siapa nia* recorded. Due to place assimilation to *p*, the onset of the second syllable in *siapa* ‘who’, *siapa punya* and *siapa mia* may be preferred over *siapa nia*. Inadvertently, the answers to these questions are utterances that replace *siapa* with a noun phrase, indicating who the possessor is, as with example (697). Example (171) is replicated here as (695).

- (695) ***Siapa punya bubor itu?***
who POSS porridge that
 ‘Whose porridge (is) that?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-032, 00:04:10.1-00:04:12.1)

⁸¹ Where *siapa* ‘who’ and the second person pronoun *lu* are concerned, it is generally more polite to ask “*Lu siapa?*” rather than “*siapa lu?*”

(696) *Ini rumah siapa mia?*
 this house **who** POSS
 ‘This house (is) whose?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-092, 00:02:01.1-00:02:03.9)

(697) *Gua mia ngkua mia.*
 1.SG POSS father-in-law POSS
 ‘My father-in-law’s.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-092, 00:03:03.0-00:03:06.3)

Berapa denotes ‘how many’ (recall that it can also be used to mean ‘some’, see section 5.1.2.4). *Berapa* can be used with noun classifiers and nouns, as with (698), or on its own, as with (699). The default meaning of the utterance concerns how much something costs, but context, noun classifiers and nouns can be used to tell what is being asked. Example (700) replicates (384), and shows how questions with *berapa* may be answered.

(698) *Berapa bijik itu lémo kat sana?*
how.many Clf.fruit that lemon PREP there
 ‘How many lemons (are) there?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-093, 00:27:14.8-00:27:17.8)

(699) *Itu lémo berapa?*
 that lemon **how.many**
 ‘That lemon (is) how much?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-093, 00:27:05.5-00:27:07.9)

(700) *Satu bijik timun*
 one CLF.small.round cucumber
 ‘One cucumber.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-149, 00:03:17.8-00:03:19.4)

Mana has two functions, one of which is ‘which’ and the other is ‘where’ (‘Where’ examples are discussed later). *Mana* is usually used when an interlocutor is presented more than one option to choose from, and asked to state a preference for a particular one. There is some individual variation. In addition to *mana*, as with example (701), a speaker also uses *apa* to indicate ‘which’, as demonstrated by example (703). Example (702) replicates (173), and example (704) shows how these questions can be answered.

- (701) *Mana satu lu suka?*
Which one 2.SG like
 ‘Which one do you like?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-089, 01:03:17.5-01:03:18.7)
- (702) *Mana baik, mana satu tak baik?*
Which good **which** one NEG good.
 ‘Which (is) good, which one (is) not good?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:23.2-00:01:26.3)
- (703) *Apa colour?*
what
 ‘What colour?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-089, 01:03:44.8-01:03:47.0)
- (704) *Gua suka merah.*
 1.SG like red
 ‘I like red.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-089, 01:04:17.2-01:04:19.4)

Whereas the above show uses of interrogative pronouns, the following examples demonstrate how interrogative adverbs are used. These may substitute adverbial clauses. In most instances, *bila* questions take the form of (705) rather than (706), even though both are well-formed to speakers. Example (638) is replicated as (707), and shows how *bila* interrogatives may be answered. Answers usually reflect some type of temporal information.

- (705) *Bila lu mo datang?*
When 2.SG want come
 ‘When (do) you want to come?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-093, 00:05:24.3-00:05:26.5)
- (706) *Dia nia sehjit bila?*
 3.SG POSS birthday **when**
 ‘Her birthday (is) when?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:43:10.9-00:43:13.6)
- (707) *hari ini, Mary jumpa Jane dekat pasair*
 day this meet PREP market
 ‘Today, Mary is meeting Jane at the pasair.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:07:20.9-00:07:25.3)

Other than being used to mean ‘which’, *mana* primarily functions to mean ‘where’. The following examples show how *mana* is used to mean ‘where’. Again, *mana* usually occurs fronted as with (708). Example (709) is a rare instance of *mana* occurring *in situ*. Example (235) is replicated here as example (708), and example (710) shows a typical answer to *mana* interrogatives. Answers to *mana* interrogatives in this sense usually indicate a particular location.

(708) *Mana pi si Mary?*
where go PERSON
 ‘Where goes Mary?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:08:42.2-00:08:43.5)

(709) *Lu pakay chanték mo pi mana?*
 2.SG wear beautiful want go **where**
 ‘You dress beautifully to go where?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:07:01.5-00:07:03.6)

(710) *Mary pi sekolah*
 go school
 ‘Mary goes to school.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-117, 00:16:26.0-00:16:27.5)

Apa pasal (refined form: *pasair*, short refined form *sair*) literally means ‘what reason’, and corresponds to ‘why’. Note that *apa pasal* is always fronted, and it has not been observed at the end of utterances. Example (713) show how these questions are usually answered, where *pasal/pasair/sair* occurs as a grammaticalized ‘because’ (see section 5.6.8.2).

(711) *Apa pasal lu marah?*
what reason 2.SG angry
 ‘Why (are) you angry?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-093, 00:23:43.8-00:23:45.1)

(712) *Apa sair si Mary bikin kuéh?*
what reason PERSON make cake
 ‘Why (is) Mary baking a cake?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:01:06.7-00:01:09.1)

- (713) *Pasal dia jahat, macham hantu.*
because 3.SG evil, like ghost
 ‘Because he (is) evil, like a ghost.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-093, 00:26:21.2-00:26:27.0)

Apa macham (short form: *cham*), literally ‘what like’, is used to indicate ‘how’ in BM. They usually occur fronted, as with example (714). The only instances where they do not occur right at the beginning of the utterance are instances wherein other elements have been fronted, as demonstrated by example (715). *Apa macham* questions can be answered by simple statements, as demonstrated by example (716).

- (714) *Apa macham pi bank?*
what like go
 ‘How (do I) go to the bank?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-103, 00:13:54.5-00:13:56.7)

- (715) *Bésok, si Mary, apa cham mo balék sekolah?*
 tomorrow Person **what like** want return school
 ‘How (does) Mary want to return to school tomorrow?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:41:02.5-00:41:04.8)

- (716) *Dia mo dudok bus.*
 3.SG want sit bus
 ‘She wants to take the bus.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:09:14.8-00:09:18.0)

Whereas the previous interrogatives are used for the purpose of seeking an answer, *mana ada* ‘where EXIST’ is used as a rhetorical question. *Mana ada* always occurs at the beginning of utterances. These are used in response to an interlocutor’s statement or question, to imply the negative. For example, (717) follows an interlocutor’s question about where a knife is, and it denies the existence of said knife. (718) on the other hand, is a response to an interlocutor’s statement, that a particular person is clever. It does not give one’s opinion directly, and is akin to asking, “How is he clever?” There is an inherent implication that the speaker does not share the same opinion as the interlocutor. Note that this structure is essentially derived from Hokkien, and will be revisited in section 7.4.4.2 Example (238) is replicated as (717).

(717) *Mana ada piso?*
where EXIST knife
 ‘Where is the knife (implying there is no knife)?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:21:02.8-00:21:05.6)

(718) *Mana ada panday?*
where EXIST clever
 ‘How is he clever (implying he is not)?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-093, 00:21:32.7-00:21:34.2)

In summary, whereas interrogative pronouns can occur fronted, or *in situ*, there appears to be a preference for interrogative adverbs to be fronted. This is in line with the nature of the adverbial clauses, which tend to be fronted in BM (see section 5.6.6).

5.6.10.2 Tag questions

Other than questions that feature interrogative pronouns and adverbs, tag questions are also commonly used in BM. Tag questions are essentially declaratives that have been converted into interrogatives with the tags attached at the end of the sentence. Tags are usually negative in nature, and common ones include adverb, *belum* ‘not yet’ (see section 5.2.5.1), noun negation marker, *bukan* (see section 5.1.4), and verb negation marker, *tak* (see section 5.2.8). Another tag is the *tak* VP *tak* option.

Belom ‘not yet’ is used at the end of sentences featuring *sudah* ‘already’ to ask if an event has already happened. Examples (719) to (721) show how the *belum* tag is used, and examples (722) and (723) demonstrate typical answers to questions with *belum*. The positive answer would comprise *sudah* ‘already’ VP, while the negative answer uses *belum* ‘not yet’ VP.

(719) *Itu kueh sudah jadi belum?*
 that cake alreadybecome **not.yet**
 ‘Has that cake formed yet?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-098, 00:19:55.5-00:19:57.7)

(720) *Itu mia lauk sudah siap belum?*
 that Rel cooked.food alreadyprepared **not.yet**
 ‘Has that cook food been prepared yet?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-099, 00:18:12.8-00:18:16.5)

- (721) *Lu sudah siap belum?*
 2.SG alreadyprepared not.yet
 ‘Are you prepared yet?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-100, 00:10:36.4-00:10:39.0)
- (722) *Itu mia lauk sumua sudah siap.*
 that Rel food all alreadyprepared
 ‘That cooked food is all prepared already.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-099, 00:18:51.2-00:18:53.9)
- (723) *Mary belum siap.*
 not.yet prepared
 ‘Mary is not prepared yet.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-099, 00:20:11.5-00:20:13.9)

Another negative tag in BM is *bukan* (short version: *kan*⁸²). The following examples show the *bukan* tag is used. Earlier, the use of *bukan* to negate noun phrases was noted (see section 5.1.4). In these instances, *bukan* is used to negate the clause. Essentially, the speaker has a view and seeks confirmation. Typical replies to these questions will comprise *ya* ‘yes’, or *bukan* ‘no’. These are demonstrated by examples (727) and (728).

- (724) *Dia ada beli apple bukan?*
 3.SG PFV buy no
 ‘He bought apples, no?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:40:03.6-00:40:07.2)
- (725) *Dia pi sekolah bukan?*
 3.SG go school no
 ‘He goes to school, no?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:40:24.4-00:40:27.4)
- (726) *Lu mo bukak pintu kan?*
 2.SG want open door no
 ‘You want to open the door, no?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-044, 00:20:58.3-00:21:00.1)
- (727) *Ya lah. Lama tak jalan sini.*
 yes EMP long.time NEG walk here
 ‘Yes. (I have) not walked here for a long time.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:05.1-00:00:08.3)

⁸² This is not to be mistaken with transitive marker *-kan*.

(728) *Bukan. Terima.*

no accept

‘No. Accept (it).’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:17:57.8-00:17:59.7)

While *bukan* negates the entire clause (and noun phrase), *tak* has been noted to negate verb phrase (see section 5.2.8). When *tak* is used as a tag marker, the speaker seeks information with no particular presupposed stance. Compare (724) to (729), and (725) to (730). In (724), the speaker is questioning the *notion* of agent having bought apples, thinking that the agent has most probably bought apples, whereas in (729), the speaker is questioning whether the agent bought apples or not, without presuming that the agent has bought apples. Similarly, in (725), the speaker questions the *notion* of whether the agent goes to school, presuming that the agent goes to school, whereas in (730), the speaker questions whether the agent goes to school, without any strong presumptions about whether the agent goes to school. Example (427) is replicated as (733).

(729) *Mary ada beli apple tak?*

PFV buy NEG

‘Did Mary buy the apples?’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:39:15.5-00:39:17.9)

(730) *Dia pi sekolah tak?*

3.SG go school NEG

‘Does he go to school?’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:40:17.5-00:40:18.6)

(731) *Lu jaga chuchu tak?*

2.SG take.care grandchildren NEG

‘Do you take care of grandchildren?’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:50.8-00:00:52.5)

(732) *Lu boléh tunjok -kan jalan tak?*

2.SG can show -TR walk NEG

‘Can you show the way?’

(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-083, 00:19:58.2-00:20:01.9)

- (733) *Rumah gua ada chanték tak?*
 house 1.SG COP beautiful NEG
 ‘My house is beautiful or not?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:02:41.9-00:02:43.7)

Other than the tag *tak*, it is also possible to form questions in the VP *tak* VP form. These have the same function as the basic *tak* tag. These can be thought of as more fully developed versions of the basic *tak* questions. Examples (734) to (738) show how the VP *tak* VP form is used. Note that (737) and (738) appear as AP *tak* AP on the surface, but they are versions of VP where the copula is not expressed, unlike in (734), where the copula is fully expressed. Example (734) replicates (426).

- (734) *Ada baik tak -a*
 COP good NEG -COP
 ‘Are (you) well or not?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:09.2-00:00:11.2)

- (735) *Lu perchaya tak perchaya*
 2.SG trust NEG trust
 ‘Do you trust or not trust (this news)?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-125, 00:02:09.2-00:02:11.8)

- (736) *Lu setuju tak setuju*
 2.SG agree NEG agree
 ‘Do you agree or not agree?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-125, 00:02:14.3-00:02:16.6)

- (737) *Chanték tak chanték*
 beautiful NEG beautiful
 ‘Beautiful or not beautiful?’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-091, 00:03:41.3-00:03:43.9)

- (738) *Lain minggu, Johntak tentu senang tak senang*
 another week NEG certain free NEG free
 ‘Next week, John is not sure if he is free or not free.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-091, 00:27:52.6-00:27:57.9)

The following examples demonstrate how *tak* and VP *tak* VP questions may be answered. Example (739) is a response to (737), while example (740) answers the question in example (731). Note that while *ya* is still used to mean ‘yes’, *bukan* ‘no’ is not used as a response to these questions (unlike the *bukan* questions). Instead, *tak ada* (short: *tak a*) ‘NEG EXIST’ is used to indicate a negative response. It is possible also for answers to be given without *ya* or *tak ada*. Declarative statements like those in examples (741) and (742) are also common responses to *tak* questions.

(739) *Ya. kasot ini chanték.*
 yes shoe this beautiful
 ‘Yes. These shoes (are) beautiful.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-091, 00:04:06.2-00:04:10.1)

(740) *Tak a lah. Tak gitu lah.*
 NEG EXIST EMP NEG like.that EMP
 ‘No. (It is) not like that.’
 (Jane Quek oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:00:52.5-00:00:53.8)

(741) *Suka lah.*
 like EMP
 ‘(I) like (response to whether or not the agent likes to play mahjong).’
 (Jane Quek oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:15:07.7-00:15:09.4)

(742) *Tak boléh.*
 Neg can
 ‘Cannot (response to whether something can or cannot be done).’
 (Jane Quek oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:05:14.9-00:05:15.4)

Besides the above two main types of tag questions, other tags also exist, such as the ones below.

(743) *Asam gugol mia kulit ya?*
 tamarind dried.fruit POSS skin, yes
 ‘The dried tamarind’s skin, yes?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:11:51.6-00:11:54.4)

(744) *Gua mia favourite ah?*
 1.SG POSS Q
 ‘My favourite?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:10:37.2-00:10:39.5)

(745) *oh lu Catholic ada chutsi eh?*⁸³
 2.SG have rebirth Q
 ‘Oh you Catholics have the concept of rebirth?’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:06:08.2-00:06:10.8)

(746) *Gua lapair tau?*
 1.Sg hungry know
 ‘I (am) hungry, (you) know?’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-002, 00:17:16.3-00:17:17.3)

The responses to these questions depend on whether the speaker has an intended answer (in which case the appropriate responses are *ya* ‘yes’ and *bukan* ‘no’), or whether the speaker is genuinely enquiring for information (in which case the appropriate responses are *ya* ‘yes’ and *tak* VP). In the above questions, only (743) can be answered with *bukan*, since it is the notion of the clause that is being questioned.

It is important to recall that both content questions and tag questions differ in intonation (see section 3.6.3). Whereas content questions are accompanied by a rise-fall when the interrogative occurs in utterance-initial position, and a rise at the end of question when the interrogative occurs in utterance-final position, tags always occur at the end of the utterance, hence tag questions are always accompanied by a rise at the end of the utterance.

5.6.11 Imperatives

Verb phrases function as imperatives on their own, except for negative imperatives that are essentially verb phrases made up of *jangan* ‘do not’ and the main verb phrase (see section 5.2.8). The following are examples of non-negative imperatives.

(747) *Pi buang ayé kechik.*
 Go throw water small
 ‘Go urinate.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-052, 00:18:51.7-00:18:53.9)

(748) *Pegi lah.*
 Go EMP
 ‘Go.’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:50:25.1-00:50:26.4)

⁸³ Intended as a joke.

(749) *Pegi buang ayé baik-baik.*
go throw water good-good
'Go urinate well.'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-037, 00:52:15.4-00:52:17.0)

(750) *Belajar chakap.*
learn speak
'Learn to speak'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:32:04.1-00:32:05.8)

(751) *Tutup pintu.*
close door
'Close the door.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:23:01.3-00:23:02.8)

By and large, it is considered impolite to use these imperatives as well as the negative imperatives discussed in section 5.2.8 with speakers who have more seniority than oneself. Requests have to be couched in questions together with the appropriate terms of address. For more information on interrogatives, please refer to section 5.6.10.

6. Differences between Singapore Baba Malay and Malacca Baba Malay

The label ‘Baba Malay’ is usually applied by researchers to the language of the Peranakans, both in Singapore and in Malacca. Thus far, research on BM has been restricted to condensed grammatical descriptions or sketches, since the aim of the researchers was mostly to determine how much of BM is due to influence from contact with speakers of Hokkien. Pakir (1986) and Lee (1999) explicitly state that their fieldwork is conducted in Singapore, Sharif (1981) and Lim (1981, 1988) state that their fieldwork took place in Malacca, and Shih (2009) expresses that she conducted interviews in both Singapore and Malacca. However, little has been said about whether the variety in Singapore is similar to or different from the variety that is spoken in Malacca. Thus far, no researcher has stated that these varieties should be different, with the exception of Pakir (1986) who suggests towards the end of her dissertation that comparisons between the two dialects could be carried out in the future.

While the focus of this dissertation has been mainly the grammar of SBM, this chapter aims to point out some differences and also selected similarities between SBM and MBM where they exist. Primary fieldwork was carried out with a key Malacca consultant and his mother. In addition, the researcher’s granduncle, an MBM speaker, was also consulted. These amount to twelve hours in audio recordings which are also archived in Kaipuleohone. The works of Sharif (1981), Lim (1981) and Lim (1988) are used to supplement this data, since both researchers carried out their fieldwork in Malacca.

6.1 Phonetic and phonological differences

Phonetically, MBM is different from SBM in that MBM speakers do not produce the vowel /ɔ/, whereas SBM speakers do. Phonologically, where words end with /e/ and /o/ in SBM, they may end with /aj/ and /aw/ in MBM, making MBM more similar to Malay.

6.1.1 Vowel /ɔ/

The main phonetic difference between SBM and MBM is the lack of vowel /ɔ/. The vowels of a proficient male MBM speaker in his sixties, are measured for their first and second formants. The tokens are extracted from twelve hours’ worth of interview sessions, from sections where

speech is naturally-occurring. The same methodology adopted for measuring vowels of SBM (see section 3.7) is used. The following vowel chart shows the vowel space of this MBM speaker, in which /ɔ/ is clearly missing.

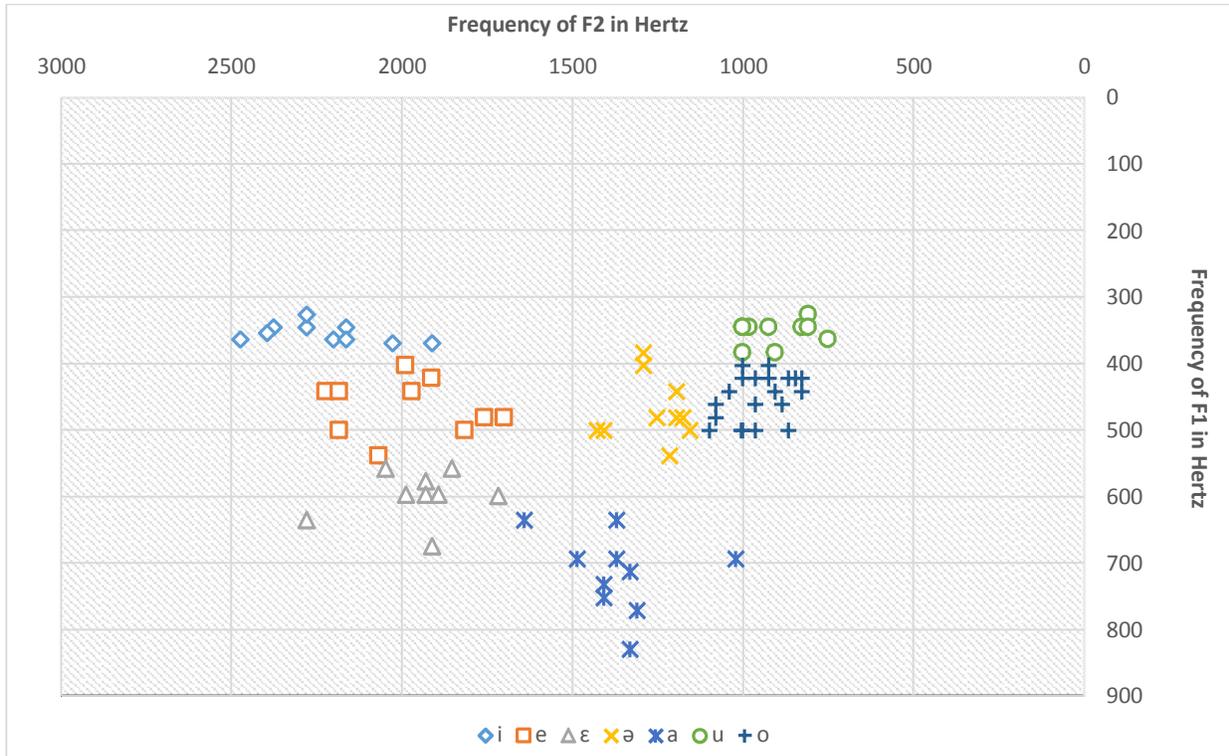


Figure 32: Vowel space of a Malacca Baba Malay speaker

This can be compared with the vowel space of a proficient SBM speaker, also a male in his sixties. The vowel space of the SBM speaker, featured earlier as figure 22, is replicated here as figure 33. The SBM speaker is observed to produce the vowel /ɔ/, while the MBM speaker does not do so.

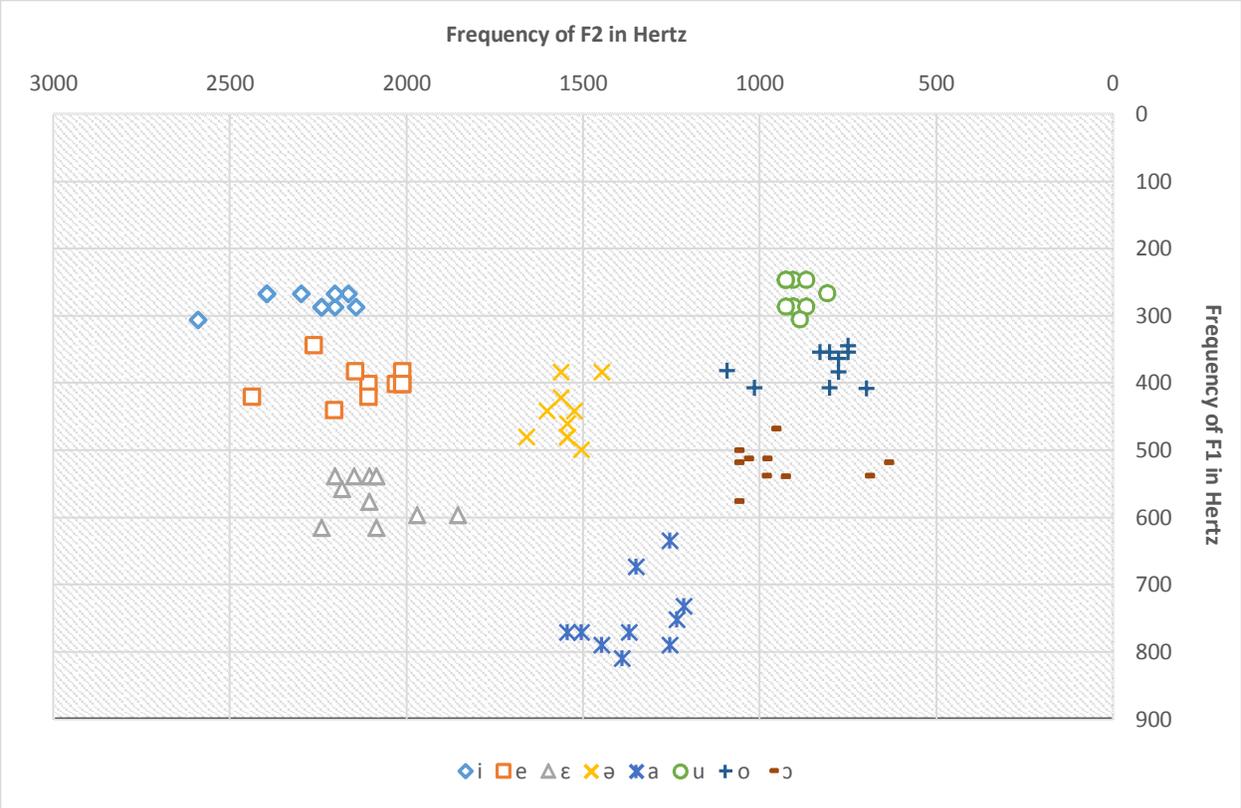


Figure 33: Vowel space of a Singapore Baba Malay speaker

That MBM has vowels /i/, /e/, /ɛ/, /ə/, /a/, /u/, and /o/ has also been observed by Lim (1981,1998). Note that Sharif (1981) only observed /i/, /e/, /ə/, /a/, /u/, and /o/. The findings of this dissertation are aligned with Lim (1981, 1988) and not Sharif (1981), since Sharif (1981) does not state that /ɛ/ occurs in MBM. Note that it is only possible to accurately identify these sounds, and recognize phonetic variation between the speakers with acoustic representations (as compared to earlier work and other traditional grammars that do not utilise acoustic information). The phoneme /ə/ is further back for the MBM speaker, and there is more variation in /a/ as well.

The vowels of MBM can thus be represented by vowel chart in table 41.

	front	central	back
	non-rounded	non-rounded	rounded
close	i		u
close-mid	e	ə	o
open-mid	ɛ		
open		a	

Table 41: Vowel chart of Malacca Baba Malay

Where consonants are concerned, MBM has the exact same consonant inventory as SBM. The consonants of MBM can be represented by the following chart. Essentially, this chart has the same components as the consonant chart of SBM (see section 3.1.1).

		Labial	Alveolar	Post- alveolar	Velar	Glottal
plosives	voiceless	p	t		k	ʔ
	voiced	b	d		g	
affricates	voiceless			tʃ		
	voiced			dʒ		
fricatives	voiceless		s			h
nasals		m	n	ɲ	ŋ	
lateral			l			
flap			ɾ			

glides: w, j

Table 42: Consonant chart of Malacca Baba Malay

It is interesting to note that these findings mostly accord with those with Lim (1988)'s work on MBM, except that the glottal plosive /ʔ/ is missing from his consonant chart. It may not have been his intention to leave it out, especially considering that he states that the phonological

system of BM is “completely congruent with that of Malay” (Lim 1988:14), by which he means the standard Bahasa Melayu. He notes the presence of glottal plosive /ʔ/ in Malay. More information on how the sound systems of Hokkien and Malay have contributed to BM can be found in section 7.4.1. Comparison with Sharif (1981)’s work on MBM shows that he treats liquid /r/ as velar fricative /ɣ/. This differs in general from anyone’s work on BM (see Lim 1981, 1988 on MBM, and Pakir 1986 on SBM).

6.1.2 /e/, /o/ versus /aj/, /aw/

As discussed in section 3.3, words that end with /aj/ and /aw/ in Malay end with /e/ and /o/ respectively in SBM. MBM is similar to both Malay and SBM, in that both /aj/ and /aw/, as well as /e/ and /o/ endings can be observed. Again, it is not accurate to generate [e] and [o] from a hypothetical monophthongization of /aj/ and /aw/ since /aj/ and /aw/ appear to be a later development due to the influence of Malay. Examples (752) and (753) demonstrate the use of /e/ word-finally, whereas examples (754) and (755) show the use of /aj/ word-finally. Comparing MBM directly with Malay, Sharif (1981) notes that [e] replaces [aj] word-finally in MBM (he represents [aj] as [ai]),⁸⁴ he provides example (755) from his transcripts as a counterpoint, stating that this may be due to “a considerable amount of exposure to the Malay society” (Sharif 1981:106).

(752) [dja **pake** matʃam kawboj]
 3.SG **wear** like cowboy
 ‘He wore (it) like a cowboy’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:01:12.3-00:01:14.3)

(753) [kaki taʔ **sampe**]
 leg NEG **reach**
 ‘Legs do not reach’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:02:52.6-00:02:54.0)

(754) [təpoʔ **pandaj** **pandaj**]
 clap **clever** **clever**
 ‘Clap cleverly’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-024, 00:15:38.3-00:15:39.4)

⁸⁴ Sharif (1981) represents these as phones instead of phonemes.

- (755) [**pakaj**]
use
 ‘Use’
 (Sharif 1981: 106)

Similarly, word-final /aw/ and /o/ both occur in MBM, whereas word-final /aw/ occurs in Malay, but not in SBM, and word-final /o/ occurs in SBM but not in Malay. In the current corpus, *piso* ‘knife’ is the only item that shows up with word-final /o/, as with example (756). Examples (757) to (761) demonstrate the use of word-final /aw/. Examples (760) and (761) are derived from Sharif (1981). Note that Sharif (1981) makes no observations regarding /aw/ and /o/.

- (756) [**piso** mana]
knife where
 ‘Where (is) the knife?’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-048, 00:14:07.8-00:14:08.6)

- (757) [**kalaw** bole]
if can
 ‘If possible’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:02:09.2-00:02:10.2)

- (758) [oraŋ **kalaw** tʃakap terus teraŋ]
 people **if** speak straight bright
 ‘People, if (they) speak straight-forwardly’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:04:13.7-00:04:16.5)

- (759) [**maw** aŋkat tu]
want carry that
 ‘Want (to) carry that’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:02:33.1-00:02:33.5)

- (760) [**kalaw**]
if
 ‘If’
 (Sharif 1981: 26)

- (761) [mesti **maw**]
 must **want**
 ‘must’
 (Sharif 1981: 20)

The best explanation again, may be that Peranakans in Malacca are highly exposed to standard Malay, since word-final /aj/ and /aw/ occur in standard Malay, but not word-final /e/ and /o/. Not only are Peranakans highly exposed to standard Malay speakers in Malacca, many of them also undergo a standard Malay medium education in Malaysia. This may be the basis for why MBM appears to be more Malay than SBM is.

In fact, Sharif (1981)'s transcripts based on three interviews carried out in Malacca, show instances of further convergence towards standard Malay. Words that phonologically end with /a/ are phonetically transcribed with [ə] endings. For example, possessive /pupa/ is transcribed as [pupə], and /dʒumpa/ for 'meet' is transcribed as [dʒumpə]. This word-final alternation is characteristic of standard Malay in both Singapore and Malacca (Omar 1977), but not noted in any work on BM thus far. This alternation may have occurred due to the fact that the interviewers in Sharif (1981)'s transcripts used Standard Malay to elicit BM responses.⁸⁵ This alternation also does not show up in the current corpuses of both SBM and MBM.

6.2 Morphological difference

In terms of morphology, MBM differs from SBM only slightly. At first glance, this variety appears to take on more affixes than SBM.

6.2.1 *Ke- -an* nominalizing circumfix

In other varieties of Malay such as Bahasa Melayu and Bahasa Indonesia, the circumfix *ke-* and *-an* may be added to adjectives and verbs to form nouns of an abstract nature. In the current SBM corpus, there is no occurrence of this circumfix, whereas in the MBM corpus, the circumfix occurs with the word *banyak* 'many' so that it becomes a noun. This is also observed to occur in Sharif (1981)'s transcripts of MBM. Also, while 'many' in SBM is *manyak*, the word is *banyak* in these observed MBM instances.

⁸⁵ While this is not outrightly stated in his thesis, this information can be gleaned from the transcripts of Sharif (1981).

- (762) *Ke-banyak-an* *chakapsaya*
 NMZ-many-NMZ speak 1.SG
 ‘Many say *saya* (for I).’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:01:06.1-00:01:08.2)
- (763) *Anjingitu* *ke-punya-an* *Mary.*
 dog that NMZ-possess-NMZ
 ‘That dog (is) Mary’s possession.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-010, 00:36:11.7-00:36:15.5)
- (764) *Ke-banyak-an* *orang tongsan*
 NMZ-many-NMZ people China
 ‘Many (were) Chinese’
 (Sharif 1981: 41)⁸⁶

However, note that *ke-* may also be omitted, as with example (765) and (766). In fact, *-an* in (765) appears as a regular nominalizer that may also be used for non-abstract items, as with the *chakap -an* ‘speech’ in (766). The general *-an* nominalizer is also found in SBM (see section 4.1.1.1).

- (765) *Orang bahasa* *China banyak-an*
 person language China many-NMZ
 ‘Many of the Chinese speakers’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:04:18.2-00:04:20.4)
- (766) *Chakap-an* *Baba*
 speak-NMZ
 ‘The speech of the Babas.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:04:10.3-00:04:11.4)

6.2.2 Other affixes

Similar to SBM, affixes in MBM may not be meaningful. For example, in (768) and (770), Malay middle voice marker makes no difference to the interpretation of the word it is attached to.

⁸⁶ Sharif’s transcripts (1981) are transcribed phonetically, but they are presented orthographically here for ease of understanding.

- (767) *bawak datang kuéh itu*
carry come cake that
 ‘Carry that cake here.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-013, 00:15:52.9-00:15:55.1)
- (768) *Dia pegi jalan jalan berbawak satu jarring*
 3.SG go walk walk **carry** one net
 ‘He went walking carrying one net.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-047, 00:00:54.8-00:00:56.9)
- (769) *Dia gonchang botol itu.*
 3.SG **shake** bottle that
 ‘He shook that bottle.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-025, 00:09:39.8-00:09:42.1)
- (770) *Bergonchang*
Shake
 ‘Shake (something).’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-025, 00:09:36.9-00:09:38.7)

In addition to nominalizer *-an*, other commonly used productive affixes in MBM are the ones that are also used in SBM, these being transitive marker *-kan* and accidental and movement prefix *ter-* (see sections 4.2.1.1). Examples (771) to (773) show usage of the transitive marker, examples (774) to (776) demonstrate how the accidental prefix is used, and examples (777) and (778) show usage of the movement prefix.

- (771) *pechah-kan*
break-TR
 ‘break (something).’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-024, 01:12:17.0-01:12:18.6)
- (772) *Lu jatoh-kan barang.*
 2.SG **fall-TR** thing
 ‘You dropped something.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-025, 00:05:13.8-00:05:15.9)
- (773) *Sandah-kan itu tangga.*
lean-TR that ladder
 ‘Lean that ladder (against something).’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-025, 00:03:22.2-00:03:26.5)

- (774) *Dia ter-tendang depan dia*
 3.SG ACD- kick front 3.SG
 ‘He accidentally kicked in front of him.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-047, 00:01:23.7-00:01:25.7)
- (775) *ter-langgair*
 ACD-crash
 ‘Accidentally crash.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:03:33.0-00:03:33.6)
- (776) *Dia mia kopiah ter-tinggal*
 3.SG POSS hat ACD-stay
 ‘His hat accidentally left behind.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:04:39.0-00:04:40.9)
- (777) *ter-pusing*
 MVT-whirl
 ‘Whirl around.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:03:35.0-00:03:38.5)
- (778) *keréta tu ter-balék.*
 car that MVT-return
 ‘That car turned upside down.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:04:06.1-00:04:10.8)

Note that while derivational morphology is limited in both MBM and SBM, MBM speakers are mostly aware of other derivational affixes used in Standard Malay and are able to use these. For example, while MBM speakers interviewed state that there is no person prefix in BM, this prefix is observed in Sharif’s data. Again, this may be due to the fact that the speaker is being interviewed by a speaker of standard Malay. Also note that MBM speakers have knowledge of standard Malay through education, as well as through contact with other speakers.

- (779) *pe-kerja dia*
 PERSON-work 3.Sg
 ‘His worker.’
 (Sharif 1981: 45)

In general, while MBM speakers may be aware of other derivational affixes in Malay, the ones that are considered to be BM to them are the abstract nominalizer circumfix *ke- -an*, general nominalizer *-an*, transitive marker *-kan*, and accidental prefix and movement prefix *ter-*.

6.3 Syntactic differences

In terms of syntax, several differences can be noted, including word order within the noun phrase, and lexical choice of function words.

6.3.1 Noun phrase

With regard to the noun phrase, MBM and SBM are mostly similar. The only major difference concerning the noun phrase is that demonstrative determiners follow the main noun phrase in current day MBM, whereas these demonstrative determiners both precede and follow the noun phrase in SBM. This appears to be a newer development considering that the older transcripts of Lim (1981, 1988) and Sharif (1981) show both word orders. There are also a couple of clarifications that have to be made where personal pronouns and relative clauses are concerned. While previous research only highlights one first person pronoun, the current corpus as well as the transcripts of Lim (1981, 1988) and Sharif (1981) show that, akin to SBM, MBM has two first person pronouns. In addition, while Lim (1981, 1988) only indicates a prenominal relativizer, and Sharif (1981), a postnominal relativizer, both are actually used in MBM, just as with SBM.

6.3.1.1 Demonstrative determiners

While demonstrative determiners are observed to both precede and follow the main noun phrase in SBM, the MBM data recorded show that determiners follow the main noun phrase rather than precede it. Examples are as follows.

(780) *Barang ini banyak.*
thing this many
'These things (are) many.'
(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:07:00.0-00:07:01.3)

(781) *Peté^k buah tu*
pluck fruit that
'Pluck that fruit.'
(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:00:36.5-00:00:38.0)

- (782) *Dia pakay kopiah itu*
 3.SG wear hat **that**
 ‘He wears that hat.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:02:49.9-00:02:50.8)
- (783) *Tangkap kodok tu*
 capture frog that
 ‘Captured that frog.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:01:19.2-00:01:20.2)
- (784) *Anjingitu pun jatoh*
 dog **that** also fall
 ‘That dog also fell.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:01:42.6-00:01:44.2)
- (785) *Kawin itu.*
 marriage **that**
 ‘That marriage.’
 (Uncle Chan, oai:scholarspace.manoa.hawaii.edu: NL1-008, 00:04:54.0-00:04:56.1)

The finding that determiner follows noun in MBM, differs from that of Lim (1991, 1998) and Sharif (1988). It is interesting to note that while they state that determiner precedes noun, their transcripts show that both word orders are common.

- (786) *Ini orang pegi mintak.*
this person go ask
 ‘This person goes and asks.’
 (Lim 1988: 53)
- (787) *Itu orang*
that person
 ‘That person.’
 (Sharif 1981: 17)
- (788) *Itu kuéh*
that cake
 ‘That cake.’
 (Sharif 1981: 22)
- (789) *Anak itu dua taon*
 child **that** two year
 ‘That child (is) two years old.’
 (Lim 1988: 54)

(790) *Malam ini mesti mau sambod*
 tonight **this** must want pray
 ‘Tonight (we) must pray.’
 (Sharif 1981: 18)

(791) *Macham kuéh pau tu*
 like cake bun that
 ‘Like that cake bun (a type of food)
 (Sharif 1981: 21)

A feasible explanation is that MBM has lost its determiner noun word order within the last thirty years or so, noting that that Sharif’s and Lim’s fieldwork appear to have been carried out before 1981. This may be due to the fact that BM speakers come into contact with Malay speakers extensively in Malacca. Residential areas are mixed, and it is also no longer the case that most Chinese settlers would be located in the area around the Malacca River, while the local Malay population live mostly in the rural areas (see Lim 1988:16). In contrast, Malay is less dominant in Singapore, and the determiner noun word order may be maintained due to the fact that the determiner noun word order occurs in English, which is the dominant language in Singapore (see section 1.2.2 for more information on the different language environment of Malacca and Singapore).

6.3.1.2 Personal pronouns

It should be clarified that there is no difference between the personal pronominal system of MBM and SBM. Although Lim (1981, 1988) and Sharif (1981) indicate *gua* as the first person pronoun, their transcripts show that both coarse form *gua* and refined form *saya* are commonly used. Example (762) is replicated here as example (792).

(792) *Ke-banyak-an chakapsaya*
 NMZ-many-NMZ speak **1.SG**
 ‘Many say *saya* (for I).’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:01:06.1-00:01:08.2)

- (793) *Saya punya mak*
1.SG POSS mother
 ‘My mother’
 (Lim 1988: 53)
- (794) *Saya niat*
1.SG intend
 ‘I intend.’
 (Sharif 1981: 19)
- (795) *Gua*
1.SG
 ‘I’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:01:01.5-00:01:03.0)
- (796) *Gua ingat*
1.SG think
 ‘I think’
 (Lim 1988: 55)
- (797) *Gua tak berapa tau ni*
1.SG NEG some know this
 ‘I (do) not know much (of) this.’
 (Sharif 1981: 38)

The system of personal pronouns in MBM is thus the same as that of SBM, and can be replicated as such. Table 29 is replicated here as table 43.

1.SG (refined) (coarse)	<i>saya</i> <i>gua</i>
1.PL	<i>kita</i>
2.SG	<i>lu</i>
2.PL	<i>lu-orang/ lu</i>
3.SG	<i>dia</i>
3.PL	<i>dia-orang/ dia</i>

Table 43: Personal pronouns in Baba Malay

6.3.1.3 Relative clauses

While Lim (1988) suggests that *punya* is used as a pre-nominal relativizer, Sharif (1981) states that post-nominal relativizer *yang* is used. It should be clarified that similar to SBM, both *punya* and *yang* are used for the purpose of relativization (see section 5.6.3). Again, while *yang* can be used to relativize subject, direct object, indirect object, and possessor, *punya* can only be used to relativize subject, direct object, and indirect object.

(798) Subject

Orang [yang jaga kucing], ati nia baik
 person REL take.care cat heart 3.SG good
 ‘The person who takes care of the cat, his heart (is) good.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-013, 00:11:54.2-00:11:58.6)

(799) [*Tarék chia punya*]orang tak ada lagik.
 pull rickshaw REL person NEG EXIST more
 ‘The people who pulled the rickshaws are not around anymore.’
 (Lim 1988: 19)

(800) Direct object

Budak itu nampak keréta [yang Mary bawak].
child that see car REL drive

‘That child saw the car that Mary drives.’

(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-013, 00:15:21.5-00:15:29.6)

(801) [*Gua pukol punya*]itu orang
1.SG hit REL that person

‘That person whom I hit.’

(Lim 1988: 19)

(802) Indirect object

Orang jumpa budak perempuan [yang dia kasi duit].
person meet child female REL 3.SG give money

‘The person met the girl whom he gave money (to).’

(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-013, 00:17:43.1-00:17:48.7)

(803) *Budak ini lupa [dia kasi makan nia] kucing tu*
child this forget 3.SG give food REL cat that

‘This child forgot that cat that he gave food (to).’

(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-013, 00:13:50.0-00:13:54.3)

(804) Possessor

Dia nampak Ali [yang kawan pukol dia].
3.SG see REL friend hit 3.SG

‘He saw Ali whose friends hit him.’

(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-013, 00:19:57.0-00:20:00.9)

6.3.2 Verb phrase

Where the verb phrase is concerned, there are lexical differences in the aspectual systems of MBM and SBM.

6.3.2.1 Progressive aspect

In SBM, the progressive aspect is expressed with the use of auxiliary verb *ada* (see section 5.2.5.4), which literally means ‘possess’, and is also an existential marker, a copula, and a perfective marker. In MBM, *tengah* functions as a progressive marker in addition to *ada*. *Tengah* literally translates to ‘middle’, and is also used in Bahasa Melayu to express the progressive.

Another progressive marker in Bahasa Melayu that is not found in BM is *sedang*. It is surprising that that *tengah*, which occurs in MBM, has been noted to be more formal than *sedang*, which does not occur in MBM (Mintz 1994). Note that while Sharif (1981) indicates *tengah* as being the progressive marker, his transcripts also demonstrate *ada* being used as a progressive marker. This is shown in examples (808) and (809). Examples (805) and (806) concern the use of *tengah* as a progressive marker, while examples (807) to (809) show *ada* being used for the same function.

- (805) *Budak ini tengah mandi sama anjing.*
 child this PROG bathe with dog
 ‘This child is bathing with the dog.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-047, 00:03:34.0-00:03:36.9)
- (806) *Kita tengah bikin.*
 1.PL PROG make
 ‘We are making (something).’
 (Sharif 1981: 53)
- (807) *Dia ada ingat.*
 3.SG PROG think
 ‘He is thinking.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:05:09.0-00:05:10.9)
- (808) *Kita ada semayang Tuapékong*
 1.PL PROG pray (name of a deity)
 ‘We are praying (to) Tuapekong.’
 (Sharif 1981: 20)
- (809) *Dia ada kerja*
 3.SG PROG work
 ‘He is working.’
 (Sharif 1981: 51)

6.3.2.2 Perfective aspect

Similar to SBM, the perfective aspect can be expressed using either *sudah* ‘already’ or *ada* (see section 5.2.5.2). There is no difference, except for the fact that *sudah* can be used in a contracted form *dah*. This differs from SBM contracted form, *sua*. Examples (810) to (812) demonstrate

this usage of *dah*, example (813) shows *sudah* used in its complete form, while (814) to (816) show how *ada* can also be used to express the perfective aspect.

- (810) *Hari pun dah petang.*
 day also **already** evening
 ‘The day has already (become) evening.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-047, 00:03:02.2-00:03:04.3)
- (811) *Buah tu dah jatoh*
 fruit that **already** fall
 ‘That fruit already fell.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-046, 00:00:58.0-00:01:00.9)
- (812) *Military tu dah mati.*
 that **already** die
 ‘That military (man) already died.’
 (Sharif 1981: 40)
- (813) *Kita sudah jadi orang besar lah.*
 1.PL **already** become person big EMP
 ‘We already became adults.’
 (Sharif 1981: 29)
- (814) *Barang ada pechah.*
 thing **PFV** break
 ‘That thing broke.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-024, 01:12:40.5-01:12:42.8)
- (815) *Jepun ada chakap*
 Japanese **PFV** talk
 ‘The Japanese talked.’
 (Sharif 1981: 40)
- (816) *Ada ikot Melayu dulu dulu.*
PFV follow Malay old old
 ‘They followed the old Malays.’
 (Sharif 1981: 26)

6.3.3 Adjectival and adverbial phrases

There are also slight differences in specific types of adjectival and adverbial phrases between MBM and SBM. Differences are found in the structures of comparatives and in the form of *dengan* ‘with’ adverbial phrases.

6.3.3.1 *Daripada* comparatives

With regard to comparatives, while the SBM speakers use the form *lagik* AP *lagik* (*lagik* literally means ‘more’ in this instance, see section 5.3.1), MBM speakers mostly use the standard Malay form *lebéh* AP *daripada*, as with examples (817) to (819). *Lebéh* translates to ‘more’ and *daripada* ‘from’.⁸⁷ Note that *lebéh* is also used to indicate ‘more’ in SBM, but it is not used in comparatives. A separate but related form is *kurang* AP *daripada*, *kurang* meaning ‘less’. This usage is demonstrated in example (820). There is no corresponding form in SBM.

(817) *Apple mérah lebéh manis daripada apple ijo.*
red more sweet than green
‘The red apple is sweeter than the green apple.’
(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-005, 00:30:38.9-00:30:45.9)

(818) *Apple ini lebéh manis daripada apple itu*
this more sweet than that
‘This apple is sweeter than that apple.’
(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-005, 00:30:33.0-00:30:37.5)

(819) *Di sini lebéh banyak apple daripada sana*
PREP here more many than there
‘Here (there are) more apples than there.’
(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-005, 00:31:18.6-00:31:22.1)

(820) *Apple ini kurang banyak daripada sana*
here less many than there
‘(There are) less apples here than there.’
(Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-005, 00:31:54.0-00:32:09.8)

6.3.3.2 *Dengan* ‘with’ adverbial phrase

Adverbial phrases of manner in MBM can be introduced with *dengan* in addition to *sama*. Their SBM counterparts are introduced by *sama*, which is also the general conjunction marker in SBM (adverbial phrases of manner that use *sama* are covered in section 5.4.3.). This use of *dengan* that is also found in Malay, is very rarely found in SBM. The following examples show how *dengan* ‘with’ is used to introduce adverbial phrases in MBM. It is interesting to note that neither

⁸⁷ *Daripada* can be further analysed in Malay as containing preposition *dari* ‘from’ and *pada* ‘on’.

suka hati in example (821) nor *gembira* in (823) are used in SBM to mean ‘happy’. In SBM, the same notion is expressed by Hokkien term *huahi*. Examples (821) to (825) demonstrate the use of *dengan*, while example (826) demonstrate the use of *sama*, which is also used in SBM.

- (821) *Orang itu pukol kucing dengan kayu.*
 person that hit cat **with** stick
 ‘That person hit the cat with the stick.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-010, 00:21:35.6-00:21:38.1)
- (822) *Gua sangat suka.hati gua chakap dengan lu.*
 1.SG very like.heart 1.SG speak **with** 2.SG
 ‘I (am) very happy I am speaking with you.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-005, 00:46:12.1-00:46:14.4)
- (823) *Dia makan apple manis dengan gembira.*
 3.SG eat sweet **with** happy
 ‘He ate the sweet apple happily.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-005, 00:36:33.0-00:36:41.0)
- (824) *Ini ada satu cherita reka –an sahja dengan se-orang budak.*
 this COP one story invent-NMZ only **with** one-person child
 ‘This is only one invented story with one child.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-047, 00:00:15.7-00:00:18.9)
- (825) **Dengan Santali**
with (company name)
 ‘With Santali.’
 (Sharif 1981: 41)
- (826) *Budak ini tengah mandi sama anjing.*
 child this PROG bathe **with** dog
 ‘This child is bathing with the dog.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-047, 00:03:34.0-00:03:36.9)

6.3.4 Conjunctions

Two conjunctions in MBM differ lexically from those used in SBM.

6.3.4.1 *Dan/sama* ‘and’ coordinating conjunction

While SBM speakers strictly use *sama* to indicate general conjunction ‘and’ (see section 5.6.8.1), MBM speakers use *sama* and *dan*. *Dan* has the same function in standard Malay, while *sama* is

used to indicate ‘with’ in that language. It thus appears that MBM is more influenced by standard Malay than SBM. Examples (827) and (828) demonstrate the usage of *dan*, and example (829) demonstrates the usage of *sama* by the same speaker. Example (830) shows that *sama* cannot be used for sequential relations. Sequential relations in both MBM and SBM are expressed by entirely separate clauses that can be mediated with the use of *habi* (see section 5.6.8.1).

- (827) *Satu ékor anjing dan jugak satu ékor kodok.*
 one CLF.animal dog **and** also one CLF.animal frog
 ‘One dog and also one frog.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-047, 00:00:23.2-00:00:28.2)
- (828) *Budak jantan ini dan anjing balék*
 Child male this **and** dog return
 ‘This boy and dog returned.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-047, 00:03:16.2-00:03:18.6)
- (829) *Apal sal budak ini sama anjing ini mo ni?*
 what reason child this **and** dog this want this
 ‘Why do this child and this dog want this?’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-047, 00:03:27.7-00:03:30.6)
- (830) *Dia chakapchakapsama nangis.*
 3.SG speak speak **and** cry
 ‘He was speaking and crying.’
 (Lim1988: 37)

6.3.4.2 *Atau/ ka* ‘or’ coordinating conjunction

While SBM speakers uses *ka* to indicate ‘or’, MBM speakers use *atau* in addition to *ka* (the usage of *ka* in SBM is explained in section 5.6.8.1.). Unsurprisingly, considering that MBM speakers come into much more contact with standard Malay than SBM speakers, *atau* is derived from Malay. *Ka* on the other hand is derived from Hokkien. Examples (831) and (832) show the usage of Malay-derived *atau* ‘or’, while example (833) shows the usage of Hokkien-derived *ka*. Example (834), which has been taken out of Sharif’s (1981) transcripts, is interesting, because it shows the use of both *ka* and *atau* in the same sentence, showing that both lexical items are available to the speaker of MBM.

- (831) *Sekarang mia murid murid atau budak budak sekolah.*
 now REL disciple disciple or child child school
 ‘The current disciples or school children.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:01:22.7-00:01:27.0)
- (832) *Dia mo buat kawan sama dia atau mo tangkap gua makan.*
 3.SG want makefriend with 3.SG or want capture 1.SG eat
 ‘He wanted to make friends with him or capture me to eat.’
 (Albert Ku, oai:scholarspace.manoa.hawaii.edu: NL1-003, 00:01:22.7-00:01:27.0)
- (833) *Lu suka chakap Melayu ka suka chakap English?*
 2.SG like speak Malay or like speak English
 ‘Do you like to speak Malay or do (you) like to speak English?’
 (Lim1988: 38)
- (834) *Chuchu ada empat ka atau lima*
 grandchild have four or or five
 ‘Granchildren (I) have four or five.’
 (Sharif 1981: 36)

In view of the above examples regarding demonstrative determiners, progressive and perfective aspects, *daripada* comparatives, *dengan* adverbial phrases, as well as *dan* and *atau* conjunctions, differences between MBM and SBM appear to be mainly of a lexical nature. A clarification made in this chapter is that both prenominal *punya* and postnominal *yang* are used for relativization in MBM (c.f. Sharif 1981, Lim 1988).

7. What type of language is BM?

The intention of this dissertation has been to value BM as a language in its own right. In this spirit, inasmuch as possible, the language's grammar has been described without referring to its possible component languages from the get-go. However, it is undeniably interesting to ask questions such as, which language particular lexical items or grammatical structures may be attributed to, and what that may portend for the language in terms of classification.

The findings will show that BM is indeed a contact language. The substratal component is Hokkien from Zhangzhou and Quanzhou, from where most Hokkiens in the Strait of Malacca originated. Although the Hokkien spoken in Singapore has often been referred to as Amoy Hokkien, it is very different from the Hokkien that is actually spoken in Amoy (see Pitcher 1912). The Zhangzhou and Quanzhou version of Hokkien, which is spoken in Singapore, has been recorded by Douglas (1873) and Bodman (1955, 1958) in some detail. Where Malay is concerned, there are differing views. Some have stated that BM was more likely influenced by colloquial Malay or Low Malay rather than the standard variety or High Malay, with the assumption that some kind of pidgin Malay was already spoken by the Chinese traders, and that the local women they married spoke colloquial Malay (see Ansaldo et al. 2007). Others are of the view that the language was more likely influenced by more standard Malay or High Malay (Lee 2009). There is no record that the local women spoke Low Malay instead of High Malay. All that is usually stated is that BM was formed via intermarriage of early Southern Chinese male settlers and local Malay women (Vaughan 1879, Clammer 1980, Chia 1983). Others have also hinted at the possibility that some of these indigenous women were slaves from the Dutch East Indies (now Indonesia), such as the Batak, Bugis, Balinese, and Siamese. (Rudolph 1998). However, no other native language has ever been mentioned in connection with these local women. In addition, BM is more similar to standard Malay than it is to colloquial Malay. For example, it utilizes noun classifiers which are also used in standard Malay but not in colloquial Malay, and *yang* as a relative clause marker occurs frequently in BM, just as it does in standard Malay but not in colloquial Malay (See Aye 2005). The version of Malay that most likely influenced the formation of BM was recorded by Marsden (1812) and Crawford (1852). The link between the language that these sources represent and BM has been recognised by Thurgood (1998). These

references, as well as Douglas (1873) and Bodman (1955, 1958)'s work on Hokkien will be utilized when comparing BM to its component languages.

7.1 Classification

The sociohistorical background of BM needs only brief discussion, having been introduced in sections 1.1 and 1.2. BM was formed through the intermarriage of Hokkien-speaking traders and Malay-speaking indigenous women, as early as the 15th century. The descendants are known as the Peranakans, the females, Nyonyas and the males, Baba. Together with their unique culture that emphasizes customs, food, and attire, BM was thus developed.

Incontrovertibly, given the special circumstances of the language's genesis, the main focus of researchers working on BM has been the classification of BM. As mentioned in section 1.3.2, researchers are divided into two camps. There are those who regard BM as a dialect of Malay (Pakir 1986, Thurgood 1998), and those who regard BM as a language that has undergone creole formation (Lim 1981, Lim 1988, Ansaldo and Matthews 1999, Ansaldo et al. 2007, Shih 2009). To complicate the matter further, BM appears to share a similar sociohistorical background with languages that typify a category labelled 'mixed languages'. Using sociohistorical information, linguistic data from this grammar, as well as observations of researchers who have previously worked on BM, this chapter aims to shed light on how BM should be classified. Essentially it maintains that a process-based approach to understanding BM's classification is preferable, rather than one that purely relies on types, these types being the traditional labels of "dialect", "pidgin", "mixed language", or "creole", to name a few.

The issue of classification in the field of contact linguistics has always been fraught with complications, in part due to the fact that some of these labels were "lay terms" before linguists began using them in more specific ways (Mufwene 1997: 39). For example, the word "creole", possibly from Latin *creāre* 'to create' and later Portuguese *criar* 'to raise (e.g. a child)' appears to have been used for children born in the colonies by Africans, and then later by Europeans, and is also used to refer to their customs and speech (Holm 2000). As these labels came into use into contact linguistics, these terms have often been interpreted differently by different researchers.

For example, to some researchers, a creole refers to a nativized pidgin (or even jargon) (Holm 2000), whereas to others, the process of vernacularization distinguishes a pidgin from a creole (Chaudenson 1992). Other factors for the difficulties in labelling and hence, identification of pidgins and creoles have also been discussed. Mühlhäusler (1986) writes that often, lower-class pidgin and creole speakers are unaware that they speak a separate language, and they may not wish to admit that they speak a separate language even in cases where they are aware of it. This is especially so when the norm is to not use these languages in front of white speakers (who presumably until today constitute most researchers working on contact languages). On a separate note, Mufwene (1997) proposes that the problems with labels stem from the fact that none of the proposed definitions for these terms takes into account the contact languages' genetic histories, and that linguistic structure cannot be a factor in naming. Whether or not this is an accurate statement will be further evaluated in this chapter.

All that being said, labels are still necessary, for they allow abstraction in the field of contact linguistics (Thomason 1997b). In effect, these labels enable comparisons to be made between different languages within the same type, or even different languages across different types. For example, within the category of creoles, there are European language-based creoles, and non-European language-based creoles, which Ansaldo et al. (2007) consider BM to be. Presuming as Ansaldo et al. do, that BM is indeed a creole, they then ask questions about whether the language shares similar structural properties to European language-based creoles, and how much of it can be attributed to the influence of the individual component languages or even universal mechanisms. Across categories, one may choose to compare BM with Bazaar Malay, which is a pidgin comprising the same lexifier language (Malay) and substrate language (Hokkien) as BM (Aye 2005), and ask questions about whether BM might have emerged from Bazaar Malay (see section 7.4.4 for discussion on this topic). Note however, that at the end of her article on the typology of contact languages, Thomason (1997: 86) urges researchers against classifying languages that fall between dichotomies, such as between pidgin vs. ephemeral and unsystematic speech form, bilingual mixed language vs. ephemeral and unsystematic codeswitching mixture, creole vs. decreolized dialect of a lexifier language, pidgin/creole vs. dialect of lexifier when all the input languages are closely related, and between pidgin vs. ad hoc simplified version of a

lexifier language. She argues that it is particularly necessary to accept fuzzy boundaries, when “there are no congruent social and linguistic correlates by which they (the languages) may be identified” (Thomason 1997:86). Likewise, even though Thomason does not directly address the dichotomies between a dialect undergoing shift and a creole, or a creole and a mixed language, these labels are only useful if they are sociohistorically or linguistically meaningful. With the above in mind, the following sections considers arguments for the classification of BM as a genetic dialect of a parent language, as a mixed language, or as a creole, since these are the three most feasible categories that BM may belong to.

7.2 BM as a “dialect” of Malay

The term “dialect” is complex, its manifold definitions ranging from social to geographical and political ones. While it is usually employed by linguists to mean a “variety (regional or social) of a language, which is mutually intelligible with other dialects of the same language” and “is no longer used to refer to a daughter language of a language family” (Campbell 2013: 173), the word has been used in the literature of BM to mean a genetic daughter language (instead of one whose genetic affiliation is difficult to establish, or one that cannot be attributed to a single genetic parent, such as a creole). Thus, the word “dialect” is used here to mean genetic daughter language.

The suggestion that BM is a dialect of Malay has been put forth by Tan (1980), Thurgood (1998) and Pakir (1986). While Thurgood’s dissertation was written later than Tan’s and Pakir’s works, she addresses what she terms as “Old Baba Malay”, a variety of the language that possibly existed before the arrival of more Chinese immigrants in the late nineteenth and early twentieth century. The source of Thurgood’s data are newspapers published in BM from 2 July 1894 to 2 July 1895, as well as *A manual of the Malay colloquial, such as is spoken by all nationalities in the Colonies of the Straits Settlements, and designed for domestic and business purposes* (Lim 1887). Tan and Pakir on the other hand, are concerned with more modern BM.

7.2.1 Thurgood (1998)

Thurgood (1998) cites Thomason and Kaufman (1988) who differentiate between language shift and creole formation. While the shifting population has full access to the target language in the case of language shift, access to the target language (or lexifier language) is restricted in cases of creole formation. With regard to BM, since there should have been full access to Malay, given the intermarriage scenario, the language is then evaluated as being a dialect of Malay, and one that is shifting towards the lexifier (Thurgood 1998). However, an important issue with this analysis is that it assumes that the speakers' target is the lexifier language. It is possible that the very initial male speakers were bilingual in the languages of the traders. Yet, it has been noted that the traders used a form of lingua franca (Lim 1981, 1988) that would not have been a full variety of the Malay lexifier. Some have gone on to postulate that BM itself became a lingua franca for trading in the region (Tan 1988). In the intermarriage scenario, it is even more unlikely that the speaker's target was the lexifier language. One can only postulate what may have happened in the minds of those speakers in the past, but it is probable that their goal was to be able to communicate with each other, and that there was no need or desire to fully acquire the lexifier language, Malay (or even the substrate language, Hokkien for that matter). In fact, the Peranakans did not want to be associated with both groups. They called the newly arriving Chinese *sinkék* 'newcomer', and preferred their children to marry other Peranakans. The Peranakans also did not want to be associated with speaking pure Malay, "calling it *Malayu hutan*—the language of the jungle" (Shellabear 1913:156, italics in original). It is relevant that the Peranakans have never fully assimilated into Malay culture. Instead they developed their own culture that emphasizes unique ways of dress, cooking, and traditions, these demonstrating a blend of Chinese, indigenous, and Western influences. Returning to the matter of what this means for the classification of BM, assuming that speakers did not treat the lexifier language as the target language, BM may be classified within the group of pidgins and creoles (albeit peripherally) based on Thomason's (1997) typology. Thomason states that "when no group has the need, the desire, and/or the opportunity to learn any of the other groups' languages", pidgins and creoles form (1997: 78). While the speakers were widely exposed to, and had the opportunity to learn either groups' language, it is probable that they had no need nor desire to.

The other argument that Thurgood (1998) puts forth for classifying BM as a product of language shift is that two languages are typically involved in instances of language shift, while more than two languages are usually involved in instances of creole formation. This appears to be based on Thomason and Kaufman (1988)'s comment that all or most creoles stem from contact among more than two languages. Note however, that using this criteria among others, Thomason (1997) also differentiates between pidgins/creoles and mixed languages. Hence, that BM has only two main languages involved in its formation, does not necessarily nor sufficiently warrant its classification as a language undergoing shift. It is possible to consider BM as a mixed language (see section 7.3). On the other hand, assuming that the number of languages involved in contact is indicative of whether a language should be classified as one undergoing shift or as a creole, it is disputable whether BM involves the contact of only two languages, considering that other languages such as English and even Cantonese had been present during the late 19th century (during which time the Old Baba Malay that Thurgood addresses existed). Earlier adstrates from the 16th to early 19th century include Dutch and Portuguese. This is evidenced by loanwords in BM, such as *mentéga* 'butter' and *menjéla* 'window' (derived from Portuguese *manteiga* and *janela* respectively) and even *apple* and *pear* from English.⁸⁸ Note however, that while these other languages were around, predominantly and overwhelmingly, only Hokkien and Malay were involved to any significant degree in the development of BM. All that being said, the criteria for "creolehood" varies among researchers, depending usually on which theory of creole formation is being endorsed. The number of languages involved is usually a peripheral concern rather than a central one. For example, there are cases where language shift involves more than a language shifting towards another. On the Upper Rio Negro in the Northwest Amazon of Brazil, language shift is going on in the direction of two additional languages at once. Speakers are shifting from Eastern Tukanoan languages to Tukano, and to Portuguese (Flemings 2010). There are also arguably creoles that do not involve multiple languages. These occur due to the vernacularization of a pidgin, where the pidgin may have only involved contact between two languages. For example, Ngatikese Men's Creole that is spoken on the Ngatik atoll in Micronesia

⁸⁸ Pakir (1986) provides a list of lexical items that are borrowed from other languages, some of which have also been borrowed into Malay. Note that she considers words of Hokkien origin to be borrowings in BM, while this current grammar treats them as words of BM that are derived from the substrate language.

comprises Sapuahfik Pohnapeian and English (Tyron et al. 1996). More discussion of creole formation ensues in section 7.4.4.

Another point that Thurgood (1998) then makes is that BM appears to be undergoing light to moderate interference through shift, since according to Thomason and Kaufman (1988), this usually results in grammatical simplification. Thurgood (1998:249) states that “the influence of Hokkien is found, not in an abundance of Hokkien-like constructions, but rather in the patterns of simplification and generalization”. As an example of simplification, Thurgood points out the lack of verbal suffix *-i* in BM, which together with *-kan* in Malay increases the number of arguments a verb can take. Note that while Thurgood observes the use of *-kan* to be unpredictable in Old BM, this grammar has shown its consistent use as BM as a transitive marker (see section 4.2.1.1). This may be a later development of BM. However, more importantly, analysis will show that the structural influence from Hokkien is considerable, and not all of the cases of this show grammatical simplification. Hokkien influence can be demonstrated in patterns of demonstrative use, the classifier system, relativization, the passive construction, and the benefactive construction, among others. Some of these have been considered by researchers such as Lim (1988), Ansaldo and Matthews (1999), Shih (2009) and Lee (2012). More discussion on Hokkien influence ensues in section 7.4.

Given the above, it is also important not to compare directly the structures of old Baba Malay with the modern-day version of the language. Thurgood (1998) reminds the reader at the end of her dissertation that the language has since undergone change due to a large influx of new Chinese immigrants as well as the growth of Singapore and other British-controlled towns in the late 19th century. However, the data used to support her arguments is from the late 19th century. The BM newspaper articles from *Bintang Timor* ‘East Star’ date from 2 July 1894 to 2 July 1895, and Lim’s manual of colloquial Malay is dated 1887. While Thurgood (1998) clarifies that Old Baba Malay that was recorded in writing appears to be different from the modern variety that Pakir (1986) describes, it is unclear when the influence of later Hokkien-speaking immigrants supposedly begins to appear in the language. Note that by 1860, twenty to thirty years before the written sources used were published, the Chinese had already numbered 50,043

out of a population of 81,734 (Ng 1961). In addition, one has to be cautious about the use of written material as representative of the spoken language, particularly formally written newspaper articles. Although that Thurgood (1998) mentions that the less formally written commentaries contributed by readers in *Bintang Timor* and the manual on colloquial Malay may exemplify the language better, it is not clear if findings from these two main sources differed in any significant way. With the limitations in mind of researching an older form of the language that existed when oral recordings could not be made, Thurgood (1998) provides a thorough analysis of the written material, and is representative of Old Baba Malay.

7.2.2 Tan (1980)

Primarily a sociologist whose dissertation was based on the Peranakan culture (Tan 1979), Tan's aim was to show that BM is not a corrupted language, but "a dialect which has developed out of a particular kind of historical process" (Tan 1980:150). Tan introduces the sociohistorical background of the Peranakans. He views BM as having evolved out of bazaar Malay, which was the colloquial language of business. (This will be further discussed in relation to the possibility of BM being a creole in section 7.4.) He then proceeds to show various patterns of differences between BM and Standard Malay. Note that while his fieldwork was carried out in Malacca, he regards the variety of BM spoken in Malacca to be similar to the one spoken in Singapore. Tan (1980) states that the most significant differences are found in the use of loanwords and in its phonology. He regards Hokkien words in BM to be loanwords, and provides a list of these, as well as some from Indonesian and English, which show interesting cultural differences between BM speakers and Malay speakers. With regard to phonology, Tan (1980) shows differences between BM and Standard Malay, for example, vowel raising to ϵ , and h -deletion, among others. These are also noted by Pakir (1986), and in this grammar (see section 3.3). Essentially, through showing systematicity in the language, Tan seeks to defend BM from the view that BM is "a corrupted version of the Malay language", which "implies that the Babas have no language of their own" (1980:165). While commendable, it must be noted that Tan (1980) does not differentiate between notions of a creole and a genetic daughter language. There are no

definitions provided for the term “dialect.” Crucially, this work positions BM as a language in its own right, and not as a broken, unsystematic version of Malay.

7.2.3 Pakir (1986)

Pakir (1986)’s dissertation also focuses on modern BM, its goal being to resolve the question of BM’s position, since the language had been identified as a creole by some (Clammer 1980, Clammer 1983, Grimes 1974, Grimes 1984) and as a dialect of Malay by others (Tan 1980). The areas of phonology, lexicon, and syntax are investigated for Hokkien influence. In terms of phonology, Pakir (1986) states that BM is similar to other dialects of Malay. While there are some differences, such as vowel raising to /e/ (recognised as a refined form by this grammar, see section 3.3), and h-deletion, among others. While Pakir observes that Hokkien is responsible for the introduction of /ɔ/ in the phoneme inventory, Hokkien loanwords “have had relatively little effect on the phonological system” (Pakir 1986:207). Where the lexicon is concerned, she points out Hokkien borrowings that are “clearly delimited to customary and cultural aspects of Baba life and to ‘ethnic’ value terms” (Pakir 1986:209). With regard to syntax, Pakir (1986) focuses on constructions with *punya* ‘possessive’, *kasi* ‘give’, *kena* ‘subjected to’, and *mo* ‘want’, whose uses are identified in this grammar as a prenominal relative clause marker (section 5.6.3), a benefactive verb used in passives as well as in ditransitive, benefactive and causative constructions (section 5.2.4), a passive marker (section 5.2.3), and as a deontic marker (section 5.2.2). These are said to have “functions analogous to the functions of some particular H[okkien] morpheme” (Pakir 1986:210). Crucially Pakir notes that there are still syntactic properties of BM that cannot be explained by Hokkien influence, and draws the conclusion that “while there has been some borrowing and other influence from H[okkien], the H[okkien] element and role in BM has been overestimated or too often reported on without much linguistics basis” (Pakir 1986:211). With regard to the position of BM, Pakir states that “[t]here should be no stigma to speaking a language like BM for it is the creation of the Baba Chinese” and that “[e]ven if BM has its origins in a type of Bazaar Malay pidgin, the fact that it has acquired native speakers for at least two centuries if not more, should not detract from its value as a “Malay dialect in its own right”” (Pakir 1986: 213).

While Pakir (1986) is commendable for being an early, in-depth study of BM, the implication is that creoles are corrupted, broken forms of a language. This has also been recognised by Ansaldo and Matthews, who comment that “[s]ome studies on BM are clearly permeated by the idea that recognising the creole-like nature of BM would somehow diminish its status, that the idea of creole as a corrupt, bastardised language would endanger the originality of BM” (1999:62). They also point out that these views may be “attributable to an outdated view of creoles, or to policies of linguistic nationalism in Malaysia, either way, it has little linguistic basis” (Ansaldo and Matthews 1999:62). Ansaldo and Matthew (1999) also state that shared features that may be found in other varieties of Malay do not invalidate arguments for Hokkien substratal influences as Hokkien was influential in the region, and that Hokkien influence in the lexicon is by no means shallow. They then provide more evidence of substratal influence from Hokkien to support the notion of BM being a creole (see section 7.4.2 for more discussion).

In sum, there is little basis for regarding BM as merely a dialect of Malay. Some may argue that BM is mutually intelligible with Malay (Chia 1994), but this is only natural given that most of BM’s lexicon is derived from Malay (Ansaldo and Matthews 1999). Note that there is an extent to how much Malay a BM speaker who has not acquired Malay formally would know, just as there is an extent to how much BM a Malay speaker would be able to understand. Even so, the criteria of being mutually intelligible with Malay does not stand in its way of creolehood, for creoles vary extensively with regard to how much they vary from their lexifier languages. For example, Louisiana Creole is closer to standard French than Haitian (Mather 2007) (keeping in mind that the lexifier language that community members were exposed to may have been non-standard varieties).

With regard to the main subject matter, some would take the view that BM is not a genetic dialect of Malay, since it would not have not undergone normal transmission (especially if one were to subscribe to the notion that BM developed from a pidgin that was earlier used for trade). The comparative method cannot apply when it would yield no systematic correspondences. Thomason and Kaufman (1988) highlight this, pointing to Tok Pisin and the Melanesian language Tolai. Thomason and Kaufman also maintain that the study of genetic relation should

be “thus based theoretically on the social facts of normal transmission rather than merely on the linguistic facts themselves” (1998:12). This view is not shared by some, such as Mufwene (2003a). In response to a similar proposal in Thomason (2002), Mufwene (2003a: 275) states that this method of classification simply dismisses the problem at hand, and that there can be no break in transmission of the lexifier when a pidgin or a creole shares much of its grammar with its lexifier, “despite modifications that are to be expected from the learning process” (Mufwene 2003a: 281-282). At the other extreme end are writers such as Dillard (1975) and Burling (1973) express that pidgins and creoles are mostly influenced by the native language of their speakers, more specifically, they suggest that African American Vernacular English can be attributed to the effects of the speakers’ native African languages. Burling (1973: 113) goes as far as to state that the slaves on the plantations “would have had little opportunity to hear or imitate the colloquial speech of the whites, but they could easily have imitated one another”. Such a stance is clearly untenable, for there is no substantial evidence that supports the notion that early speakers of African American Vernacular English had such limited contact with white speakers that they could not approximate the speech of the latter group. With regard to Thomason and Kaufman’s (1988: 12) assertion that that the study of genetic relation should be “thus based theoretically on the social facts of normal transmission rather than merely on the linguistic facts themselves”, this grammar agrees with Mufwene (2003a) that dismissing the genetic affiliations pidgins and creoles have to their source language(s) is simply a convenient proposal that does not add value to creolistics, it takes the view that a creole cannot be traced back to a single source. On separate occasions, Mufwene has also argued for the “complementary hypothesis”, which proposes that “the only influences in competition are structures of the lexifier and of the substrate languages” (Mufwene 2001: 34), hence suggesting that pidgins and creoles do not have to be strictly traced back to one source. Given the sociohistorical facts that gave rise to BM, it is not viable to regard BM as having been derived from Malay only.

7.3 BM as a mixed language

According to Thomason (1997), a bilingual mixed language usually develops in a two-language contact situation wherein at least one of the speaker groups is bilingual in both source languages.

Among other names given to mixed languages are ‘split languages’ (Myers-Scotton 2003) and ‘fused lects’ (Auer 1999). These mixed languages can usually be separated by its source languages, and these components are typically not simplified, given the widespread bilingualism on the part of at least one of its initial speaker groups (Thomason 1997).

Sociohistorically, mixed languages have been related to both the creation and maintenance of ethnic identity. For example mixed languages can arise from intermarriages, such as with Mednyj Aleut (also known as Copper Island Aleut), which was formed via the marriages of Russian fur seal hunters and Aleutian women on the Mednyj Island in the Bering Strait (Golovko 1994), as well as Michif, which was formed via the marriages of French fur traders and Cree women in the Red River Colony area (corresponding to present-day Manitoba and North Dakota) (Bakker 1997). Mixed languages can also arise through language identity maintenance, when groups undergoing language shift resist complete assimilation (Winford 2003). For example, Gurindji Kriol halts a complete shift to Kriol. The language is spoken by younger Gurindji people in Kalkaringi (Northern Territory, Australia), where Gurindji is still spoken by older people, but mostly in code-switching with Kriol, which is an English-lexifier creole (Meakins 2008). Similarly, Anglo-Romani was formed by Romani-speaking gypsies shifting towards English, and it comprises Romani structural and lexical insertions into a predominantly English discourse, the Romani material being used to express solidarity or even affection (Matras et al. 2007). However, not all mixed languages are formed via intermarriage, or out of attempts to maintain an ancestral language. Light Warlpiri, which is spoken in Lajamanu (Northern Territory, Australia) is spoken by younger Warlpiri speakers who continue to speak both their heritage language and English, and can be construed as an expression of their particular group identity (O’Shannessy 2005). In all instances, the relationship between mixed language and group identity is incontrovertible. Given the fact that commonly cited mixed languages such as Michif, Mednyj (Copper Island) Aleut, and Media Lengua emerged as vernaculars used to symbolize a separate ethnic identity, the proposition to also classify BM as a mixed language is ostensibly appealing. Notably, some have identified a separate Peranakan language spoken in

Indonesia (a blend of Malay and Javanese) (see section 1.2.1) as a mixed language (Dreyfuss and Oka 1979, Matras and Bakker 2003, Wolff 1983).⁸⁹

Structurally, mixed languages are a diverse category, contrary to Bakker's (2003: 108) notion that mixed languages are still recognised to "differ so radically from other results of language contact, that they show more similarities with each other than differences". The category includes languages that have different types of subsystem splits. There are mixed languages whose sources are split across grammar and lexicon, such as *Media Lengua*, spoken in Central Ecuador, which has Quechua grammar and almost 90 percent⁹⁰ Spanish lexicon (Muysken 1981). Another language that is split across grammar and lexicon is *Ma'a* (also known as *Mbugu*), spoken in Tanzania by a nomadic group that has moved into a region surrounded by Bantu languages. *Ma'a* is said to combine Bantu grammar with Cushitic lexicon (Mous 2003). There are also other languages such as *Anglo-Romani*, whose grammar comes from English and whose lexicon comprises of a restricted set of Romani words (in an otherwise English set of words) (Matras et al. 2007). A separate type of mixed languages is one whose sources are split between noun phrase and verb phrase. One of the most popular examples of these is *Michif*, which comprises predominantly French noun phrases (with French phonology intact) and Cree verb phrases (with Cree phonology intact) (Bakker 1997). A more recent example of a mixed language split across noun phrase and verb phrase is *Light Warlpiri*. While most verbs are derived from *Kriol/ Australian Aboriginal English*, nouns are derived from both *Warlpiri* and *Kriol/ Australian Aboriginal English* (O'Shannessy 2005). Again, the type of mixture may vary a lot from a language to the next. For example, instead of being a language split between noun-phrase and verb phrase, *Mednyj Aleut's* nouns and nonfinite verb inflections are derived from Aleut, while its finite verb inflections are derived from Russian, the lexicon in this instance being mostly derived from Aleut (Golovko 1994).

⁸⁹ According to Dreyfuss and Oka (1979), 88 percent of grammatical affixes in this language that researchers term "Chindo", are Javanese, while function words, adjectives and adverbs are almost equivalently distributed between Javanese and Malay. Among themselves, the Peranakans of Indonesia often refer to their language as "Peranakan", just as the Peranakans of Malacca and Singapore refer to *Baba Malay* as "Peranakan".

⁹⁰ However, based on a separate Swadesh list count, only 65 percent of the stems are from Spanish (Muysken 1997)

Just as there are structurally diverse types of mixed languages, there are also diverse ways in which mixed languages could have come about. These include notions of extensive borrowing (Thomason and Kaufman 1988), code-switching (Myers-Scotton 2003), relexification (Muysken 1981), language intertwining (a situation in which languages are combined, instead of one wherein a language replaces the functions of another) (Bakker 1997), as well as metatypy, which occurs when the syntax of one language becomes restricted based on the syntax of another (Ross 1996). It is however difficult to ascertain which process was the most responsible for the mixed language outcomes (since most of these recognized languages were formed some time in the past). For this reason, only sociohistorical and structural factors are considered in the assessment of BM as a possible mixed language.

Thus far, there have been no attempts at classify also BM as a mixed language, although Ansaldo et al. (2007) do also iterate the fact that it is tempting to compare BM's formation with typical mixed languages. However, for several reasons, Ansaldo et al. (2007) do not do so. First, they question if mixed languages do form a valid class that is different in its genesis and type than other types of contact languages, and they also state that the same mechanisms of formation apply in all contact-induced change, citing Mufwene (2001, 2007). Next, they comment that pure bilingualism is hard to prove, and that the Peranakans showed high levels of multilingualism, and they also reference Siegel (1997), who proposes that features are transferred to the contact language in stages of interlanguage, and not directly from component languages to the contact language. The assumption here appears to be that the mixed languages are formed by groups that are highly bilingual in both component languages, and that features should not have to be transferred in stages to interlanguages, but all at once. However, there is no reason for why this should be so. Finally, they state that the structure of BM is not an "even distribution of lexical versus grammatical features of two different (groups of) languages", and that the "lexicon is mixed, and a combination of typological congruence and innovations can be found, similar to the type of restructuring we find in many creoles [and non-creole] languages" (Ansaldo et al. 2007:213). These concerns will be addressed below in the current assessment of whether BM should be categorized as a mixed language.

Before addressing how BM may or may not be like mixed languages in sociohistorical and structural ways, it appears to be necessary to validate this category of contact languages. As mentioned in the earlier paragraph, there are those who question if mixed languages do form a valid class that is different in its genesis and type than other types of contact languages (Ansaldo et al. 2007). For example, Mufwene (2001) is of the opinion that regardless of the different names proposed for the different mixed systems, the levels and types of mixedness can be explained by the same principles of a language speciating in an ecology⁹¹. In relation to this, other researchers such as Thomason (2003) are of the opinion that mixed languages differ from pidgins and creoles since the genesis of pidgins and creoles involve imperfect learning. Thomason then goes on to postulate that the main mechanism that may have contributed to the genesis of mixed languages is “change by deliberate decision” (2003:34). Change by deliberate decision can involve intentionally making a language different than the neighbouring dialects, so as to be distinct as a group (Campbell 2013: 325). Thurston (1989) and Ross (1996) describe “esoterogeny” as the way in which speakers of a particular language make their language more complex, so as to distinguish it from the languages of their neighbours.⁹² Ross (1996) states that a language can be “emblematic” of ethnic identity in a multilingual environment. For example, a language can become more different than its neighbours with “an increase in the frequency of opaque idioms” (Ross 1996: 183). However, the notion of esoterogeny is fraught with difficulties (Campbell and Posner 2008: 352). Chief among which, “it is not clear how this hypothesized cultural motive for these changes” “could be tested” (Campbell and Posner 2008: 352). Yet, others such as Blust (2005) have shown that there are some sound changes in historical linguistics that only can be explained by change by deliberate decision. Blust (2005) specifically discusses deliberate sound changes that occur in several Austronesian languages. While

⁹¹ For example, features may be adapted to suit the ecology, and features that are compatible with the features of other languages in the same ecology are more likely to be maintained than features that are not. For more discussion, see Mufwene (1996, 2001, 2008).

⁹² The opposite of esoterogeny is exoterogeny (Thurston 1989; Ross 1996), which usually occurs when a language is simplified for the reason that it is spoken as a “contact language” by members of other communities, as its original community is strongly bound to these other communities. Campbell and Posner (2008: 353) state that while this claim may be useful, it is may not be accurate, as there is no substantial evidence that shows that a language will become simplified if it is used across communities. They also provide counterexamples with languages that maintain their complexity even though they are used across different communities (such as Arabic, Turkic, Mongolian and Georgian).

deliberate language change cannot be confirmed for certain, evidence such as that of Blust (2005) shows that it cannot be easily dismissed.

Given how the components in a mixed language are usually not simplified, it is not implausible that change by deliberate decision may account for the genesis of mixed languages rather than imperfect learning. Mechanisms of genesis aside, mixed languages are a structurally definable category. The current view taken here is, regardless of whether it is the entire lexicon, or at the level of inflectional morphemes, it should be possible to attribute some identifiable different subsystems in the mixed language clearly to different individual source languages. Given the plausibility of change by deliberate decision contributing to the genesis of this particular category and the distinct difference in its structural characteristics (compared to other types of contact languages), it is valid to ask if BM is a mixed language.

7.3.1 Sociohistorical factors

While I take no stance on whether imperfect learning accounts for the genesis of pidgins and creoles given the views of some that these languages arose from contact with non-standard speakers (Mufwene 1996), it is not an unreasonable proposition that BM arose out of change due to deliberate decision. This is particularly so since the Peranakans view themselves to be unlike other Chinese and Malays in Singapore. A heavily Malay-based lexicon differentiates the Peranakans from the Chinese, and features such as selected Hokkien personal pronouns (see section 4.1.3.1) and refined [ɛ] forms (see section 3.7) to differentiate themselves from standard Malay speakers. Again, it is not possible to look into the minds of the two original groups that formed BM, and consequently, BM's genesis cannot be attributed to change by deliberate decision with an absolute level of certainty, but particular extrapolations may be considered. BM is by no means a contact language with simplified components. To illustrate this, the language has two patterns of relativization, a prenominal strategy derived from Hokkien and a postnominal-strategy derived from Malay (see section 6.3.1.3). The implications of this for the structural-matchedness of BM with mixed languages will be discussed later, but it is clear that the initial groups must have had unimpeded exposure to both Hokkien and Malay in the

intermarriage scenario. This begs the question of whether “pure bilingualism” is indeed a necessary condition for the genesis of mixed languages (Ansaldo et al. 2007).

Notwithstanding the usual difficulties in defining bilingualism (Romaine 1995), pure bilingualism or full bilingualism refers here to being fluent in both relevant languages. That such bilingualism is involved in mixed languages has been refuted by researchers such as Matras (2003), who states that this puts varieties of recognised mixed languages formed via language attrition (or shift towards another language) such as Ma’á and Para-Romani in the periphery.⁹³ Regardless, it is necessary to examine the implications of why mixed languages are supposed to have emerged quickly in situations of full bilingualism, especially since it is an often-cited claim (Bakker 1997, Thomason 1997a). The implications are two-fold. First, lack of proficiency in any one of the two source languages may not result in a mixed language, since certain proficiency on the part of early speaker groups is essential for the development of mostly intact subsystems from both source languages. Second, knowledge in other languages (aside from the two main languages concerned) may affect the final output, so that entire individual subsystems of the language cannot have been neatly derived from two languages. In the case of the Peranakans, the early Chinese traders must have learnt some basic Malay vernacular to trade in the region, although it is not clear if this was to the extent of widespread bilingualism. Again, that being said, there would still have been little obstruction to the exposure of both component languages in the creation of BM, given the intermarriage scenario. The opinion here is that unimpeded access to both languages is sufficient to warrant the genesis of a mixed language, not necessarily “pure bilingualism”. The second implication appears to be more relevant in the case of the Peranakans, who showed high levels of multilingualism (Ansaldo et al. 2007). Note that the Peranakan community was the very first to acquire English in the Malacca/ Singapore region when it was introduced in the colonial period (Lim 2010), and they have been proficient enough in the language to have shifted towards it so much that their own language has become endangered (Chia 1983). It is however less plausible that the initial speaker groups that formed BM had been proficient at English, given that the community had existed before the period of

⁹³ Both Ma’á and Para-Romani are said to be structurally prototypical for having a lexicon that comes from one language, and a grammar that comes from another.

British colonialism in the region during the late 18th to early 20th century (see section 1.1). There is also little evidence that the initial groups speaking BM had been proficient at other languages. Thus, it is unlikely that a language other than Hokkien and Malay, would have contributed significantly to the language, besides perhaps the adstratal contributions of borrowings from languages such as Portuguese and English. Thus, BM would appear to have met the bilingual condition that gives rise to mixed languages (however tenuous this requirement is).

Also on the topic of the BM's sociohistorical background, it is necessary to state that there is no evidence for or against the quick formation of BM. While researchers who focus on mixed languages that evolved through the intermarriage scenario propose that mixed languages form rapidly (see Bakker 1997), other researchers state that this is not the case for languages associated with resistance to complete language shift (Matras 2003). However controversial, the speed of genesis may have implications for whether or not mixed languages are affected by the proposition that features are transferred to the contact language in stages of interlanguage, and not directly from component languages to the contact language (Siegel 1997). The faster the formation of the language, the less likely it is to have undergone several of these stages, and vice versa. In line with this, it is important to note that Siegel (1997) states that feature transfers that occur in stages of interlanguage affects pidgin and creoles, and indigenous varieties of languages, but does not mention mixed languages. While Ansaldo et al. (2007) appear to indicate that it is possible for features from Malay or Hokkien to have existed in states of interlanguage that may have preceded the resultant form of BM, there would technically have been nothing that prevented a quick genesis of BM given the intermarriage scenario, as in the case of Michif (Bakker 1997).⁹⁴

The speed of BM's genesis is unclear but other sociohistorical factors, such as the probability that the language arose out of a deliberate decision that it should be representative of an ethnic identity, and the widespread bilingualism in the early ecology of BM, corroborate the idea that

⁹⁴ It is interesting to note that differences can arise between researchers based on the type of process involved. For example Sri Lanka Malay is regarded to have been formed via rapid convergence by Bakker (1995), but via gradual metatypy by Ansaldo (2011). Whether or not Sri Lanka Malay as a mixed language is still being debated (see Nordhoff 2012).

BM may indeed be a mixed language. Having discussed sociohistorical factors that affect the assessment of BM as a mixed language, it is necessary to examine the structure of BM before drawing any conclusions.

7.3.2 Structural factors

Structurally, researchers have been interesting in distinguishing mixed languages from pidgins and creoles. Thomason (2003) states that in pidgins and creoles, the lexicon may come from one language, while the grammatical subsystems in these languages can be viewed as a cross-language compromise. Mixed languages (which she terms “bilingual mixed language”) on the other hand, adopt intact lexical and structural subsystems from the source languages, sometimes with some distortion or adaptation (Thomason 2003). In a similar vein, Bakker (2003) asserts that it is usually not impossible to identify a clear source for the grammatical component in pidgins and creoles, whereas mixed languages should have numerically equal and identifiable components from two languages. Whether or not only one component is clearly identifiable to a pidgin or creole’s source language is debatable and will be discussed in the next section.

Returning to the topic of mixed languages, mixed languages should clearly be identified by their structure, which comprises two component languages that are individually compartmentalized in separate subsystems, and BM’s structural status will be assessed accordingly. Another interesting supposition is that in the instance of intermarriages, mothers provide the grammar language when raising their children (see Bakker 1997). This is not as strong as the earlier claim, but will be brought up as a point of interest in the following discussion.

Structural findings do not match up with sociohistorical ones. The findings here underscore Ansaldo et al.’s position, that “BM cannot be captured as an even distribution of lexical versus grammatical features of two different (groups of) languages” (2007:213). Even at the narrower level of individual subsystems, BM does not show compartmentalization of individual languages. To examine if there is any systematicity in the lexicon, three word lists are presented, the first showing a completely Malay lexicon, the second shows an almost systematic Malay-Hokkien split, and the third, shows an unsystematic mixture. Where the grammar is concerned,

the tense and aspectual system, as well as relativization patterns are used to show that there is no systematic mixture in BM.

Most words in BM are derived from Malay, and this is reflected in a Swadesh 100 word list (the complete list is appended in Appendix C). The first twenty words on this list are shown in the following table.

1.	I	<i>saya</i> (refined) <i>gua</i> (coarse)
2.	you	<i>lu</i>
3.	we	<i>kita</i>
4.	this	<i>ini</i>
5.	that	<i>itu</i>
6.	who?	<i>siapa?</i>
7.	what?	<i>apa?</i>
8.	not	<i>bukan</i> (noun negator) <i>tak</i> (verb negator)
9.	all	<i>sumua</i>
10.	many	<i>manyak</i>
11.	one	<i>satu</i>
12.	two	<i>dua</i>
13.	big	<i>besar</i> (coarse)

		<i>besair</i> (refined)
14.	long	<i>panjang</i>
15.	small	<i>kechik</i>
16.	woman	<i>perompuan</i>
17.	man	<i>jantan</i> ⁹⁵
18.	person	<i>orang</i>
19.	fish	<i>ikan</i>
20.	bird	<i>burong</i>

Table 44: List of first 20 words on Swadesh basic word list.

Given that almost the words on such a list are derived from Malay, it is indeed attractive to think of BM as a mixed language that has a Malay lexicon and a Hokkien grammar. But this is not the case. A substantial number of Hokkien words can be found in the language, particularly in areas of kinship terminology, religious customs, rituals and celebrations, as well as in terms of ethnic value and emotive import (Pakir 1986). In view of this information, the following list is provided. The words on this list represent a very specific domain, roles in a wedding ceremony.⁹⁶⁹⁷

⁹⁵ Note that MBM speakers prefer to use *laki* to mean ‘man’, since *jantan* refers to male animals in the source language, Malay. *Laki* also refers to ‘husband’ in both MBM and SBM.

⁹⁶ These are used both in SBM and MBM.

⁹⁷ For more information on the complex Peranakan wedding ceremonies, Cheo (1983) is a good source.

Role	Word in BM	Source of word in BM
‘bride’	<i>kemantin</i>	Malay
‘groom’	<i>kiasai</i> ⁹⁸	Hokkien
‘bride’s girl assistant’	<i>pengapék</i>	Malay
‘groom’s boy assistant’	<i>kuya</i>	Hokkien
‘bride’s assistant for kneeling and general help’	<i>bukak kun</i>	Malay-Hokkien
‘groom’s family’s host’	<i>chu-lang</i>	Hokkien
‘bride’s ceremonial assistant’	<i>sangkék-em</i>	Hokkien
‘groom’s ceremonial assistant’	<i>pak-chindék</i>	Malay

Table 45: Key roles in a Peranakan wedding

If BM is a mixed language, one might expect two scenarios. In the first scenario, the lexicon comprises almost entirely of words from one language. This is not the case, as seen in the table

⁹⁸ Also son-in-law

above. In the second scenario, the lexicon might be mixed in a systematic fashion. Thus, concerning the domain of the Peranakan wedding, words to do with the groom might be derived from Hokkien, while words concerning the bride might be derived from Malay. Again, this is not the case. While the words for the bride and the bride's girl assistant (akin to a flower girl) are of Malay origin, and the words for the groom, the groom's boy assistant (akin to a page boy) and the groom's family's *chu-lang* (literally 'host person') who does the job of hosting the main ceremony) are of Hokkien origin, the labels for the other roles are less systematic. The assistant who quickly lifts the bride's heavy skirt for her when the bride is required to kneel is called the *bukak kun* (literally 'open skirt'), the word *bukak* is derived from Malay while *kun* is derived from Hokkien (since the type of skirt the bride wears is a Chinese-style skirt). In addition, the word for the bride's ceremonial assistant (whom the *bukak kun* also answers to) is derived from Hokkien, while the word for the groom's ceremonial assistant is derived from Malay. These have to do with cultural underpinnings. The *sangkék-em* (literally 'deliver marriage aunty) dresses and performs rituals for the bride, while the *pak-chindék* (literally 'uncle' Chindek), traditionally a Boyanese man, looks after the groom. One might expect that the groom's valet, being under the groom's employment would have been Chinese, while the mistress of ceremony, usually being under the bride's family's employment, would have been local. Yet, as the lexicon indicates, the *sangkék-em* is a Hokkien tradition, while the *pak-chindék* is a Malay tradition (these traditions are very rarely practised by any group today in Singapore and Malacca). Hence, what this domain shows is earlier ceremonial choices made by the Peranakans, but not so much systematicity in the sense of mixed languages. In the same vein, one might argue that cultural domains are not reflective of the language's true mixture, since they usually reflect conscientious cultural choices.

Bearing that in mind, the following table of personal pronouns (which are arguably more basic than cultural terms) is presented.

Gloss	BM form	Hokkien form	Malay form
1.SG	<i>saya</i> ‘refined’	<i>goá</i>	<i>saya</i> ‘implies slave, but not literally so, polite, formal’ <i>patek</i> ‘implies even more humility than <i>saya</i> , seldom used’ <i>amba</i> ‘servant’s speech’ <i>beta</i> ‘domestic servant’s speech’ <i>aku</i> ‘informal’
	<i>gua</i> ‘coarse’		
1.PL	<i>kita</i>	<i>goán (goá-lâng)</i> ‘inclusive’	<i>kita</i> ‘inclusive’
		<i>lán-lâng</i> ‘exclusive’	<i>kami</i> ‘exclusive’
2.SG	<i>lu</i>	<i>lí</i>	<i>angkau</i> ‘used by both superiors and inferiors’ <i>kamu</i> ‘superior to inferior’ <i>tuan</i> ‘literal meaning: master’
2.PL	<i>lu-orang/ lu</i>	<i>lín (lín-lâng)</i>	
3.SG	<i>dia</i>	<i>i</i>	<i>diya</i>
3.PL	<i>dia-orang/ dia</i>	<i>in (i-lâng)</i>	<i>marika</i>

Table 46: Personal pronouns in Baba Malay, Hokkien, and Malay

The list of Hokkien personal pronouns that appears in table 46 is mostly based on Douglas (1873), while the list of Malay personal pronouns used is from Marsden (1812), since BM’s formation is more likely influenced by these forms of the component languages, rather than current-day forms. Note however, that the forms in parenthesis featuring *lâng* ‘people’ in the Hokkien column have not been mentioned by Douglas (1873). Ansaldo and Matthews (1999)

state that these forms may be later fused forms of the plural pronouns, given that *lang* is used widely in Hokkien to indicate an indefinite sense meaning ‘people’ (Douglas 1873), and that the forms of singular pronoun plus *lang* still exist in other varieties of Southern Min, such as the one used in Shantou. Given this explanation, the plural forms in BM do appear to look like calques of the Hokkien forms, and this has been recognized by Lim (1981), Pakir (1986), and Ansaldo and Matthews (1999). There is less of a match between the Malay forms cited by Marsden (1812) and the BM forms, except for the fact that besides *gua* ‘1.SG coarse’ and *lu* ‘2.SG’, Malay provides most of the BM lexicon. Semantically however, the distinction between refined and coarse speech may have developed with the derivation of *saya* from Malay, the use of *saya* implies that the speaker is a slave, although it is more often used metaphorically to express politeness or formality (see Marsden 1812). This is been reinterpreted in BM so that *saya* becomes a refined version of the first person while Hokkien first person *gua* has taken on implications of coarseness. If BM is a mixed language, one may expect systematicity, that the personal pronominal system might comprise entirely of labels from one language, and semantic structure from another. However, this is not the case, BM’s personal pronominal system uses labels from both Hokkien and Malay, and is structurally influenced by both calques from Hokkien and as well as by the formality and politeness denoted by Malay *saya*.

With regard to all lexical items derived separately from Malay and Hokkien, it must be noted from previous chapters that the separate sound systems do not stay intact. The sound systems appear to converge (see section 7.4.1), with the introduction of /ɛ/ and /ɔ/ from Hokkien, which are phonemes also used in words of Malay origin (see section 3.1.2.2). In addition to losing tonal contrasts on Hokkien words (see section 3.6.2), word-final stress in BM (see section 3.6.1) also appears to follow (but not exactly so) phrase-final high boundary tone in Malay (Ng 2012). BM’s phonological system has elements from both Malay and Hokkien, but it appears more Malay. This is unlike mixed languages that contain more unmixed words such as Michif (van Gijn 2009). Michif has an intact Cree phonological system for its verb phrases and an intact French phonological system for its noun phrases, as compared to Media Lengua, which is characterized by mixing within words, and has Quechua phonological system to accompany its mainly Quechua grammar (van Gijn 2009). BM should be compared to mixed languages with unmixed

words, given its more isolating nature, and this comparison shows that it is unlike any recognized mixed language.

While the above examples show that BM is neither lexically nor semantically close to being a mixed language, it is also necessary to examine other aspects of BM's grammatical system for a more complete picture. An area that can be examined is the tense and aspectual system of BM. Table 47 compares the tense and aspect systems of BM and its component languages. Again, Hokkien data is taken from Douglas (1873), Chiang (1940), and Bodman (1955, 1958), while the Malay data is gleaned from Marsden (1812) and Crawfurd (1852). Note that while Chiang (1940) and Douglas (19874) represent future 'want' in Hokkien as *beh*, and 'have' as *ū*, Bodman (1995, 1958) represents these as *bèq* and *ũ* respectively⁹⁹.

⁹⁹ These differences may be a matter of auditory interpretation. While *beh* has no tone, *bèq* has a low tone, and while *ū* has a mid tone, *ũ* has a rising tone.

Tense/ Aspect	BM form	Hokkien form	Malay form
future ‘not yet’	<i>belum</i> VP	<i>bōe</i> VP	<i>belum</i> VP
future ‘later’	<i>nanti</i> VP		<i>nanti</i> VP
future ‘want’	<i>mo</i> VP	<i>beh/bèq</i> VP	(<i>mo</i> ‘want’ is not recognised to indicate the future)
future ‘will’			<i>akan</i> VP
progressive ‘literal: have’	<i>ada</i> VP	<i>ū/ũ</i> VP	<i>ada</i> VP
habitual ‘literal: have’	<i>ada</i> VP	<i>ū/ũ</i> VP	
tentative	verb reduplication	verb reduplication	
perfective ‘have’	<i>ada</i> VP	<i>ū/ũ</i> VP	
perfective ‘already’	<i>sudah</i> VP	<i>i-keng</i> VP	<i>sudah</i> VP
			<i>telah</i> VP
recent perfect ‘just’	<i>baru</i> VP		<i>baharu</i> VP
experiential perfect ‘ever’	<i>pernah</i> VP	VP <i>kè</i>	<i>pernah</i> VP

Table 47: Tense and aspect of Baba Malay, Hokkien, and Malay.

While Marsden (1812) states that *sudah* and *telah* are used to indicate the past, examples provided show perfective use. *Mo* ‘want’ is recognized to be the equivalent of *beh* as BM has a

construction *mesti mo* that appears to be a calque of ‘*beh ai*’ in Hokkien (see 5.2.2 and Pakir 1986). Also, tentative aspect via verbal reduplication is characteristic of Hokkien (Tsao 2004). A quick glance at the table shows that while Malay contributes all the lexicon of the tense and aspect system of BM, grammatical functions can be attributed to both Hokkien and Malay. All functions from Hokkien appear to be transferred over to BM by being transferred over directly from BM, and by virtue of having compatible functions and forms in Malay (except for the fact that experiential perfect *kè* is postverbal rather than preverbal). However not all Malay forms and functions are found in BM (*akan* ‘will’, *nanti* ‘later’, and *telah* ‘already’). Regarding BM’s possible mixed language status, findings show that the two component languages are not individually compartmentalized in separate subsystems, and hence BM cannot be a mixed language.

Similar findings can be made by examining other subsystems in BM, such as its relativization strategies. There are two relative markers in BM, one being *punya* and the other being *yang* (see section 5.6.3). The connection between *yang* in BM and *yang* in Malay is straightforward enough. Example (605) is replicated here as (836). In both instances, *yang* serves as a postnominal relative clause marker.

(835) Malay

Musim [yang datang].
 season REL come
 ‘The season that is coming.’
 (Marsden 1812:50)

(836) BM

Ini budak [nang ter-teriak wolf].
 this child REL ACD-call.out
 ‘This boy that accidentally cried wolf.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:34.9-00:03:37.3)

On the other hand, even though *punya* in BM is lexically Malay in origin, it is clearly related to Hokkien *ê* (see Lim 1981, Pakir 1986, Lim 1988, Ansaldo and Matthews 1999, Lee 2012).

Malay *punya* and Hokkien *ê* are related in that both are possessive verbs. This is reflected in BM *punya* as well. Example (18) is replicated below as example (839).

(837) Malay

orang china punya arta
 person China **POSS** belongings
 ‘a Chinese person’s belongings.’
 (Marsden 1812:32)

(838) Hokkien

lâng ê kha
 person **POSS** foot
 ‘a person’s foot.’
 (Douglas 1873:99)

(839) BM

Peter punya bapak,
 Peter **POSS** father
 ‘Peter’s father,’
 (Victor, oai:scholarspace.manoa.hawaii.edu: NL1-009, 00:45:01.7-00:45:06.9)

By extension, Hokkien *ê*’s function as a prenominal relative clause marker appears to have also been transferred to BM as a function of *punya* (Lim 1988, Lee 2012).

(840) Hokkien

Goá hō i [goá thān ê] lui.
 1.SG give 3.SG 1.SG earn **REL** money
 ‘I gave him/her the money that I earned.’
 (Lee 2012)

(841) BM

Gua nampak [orang tarék punya] chia
 1.SG see person pull **REL** car
 ‘I saw the car (rickshaw) that the man pulled.’
 (Lee 2012)

Thus both postnominal and prenominal relativization strategies are available in BM. Based on section 5.6.3 and Lee (2012), the relativization patterns in BM, Hokkien and Malay can be represented in the following table. It is clear from the table that not all patterns of relativization are transferred over from the component languages to BM. While Hokkien *ê* can be used to relativize possessor, BM *punya* is not able to do so.

	Subject	Direct Object	Indirect Object	Possessor
BM <i>punya</i> (prenominal)	+	+	+	-
Hokkien <i>ê</i> (prenominal)	+	+	+	+
BM <i>yang</i> (postnominal)	+	+	+	+
Hokkien <i>yang</i> (postnominal)	+	+	+	+

Table 48: Relativization patterns in Baba Malay, Hokkien, and Malay

In terms of the current mixed language debate, the patterns of relativization show that BM cannot be a mixed language structurally. There is not one but two component languages that underlie the language's relativization system. In addition, not all patterns of relativization from both component languages are found in BM. One of the component language's subsystem, namely that of Hokkien, is no longer intact in the resultant language. This makes the case for BM being a mixed language even less convincing.

Regarding processual matters, the relexification hypothesis that Muysken (1981) proposes for the mixed language, Media Lengua, cannot apply directly in any of the above instances. More discussion ensues in the next section. Essentially, while BM may appear to share similar

sociohistorical characteristics as other mixed languages, the lexical and grammatical examples here show that the language is structurally dissimilar to any recognized mixed language.

7.4 BM as a creole

It is the view of this chapter that although the notion of a creole is a tenuous one, it is useful if it refers to a group of languages that have more in common with each other than differences (Bakker 2003 on mixed languages). That being said, while it will be shown that BM does not fit the most traditional notions of a creole, the concept of a creole is still an evolving one, and BM best fits into these expanding if fuzzy boundaries.

Thus far, different modern definitions have been proposed for creoles. Some of these consider sociohistorical factors, others take into account structural features. These differences arise in no small part due to the varying approaches to creole formation—the universalist approach (e.g. Bickerton 1981, Bickerton 1984, Coelho 1880–6), the substratist approach (e.g. Adam 1883, Muysken 1981, Lefebvre 1998, Lumsden 1999, Siegel 1999) and the superstratist approach (Faine 1936, Chaudenson 1974, Mufwene 1996, Mufwene 2001, DeGraff 2001, DeGraff 2003). In brief, universalists usually focus on shared properties of creoles that arise because of natural mechanisms; substratists attribute creole features to the substrate languages; while superstratists typically maintain that the creole is a variety of its lexifier language. A creole would consequently mean different things to different types of creolists. To a universalist, a creole would be a contact language that shows particular features; to a substratist, a creole would be a language that has derived structural or semantic information from another; and to a superstratist, a creole may invoke a particular sociohistorical background. These inferences are but brief generalizations. More detailed discussions regarding these various theories ensue in sections 7.4.4.1, 7.4.4.2, and 7.4.4.3.

Leaving the different approaches to creole formation aside for now, another area of disagreement involves the distinction between pidgin and creole. For researchers such as Holm, a “*creole* has a jargon or a pidgin in its ancestry” (italics in original), and is “spoken natively by an entire speech community” (2000: 6). For others, nativization does not differentiate a pidgin from a creole.

Mufwene (2001) states that pidgin and creoles have separate paths of development, while Chaudenson (1986) and Mühlhäusler (1992) are of the view that creoles expand from earlier stages after speakers begin using them as the primary mode of communication. Recognising these fuzzy boundaries that inadvertently surround these labels of contact languages, Thomason (1997b) proposes a prototypical definition of creoles. She states that these develop in situations that comprise more than two speaker groups, when neither group has “the need, the desire, and/or the opportunity to learn any of the other groups’ languages” (Thomason 1997b:78). The difference between prototypical creoles and prototypical pidgins is the “social feature of primary vs. restricted communicative functions” (Thomason 1997b:79). Consequently, a prototypical creole would be one that is the main language of the speech community and learnt as a native language, while a prototypical pidgin serves only limited functions. Hence, a prototypical creole would usually have more extensive structure and lexicon (similar to a full-fledged language) (Thomason 1997b). Such a view differs from that of researchers such as McWhorter (2005) who focuses mainly on the structural traits of creoles.

McWhorter (2005) states that prototypical creoles are unique in their combination of three traits, which are characteristic of languages that are young, the main assumption here being that creoles evolved more recently from pidgins. These traits are little or no inflectional affixes, little or no tone for the purpose of lexical and morphosyntactic differentiation, and non-compositional derivation. While by these criteria, BM is included in McWhorter’s (2005:52) list of prototypical creoles, his definition is not uncontentious. First of all, the assumption that creoles emerged from pidgins has been debated intensively (Mufwene 2001). Second, McWhorter (2005) does not recognize substrate or superstrate influence in creolization itself, and states that these types of influence at the time of genesis and after the time of genesis may cause the creole to not show these three traits. This is problematic for it is extremely difficult to prove in many instances that many creoles had all three traits at one point. Third, McWhorter’s (2005:140-141) list of sociohistorical circumstances, which according to him would result in these prototypical creoles with particular structural traits, cannot apply to some of his example languages neatly (such as BM itself). These sociohistorical requisites are that the creole should not be a compromise between closely related languages, that it should have emerged in relative isolation from its

lexifier, and that it should not be in a close diglossic relationship with its lexifier; presumably these factors would introduce too much structural influence from the component languages. This sociohistorical narrative cannot apply to BM since BM emerged where its lexifier, Malay was and is spoken, and BM is in a diglossic relationship with Malay in Malacca, where BM is a language spoken at home and among friends and Malay is an official language used for administrative purposes and as a medium of education in Malay schools. Overall, the role of McWhorter's (2005) sociohistorical narrative appears to be secondary to his use of structural traits in the identification of a prototypical creole, and it is strange that he would consider BM to be a prototypical creole according to his criteria (2005:141).

Hence, while the notion of a prototype is useful, it is difficult to define creoles strictly by structural traits (see Mufwene 1977). Instead, a view that takes into account both sociohistorical and structural factors, while still acknowledging fuzzy boundaries, may be preferable. Further discussion ensues in section 7.4.4. In the meantime, it is useful to note that the conceptualization of BM as a creole is not new. Such a notion has been proposed by Lim (1981, 1988), Ansaldo and Matthews (1999), Ansaldo et al (2007), and more recently by Shih (2009).

7.4.1 Lim (1991, 1988)

Based mostly on his fieldwork with MBM, Lim (1981, 1988) posits a possible connection between Bazaar Malay (then a lingua franca used for trade since the 17th and 18 century) and BM. He points out both languages had been previously conflated as 'Low Malay' by Shellabear (1913). While noting that Bazaar Malay has a more variable structure and a smaller lexicon than BM, Lim (1981, 1988) supports the notion that the Chinese had a large part to play in the creation of both Bazaar Malay and BM. He is also of the opinion that since some form of reduced Malay must have existed in both Chinese homes as well as in the trading community, hence BM would have not likely developed from the pidgin now known as Bazaar Malay. Instead, BM may have arisen quickly from an unstable and variable pre-pidgin continuum. Based on Hyme's (1971) notions that creolization involves complication, expansion and extension, in addition to convergence, BM's status as a possible creole is evaluated. In this instance, convergence "refers to the mixture of linguistic elements that is found in pidgin and creoles at

each of the phonetic, the lexical, the syntactic and the semantic levels” (Lim 1988:12). Given the fact that there is no prior documentation of a pidgin or a pre-pidgin continuum, it is difficult to demonstrate whether it is complication or simplification, or expansion or reduction that takes place. Lim however identifies several areas of convergence.

At the phonological level, Lim compares the phonemic inventories of MBM with those of Hokkien and Malay. With regard to consonants, Lim notes that there is no sign of admixture as the consonantal system is “almost exactly congruent with that of Malay, and there is no interference from the Hokkien system whatsoever (1988:15). Note again that he does not include glottal plosive [ʔ] in his consonant chart of BM, and that it may not have been his intention to leave it out, especially considering that he states that the system of BM is congruent with that of Malay, by which he means the standard Bahasa Melayu (see section 6.1) — the glottal stop exists both in Hokkien and in Malay. The consonant chart Lim (1988) presents of MBM is entirely similar to that of both SBM and MBM in this grammar (see sections 3.1.1 and 6.1).

Hokkien			Labial	Dental	Velar	Glottal
Stops	unasp.	v'less	p	t	k	ʔ
		voiced	b	d	g	
	asp.	v'less	p ^h	t ^h	k ^h	
Affricates	unasp.	v'less		tʃ		
		voiced		dʒ		
	asp.	v'less		tʃ ^h		
Fricatives		v'less		s		h
Nasals		voiced	m	n	ŋ	
Laterals		v'less		l		

Malay		Labial	Dental	Palatal	Velar	Glottal
Stops	v'less	p	t		k	ʔ
	voiced	b	d		g	
Affricates	v'less			tʃ		
	voiced			dʒ		
Fricatives	v'less	f	s	ʃ		h
	voiced		z			
Liquids	voiced		l	r		
Nasals	voiced	m	n	ɲ	ŋ	
Semi-vowels		w		j		

Table 49: Consonant charts of Hokkien and Malay (Lim 1988)

Baba Malay		Labial	Dental	Palatal	Velar	Glottal
Stops	v'less	p	t		k	(?)
	voiced	b	d		g	
Affricates	v'less			tʃ		
	voiced			dʒ		
Fricatives	v'less		s			h
Liquids			l	r		
Nasals		m	n	ɲ	ŋ	
Semi-vowels		w		j		

Table 50: Consonant chart of Malacca Baba Malay (Lim 1988)

Where vowels are concerned, Lim presents the following comparison. The vowel inventory of MBM is almost congruent with that of Malay, except for [ɛ], which results from raising discussed in section 3.3.5, although Lim (1981, 1988) attributes this to urbanity, and not “refinement”. While there are nasalized vowels in Hokkien, no nasal vowels are reported for BM. For example, /t^hĩa/ for ‘living room’ is reported as /tia/ (Lim 1988:15).

Hokkien	Front	Central	Back
High	i ĭ		u ũ
Mid	e ě		o õ
Low	a ã		ɔ õ

Malay			
	Front	Central	Back
High	i		u
Mid	e	ə	o
Low	(ɛ)	a	

Table 51: Vowel charts of Hokkien and Malay (Lim 1988)

Baba Malay			
	Front	Central	Back
High	i		u
Mid	e	ə	o
Low	(ɛ)	a	

Table 52: Vowel chart of Malacca Baba Malay (Lim 1988)

Thus, Lim (1988) states that unlike other creoles that may show substrate influence in phonology, BM shows little influence from Hokkien. Note however, that Lim's findings regarding MBM differ from this grammar's findings on SBM, where ϵ and also ɔ both exist in the vowel inventory, and are attributed to Hokkien substrate influence (see section 3.1).

Influence from Hokkien is syntactically and semantically more prevalent.

Lim (1988) attributes the relativizer form and function of *punya* to that of Hokkien, as this dissertation and Lee (2012) also do (discussed in sections 5.6.3 and 7.3.2), although he also identifies it as a possessive, temporal and locative marker. These can be interpreted as extensions of the relative marker (as this grammar has done). Alternate glosses are provided for Lim's

examples below. Free translation is only provided for the BM examples, since the Hokkien and Malay examples have equivalent meanings.

(842) BM

*Sek Po **punya** keréta*
REL car
'Sek Po's car'
'The car that is Sek Po's (alternate gloss).'

Hokkien

Sek Po ê chia
REL car

Malay

Keréta Sek Po
car

(Lim 1988: 17)

(843) BM

*Sini **punya** orang*
here REL people
'The people of this place.'
'The people that are here.' (alternate gloss)'

Hokkien

Chit-tau ê lang
here REL people

Malay

Orang yang di sini
person REL Prep here

(Lim 1988: 17)

Note however, that not all relativization functions that exist in Hokkien do so in BM. While Hokkien *ê* can be used to relativize possessor, BM *punya* is not able to do so (see section 7.3.2). Other functions that are transferred over include the *kasi* benefactive and causative constructions (see section 5.2.4.2) among others. The following benefactive and causative examples are presented in examples (844) and (845).

(844) BM

Dia-orang ***kasi*** *gua* *tau*
 3.PL **let** 1.SG know
 ‘They let me know.’

Hokkien

in-lâng ***hō*** *gúa* *chai*
 3.PL **give** 1.SG know

Malay

Mereka *memberi* *tau* *kepada saya*
 3.PL inform know to 1.SG
 (Lim 1988: 20)

(845) BM

Dia *pekék-pekék* ***kasi*** *gua* *terperanjat*¹⁰⁰
 3.SG shout-shout **cause** 1.SG shock
 ‘She shouted about, shocking me.’

Hokkien

i *jióng* ***hō*** *gúa* *chuak*
 3.SG shout **cause** 1.SG shock

¹⁰⁰ The glosses I have used for these examples differ from Lim (1988)’s, wherein *pekék* is glossed ‘scream’ rather than ‘shout’, and where *dia pekék-pekék* is interpreted as ‘her screams’. Since *pekék-pekék* follows a subject directly, and these two elements are not mediated by possessive *punya*, it is preferable to treat *pekék-pekék* as a tentative verb form, rather than a plural form of a noun.

Malay

Pekek-nya memeranjatkan saya
shout-3.SG shock 1.SG
(Lim 1988: 21)

In addition to substratal influence in the form of the *punya* relative marker, and the *kasi* causative and benefactive constructions, Lim (1988) addresses the fact that the adversative passive *kena*, which has been derived from Malay, is also used as a non-volition marker, and that this use appears to have been derived from Hokkien *tioq*. It is important to note that Hokkien *tioq* is also a passive marker, and its congruence with Malay *kena* may have triggered the use of *kena* as a non-volition marker. Compare the congruence between the passive function in BM, Malay and Hokkien in examples (846) and (847), as well as the non-volition function of *kena* and *tioq* in examples (848) and (849). It is clear that the non-volition function of *kena* in BM is related to that of Hokkien *tioq*.

(846) Malay (also BM)

Dia kena pukul
3.SG PASS hit
'He was hit.'

(847) Hokkien

i tioq phah.
3.SG PASS hit
'He was hit.'

(848) BM (but not Malay)

Gua kena pegi.
1.SG **subjected.to** go
'I had to go (i.e. I had no choice).'

(849) Hokkien

Gúa tioq khi
1.SG **subjected.to** go
'I had to go.'
(Lim 1988: 21-22)

Aside from these Hokkien influences in syntax-semantics, which Lim takes to be evidence of convergence between Hokkien and Malay, he also presents evidence of convergence in terms word order, citing that the topic-comment order that is common in Hokkien is also common in BM (see section 5.6.9). Where the lexicon is concerned, BM's pronominal system is presented as evidence of convergence, since it derives words from both Hokkien and Malay and has calques from Hokkien. (see section 4.1.3 and 7.3.2). For example, third person plural *dia-orang*, literally '3.SG-person' is derived from *i-lâng* '3.SG-person' in Hokkien.

Hence, identifying convergence as a dominant phenomenon in BM, Lim (1988) states that BM as a creole.

7.4.2 Ansaldo and Matthews (1999) and Ansaldo et al. (2007)

While Ansaldo and Matthews (1999) provides evidence of substratal elements from Hokkien, showing that structures in BM are in accordance with Thomason and Kaufman's (1988) notion of substratum interference, Ansaldo et al. (2007) disagree with their views that contact varieties such as BM (which they regard to being similar to a creole) have to be associated with notions of imperfect learning or break in the transmission of the lexifier. They also argue that the emergence of such contact varieties do not have to be associated with "extraordinary social conditions of dominated, oppressed, or underprivileged populations" (Ansaldo et al. 2007: 204). Instead, the features in BM can be explained by the language's ecology, in which an admixture of language materials from different sources compete for selection.

The data that Ansaldo and Matthews (1999) draw on are oral data recorded by Lau (1984) and Pakir (1986), Gwee's glossary (1988), as well as a play written by Chia (1989). They cite Thomason and Kaufman (1988), stating that substratum influence can be differentiated from borrowing, and that borrowing begins with the lexicon, while in substratum influence, structural interference and borrowing in the lexicon can overlap. Ansaldo and Matthews (1999) then show examples of structural influence, in areas such as passivization, the attributive construction (analyzed here as relative clauses, see section 7.4.1 for discussion of Hokkien substrate influence

in passivization and relative clause formation), existential *ada* constructions, and the tense-aspect system.

While Ansaldo and Lim (1999) state that *ada* and its Hokkien equivalent, *ǔ*, can mean ‘have’ or ‘exist’ in both Hokkien and Malay, they attribute its use as a preverbal auxiliary to Hokkien.

Ansaldo and Lim (1999) cite examples such as the following.

(850) BM

Lu ingat sekarang punya orang ada ka berchaya ini chiong-chiong mia hal?
2.SG think now POSS people **have** Q¹⁰¹ believe this taboo-taboo POSS thing
‘Do you think today’s generation actually believe this stuff about taboos?’

(851) Hokkien

Lì ǔ siū bèq khǐ bou?
2.SG **have** think will go NEG
‘Are you thinking of going?’
(Bodman 1958:2)

Note however, that while this chapter agrees that the uses of *ada* are more alike that of Hokkien than Malay, *ada* as a progressive marker, has been noted in Malay (see Marsden 1812). Instead, the perfective use of *ada* may demonstrate Hokkien influence better, since this function does not exist in Malay. This can be demonstrated by the following examples. Example (489) is replicated here as (852).

(852) BM

Gua ada tutop.
1.SG **PFV** CLOSE
‘I closed (the door).’
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-043, 00:22:15.8-00:22:16.7)

(853) Hokkien

Guà ǔ khì hê-chiā-cām
1.SG **PFV** go train-station
‘I went to the train station.’
(Bodman 1955:18)

¹⁰¹ The use of *ka* as a question marker has not been noted in any data elicited for this grammar.

Other uses of *ada* and the implications of these substratum influence is discussed in section 7.4.4.2. Where the tense and aspect system is concerned, Ansaldo and Matthews (1999) state that besides being consistent with typical creole typology, particular constructions appear to have been influenced by the Hokkien tense and aspect system. They cite the data in examples (854) and (855). While both *mo* and *bèq* [transcribed as *beh* by others such as Douglas (1873)] share a common meaning ‘want’ in Malay and Hokkien respectively, only Hokkien *bèq* has a future meaning. This future meaning is transferred to BM *mo*, even though it does not exist in Malay.

(854) *Apa mo jadi*
 what **FUT** happen
 ‘What will happen?’

(855) *Guà bèq khi hê-chiā-cām*
 1.SG **will** go train-station
 ‘I will (am expected) to go to the train station.’
 (Bodman 1955:97)

This observation has also been made by Lim (1988) and in this grammar. Ansaldo and Matthews (1999) go a step further in proposing that there is a difference between *nanti* ‘later’ future constructions and *mo* ‘want’ future constructions (see section 5.2.5.1), stating that *nanti* is a general future construction, while *mo* is immediate/intentional future. The data covered in this dissertation does not show that *mo* indicates immediacy, although examples with *mo* do signal intention, as *mo* is also a deontic modality marker (see section 5.2.2).

Given the multiple structural inputs from Hokkien, Ansaldo and Matthews (1999) conclude that BM can be accounted for by substratal influence rather than by borrowing, based on distinctions made by Thomason and Kaufman (1988). They also propose that BM underwent creolization, due to the historical circumstances in which Malay women and Chinese men formed a community, as well as due to characteristics that show processes associated with creolization, such as morphological reduction (as with the loss of verbal prefixes) and grammaticalization (for example, the expansion of possessive *punya*’s use). They note that BM appears to be atypical, being a dual-input creole instead of one based on at least three languages. They name Pitcairnese as another dual-input creole (English-Tahitian), as well as dual-input pidgins such as Pidgin Delaware and Ndyuka Trio.

Whereas Ansaldo and Matthews (1999) are more concerned with structural matters, and demonstrate the validity of particular concepts put forth by Thomason and Kaufman (1988), Ansaldo et al. (2007) are focused on sociohistorical issues, and they propose that contrary to Thomason and Kaufman (1988), notions of imperfect learning or break in the transmission of the lexifier may not be necessary in accounting for the formation of restructured vernaculars. In the same vein, Ansaldo et al. note that “the emergence of such contact varieties [do not have to] be associated with extraordinary social conditions of dominated, oppressed, or under-privileged populations” (2007:204).

While Ansaldo and Matthews (1999) identify BM as a creole, Ansaldo et al. do not outrightly do so. They do, however, draw parallels between BM and other varieties that emerged in situations of “combined trade and settlement colonization”, stating that BM “offers useful comparative information in relation to ‘classic creoles’ of especially the American-Caribbean region and the Indian Ocean” (2007: 206). In contrast to populations associated with these prototypical creoles, speakers of BM were neither dominated by others, oppressed or underprivileged, nor had they ever undergone traumatic geographical displacement. Instead, citing Tan (1988), Ansaldo et al (2007) note that speakers of BM were an influential class of Chinese capitalists. Referred to as the “King’s Chinese”, the Peranakans were among the first to be educated in English, and they played significant roles as intermediaries between the locals, the Asian newcomers, and the British administrators and merchants (Tan 1988, Ansaldo et al. 2007). The Peranakans were therefore “neither and economically dominated nor an underprivileged group” (Ansaldo et al. 2007: 212). In addition, Ansaldo et al. argue that there was no imperfect learning or break in the transmission of the lexifier, as “the creators of BM were always minorities surrounded by, and in many cases cohabitating with, speakers of colloquial Malay” (2007: 218). In relation to this, there is structurally little that would suggest “an evolution characterized by simplification (Ansaldo et al. 2007: 218). According to the authors, the structures in BM are better explained as having been selected and adapted from various sources in its ecology. For example, while kinship terms are derived from Hokkien (see Appendix A: Kinship terms), the pronominal system has both Hokkien and Malay influences, as also mentioned by Pakir (1986), Lim (1988),

and explained in section 7.3.2.¹⁰² While these are interesting observations, note that there is no defining explanation for why particular structures or lexicon are influenced by Hokkien, and why others are influenced by Malay.

7.4.3 Shih (2009)

The most recent study of BM is that of Shih (2009), who conflates both SBM and MBM as a single object of study, and shows that there is Hokkien substratal influence in BM. She observes specific aspects of the language's phonology, morphology, lexicon, and syntax, based on interviews conducted with seven speakers of BM, both in Malacca and Singapore, as well as based on the *Kitab Perjanjian Bharu*, literally 'scripture testament new', originally written by Shellabear and other members of the British and Foreign Bible Society in 1913.

With regard to phonology, Shih (2009) states that BM is more similar to Malay on the surface, but it shows Hokkien influence, through the lack of several fricatives in its consonant system (for example [f], [x] and [z]), and that BM words are at most trisyllabic, similar to Sinitic languages. These points are disputable since [f] is a much more recent development in Malay, the usage of which may have expanded alongside the borrowing of Arabic loanwords into Malay¹⁰³; and [x] and [z] were also introduced into Malay via Arabic loanwords. In addition, there are quadrisyllabic words in BM, as with affixed forms such as *ter-balék-kan* 'MVT-turn over-TR' and reduplicated forms such as *kupukupu* 'butterfly'. There are inaccuracies in the section on morphology. Shih (2009) states that there is a lack of affixation, in line with Hokkien, but this is untrue. As noted in this grammar, affixes such as transitive marker *-kan*, and accidental and movement marker *ter-* are productive, and commonly found in BM data. With regard to conversion processes, Shih (2009) notes that Malay roots can be used for different categories in BM, and that this is also true of Hokkien roots, but not Malay and Indonesian ones. This is more accurate, since *jalan* would mean both 'walk' and 'road' in BM, whereas *berjalan* would mean

¹⁰² Some of these observations differ from those of this grammar, for example, while they state that demonstrative precedes noun in BM (following Hokkien rather than Malay), SBM data shows that both demonstrative noun and noun demonstrative orders are acceptable, and MBM data shows a preference for the noun demonstrative order.

¹⁰³ While *f* occurs now in modern Malay words such as *fikir* 'think', the only words that utilise *f* in older High Malay were loanwords from Arabic. This is attested to by Crawfurd's (1852) dictionary.

‘walk’, but not ‘road’ in Malay and Indonesian. Significantly more Hokkien influence is found in the lexicon. Similar to Pakir (1986), Shih (2009) notes that kinship terms (see Appendix A: Kinship terms) and cultural terms¹⁰⁴ have been derived from Hokkien.

In terms of syntax, Shih (2009) focuses on the usages of *kasi* ‘give’ derived from Hokkien *hoo*. Shih (2006) states that *kasi* can be used as a ditransitive verb meaning ‘give’, as a goal marker in a ditransitive construction, as a causative verb, in special pivotal constructions, as a verb in imperative verb-complement constructions, and as an agent marker in passive constructions. The special pivotal construction mentioned by Shih (2006) is the same as the benefactive construction recognized by the grammar, and the usage of *kasi* in imperative verb-complement construction is basically an agentless imperative causative construction. Ditransitive, benefactive, and causative constructions are described in section 5.2.4. The notion that the benefactive and causative constructions have been derived from Hokkien are also described by Lim (1988) and in section 7.4.1. The *kasi* passive construction is described in section 5.2.3. Shih (2006) provides the following comparison in examples (856) and (857). Essentially, other structural extensions of *kasi* beyond the ditransitive verb meaning ‘give’, such as the causative, the benefactive, and the passive, are relexified versions of Hokkien constructions.

(856) BM

Gua kasi dia tipu
 1.SG PASS 3.SG cheat
 ‘I was cheated by him.’
 (Shih 2009)

(857) Hokkien

Guà hō i phiǎn
 1.SG PASS 3.SG cheat
 ‘I was cheated by him.’
 (Shih 2009)

Hence, based mostly on structural evidence that there is substratal influence from Hokkien, Shih concludes that BM is a creole, instead of a dialect of Malay (in the sense of a related language).

¹⁰⁴ See Pakir (1986) for an in-depth list.

In addition, Shih (2009) goes further in saying that BM is a creole that developed from the lingua franca, Bazaar Malay, since it was colloquial Malay that immigrants that arrived in the region were exposed to, before any intermarriage with the Malays. This last conjecture is reminiscent of Lim (1988)'s notion that BM may have arisen quickly from an unstable and variable pre-pidgin continuum. This raises a disputed issue of whether the founder effect would be stronger or if swamping by later migrants would have taken place (Gordon et al. 2004: 247). With founder effect, the structural characteristics of the language would have been mostly determined by the founder population (Mufwene 1996), whereas with swamping by later migrants, the effect of later migrants is so significant that traits from earlier varieties are swamped out (Lass 1997: 206). In BM's scenario, early speakers who were exposed to Bazaar Malay might have been overwhelmed by later numbers of those marrying Malay speakers. The view taken by this grammar is that this is not entirely implausible, since it is congruent with the narrative that BM was formed by the intermarriage of Hokkien-speaking traders (who must have had some knowledge of Bazaar Malay for the purpose of trade, prior to settling in the Malay Peninsula) and Malay-speaking women.

7.4.4 Creole theories and BM

Thus far, arguments regarding whether or not BM is a creole have focused mostly on structural traits that show Hokkien substratal influence, except for Ansaldo et al. (2007), who have used BM to challenge sociohistorical notions of what a creole is. This appears to put BM in the periphery of "creole-hood". In effect, BM appears to be a less than prototypical creole by other standards. For example, Thomason's notion of prototypical pidgins and creoles include languages that develop in situations that comprise more than two speaker groups, when neither group has "the need, the desire, and/or the opportunity to learn any of the other groups' languages" (1997b:78). The dominant language in trading was colloquial Malay, and not the Malay spoken as anyone's first language (Lim 1981, 1988). As mentioned in section 7.2.1, while early speaker groups most likely had the opportunity to learn either groups' language, it is highly likely that they had neither need nor desire to in the intermarriage scenario. However, this comes

with a caveat that it is impossible to establish with certainty if the Malay-speaking women did not learn Hokkien, or if the Hokkien-speaking men did not learn Malay.

In addition, it is disputable if BM involves the contact of only two languages, or more than two languages (see section 7.2.1). Earlier adstrates include Dutch and Portuguese, while later ones include English and Cantonese, among other languages. Yet, it is clear from BM, that these other languages left very little imprint on it, no more than casual borrowing in other language situations. Again, the number of languages involved is usually a peripheral concern rather than a central one, and as a case in point, Ansaldo and Matthews (1999) highlight Pitcairnese as a dual-input creole comprising English and Tahitian.

Next, it is interesting to consider if BM did indeed arise from Bazaar Malay. Thomason states that the difference between prototypical creoles and prototypical pidgins is the “social feature of primary vs. restricted communicative functions” (1997b: 79). In an ideal theoretical world, Bazaar Malay would be the pidgin that BM originates from—Bazaar Malay is a pidgin that was used in the limited domains of trade and inter-ethnic communication, whereas BM is used as the main home language of the Peranakans. If this was the case, it would make sense to compare the structures of BM and Bazaar Malay as if one evolved from the other, since a prototypical creole would usually be more extensive in structure and lexicon than a prototypical pidgin (Thomason 1997b). However, there is no substantial evidence to suggest that BM did evolve from Bazaar Malay, or even a pre-pidgin continuum. On a separate but related note, one might ask which variety of Malay influenced BM most. It is necessary to consider the fact that in an intermarriage scenario between Chinese traders and local women, there would have been no impediment to non-colloquial High Malay, especially if many of these women spoke Malay natively. It is easy to find evidence of High Malay structural influence in BM. For example, as mentioned in section 7, BM uses noun classifiers and *yang* relative clauses. These occur in High Malay but not in colloquial Malay (Aye 2005). The structural evidence shows that it is necessary to compare BM with High Malay, and that it is not possible to adopt a strong stance regarding whether or not BM evolved from Bazaar Malay. That these Hokkien-speaking males who intermarried with Malay-

speaking females were heavily exposed to Bazaar Malay through trade before any intermarriage occurred remains an interesting conjecture.

At the end of the day, definitions appear to be somewhat arbitrary. For example, creolehood depends on whether or not one accepts a dual-input contact language to be a creole (see Ansaldo and Matthews 1999, Thomason 1997b), and whether or not a creole is a nativized version of a pidgin (Holm 2000), or a contact language that has more communicative functions than a pidgin (Chaudenson 1986, Mühlhäusler 1992, Thomason 1997b). Considering that creolists who subscribe to different views of creole formation have different ideas about what constitutes a creole, it is immensely more interesting to question how the structural and sociohistorical nature of BM fits in with theories of creole formation. Findings show that BM is far more like a creole when processes of creole-formation are considered, and that BM can also be used to provide perspectives on the various creole formation theories.

7.4.4.1 The universalists

It is said that “creole studies blossomed in the 1880s” (Holm 2000:27). One of the earliest theoretical positions taken on the origin of creoles was that of Coelho who attributed the form of creoles to certain universal tendencies in second language learning by adults. Coelho states that the “Romance and creole dialects, Indo-Portuguese and all the similar formations represent the first stage or stages in the acquisition of a foreign language by a people that speaks or spoke another” (1880–6:193), and that these languages “owe their origin to the operation of psychological or physiological laws that are everywhere the same, and not to the influence of the former languages of the peoples among whom these dialects were found” (1880–6:195). Some of the evidence Coelho provides are the lack of overt functional categories, as well as the use of preverbal progressive marker generally found in creoles. It is interesting that one of the characteristics that Coelho points out is the use of *misti* to mean ‘need’ in several creoles, because BM uses *mesti* to mean ‘must’ (see section 5.2.2). Not surprisingly, Holm (2000) highlights that these were mainly Portuguese-based creoles, Portuguese being one of the adstrates spoken in the same environment as BM. Crucially, Coelho does not provide any

explicit explanation of how psychological or physiological laws may constrain creole formation. It is thus difficult to evaluate the strengths of such an approach. Even if the preverbal progressive marker does occur in BM (see section 5.2.5.4), there is no evidence that this feature is due to some psychological or physiological law that is constraining second language acquisition.

Another more widely-known 20th century version of a universalist-type theory is Bickerton's Language Bioprogram Hypothesis (LBH). Whereas Coelho views creole formation to be the result of second language acquisition, Bickerton views creole formation as a product of first language acquisition. More specifically, the "LBH claims that the innovative aspects of creole grammar are inventions on the part of the first generation of children who have a pidgin as their linguistic input, rather than features transmitted from existing languages" (Bickerton 1984:173). Drawing from his observations of Hawai'i Creole English, also known as Pidgin to its speakers, Bickerton explains that creoles are created abruptly, pointing to language acquisition by children of imported plantation laborers and slaves. The parents of these children spoke different languages, and could only communicate with each other in a very limited fashion. Consequently, the children did not have much linguistic material to work with, and they would have had to rely on their innate linguistic faculty to develop this unstable language into a full-fledged language (Bickerton 1981, 1984). As such, similar to Coelho's notion of universality, the LBH is based on the idea that the "inventions show a degree of similarity, across wide variations in linguistic background, that is too great to be attributed to chance" (Bickerton 1984:173). Thus, in support of the LBH, Bickerton (1981) presents a set of features that are presumably found in creoles, but not in the languages that precede the creoles. These include for example, the Subject Verb Object (SVO) word order, a category of adjectives that form a subcategory of stative verbs, relativization that contains no surface marker of relativization, as well as verb serialization among others.

There are several weaknesses with the basic assumptions of Bickerton's LBH. First, in its narrow view, a creole would necessarily have been formed on a plantation. This excludes then, any creoles that would have been formed in other domains, such as through trade, colonization, and, intermarriage between two different ethnic groups, as in the case of BM, which is a language that

was formed via the intermarriage of Chinese traders and Malay indigenous women in the Malay between the 15th to 19th century (Pakir 1986). Next, it would be necessary to evaluate the claim that the structures in creole are “shared by all locally born (i.e. creole) speakers and no immigrant (i.e. pidgin) speaker” (Bickerton 1984:174). Roberts (1998) demonstrates that some of the so-called universal features in Hawai‘i Creole English, which this particular universalist theory is mostly founded upon, are connected with the first generation speakers’ languages. In addition, some of these features exist in contact languages that have not become nativized, such as Ghanaian Pidgin English (Huber 1999). Where BM is concerned, any direct link between the creole and Bazaar Malay is purely speculative, and there is no known documentation of a pidgin version of BM. The language was not written until the late 19th century when the Peranakans began publishing newspapers, magazines, and translating Chinese stories (Yoong and Zainab 2002). While it is not possible to evaluate whether there are universal features in BM that can or cannot be connected to a pre-nativized version, it is possible to evaluate if these universal features can be found in BM, and if they are found, whether or not they can be linked to the original languages that were spoken by the original populations, in this case, Malay and Hokkien. Out of the universal features listed at the end of the last paragraph, two are not found in BM, while two can easily be attributed to the languages that were spoken by the original populations. BM does not have a special category of adjectives that form a subcategory of stative verbs, neither does it have relativization that does not make use of a surface marker for relativization. In fact, relativization in BM makes use of two different markers of relativization, depending on whether relativization is pre-nominal or post-nominal (see sections 5.6.3 and 7.3.2). The following are examples of pre-nominal and post-nominal relativization in BM. Examples (610)/(841) and (836) are replicated here as (858) and (859).

(858) BM

<i>Gua</i>	<i>nampak</i>	[<i>orang tarék</i>	<i>punya</i>]	<i>chia</i>
1.SG	see	person pull	REL	car

‘I saw the car (rickshaw) that the man pulled.’
(Lee 2012)

- (859) *Ini budak [nang ter-teriak wolf].*
 this child REL ACD-call.out
 ‘This boy that accidentally cried wolf.’
 (Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-030, 00:03:34.9-00:03:37.3)

The surface markers *punya* and *yang* are used in prenominal and postnominal relativization respectively, and as mentioned in section 7.3.2, these constructions have counterparts in Hokkien and Malay. Examples include (840) and (835), replicated here as (860) and (861).

- (860) Hokkien

Goá hō i [goá thān ê] lui.
 1.SG give 3.SG 1.SG earn REL money
 ‘I gave him/her the money that I earned.’
 (Lee 2012)

- (861) Malay

Musim [yang datang].
 season REL come
 ‘The season that is coming.’
 (Marsden 1812:50)

Thus, in addition to not conforming to Bickerton’s universal, the relativization data also suggests that the contact language’s structure may be attributed in some way to languages spoken by the original population. This is something that Bickerton’s LBH does not acknowledge. In a similar vein, that a contact language has the SVO word order does not necessarily mean that this is a universal rule. In the case of BM, the SVO word order could be attributed to the fact that the SVO word order exists in both Malay and Hokkien. Similarly, the fact that verb serialization, as exemplified by Bickerton’s (1981:275) example of “Ug take meat give Og”, occurs in BM can simply entail the fact that it occurs in Hokkien too.

- (862) BM

Gua suap tu baby minum susu.
 1.SG feed that drink milk
 ‘I feed that baby milk to drink.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-045, 00:00:17.6-00:00:20.5)

(863) Hokkien

Líê chheh chioh goá khoà chit ē
2.SG book lend 1.SG see one time
'Your book, lend it to me to read once.'
(Chiang 1940: 31)

The above examples show that it may be necessary to consider the input of other languages in the formation of a contact language. This is an issue that the substratists and the superstratists contend with.

7.4.4.2 The substratists

Similar to the universalist approach, the substratist approach was also first articulated in the 1880s. Adam (1883) compared Guiana Creole French to that of Trinidad and several West African languages, and compared Mauritian Creole French to the Malagasy language of Madagascar. Inaccurate as the second comparison is, he states that “the Guinea Negroes, transported-to those [Caribbean] colonies, took words from French, but retained as far as possible the phonology and grammar of their mother tongues” (Adam 1883:4).

A more modern day approach exists in the notions of relexification and substrate transfer. The term ‘relexification’ was introduced by Stewart (1962) who noted the relationship between Caribbean creoles and particular pidgin and creole languages spoken in Asia and Africa. Stewart (1962: 46) defines the process of relexification as one wherein the vocabulary of a source language of a creole is replaced (through widespread borrowing) by the vocabulary of another language, “while the original grammatical structure is preserved practically unchanged.” In a similar vein, Whinnom (1965: 522) suggests that:

Certain pidgins and creoles are relexifications of an advanced Portuguese pidgin, others may be relexifications of a more primitive Portuguese pidgin which originated as a relexification of Sabir, or possibly, directly of Sabir itself.

Another proponent of relexification is Muysken (1981: 61) who proposed that *Media Lengua*, a mixed language, arose due to a “process of vocabulary substitution in which the only information adopted from the target language in the lexical entry is the phonological representation.” Such an approach has since been adopted by Lumsden (1999) and Lefebvre (1998) in explaining creole formation. Lefebvre (1998: 16) explains that in the process of relexification, the syntactic and semantic features of a lexeme in a substrate lose their original phonological label and are relabelled with a “phonetic string” from the lexifier. Crucially, this takes place due to the influence of the creator’s first language (L1) or languages. This differs markedly from Bickerton’s LBH in which the role of creole formation is left to children, and influence from the first language is overtly not taken into account. Relexification is often compared to substrate transfer, which Siegel (2000) uses to account for the morphosyntactic properties of Hawai‘i Creole English. Siegel (1999) argues that the selection of superstrate features depends on factors such as perceptual salience, transparency and frequency. Very little else is said of the superstrate language. Conceptually, relexification views the process more from the perspective of L1 input, focusing on how second language (L2) items are incorporated into the learner’s system as labels for L1 derived semantic/ functional categories, while transfer views the effects of L1 influence from the perspective of the L2, focusing on how the input is changed under the L1 influence (Winford 2003:345). The two are essentially similar, as they emphasize the importance of L1 influence, and differ markedly from the LBH by focusing on the notion of L2 acquisition. Substrate transfer in the following discussion thus also refers to relexification.

BM is recognized as comprising Hokkien as a substrate language, and Malay as a lexifier language (Lim 1981, 1988, Holm 1989, Ansaldo and Matthews 1999, Shih 2009). The substrate language here corresponds to L1 in the above discussion, while the superstrate language corresponds to L2. It is not immediately clear why Hokkien would constitute L1, and Malay L2, in this instance. Usually, it is the language that is external to the local community that would have L2 status, as with English in Hawai‘i Creole English, or even French in Haitian Creole. In the case of BM, it was the Hokkien speakers who journeyed south to the Malay Peninsula, and yet, Hokkien corresponds to the substrate language, and hence L1, rather than L2 or the superstrate language. It would be necessary to postulate reasons as to why Malay was the target

language and not Hokkien. This may be due to the fact that the Chinese traders had to learn a variety of Malay for the purpose of trade, since it was a dominant lingua franca in the region between the 14th to 19th century (Sneddon 2003) It is more likely that the Hokkien-speaking traders had to learn the local language, rather than the local women having to learn the Hokkien. Unfortunately, the substratists have not articulated clearly how the sociohistorical background of a creole contributes towards its structure. This remains one of the main criticisms of these substratal approaches (Winford 2003:337), and it is an aspect that is better addressed by superstratists (see section 7.4.4.3).

Next, it is necessary to evaluate if substrate transfer can account for the data in BM. Fundamentally, researchers who have identified BM as a creole, have done so based on substratal transfer, even if the exact term is not used. For example, as demonstrated above in section 7.4.1, Lim (1988) attributes the prenominal relativizer *punya* construction, the causative/benefactive *kasi* constructions, and the topic-comment structure to the influence of Hokkien prenominal relativizer *ê*, causative/benefactive *hō*, and the topic-comment structure in the language, while Ansaldo and Matthews (1999) state that future marker *mo* is derived from Hokkien *bèq*. There are a number more structures that can immediately be attributed to substrate transfer. Within the aspectual system itself, tentative aspect verbal reduplication and perfective aspect *ada* have also been derived through substrate transfer.

As discussed in sections 5.2.5.7 and 7.3.2, tentative aspect in BM is indicated by verbal reduplication. Examples (503) and (504) are replicated below as (864) and (865).

(864) *Jalan jalan*
walk walk
'Take a walk.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:13.1-00:02:14.6)

(865) *Téngok téngok*
look look
'Take a look.'
(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:02:14.6-00:02:16.2)

While tentative aspect through verbal reduplication does not appear to occur in Malay, it is fairly common in Hokkien and other Sinitic languages (See Tsao 2004). Compare examples (864) and (865) with (866). These instances of verbal reduplication express the tentative aspect where the situation is “of short duration, and of little importance” (Smith 1991: 356).

(866) Hokkien

Goá beh khi kiá" kiá"
 1.SG FUT go walk walk
 ‘I want to go take a walk.’
 (Chiang 1940: 13)

Note that in addition to verbal reduplication signalling tentative aspect, other word classes can be reduplicated for metaphorical tentativeness in BM. Example (867) shows how adjectival reduplication is used to express a metaphorical tentativeness.

(867) *Tawair tawair*
 Tasteless tasteless
 ‘Kind of tasteless.’
 (Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-142, 00:11:12.7-00:11:15.0)

Adjectival reduplication that expresses metaphorical tentativeness is essentially also found in Hokkien. Note that this is different from reduplication that forms adverbs (see section 4.4.1). Tsao provides the following example.

(868) Hokkien

In lāupě sán sán a
 3.SG father thin thin PRT
 ‘His father is kind of thin.’
 (Tsao 2004: 295)

Hence in BM, tentativeness in aspect, and in adjectival reduplication result from substrate transfer from Hokkien. Similarly, substrate transfer accounts for the use of *ada* as a perfective aspect marker in BM. *Ada* is not used as a perfective marker in Malay (Marsden 1812). In their

most basic form, Malay *ada* and Hokkien *ū*¹⁰⁵ are possessive verbs. Unsurprisingly BM *ada* is also a possessive verb.

(869) BM possessive *ada*

Kambing dia ada lochéng.
Goat 3.SG has bell
'His goat has a bell.'

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-022, 00:01:44.8-00:01:47.3)

(870) Hokkien

Goá ū tāmpoh sū
1.SG has little matter
'I have a little matter.'

(Chiang 1940: 27)

(871) Malay

Raja itu ada s' orang anak nia perempuan
king that has one person child 3.SG female
'That king had one daughter.'

(Marsden1812: 58)

Other shared features between BM *ada*, Hokkien *ū* and Malay *ada* are the existential/location function, the copula function, and the progressive aspect function (Lee 2009). The following are examples of these share functions across the contact language, its substrate language, and its lexifier language. Examples (872) to (874) compare existential/location usage of *ada* and *ū*. The use of *ada* as an existential marker in BM is also mentioned in section 5.1.4.

(872) BM existential/location *ada*

Di tepi library ada satu keday jual kuay
PREP next EXIST one shop sell cake
'Next to the library, there is a shop selling cake.'

(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-107, 00:09:13.1-00:09:16.6)

¹⁰⁵ Following Douglas (1873) and Chiang (1940)'s orthographic transcriptions

(873) Hokkien

ēnhng ũ chaikuán
Amoy **EXIST** restaurant
'There are restaurants in Amoy.'
(Bodman 1955: 18)

(874) Malay

Ada s' orang raja
EXIST one person king
'There was a king.'
(Marsden 1812: 58)

Examples (875) to (877) demonstrate copula use of *ada* and *ū*, example (427) is replicated here as (875). See section 5.2.1 for more information on copula constructions using *ada*.

(875) BM copula *ada*

Rumah gua ada chanték tak
house 1.SG **COP** beautiful NEG
'My house is beautiful or not?'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-035, 00:02:41.9-00:02:43.7)

(876) *chit keng lūtiám ũ hó bô*
this CLF.BUILDING hotel **COP** good Neg
'Is this hotel good or not?'
(Bodman 1955: 61)

(877) *rumah ini ada baik*
house this **COP** good
'This house is good.'
(Lee 2009)

Examples (878) to (880) show how *ada* and *ū* are used to express the progressive aspect, example (496) being replicated here as (878). The use of *ada* as a progressive marker is mentioned in section 5.2.5.4.

(878) BM progressive *ada*

Lu ada bikin apa?
2.SG **PROG** make what
'You are making what?'
(Peter Wee, oai:scholarspace.manoa.hawaii.edu: NL1-042, 00:12:50.2-00:12:51.8)

(879) Hokkien

Lì ǎ sîu bəq khǐ bou?
2.SG **PROG** think want go NEG
'You are thinking of going, or not?'
(Bodman 1958: 2)

(880) Malay

Orang ada makan
people **PROG** eat
'The people are eating.'
(Marsden 1812: 58)

While all the possessive verb function, the existential/location function, the copula function, and the progressive aspect function are shared by BM *ada* and its lexifier and substrate equivalents, the perfective function is only shared by BM *ada* and its substrate equivalent *ū*. Example (487) appears here as (881), (853) as (882), and (880) as (883). The perfective use of *ada* is noted in section 5.2.5.2.

(881) BM perfective *ada*

Dia ada beli apple, bukan?
3.SG **PFV** bought no
'She bought an apple, no?'
(Victor, oai:scholarspace.manoa.hawaii.edu: NL1-051, 00:40:03.5-00:40:06.9)

(882) Hokkien

Guà ǎ khì hê-chiā-cām
1.SG **PFV** go train-station
'I went to the train station.'
(Bodman 1955:18)

(883) Malay

Orang ada makan
people **PROG** eat
* 'The people ate.'
'The people are eating.'
(Marsden 1812: 58)

In Malay, *sudah* ‘already’ is used to mark the perfective. This is also used in BM (see section 5.2.5.2). Hokkien *i-keng* which also mean ‘already’ is congruous in function. Returning to the subject matter of substrate transfer, perfective use has been transferred from Hokkien *ū* to BM *ada*, since it is not a feature of Malay *ada* (Lee 2009).

A final example of the influence of the Hokkien substrate is the form *mana ada* ‘where EXIST’; its usage is discussed in section 4.4.2. Example (718) is replicated here as (884). Compare this example with example (885)

(884) BM *mana ada*

Mana ada panday?

where EXIST clever

‘How is he clever (implying he is not)?’

(Jane Quek, oai:scholarspace.manoa.hawaii.edu: NL1-093, 00:21:32.7-00:21:34.2)

(885) Hokkien

Tó-loh ũ lâ-sâm?

where EXIST dirty

‘Where is it dirty (implying it is not dirty)?’

Given translation: I don’t think it’s dirty

(Teo 1996: 75)

The form *mana ada* appears to be a calque of Hokkien form *tó-loh ũ*, both literally indicating ‘where EXIST’, and connoting that the speaker thinks the listener’s earlier assertion is not correct. Evidence such as this appear to support the substratist position.

However, matters are often more complicated. Not all substratal semantic-syntactic structures are transferred to BM. Recall for example, from section 7.3.2 that not all patterns of relativization are transferred over from prenominal relativizer, Hokkien *ê* to BM *punya*. Specifically, while Hokkien *punya* is able to relativize possessor, BM *punya* is not able to do so. Interestingly, BM *yang* shows all relativization patterns associated with postnominal relativizer, Malay *yang* (see section 5.6.3 and 7.3.2). Table 48 is reproduced here as table 53.

	Subject	Direct Object	Indirect Object	Possessor
BM <i>punya</i> (prenominal)	+	+	+	-
Hokkien <i>ê</i> (prenominal)	+	+	+	+
BM <i>yang</i> (postnominal)	+	+	+	+
Hokkien <i>yang</i> (postnominal)	+	+	+	+

Table 53: Relativization patterns in BM, Hokkien, and Malay

The following examples contrasts BM *punya* with Hokkien *ê* in terms of possessor relativization. Example (616) appears here as (887).

(886) Hokkien *ê* possessor relativization

hit ê [pêngiú phah goá ê lāng]
 that REL friend hit 1.SG REL person
 ‘That person whose friends hit me.’
 (Lee 2012)

(887) BM *punya* possessor relativization

**itu (orang) punya kawan-kawan pukol gua punya orang*
 that (person) possess friend-friend hit 1.SG REL person
 ‘that person whose friends hit me.’
 (Lee 2012)

Lee (2012) constructed sentences such as example (887) based on what is expected of BM. It conforms to the language as it is necessary for the possessor to occur before *punya* ‘to possess’ in order to express that it was the ‘person’s friends’ who hit ‘me’. However, this would not be

compatible with relativization using *punya*, since *punya* has to precede the head it modifies in its most basic form in Malay. This is demonstrated by example (888).

- (888) Malay
- | | | |
|------------|---------------------|-------------|
| <i>Ali</i> | <i>punya</i> | <i>kuéh</i> |
| | possess | cake |
- ‘Ali’s cake.’
(Lee 2012)

Example (887) is further incompatible with the gap-type strategy that is used in relative clause constructions containing Hokkien *e* or BM *punya*, the gap-type strategy being one that does not provide any overt indication of the role of the head within the relative clause (Comrie 1989). In (887), *orang* ‘person’ in parenthesis would overtly indicate what the role of the head is within the relative clause. Due to these misalignments, BM speakers judge this sentence to be ill-formed, stating that it is not possible to produce a sentence with such a meaning in the language. Hence, properties of both the lexifier language and the substrate language prevent the transfer of the feature that would otherwise allow for the relativization of possessors. Where major approaches to creole formation are concerned, substrate influence on its own cannot fully account for resultant structures.

Outside of syntactic structure, the same is also true. Substrate transfer on its own cannot explain all phenomena in BM. For example, it fails to account for why most words or ‘phonetic strings’ for kinship in BM are derived from Hokkien. The paternal uncle who is older than one’s father is called *pék*, the one who is younger than one’s father is called *chek*, while the maternal uncle is called *ku* (see Appendix C: Kinship terms). These differs drastically from Malay, wherein these various relations can be subsumed under the word *pakcik* for addressing one’s uncle. Without discussing the sociohistorical background of these speakers—that these Peranakans held on to Chinese culture and tradition (for example, in ancestral worship, marriage rites, and in sending their sons for education in China before British colonialism made English-medium education popular), it would be difficult to understand why the Peranakans observe Chinese kinship terms.

An account of transfer that is solely focused on substrate transfer is limited. Lumsden acknowledges that “the RH [Relexification Hypothesis] is NOT a claim that relexification is the one and only process in creole genesis, nor is it argued that relexification can account for all the properties of creole languages” (1999:230, fn.7). That being said, the substratists have not provided processes that can fully account for all the properties of creole languages. Hence, in order to more fully understand data from creole, or even BM, it is necessary to refer to yet another opposing point of view.

7.4.4.3 The superstratists

The insufficiency of substrate influence to explain for creole features has been discussed by scholars since the 1930s. Göbl-Gáldi (1934) for instance, highlights that aside from substrate influence, there is also evidence of retentions from archaic and regional French, in addition to creole-internal developments. Going a step further, Faine claims negligible African influence on Haitian Creole French, and that “at least three quarters [of the creole] is from the Norman dialect of the sixteenth and seventeenth centuries, which has been preserved in a very pure state” (1936:1). Based on his work with Réunionnais, one of the more famous superstratists, Chaudenson (1974) argues that creoles evolved from a variety of colloquial French (*français avancé*), which was ahead of the standard language in its natural evolutionary tendencies such as the loss of inflections. His view that a creole is a variety of its lexifier is also found in a more modern day approach, that of Mufwene (1996, 1997, 2001, 2003, 2007, 2008) and DeGraff’s (2001, 2003).

One of the biggest differences between the works of these superstratists and the other approaches that have been discussed thus far, is that they do not view creoles as being ‘exceptional’ (De Graff 2001, 2003). DeGraff highlights terms that have been used to discuss creoles— “DEGENERATE DESCENDENTS of their European ancestors”, “ABNORMAL TRANSMISSION or a BREAK IN TRANSMISSION”, “LIVING LINGUISTIC FOSSILS” and “SPECIAL HYBRIDS with exceptional genealogy” (2003: 393, emphasis in original). He goes as far to say that any approach that treats creoles as being exceptional is discriminatory. DeGraff

(2003) contends with Lefebvre's substratist approach (1998), stating that such a view implies that the creole creators were not able to abstract structure from the target European language, which is inaccurate as the lexicon and morphology of Haitian Creole demonstrates that these creators were able to segment and parse target speech down to the phonetic forms of many affixes. Likewise, Mufwene (1996, 2001, 2008) maintains that the word 'creole' is not a valid term for classifying languages, and that creoles are simply approximations of earlier regional dialects spoken by the founders of the colonies.

Similar to Bickerton's LBH, Mufwene's founder effect principle (1996, 2001, 2008) appears to be largely applied to plantation creoles, although Mufwene does not seem to be averse to applying it to creoles of different sociohistorical backgrounds (2003). It asserts that the structural features of creoles have been predetermined to a large extent (but not exclusively) by characteristics of the vernaculars spoken by the populations that founded the colonies in which they developed (Mufwene 1996, 2001, 2008). These vernaculars would often be regional vernaculars. Mufwene states that "European colonies often started with large proportions of indentured servants and other low-class employees of colonial companies, thus speakers of nonstandard varieties of creoles' lexifiers" (1996:84). This is reminiscent of Göbl-Galdi (1934) and Chaudenson's (1974) that the French creoles emerged from regional varieties of French. In addition, basilectalization takes place. Slaves, having come into contact with the speech of the Europeans began making approximations of it. As "restructured varieties" containing approximations "became the models for some of the newcomers", creoles diverged more and more from the lexifiers (Mufwene 2001: 51). Mufwene (1996, 2001, 2008) also uses notions from population genetics to explain for the selection of creole features. Features can be selected by mutation (this does not receive very much attention), be favoured by new ecological conditions in the colony, or the colony may have received significant proportions of carriers of the features/genes, a situation which maximized the chances for their successful reproduction. The features may then be adapted to suit the ecology, so that features that are compatible with the features of the other languages in the ecology are more likely to be maintained than features that are not. (Mufwene 1996, 2001, 2008). In addition to ethnographic factors such as the demographic proportion of the newcomers relative to the local populations,

their attitudes towards each other, and their relative social statuses, Mufwene also mentions regularity, transparency and perceptual salience as important factors that help decide which feature is to be selected if there is competition (1996, 2001, 2008). This is interestingly reminiscent of Siegel (1999) who argues that the selection of superstrate features depends on factors such as perceptual salience, transparency and frequency.

With regard to evaluating the superstratist approach, especially Mufwene's founder effect, some considerations that have to be made are whether an approach that treats creoles as 'unexceptional' is necessary or worthwhile, and how much such an ecological account (one that takes into consideration the environment in which the lexifier comes into contact with other languages whose features then compete with its own) (see Mufwene 1996) contributes towards the understanding of a non-prototypical creole such as BM. While there is no contesting the undesirability of discrimination, it is important to note that modern day creole formation theories are not set out to discriminate against any particular group of people. There is hence no particular impetus to align one's work with this approach if the only reason is to be anti-discriminatory. In terms dealing with sociohistorical background, this particular substratist approach differs from Bickerton's LBH which focuses narrowly on the plantation scenario (and even then, ignores how sociohistorical factors may affect creole structure), as well as from the substratists who do not emphasize sociohistorical factors in determining creole structure.

An approach that focuses on the sociohistorical background of BM is interesting. As seen in the above section on the substratist approach, there is no way of accounting for why Malay is the lexifier language, while Hokkien is the substrate language of BM. There are debates over whether mothers as primary caregivers might provide more lexicon, or more grammar to their children. For example, Bakker (1997: 207) who focuses on mixed languages, suggests that as the primary caregiver, mothers would provide primary input for their children's language acquisition, and this would correspond to grammatical input (presumably grammar is harder to acquire than lexicon). In BM, this is not the case, for the mother's language clearly provides more lexicon than grammatical elements. While it remains unclear if there is substantial support for such a stance, Bakker (1997: 207) also suggests that the mother's language is usually

the better-known language of the region. This is not entirely incongruent with the founder effect principle. So, under the founder effect principle, why should the Malay speaking population then be treated as the founder population, rather than the Hokkien- speaking traders? This may be plausible if one acknowledges that the Malays had arrived prior to these Chinese traders and were part of the flourishing Srivijaya empire attested to exist between late 7th century and early 12th century (Bronson and Wisseman 1976) whereas the earliest Chinese settlers arrived in the 15th century (Wade 1994, Wang 1964, Widodo 2002/2003). Bearing this in mind, it will be necessary to consider how the language ecology may have affected the selection of features. When the Chinese traders arrived after the beginning of the 15th century, the Malay language was already a dominant lingua franca in the region, having been established as one during the Srivijaya period. In the same vein, Portuguese and Dutch, which were the other languages spoken by the colonizers of Malacca, did not affect BM as much as Malay did, since these languages (especially Portuguese, which was also a dominant lingua franca in the region) cannot be treated as the founder population (see section 1.3.2). Thus, as with the discussion in section 7.4.4.2, even in intermarriage, it is more likely that the Hokkien traders had to learn the local language, Malay, rather than the local women having to learn the Hokkien. This language ecology appears to be largely maintained. Currently, with Singapore situated among Malay-speaking neighbours (Indonesia, Malaysia, and Brunei), and Malacca being part of Malaysia, it may be no surprise as to why on the surface, BM appears to be most congruent with Malay. However, for this to be a plausible account of BM, it must be shown that basilectalization has taken place, so that the creole became more and more divergent from the initial lexifier language.

It is not entirely clear what the most influential variety of Malay was at the initial stages of creole formation (whether it was High Malay or Low Malay), and there are no manuscripts of early BM. However, it is possible to examine early British accounts of the Malay language that they are said to have encountered in trade, and to assess how BM may differ from this recorded variety of Malay. For example, Marsden (1812) records the form ‘Noun Determiner’, whereas in BM as spoken in Singapore, both ‘Noun Determiner’ and ‘Determiner Noun’ patterns are permissible (see section 5.1.2.1). This differs from the current BM as spoken in Malacca, where ‘Determiner Noun’ does not usually occur (see section 6.3.1.1). Noting that ‘Determiner Noun’

occurs in Hokkien, it may be possible to compare the ecologies of Singapore and Malacca. The population in Singapore is 74 % Chinese, a majority of whom are Hokkien (Singapore 2011). This can be compared to Malacca where 32 % are Chinese (Malaysia 2011). It is unclear as to how many Hokkien speakers there are in Malacca, but based on the aggregate proportions, it is conceivable as to why basilectalization can be seen in the ‘Determiner Noun’ structure of SBM, as compared to MBM, in which ‘Noun Determiner’ appears to be the norm.

The founder effect and the notion of a linguistic ecology are useful, but questions regarding linguistic features remain unanswered. For example, why do relativization patterns in SBM and MBM follow that of both Hokkien and Malay, when the ‘Determiner Noun’ structure is not popularly found in MBM?

7.5 Conclusion

Given the sociohistorical circumstances that led to the development of the Peranakan ethnicity and culture, BM is far from being a “genetic dialect” of Malay. While its sociohistorical circumstances are reminiscent of mixed languages that have been formed via intermarriage, the structural data shows that unlike mixed languages, BM's two component languages are not individually compartmentalized in separate subsystems (although it must be said that there is more Hokkien grammar and more Malay lexicon). Even though it is disputable whether or not BM should be considered as having been formed from more than two languages, and whether or not the language should be associated to and compared with the pidgin, Bazaar Malay, BM appears to fit within the fuzzy boundaries of creolehood. By structural accounts, BM demonstrates heavy substrate influence from Hokkien, and by sociohistorical ones, BM demonstrates an important principle, the founder effect.

Definitions of what a mixed language is, and what a creole is, are inherently fraught with some level of arbitrariness and difference of points of views. More interesting questions can be asked regarding how BM fits in with theories of creole formation, and how BM can also be used to provide perspectives on the various creole formation theories. The universalist approach, the substratist approach, and the superstratist approach differ in many ways. Clearly, the universalist

approach focuses on universal principles guided by perceived inherent characteristics of a human's genetic makeup, while the substratists and the superstratists emphasize the importance of the substrate language, and the lexifier language respectively. In addition, while the locus of the LBH is children, the substratists and the superstratists view processes natural to second language acquisition as being crucial in explaining creole structure. Finally, Mufwene's founder effect is laudable for emphasizing the influence of the sociohistorical background of the creole and its linguistic ecology on the form of the creole itself, as opposed to Bickerton's LBH and the substratists' notions of relexification and substrate transfer.

While this dissertation aims to be a comprehensive grammar, with an extra focus on sociophonetics, it has provided additional information regarding some differences between SBM and MBM, as well as how BM would fit into the greater environs of language contact. However, some questions remain unresolved. Given that the focal language of this grammar is an endangered contact language, the prime question should be, whether or not the endangerment of a contact language is different from the endangerment of a normally transmitted language. The endangerment of a contact language is generally not viewed as something that is different from the endangerment of a normally transmitted language. Wurm (2001:8) states that "some contact languages have become endangered, and some of them extinct. This is a fate that affects many other languages of the world too, not just contact languages." Mufwene (2003b) is of a similar opinion. He goes a step further, stating that there are "highly stigmatized language varieties such as Appalachian English, African-American English, and several creoles around the world" and they "do not seem to be particularly endangered by the more prestigious varieties with which they have coexisted and in which their speakers acquire literacy" (Mufwene 2003:330). He postulates that "despite linguists' common claim that creoles are separate languages related to their lexifiers, speakers of all these stigmatized vernaculars think that they speak the same language in which they are provided literacy" (Mufwene 2003:330-331). Providing a separate perspective, Matras (2005) holds that language contact does not necessarily lead to language death, and that extreme contact can be prolonged and even be a trigger for language birth.

It is not the case that a creole is always an extension of a less stigmatized vernacular (see Ansaldo et al. 2007), and hence, less likely to be endangered by the more prestigious variety with which it coexists. In the case of BM, Peranakans perceived their language to be of “the refined and wealthy class of Malay-speaking Chinese” and disdained pure Malay, “calling it *Malayu hutan*—the language of the jungle” (Shellabear 1913:156, italics in original). Beyond that, it is necessary to consider decreolization as a means through which a creole can be endangered. This can be differentiated from language shift (Fishman 1964), which can affect normally transmitted languages. Decreolization (by definition of Whinnom 1971) occurs when then creole begins to resemble one of the standard languages from which it originally derived or the one language from which it originated in some views, in a post-creole continuum with erosion of the more basilect varieties and convergence towards the acrolect varieties and towards the standard language, until the varieties exhibiting the most creole-like features just disappear. In Malacca, the trend is clear. The dominant language is Malay, and decreolization towards the lexifier may occur. In Singapore, the trend is less clear. The dominant language is Colloquial Singapore English (Singlish), which has the same substrate language as BM (Lee et al. 2009). Speakers are also heavily exposed to Standard English and Mandarin. Standard English is the official language of administration in Singapore, while most Peranakans, being recognised as Chinese, would have to learn Mandarin in school as a ‘mother tongue’. Other officially recognised languages are Malay and Tamil. For BM in Singapore then, the situation of language endangerment is more complicated than a shift towards a single dominant language. Would BM (or any other contact language) be inclined to decreolize towards its lexifier language or substrate language, or shift towards another dominant language that is spoken in its environment? This grammar is a step towards the resolution of the issue of contact language endangerment, still in its nascency.

REFERENCES

- ADAM, LUCIEN. 1883. *Les idioms négro-aryen et maléo-aryen: essai d'hybridologie linguistique*. Paris: Maisonneuve.
- ADANK, PATTI; ROEL SMITS; and ROELAND VAN HOUT. 2004. A comparison of vowel normalization procedures for language variation research. *Journal of the Acoustical Society of America*. 116:3009–107.
- AIKHENVALD, ALEXANDRA Y. 2005. Serial verb constructions in typological perspective. In *Serial verb constructions: a cross-linguistic typology*, ed. by A.Y. Aikhenvald and R.M.W. Dixon, 1–68. Oxford: Oxford University Press.
- AINSWORTH, W. A. 1975. Intrinsic and extrinsic factors in vowel judgements. In *Auditory Analysis and Perception of Speech*, ed. by G. Fant and M.A.A. Tatham. London: Academic Press
- ANDAYA, BARBARA WATSON, and LEONARD Y ANDAYA. 2001. *A History of Malaysia*, 2nd edition. London: Macmillan Education.
- ANSALDO, UMBERTO. 2011. Metatypy in Sri Lanka Malay. In *Annual Review of South Asian Languages and Linguistics: 2011*, ed. by R. Singh and G. Sharma, 3–16. Berlin: Mouton de Gruyter.
- ANSALDO, UMBERTO; LISA LIM; and SALIKOKO S. MUFWENE. 2007. The sociolinguistic history of the Peranakans: What it tells us about “creolization.” In *Deconstructing Creole. Typological Studies in Language 73*, ed. by U. Ansaldo, S. Matthews and L. Lim, 203–26. Amsterdam: John Benjamins.
- ANSALDO, UMBERTO, and STEPHEN MATTHEWS. 1999. The Minnan substrate and creolization in Baba Malay. *Journal of Chinese Linguistics*. 27:38–68.
- AUER, PETER. 1999. From codeswitching via language mixing to fused lects: Towards a dynamic typology of bilingual speech. *Journal of Bilingualism*. 3:309–32.
- AUSTIN, PETER. 2006. Data and language documentation. In *Essentials of Language Documentation*, ed. by J. Gippert, N.P. Himmelmann and U. Mosel, 87–112. Berlin: Mouton de Gruyter.
- AYE, KHIN KHIN. 2005. *Bazaar Malay: History, grammar and contact*. Department of English Language and Literature, National University of Singapore, Singapore: PhD dissertation.
- BABEL, MOLLY. 2009. The phonetic and phonological obsolescence in Northern Paiute. In *Variation in indigenous minority languages*, ed. by J.N. Stanford and D.R. Preston, 23–46. Amsterdam: John Benjamins.
- BAKKER, PETER. 1997. *A Language of Our Own: The Genesis of Michif, the Mixed Cree-French Language of the Canadian Métis*. Oxford: Oxford University Press.
- . 2003. Mixed languages as autonomous systems. In *The Mixed Language Debate: Theoretical and Empirical Advances*, ed. by Y. Matras and P. Bakker, 107–50. Berlin: Mouton de Gruyter.
- . 2006. The Sri Lanka Sprachbund: The newcomers Portuguese and Malay. In *Language Areas*, ed. by Y. Matras, A. McMahon and N. Vincent, 135–59. Houndsmill Basingstoke: Palgrave MacMillan.
- BARNARD, TIMOTHY P. (ed.) 2004. *Contesting Malayness: Malay Identity Across Boundaries*. Singapore: Singapore University Press.

- BATES, DOUGLAS; MARTIN MAECHLER; BEN BOLKER; and STEVEN WALKER. 2013. lme4: Linear mixed effect models using Eigen and S4 [Computer program]. R package version 1.0-4. <http://CRAN.R-project.org/package=lme4>.
- BENJAMIN, GEOFFREY. 2009. Affixes, Austronesian and iconicity in Malay. *Bijdragen tot de Taal-, Land-en Volkenkunde*. 165:291–323.
- BEREZ, ANDREA. 2011. *Directional reference, discourse and landscape in Ahtna*. Department of Linguistics, University of California Santa Barbara, California: PhD dissertation.
- BICKERTON, DEREK. 1981. *Roots of Language*. Ann Arbor: Karoma Publishers, Inc.
- BLUST, ROBERT. 1998. In Defense of Dempwolff: Austronesian Diphthongs Once Again. *Oceanic Linguistics*. 37(2):353–62.
- BLUST, ROBERT. 2005. Must sound change be linguistically motivated? *Diachronica*. 22:219–69.
- BODMAN, NICHOLAS C. 1955. *Spoken Amoy Hokkien*, volume 1. Kuala Lumpur: The Government of the Federation of Malaya.
- . 1958. *Spoken Amoy Hokkien*, volume 2. Kuala Lumpur: The Government of the Federation of Malaya.
- BOERSMA, PAUL, and DAVID WEENINK. 2013. Praat: doing phonetics by computer [Computer program]. Version 5.3.59. <http://www.praat.org/>.
- BORSCHBERG, PETER. 2010. Ethnicity, Language and Culture in Melaka after the Transition from Portuguese to Dutch Rule (Seventeenth Century). *Journal of the Malaysian Branch of the Royal Asiatic Society*. 83(2): 93-117.
- BRONSON, BENNETT, and IAN WISSEMAN. 1976. Palembang as Srivijaya: The lateness of early cities in Southern Southeast Asia. *Asian Perspectives*. 19:220–39.
- BRUNELLE, MARC. 2009. Diglossia and monosyllabization in Eastern Cham: A sociolinguistic study. In *Variation in indigenous minority languages*, ed. by J.N. Stanford and D.R. Preston, 47–76. Amsterdam: John Benjamins.
- CAMPBELL-KIBLER, KATHRYN. 2007. Accent, (ING), and the social logic of listener perceptions. *American Speech*. 82:32–64.
- . 2009. The nature of sociolinguistic perception. *Language Variation and Change*. 21:135–56.
- CAMPBELL, LYLE. 2013. *Historical Linguistics: An Introduction*, 3rd edition. Cambridge, MA: The MIT Press.
- CAMPBELL, LYLE; NALA HUIYING LEE; EVE OKURA; SEAN SIMPSON; KAORI UEKI; and JOHN VAN WAY. 2013. New knowledge: findings from the Catalogue of Endangered Languages (ELCat). Paper presented at the 3rd International Conference on Language Documentation and Conservation, University of Hawai‘i at Mānoa, USA.
- CAMPBELL, LYLE, and MARTHA C. MUNTZEL. 1989. The structural consequences of language death. In *Investigating obsolescence: Studies in language contraction and death*, ed. by N. Dorian, 181–96. Cambridge: Cambridge University Press.
- CAMPBELL, LYLE, and WILLIAM J. POSER. 2008. *Language Classification: History and Method*. Cambridge: Cambridge University Press.
- CHAO, YUEN-REN. 1930. A system of tone-letters. *Le Maître Phonétique*. 45:24–27.
- CHAUDENSON, ROBERT. 1974. *Le lexique du parler créole de la Réunion*. Paris: Champion.
- . 1992. *Des îles, des hommes, des langues: essais sur la créolisation linguistique et culturelle*. Paris: L'Harmattan.

- CHEN, XIAOJIN. 2003. *马来西亚的三个汉语方言* [*Three dialects in Malaysia*]. Beijing: 中國社會科學出版 [China Social Sciences Publishing].
- CHEO, KIM BAN. 1983. *A Baba Wedding*. Singapore: Eastern Universities Press Sdn. Bhd.
- CHIA, FELIX. 1983. *Ala Sayang!* Singapore: Eastern Universities Press.
- . 1989. *Pileh Menantu* [*Choosing a Daughter-in-law*], edited with English translation Singapore: Magro Publishing.
- . 1994. *The Babas revisited*. Singapore: Heinemann Asia.
- CHIANG, KER CHIU. 1940. *Hokkien (Amoy) for beginners*, book three. Singapore: Chin Fen Book Store.
- CHUA, BENG HUAT. 2003. Multiculturalism in Singapore: an instrument of social control. *Race and Class*. 44:58–77.
- CHUNG, SANDRA. 1976. An object creating rule in Bahasa Indonesia. *Linguistic Inquiry*. 7:41–87.
- CHUNG, SIAW-FONG. 2011. Uses of *ter-* in Malay: A corpus-based study. *Journal of Pragmatics*. 43:799–813.
- CLAMMER, JOHN R. 1980. *Straits Chinese Society*. Singapore: Singapore University Press.
- . 1983. The Straits Chinese in Melaka. Melaka. In *The Transformation of a Malay Capital, circa 1400-1980*, volume 2, ed. by K.S. Sandhu and P. Wheatley, 156-73. Kuala Lumpur: Oxford University Press.
- CLYNES, ADRIAN, and DAVID DETERDING. 2011. Illustrations of the IPA: Standard Malay (Brunei). *Journal of the International Phonetic Association*. 41:259–68.
- COELHO, FRANCISCO ADOLFO. 1880-6. Os dialectos românicos ou neolatinos na Africa, Asia e América. *Bolletim da Sociedade de Geographica de Lisboa*. (Republished in 1967, *Estudios lingüísticos crioulos*, ed. by J. Morais. Lisbon: Acadêmia Internacional de Cultura Portuguesa.)
- COLE, PETER, and MIN-JEONG SON. 2004. The Argument Structure of Verbs with the Suffix *-kan* in Indonesian. *Oceanic Linguistics*. 43:339–64.
- COMRIE, BERNARD. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. Cambridge: Cambridge University Press.
- . 1989. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.
- CRAWFURD, JOHN. 1852. *A Grammar and Dictionary of the Malay Language*. London: Smith, Elder and Company.
- CRUTTENDEN, ALAN. 1997. *Intonation*, 2nd edition. Cambridge: Cambridge University Press.
- CRYSTAL, DAVID. 1997. *A Dictionary of Linguistics and Phonetics*, 4th edition. Oxford: Blackwell Publishing.
- CRYSTAL, DAVID. 2000. *Language death*. Cambridge: Cambridge University Press.
- DAVID, MAYA KHEMLANI, and SUBRAMANIAM GOVINDASAMY. 2005. The construction of national identity and globalization in multilingual Malaysia. In *Language policy, culture, and identity in Asian contexts*, ed. by A. Tsui and J.W. Tollefson, 55–72. New Jersey: Lawrence Erlbaum.
- DE WITT, DENNIS. 2008. *History of the Dutch in Malaysia*. Kuala Lumpur: Nutmeg Publishing.
- DEGRAFF, MICHAEL. 2001. On the origin of Creoles: A Cartesian critique of Neo-Darwinian linguistics. *Linguistic Typology*. 5:213–310.
- . 2003. Against Creole exceptionalism (Discussion note). *Language*. 79:391–410.

- DEMPWOLFF, OTTO. 1934-38. *Vergleichende Lautlehre des austronesischen Wortschatzes*. Zeitschrift für Eingeborenen-Sprachen. 1. Induktiver Aufbau einer indonesischen Ursprache, Supplement 15 (1934); 2. Deduktive Anwendung des Urindonesischen auf austronesische Einzelsprachen, Supplement 17 (1937); 3. Austronesisches Wörterverzeichnis, Supplement 19 (1938). Berlin: Reimer.
- DILLARD, JOEY LEE. 1975. A Sketch of the History of Black English. In *Black American English: Its Background and Its Usage in the Schools and in Literature*, ed. by P. Stoller, 19–48. New York: Dell Publishing.
- DORIAN, NANCY. 2001. Surprises in Sutherland: linguistic variability amidst social uniformity. In *Linguistic Fieldwork*, ed. by P. Newman and M. Ratcliffe, 133–51. Cambridge: Cambridge University Press.
- DOUGLAS, CARSTAIRS. 1873. *Chinese-English Dictionary of the Vernacular or Spoken Language of Amoy: With the Principal Variations of the Chang-chew and Chin-chew Dialects*. London: Trübner and Co.
- DRAGER, KATIE. 2011. Speaker Age and Vowel Perception. *Language and Speech*. 54:99–121.
- . 2013. Experimental Methods in Sociolinguistics. In *Research Methods in Sociolinguistics: A practical guide*, ed. by J. Holmes and K. Hazen, 58-73. Oxford: Wiley Blackwell.
- DREYFUSS, GAIL RAIMI, and DJOEHANA OKA. 1979. Chinese Indonesian: A new kind of hybrid language. *Papers in pidgin and creole linguistics*, No. 2. Canberra: Pacific Linguistics A-57:247–74.
- DU BOIS, JOHN; SUSANNA CUMMING; STEPHAN SCHUETZE-COBURN; and DANAE PAOLINO (eds) 1992. *Discourse Transcription*. California: University of California Santa Barbara.
- DUBOIS, SYLVIE, and BARBARA HORVATH. 1998. Let's tink about dat: interdental fricatives in Cajun English. *Language Variation and Change*. 10:245–61
- ECKERT, PENELOPE. 1988. Adolescent social structure and the spread of linguistic change. *Language in Society*. 17:183–207.
- EMERSON, RALPH WALDO. 1855. *Essays: Second Series*, 2nd edition Boston: Phillips, Sampson, and Company.
- ENDANGERED LANGUAGES. 2012. The Linguist List at Eastern Michigan University and The University of Hawai'i at Mānoa. <http://www.endangeredlanguages.com>.
- EVERS, VINCENT; HENNING REETZ; and ADITI LAHIRI. 1998. Crosslinguistic acoustic categorization of sibilants independent of phonological status. *Journal of Phonetics*. 26:345–70.
- FAINE, JULES. 1936. *Philologie creole: etudes historiques et etymologiques sur la langue creole d'Haïti*. Port-au-Prince: Imprimerie de l'Etat.
- FEI, HSIN. 1436. *Hsing-ch'a sheng-lan: the overall survey of the Star Raft*. (Republished in 1996, South China and Maritime Asia, volume 4. Translated by John Vivian Gottlieb Mills. Edited and annotated by Roderich Ptak Wiesbaden: Harrassowitz Verlag.)
- FIRTH, DAVID. 1993. Bias reduction of maximum likelihood estimates. *Biometrika*. 80:27–38.
- FISHMAN, JOSHUA. 1964. Language maintenance and language shift as fields of inquiry. *Linguistics*. 9:32–70.
- . 1991. *Reversing Language Shift*. Clevedon: Multilingual Matters.

- FLEMINGS, LUKE. 2010. *From patrilects to performatives: Linguistic exogamy and language shift in the Northwest Amazon*. Department of Linguistics, University of Pennsylvania, Pennsylvania: PhD dissertation.
- FOULKES, PAUL; JAMES M. SCOBIE; and DOMINIC WATT. 2010. Sociophonetics. In *Handbook of Phonetic Sciences*, ed. by W.J. Hardcastle, J. Laver and F. Gibbon, 703–54. Oxford: Blackwell.
- FREEDMAN, MAURICE. 1979. *The study of Chinese societies: essays*. Stanford: Stanford University Press.
- GOLOVKO, EVGENIY. 1994. Mednyj Aleut or Copper Island Aleut: An Aleut-Russian mixed language. In *Mixed languages: 15 case studies in language intertwining*, ed. by P. Bakker and M. Mous, 113–21. Amsterdam: Uitgave IFOTT.
- GORDON, ELIZABETH; LYLE CAMPBELL; JENNIFER HAY; MARGARET MACLAGAN; ANDREA SUDBURY; and PETER TRUDGILL. 2004. *New Zealand English: Its Origin and Evolution*. Cambridge: Cambridge University Press.
- GREENBERG, JOSEPH. 1963. *Universals of language*. Cambridge, Massachusetts: MIT Press.
- GRIMES, BARBARA F. (ed.) 1974. *Ethnologue*, 8th edition. Huntington Beach, California: Wycliffe Bible Translators.
- (ed.) 1984. *Ethnologue*, 10th edition. Huntington Beach, California: Wycliffe Bible Translators.
- GROENEVELDT, WILLIAM PIETER. 1876. Notes on the Malay Archipelago and Malacca: Compiled from Chinese Sources. (Republished in 1880, *Verhandelingen van het Bataviaasch Genootschap van Kunsten en Wetenschappen*. 39: i-x. 1-144)
- GWEE, WILLIAM THIAN HOCK. 1998. *Mas sepuloh: Baba conversational gems*. Singapore: Armour Publishing.
- . 2006. *A Baba Malay Dictionary: The First Comprehensive Compendium of Straits Chinese Terms and Expressions*. Singapore: Tuttle Publishing.
- GÖBL-GÁLDI, L. 1934. Esquisse de la structure grammaticale des patois français-créoles. *Zeitschrift für französische Sprache und Literatur*. 58:257–95.
- HALE, KENNETH. 1992. Language endangerment and the human value of linguistic diversity. *Language*. 68:35–42.
- HALL, KENNETH R. 2011. *A History of Early Southeast Asian: Maritime Trade and Societal Development, 100 – 1500*. Plymouth: Rowman and Littlefield Publishers, Inc.
- HARDWICK, PATRICIA ANN. 2008. "Neither fish nor fowl": constructing Peranakan identity in colonial and post-colonial Singapore. *Folklore of East Asia*. 31:36–55.
- HAY, JENNIFER, and KATIE DRAGER. 2007. Sociophonetics. *Annual Review of Anthropology*. 36:89–103.
- HILL, JANE. 2006. The ethnography of language and language documentation. *Essentials of language documentation*, ed. by J. Gippert, N.P. Himmelmann and U. Mosel, 113–28. Berlin: Mouton de Gruyter.
- HIMMELMANN, NIKOLAUS P. 2006. Language Documentation: What is it and what is it good for? In *Essentials of Language Documentation*, ed. by J. Gippert, N.P. Himmelmann and U. Mosel, 1–30. Berlin: Mouton de Gruyter.
- HOLM, JOHN. 1989. *Pidgins and creoles: a reference survey*. Cambridge: Cambridge University Press.

- . 2000. *An Introduction to Pidgins and Creoles*. Cambridge: Cambridge University Press.
- HOLST, FREDERIK. 2012. *Ethnicization and Identity Construction in Malaysia*. New York: Routledge.
- HOPPER, PAUL J., and SANDRA A. THOMPSON. 1980. Transitivity in Grammar and Discourse. *Language*. 56:251-99.
- HORNE, ELINOR M. 1974. *Javanese-English dictionary*. New Haven: Yale University Press.
- HUBER, MAGNUS. 1999. *Ghanaian Pidgin English in its West African Context*. Amsterdam: John Benjamins.
- HUNG, FENG-SHENG. 1996. *Prosody and the acquisition of grammatical morphemes in Chinese languages*. Department of Linguistics, University of Hawai'i, Hawai'i: PhD dissertation.
- HYMES, DELL. 1971. Introduction to Section III: General conceptions of processes. In *Pidginization and creolization of a language*, ed. by D. Hymes, 65–85. Cambridge: Cambridge University Press.
- IKRANAGARA, KAY. 1980. *Melayu Betawi grammar*. NUSA, volume 9 Jakarta: Universitas Atma Jaya.
- JOHNSON, KEITH. 2012. *Acoustic and Auditory Phonetics*. Third edition. Oxford: Wiley-Blackwell.
- JOHNSON, KEITH; PETER LADEFOGED; and MONA LINDAU. 1993. Individual differences in vowel production. *Journal of the Acoustical Society of America*. 94:701–14.
- KENDALL, MAURICE GEORGE. 1955. *Rank Correlation Methods*. New York: Hafner Publishing Co.
- KHOO, JOO EE. 1996. *The Straits Chinese: A Cultural History*. Kuala Lumpur: Pepin Press.
- KOOPS, CHRISTIAN. 2011. *Local Sociophonetic Knowledge in Speech Perception*. Department of Linguistics, Rice University, Houston: PhD dissertation.
- KRAUSS, MICHAEL. 1992. The world's languages in crisis. *Language*. 68:4–10.
- KRISTIANSEN, TORE. 2010. Attitudes, ideology and awareness. In *The SAGE Handbook of Sociolinguistics*, ed. by R. Wodak, B. Johnstone and P. Kerswill, 265-78. Los Angeles: Sage Publications, Incorporated.
- LABOV, WILLIAM. 1963. The social motivation of a sound change. *Word*. 18:1–42.
- . 1972. *Language in the inner city*. Philadelphia: University of Pennsylvania Press.
- LABOV, WILLIAM; MARK KAREN; and COREY MILLER. 1991. Near-mergers and the suspension of phonetic contrast. *Language Variation and Change*. 3:33–74.
- LADEFOGED, PETER. 2003. *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*. Oxford: Blackwell Publishing.
- LADEFOGED, PETER; JOSEPH DECLERK; MONA LINDAU; and GEORGE PAPCUN. 1972. An auditory-motor theory of speech production. *UCLA Working Papers in Phonetics*. 22.
- LADEFOGED, PETER; AND KEITH JOHNSON. 2011. *A Course in Phonetics*, 6th edition, international edition. Canada: Wadsworth Cengage Learning.
- LAMBERT, WALLACE E. 1967. A social psychology of bilingualism. *Journal of Social Issues*. 23:91–108.
- LAMBERT, WALLACE E.; RICHARD HODGSON; ROBERT C. GARDNER; and SAMUEL FILLENBAUM. 1960. Evaluational reactions to spoken languages. *Journal of Abnormal and Social Psychology*. 60:44–51.

- LASS, ROGER. 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.
- LAU, AILEEN GUEK LIN. 1984. *Language death with particular reference to the Baba Malay of Singapore*. Department of Language and Linguistic Science, University of York, York: Master of Arts thesis.
- LAWSON, ROBERT. 2014. 'Don't even [θ/f/h]ink about it': An ethnographic investigation of social meaning, social identity and (θ) variation in Glasgow. *English World-Wide*. 35:68–93.
- LEE, GWYNETH A. MAE-EN. 1999. *A descriptive grammar of spoken Peranakan*. Department of English Language and Literature, National University of Singapore, Singapore: Honour's thesis.
- LEE, NALA HUIYING. 2009. *One substrate, two lexifiers and the lexifier effect*. Department of English Language and Literature, National University of Singapore, Singapore: Master of Arts thesis.
- . 2012. Relativization and the Lexifier Effect. Paper presented at the Winter Meeting of the Society of Pidgin and Creole, Portland, Oregon.
- LEE, NALA HUIYING; HIROKI NOMOTO; and AI-PING LING. 2009. Colloquial Singapore English got: functions and substratal influences. *World Englishes*. 28:293–318.
- LEE, SU KIM. 1998. *Manglish*. Singapore: Marshall Cavendish Press.
- LEFEBVRE, CLAIRE. 1998. *Creole genesis and the acquisition of grammar: the case of Haitian creole*. Cambridge: Cambridge University Press.
- LEWIS, M. PAUL, and GARY F. SIMONS. 2010. Assessing endangerment: expanding Fishman's GIDS. *Revue Roumaine de Linguistique*. 55:103–20.
- LEWIS, M. PAUL; GARY F. SIMONS; and CHARLES D. FENNIG (eds) 2013. *Ethnologue: Languages of the World*, 17th edition. Dallas, Texas: SIL International. <http://www.ethnologue.com>.
- LI, CHARLES, and SANDRA A. THOMPSON. 1976. Subject and topic: a new typology of language. In *Subject and topic*, ed. by C. Li, 457-91. New York: Academic Press.
- LIM, HIONG SENG. 1887. *A manual of the Malay colloquial such as is spoken by all nationalities in the Colonies of the Straits Settlement, and designed for domestic and business purposes*. Singapore: Koh Yew Hean Press.
- LIM, LISA. 2010. Peranakan English in Singapore. In *The lesser-known varieties of English: an introduction*, ed. by D. Schreier, P. Trudgill and E.W. Schneider, 327-47. Cambridge: Cambridge University Press.
- LIM, SONNY. 1981. *Baba Malay: the language of the "Straits-born" Chinese*. Department of Linguistics, Monash University, Melbourne: Master of Arts thesis.
- . 1988. Baba Malay: The language of the "Straits-born" Chinese. *Papers in Western Austronesian Linguistics*. 3:1–61.
- LOBANOV, BORIS M. . 1971. Classification of Russian vowels spoken by different listeners. *Journal of the Acoustical Society of America*. 49:606–08.
- LUMSDEN, JOHN S. 1999. Language acquisition and creolization. In *Language creation and language change: Creolization, diachrony and development*, ed. by M. DeGraff, 129-57. Cambridge: MIT Press.
- LYONS, JOHN. 1977. *Semantics*. Cambridge: Cambridge University Press.

- MALAYSIA, DEPARTMENT OF STATISTICS. 2011. *Laporan kiraan permulaan [preliminary count report]*. http://www.statistics.gov.my/portal/download_Population/files/BPD/Laporan_Kiraan_Permulaan2010.pdf.
- MARSDEN, WILLIAM. 1812. *A grammar of the Malayan language*. London: Cox and Baylis.
- MATHER, PATRICK-ANDRÉ. 2007. Creole studies. In *French Applied Linguistics*, ed. by D. Ayoun, 401-24. Amsterdam: John Benjamins.
- MATRAS, YARON. 2003. Mixed languages: Re-examining the structural prototype. In *The Mixed Language Debate: Theoretical and Empirical Advances*, ed. by Y. Matras and P. Bakker, 151-76. Berlin: Mouton de Gruyter.
- . 2005. Language contact, language endangerment, and the role of the 'salvation linguist'. In *Language Documentation and Description*, volume 3, ed. by P. Austin, 225-51. London: School of Oriental and African Studies.
- MATRAS, YARON, and PETER BAKKER. 2003. The study of mixed languages. In *The Mixed Language Debate: Theoretical and Empirical Advances*, ed. by Y. Matras and P. Bakker, 1-20. Berlin: Mouton de Gruyter.
- MATRAS, YARON; HAZEL GARDNER; CHARLOTTE JONES; and VERONICA SCHULMAN. 2007. Angloromani: A different kind of language? *Anthropological Linguistics*. 49:142-64.
- MCLEOD, A. IAN. 2014. Kendall [Computer program]. R package version 2.2. <http://CRAN.R-project.org/package=kendall>.
- MCWHORTER, JOHN H. 2005. *Defining Creole*. Oxford: Oxford University Press.
- MEAKINS, FELICITY. 2008. Land, language and identity: The socio-political origins of Gurindji Kriol. In *Social Lives in Languages*, ed. by M. Meyerhoff and N. Nagy, 69-94. Amsterdam: John Benjamins.
- MINTZ, MALCOM W. 1994. *A student grammar of Malay and Indonesia*. Singapore: EPB Publishers.
- MOUS, MAARTEN. 2003. *The Making of a Mixed Language: The Case of Ma'a/Mbugu*. Amsterdam: John Benjamins.
- MUFWENE, SALIKOKO S. 1996. The founder principle in creole genesis. *Diachronica*. 13:83-134.
- . 1997. Jargons, pidgins, creoles, and koines: what are they? In *The structure and status of pidgins and creoles*, ed. by A.K. Spears and D. Winford, 35-70. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- . 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press.
- . 2003a. Genetic linguistics and genetic creolistics: a response to Sarah G. Thomason's "Creoles and genetic relationships". *Journal of Pidgin and Creole Languages*. 18:273-88.
- . 2003b. Language endangerment: What have pride and prestige got to do with it? In *When languages collide: Perspectives on language conflict, language competition, and language coexistence*, ed. by B. Joseph, J. Destefano, N.G. Jacobs and I. Lehiste, 324-46. Ohio: Ohio State University.
- . 2007. Population movements and contacts in language evolution. *Journal of Language Contact*. 1:63-92.
- . 2008. *Language Evolution: Contact, Competition and Change*. London/ New York: Continuum International Publishing Group.
- MÜHLHÄUSLER, PETER. 1986. *Pidgin and creole linguistics*. Oxford: Basil Blackwell.

- MUYSKEN, PIETER. 1981. Halfway through Quechua and Spanish: the case for relexification. In *Historicity and variation in creole studies*, ed. by A. Highfield and A. Valdman, 52–78. Ann Arbor: Karoma.
- . 1997. Media Lengua. In *Contact languages: A wider perspective*, ed. by S.G. Thomason, 365–426. Amsterdam: John Benjamins.
- MUZZI, GERALDO A. 2002. *The Portuguese in Malay land: a glossary of Portuguese words in the Malay language*. Kuala Lumpur: Geraldo A. Muzzi.
- MYERS-SCOTTON, CAROL. 2003. What lies beneath: Split (mixed) languages as contact phenomena. In *The mixed language debate: Theoretical and empirical advances*, ed. by Y. Matras and P. Bakker, 73–106. Berlin: Mouton de Gruyter.
- NATHAN, J. 1922. *The Census of British Malaya*. London: Waterloo and Sons.
- NEAREY, T. M. 1989. Static, dynamic, and relational properties in speech perception. *Journal of the Acoustical Society of America*. 85:2088-113.
- NG, E-CHING. 2012. Chinese meets Malay meets English: Origins of Singaporean English word-final high tone. *International Journal of Bilingualism*. 16:83–100.
- NG, SIEW YOONG. 1961. The Chinese Protectorate in Singapore 1877-1900. *Journal of Southeast Asian History*. 2:76–99.
- NOMOTO, HIROKI, and ISAMU SHOHO. 2007. Voices in relative clauses in Malay: A comparison of written and spoken language. In *Corpus-Based Perspectives in Linguistics*, ed. by Y. Kawaguchi, T. Takagaki, N. Tomimori and Y. Tsuruga, 353–70. Amsterdam: John Benjamins.
- NORDHOFF, SEBASTIAN (ed.) 2012. *The Genesis of Sri Lanka Malay: A Case of Extreme Language Contact*. Leiden: Koninklijke Brill NV.
- NUYTS, JAN. 2006. Modality: Overview and linguistic issues. In *The Expression of Modality*, ed. by W. Frawley, E. Eschenroeder, S. Mills and T. Ngyuen, 1–6. Berlin: Mouton de Gruyter.
- O'SHANNESSY, CARMEL. 2005. Light Warlpiri: A New Language. *Australian Journal of Linguistics*. 25:31–57.
- OMAR, ASMAH HAJI. 1977. *The Phonological Diversity of the Malay Dialects*. Kuala Lumpur: Dewan Bahasa dan Pustaka.
- . 1989. The Malay Spelling Reform. *Journal of the Simplified Spelling Society*. 2:9–13.
- . 1993. The First Congress for Malay. In *The Earliest Stage of Language Planning: The "First Congress" Phenomenon*, ed. by J. Fishman, 181–98. Berlin: Mouton de Gruyter.
- ONN, FARID M. 1980. *Aspects of Malay Phonology and Morphology: A Generative Approach*. Kuala Lumpur: Penerbit University Malaya.
- OTHMAN, MARIAM, and DANIEL MUIJS. 2013. Educational quality differences in a middle income country: the urban-rural gap in Malaysian primary schools. *School Effectiveness and School Improvement*. 24:104–22.
- PAAUW, SCOTT HORAN. 2009. *The Malay Contact Varieties of Eastern Indonesia: A Typological Comparison*. Department of Linguistics, State University of New York at Buffalo, New York: PhD dissertation.
- PAKIR, ANNE. 1986. *A linguistic investigation of Baba Malay*. Department of Linguistics, University of Hawai'i, Hawai'i: PhD dissertation.

- . 1994. Educational linguistics: looking to the East. In *Georgetown University Round Table on Languages and Linguistics*, ed. by J. Alatis, 370–83. Washington DC: Georgetown University Press.
- PITCHER, PHILIP WILSON. 1912. *In and about Amoy: Some historical and other facts connected with one of the first open ports in China*. Shanghai and Foochow: The Methodist Publishing House in China.
- PLATT, JOHN T., and MIAN LIAN HO. 1989. Discourse particles in Singapore English: substratum influences and universals. *World Englishes* 8:215-21.
- POEDJOSOEDARMO, SOEPOMO. 1968. Javanese Speech Levels. *Indonesia*. 6:54-81.
- PURCELL, VICTOR. 1980. *The Chinese in Southeast Asia*. Issued under the auspices of the Royal Institute of International Affairs [by] Oxford University Press.
- R DEVELOPMENT CORE TEAM. 2013. R: A language and environment for statistical computing [Computer program]. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org>.
- RAKERD, BRAD, and BARTŁOMIEJ PLICHTA. 2003. More on perceptions of /a/ fronting. Paper presented at New Ways of Analyzing Variation 32: Philadelphia, October 9-12.
- REULAND, ERIC. 1999. The fine structure of grammar: Anaphoric relations. In *Reflexives: Forms and Functions*, volume 1, ed. by Z. Frajzyngier and T.S. Curl, 1–40. Amsterdam: John Benjamins.
- RHODES, RICHARD A.; LENORE A. GRENOBLE; ANNA BERGE; and PAULA RADEZKY. 2006. *Adequacy of documentation: A preliminary report to the CELP [Committee on Endangered languages and their Preservation, Linguistic Society of America]*.
- ROBERTS, SARAH. 1998. The role of diffusion in the genesis of Hawaiian Creole. *Language*. 74:1–39.
- ROMAINE, SUZANNE. 1995. *Bilingualism*. Malden, MA: Blackwell Publishers.
- ROSS, MALCOM. 1996. Contact-induced change and the comparative method: cases from Papua New Guinea. In *The Comparative method reviewed: regularity and irregularity in language change*, ed. by M. Durie and M. Ross, 180–217. Oxford: Oxford University Press.
- RUDOLPH, JURGEN. 1998. *Reconstructing identities: a social history of the Babas in Singapore*. Aldershot: Ashgate Publishing Ltd.
- SALLEH, HOOD. 2006. *Peoples and traditions*, Encyclopedia of Malaysia 12. Kuala Lumpur: Archipelago Press.
- SEBBA, MARK. 1987. *The syntax of serial verbs*. Amsterdam: John Benjamins.
- SHARIF, ABDUL AZIZ MOHD. 1981. *The Baba language*. Department of Language and Linguistic Science, University of York, York: Master of Arts thesis.
- SHELLABEAR, WILLIAM G. 1913. Baba Malay: an introduction to the language of the Straits-born Chinese. Attached as appendix in J.R. Clammer (1980), *Straits Chinese Society*, 153–65. Singapore: Singapore University Press.
- SHELLABEAR, WILLIAM G., and British and FOREIGN BIBLE SOCIETY. 1913. *Kitab Perjanjian Bharu [The New Testament]*. Singapura: Methodist Publishing House.
- SHIH, MENG-YI. 2009. *The Hokkien substrate in Baba Malay*. Department of Linguistics, National Tsing-Hua University, Hsinchu: Master of Arts thesis.

- SIEGEL, JEFF. 1997. Mixing, levelling, and pidgin/ creole development. In *The Structure and Status of Pidgins and Creoles*, ed. by A.K. Spears and D. Winford, 111-50. Amsterdam/ Philadelphia: John Benjamins.
- . 1999. Transfer constraints and substrate influence in Melanesian Pidgin. *Journal of Pidgin and Creole Languages*. 14: 1-44.
- . 2000. Substrate influence in Hawai'i Creole English. *Language in Society*. 29:197-236.
- SIMONS, GARY F., and M PAUL LEWIS. 2013. The world's languages in crisis. *Responses to language endangerment, In honor of Mickey Noonan*, Studies in Language Companion Series 142., ed. by E. Mihas, B. Perley, G. Rei-Doval and W. Kathleen, 3-19. Amsterdam: John Benjamins.
- SINGAPORE, DEPARTMENT OF STATISTICS. 2011. *Census of population 2010. Statistical Release 1: Demographic Characteristics, Education, Language and Religion*. http://www.singstat.gov.sg/publications/publications_and_papers/cop2010/census_2010_release1/cop2010sr1.pdf.
- SKINNER, G WILLIAM. 1996. Creolized Chinese Societies in Southeast Asia. In *Sojourners and settlers: Histories of Southeast Asia and the Chinese*, ed. by A. Reid and K. Aililunas-Rodgers, 51-93. Hawai'i: University of Hawai'i Press.
- SMITH, CARLOTA S. 1991. *The parameter of aspect*. Dordrecht: Kluwer.
- SNEDDON, JAMES NEIL. 1996. *Indonesian Reference Grammar*. St. Leonards: Allen and Unwin.
- . 2003. *The Indonesian language: Its history and role in Modern Society*. New South Wales: University of New South Wales Press.
- SON, MIN-JEONG, and PETER COLE. 2008. An event-based account of *-kan* constructions in Standard Indonesian. *Language*. 84:120-60.
- STANFORD, JAMES N., and DENNIS R. PRESTON (ed.) 2009. *Variation in indigenous minority languages*. Amsterdam: John Benjamins.
- STEBBINS, TONYA. 2004. Mali Baining perspectives on language and culture stress. In *Small languages and small language communities*, 46, ed. by N. Dorian, *International Journal of the Sociology of Language*. 169: 161-175.
- STEWART, WILLIAM A. 1962. An outline of linguistic typology for describing multilingualism. In *Study of the Role of Second Language in Asia, Africa and Latin America*, ed. by F.A. Rice, 15-26. Washington: Center for Applied Linguistics.
- STRAND, ELIZABETH A. 2000. *Gender stereotype effects in speech processing*. Department of Linguistics, Ohio State University, Ohio: PhD dissertation.
- TAN, CHEE BENG. 1979. *Baba and Nyonya: a study of the ethnic identity of the Chinese Peranakan in Malacca*. Department of Sociology, Cornell University, Ithaca: PhD dissertation.
- . 1980. *Baba Malay Dialect*. *Journal of the Malaysian Branch of the Royal Asiatic Society*. 53:150-66.
- . 1988. *The Baba of Melaka: culture and identity of a Chinese Peranakan Community in Malaysia*. Petaling Jaya: Pelanduk Publications.
- . 1993. *Chinese Peranakan Heritage in Singapore and Malacca*. Kuala Lumpur: Penerbit Fajar Bakti Sdn. Bhd.
- TAN, CHOON HOE. 2001. *Learn to speak Penang Hokkien Dialect*. Penang: Areca Books.

- TEO, PENG SUAN. 1996. *The Syntax and Semantics of Got in Colloquial Singapore English*. Department of English Language and Literature, National University of Singapore, Singapore: Honour's thesis.
- TEOH, BOON SEONG. 1994. *The Sound System of Malay Revisited*. Kuala Lumpur: Ministry of Education Malaysia.
- THE COMMISSIONER OF LAW REVISION. 2010. *Federal Constitution: Reprint*. Printed With The Consent Of The Yang Di-Pertuan Agong Pursuant To Article 160A Of The Federal Constitution. Malaysia: The Commissioner of Law Revision.
- THOMAS, ERIK R., and TYLER KENDALL. 2014. The Vowel Normalization and Plotting Suite NORM [Computer program]. Version 1.1. <http://ncslaap.lib.ncsu.edu/tools/norm/>.
- THOMASON, SARAH G. 1997a. Introduction. In *Contact Languages: A Wider Perspective*, ed. by S.G. Thomason, 1–7. Amsterdam: John Benjamins.
- . 1997b. A typology of contact languages. In *The structure and status of pidgins and creoles*, ed. by A.K. Spears and D. Winford, 71–90. Amsterdam: John Benjamins Publishing Company.
- . 2002. Creoles and genetic relationship. *Journal of Pidgin and Creole Languages*. 17:101–109
- . 2003. Social factors and linguistic processes in the emergence of stable mixed languages. In *The Mixed Language Debate: Theoretical and Empirical Advances*, ed. by Y. Matras and P. Bakker, 21–40. Berlin: Mouton de Gruyter.
- THOMASON, SARAH G., and TERRENCE KAUFMAN. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley, Los Angeles: University of California Press.
- THURGOOD, ELZBIETA. 1998. *A description of nineteenth century Baba Malay: a Malay variety influenced by language shift*. Department of Linguistics, University of Hawai'i, Hawai'i: PhD dissertation.
- THURSTON, WILLIAM R. 1989. How exoteric languages build a lexicon. In *VICAL 1: Oceanic languages: papers of the 5th International Conference of Austronesian Linguistics*, ed. by R. Harlow and R. Hooper, 555–79. Auckland: Linguistic Society of New Zealand.
- TRAUGOTT, ELIZABETH CLOSS. 1989. On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantic Change. *Language*. 65:31–55.
- TRAUGOTT, ELIZABETH CLOSS; ALICE TER MEULEN; JUDY SNITZER REILLY; and CHARLES A. FERGUSON (eds) 1986. *On Conditionals*. Cambridge: Cambridge University Press.
- TSAO, FENG-FU. 2004. Semantics and Syntax of Verbal and Adjectival Reduplication in Mandarin and Taiwanese. In *Sinitic grammar: Synchronic and diachronic perspectives*, ed. by H. Chappell, 285–308. New York: Oxford University Press.
- TSUNODA, TASAKU. 2005. *Language endangerment and language revitalization: an introduction*. Berlin: Mouton de Gruyter.
- TYRON, DARRELL; PETER MÜHLHÄUSER; and PHILIP BAKER. 1996. English-derived contact languages in the Pacific in the 19th century. In *Atlas of Languages of Intercultural Communication in the Pacific, Asia, and the Americas*, volume 2, ed. by S.A. Wurm, P. Mühlhäusler and D. Tyron, 471–96. Berlin, New York: Mouton de Gruyter.
- UNESCO AD HOC EXPERT GROUP ON ENDANGERED LANGUAGES (Matthias Brenzinger, Arienne M. Dwyer, Tjeerd de Graaf, Collette Grinevald, Michael Krauss, Osahito Miyaoka, Nicholas Ostler, Osamu Sakiyama, María E. Villalón, Akira Y. Yamamoto, Ofelia

- Zapeda). 2003. Language vitality and endangerment. Document submitted to the International Expert Meeting on UNESCO Programme Safeguarding of Endangered Languages, Paris, 10-12 March 2003. <http://www.unesco.org/culture/ich/doc/src/00120-EN.pdf>.
- VALDÉS, GUADALUPE. 2005. Bilingualism, Heritage Language Learners, and SLA Research: Opportunities Lost or Seized? *The Modern Language Journal*. 89:419–26.
- VAN BEZOOIJEN, RENÉE. 2005. Approximant /r/ in Dutch: routes and feelings. *Speech Communication*. 47:15–31.
- VAN GIJN, RIK. 2009. The phonology of mixed languages. *Journal of Pidgin and Creole Languages*. 24:93–119.
- VAN WAY, JOHN, and NALA HUIYING LEE. 2013. Assessing levels of endangerment in the Catalogue of Endangered Languages (ELCat). Poster presented at the 3rd International Conference on Language Documentation and Conservation. 28 Feb-3 Mar. University of Hawai‘i at Mānoa, USA.
- VAUGHAN, J. D. 1879. *Manners and customs of the Chinese in the Straits Settlements*. Singapore: The Mission Press.
- WADE, GEOFFREY PHILIP. 1994. *The Ming-Shi-lu [veritable records of the Ming dynasty]*. Department of History, University of Hong Kong, Hong Kong: PhD dissertation.
- WAN, ASLYNN. 2012. *Instrumental phonetic study of the rhythm of Malay*. Department of Language and Linguistics, Newcastle University, Newcastle: PhD dissertation.
- WANG, GUNGWU. 1964. The opening of relations between China and Malacca, 1403-5. In *Malayan and Indonesian Studies: essays presented to Sir Richard Windstedt on his 85th birthday*, ed. by J. Bastin and R. Roolvink, 87–104. Oxford: Clarendon Press.
- WEE, HOCK ANN LIONEL. 1995. *Cognition in Grammar: The Problem of Verbal Prefixation in Malay*. Department of Linguistics, University of California, Berkeley: PhD dissertation.
- WEE, KIM SOON GABRIEL. 2000. *Intonation of the Babas: an auditory and instrumental study*. Department of English Language and Literature, National University of Singapore, Singapore: Honours thesis.
- WEE, LIONEL. 2010. 'Burdens' and 'handicaps' in Singapore's language policy: On the limits of language management. *Language Policy*. 9:97–114.
- WEE, PETER. 2013. Memories relieved. Paper presented at the In conversation: all things Peranakan, 19-20 Apr. Asian Civilisations Museum, Singapore.
- WHINNOM, KEITH. 1965. Origin of European-based creoles and pidgins. *Orbis*. 14:510–27.
- . 1971. Linguistic hybridization and the 'special case' of pidgins and creoles. In *Pidginization and creolization of languages*, ed. by D. Hymes, 91–115. Cambridge: Cambridge University Press.
- WIDODO, JOHANNES. 2002/2003. A celebration of diversity: Zheng He and the origins of the pre-colonial coastal urban pattern in Southeast Asia. *Journal of Southeast Asian Architecture* 6:11–22.
- WINDSTEDT, R. O. 1927. *Malay Grammar*, 2nd edition, revised. Oxford: Clarendon Press.
- . 1945. *Colloquial Malay: A Grammar*. London: Kegan Paul, Trench, Trubner and Co. Ltd.
- . 1948. The Malay founder of medieval Malacca. In *Bulletin of the School of Oriental and African Studies. Presented to Lionel David Barnett by his colleagues, past and present*. 12:726–29.

- WINFORD, DONALD. 2003. *An Introduction to Contact Linguistics*. Oxford: Blackwell Publishing.
- WOLFF, JOHN U. 1983. The Indonesian spoken by the Peranakan Chinese of East Java: a case of language mixture. In *Essays in honor of Charles F. Hockett*, ed. by F.B. Agard and C.F. Hockett, 590–601. Leiden: E.J. Brill.
- WOODBURY, ANTHONY C. . 2003. Defining documentary linguistics. In *Language Documentation and Description*, volume 1, ed. by P. Austin, 35–51. London: School of Oriental and African Studies.
- WRIGHT, MARTHA SUSAN. 1983. *A metrical approach to tone sandhi in Chinese dialects*. Department of Linguistics, University of Massachusetts, Amherst: PhD dissertation.
- WURM, STEPHEN A. 2001. Contact languages and the preservation of endangered languages. *Estudios de Sociolingüística* 2:1–12.
- YIP, MOIRA. 2002. *Tone*. Cambridge: Cambridge University Press.
- YOONG, S.K., and A.N. ZAINAB. 2002. Chinese literary works translated into Baba Malay: A bibliometric study. *Malaysian Journal of Library and Information Science*. 7:1–23.

Appendix A: Word lists

Swadesh 100 word list

The following list is based on Swadesh (1955)'s 100 word list.

No.	Word	Gloss	No.	Word	Gloss
1.	<i>saya</i>	I (refined)	2.	<i>lu</i>	you
	<i>gua</i>	I (coarse)			
3.	<i>kita</i>	we	4.	<i>ini</i>	this
5.	<i>itu</i>	that	6.	<i>siapa?</i>	who
7.	<i>apa?</i>	what?	8.	<i>bukan</i>	no, negates noun
				<i>tak</i>	negates verb
9.	<i>sumua</i>	all	10.	<i>manyak</i>	many
11.	<i>satu</i>	one	12.	<i>dua</i>	two
13.	<i>besar</i>	big (coarse) big (refined)	14.	<i>panjang</i>	long
15.	<i>kechik</i>	small	16.	<i>perempuan</i>	woman
17.	<i>jantan</i>	man, male	18.	<i>orang</i>	person
19.	<i>ikan</i>	fish	20.	<i>burong</i>	bird
21.	<i>anjing</i>	dog	22.	<i>kutu</i>	louse
23.	<i>pokok</i>	tree	24.	<i>bijik</i>	seed
25.	<i>daoun</i>	leaf	26.	<i>akar</i>	root (coarse)
				<i>akair</i>	root (refined)
27.	<i>kulit kayu</i> 'skin wood'	bark	28.	<i>kulit</i>	skin
29.	<i>isi</i>	flesh	30.	<i>darah</i>	blood
31.	<i>tulang</i>	bone	32.	<i>minyak</i>	grease, oil
33.	<i>teloh</i>	egg	34.	<i>tandok</i>	horn

35.	<i>buntot</i>	tail, backside	36.	<i>bulu</i>	feather, fine hair
37.	<i>rambot</i>	hair, coarse hair	38.	<i>kepala</i>	head
39.	<i>kuping</i>	ear	40.	<i>mata</i>	eye
41.	<i>hidong, idong</i>	nose	42.	<i>mulot</i>	mouth
43.	<i>gigi</i>	tooth	44.	<i>lidah</i>	tongue
45.	<i>kuku</i>	finger nail	46.	<i>kaki</i>	foot, leg
47.	<i>lutot</i>	knee	48.	<i>tangan</i>	hand
49.	<i>perot</i>	belly, stomach	50.	<i>léihéi</i>	neck
51.	<i>téték</i>	breast	52.	<i>hati, ati</i> <i>jantung</i>	heart heart (when referring to emotions)
53.	<i>hati, ati</i>	liver	54.	<i>minum</i>	drink
55.	<i>makan</i>	eat	56.	<i>gigit</i>	bite
57.	<i>nampak, téngok</i>	see	58.	<i>dengar</i> <i>dengair</i>	hear (coarse) hear (refined)
59.	<i>tau</i>	know	60.	<i>tidor</i>	sleep
61.	<i>mati</i>	die	62.	<i>bunuh</i>	kill
63.	<i>berenang</i>	swim	64.	<i>terbang, trebang</i>	fly
65.	<i>jalan</i>	walk	66.	<i>datang</i> <i>mari</i>	come come let us do something
67.	<i>baring</i>	lie down	68.	<i>dudok</i>	sit, stay
69.	<i>diri</i>	stand	70.	<i>kasi</i>	give, let, cause, PASS
71.	<i>chakap</i> <i>kata</i>	speak say	72.	<i>mata ari, mata hari</i> 'eye sun'	sun

73.	<i>bulan</i>	moon, month	74.	<i>bintang</i>	star
75.	<i>ayé</i>	water	76.	<i>hujan, ujan</i>	rain
77.	<i>batu</i>	stone, rock, cave	78.	<i>paser</i>	sand
79.	<i>tanah</i>	soil, ground, earth	80.	<i>awan</i>	cloud
81.	<i>asap</i>	smoke	82.	<i>api</i>	fire
83.	<i>abu</i>	ash	84.	<i>bakar</i> <i>bakair</i>	burn (coarse) burn (refined)
85.	<i>jalan</i> <i>tengah jalan</i> middle path	path road	86.	<i>bukit</i> <i>gunong</i>	hill mountain
87.	<i>mérah</i>	red	88.	<i>hijo, ijo</i>	green
89.	<i>kuning</i>	yellow	90.	<i>putéh</i>	white
91.	<i>hitam, itam</i>	black	92.	<i>malam</i>	night
93.	<i>panas</i> <i>panair</i>	hot (coarse) hot (refined)	94.	<i>sejok</i>	cold
95.	<i>penoh</i> <i>kerniang</i>	full full (from food)	96.	<i>baru</i>	new
97.	<i>baik</i> <i>bagus</i>	good good (EXCLAM)	98.	<i>bulat</i>	round
99.	<i>kering</i>	dry, hay	100.	<i>nama</i>	name

Table 54: Swadesh 100 word list in Baba Malay

Kinship terms

Some of these kinship terms can be used together with numerals derived from Hokkien. Please refer to section 4.5.3 for a list of relevant numerals. Also, the term ‘cousin’ is used here to refer to the child of a parent’s sibling.

No.	Kinship term	Gloss	No.	Kinship term	Gloss
1.	<i>mak</i> <i>nya-nya</i>	mother	2.	<i>bapak</i> <i>baba</i>	father
3.	<i>mak-yi</i> ‘mother-mother’s sister’	second mother (in the case of a second marriage or a clash in horoscopes)	4.	<i>tio-tio</i> ‘parent’s eldest.sister’s husband’ <i>eng-chék</i> ‘honorific-father’s younger.brother’	second father (in the case of a second marriage or a clash in horoscopes)
5.	<i>nya besar/ besair</i> ‘Nyonya-big’	eldest daughter	6.	<i>ba-tengah</i> ‘Baba-middle’	middle son
7.	<i>nya bongsu</i> ‘Nyonya-youngest’	youngest daughter	8.	<i>ba-chik</i> ‘Baba-small’	youngest son
9.	<i>bini</i>	wife	10.	<i>laki</i>	husband
11.	<i>chingkay</i>	parent-in-law, father-in-law (indirect address)	12.	<i>chay.em</i>	mother-in-law (indirect address)
13.	<i>kua</i> <i>eng-kua</i> ‘honorific-father.in.law’	father-in-law (direct address)	14.	<i>nio</i>	mother-in-law (direct address)
15.	<i>tachi</i>	elder sister, also used for cousins	16.	<i>hia</i>	elder brother, also used for cousins
17.	<i>tachi besar/besair</i> ‘elder.sister big’	eldest sister	18.	<i>hia besar/besair</i> elder.brother big’ <i>tua-hia</i> ‘big-elder.brother’	eldest brother

19.	<i>tachi nombor dua</i> 'elder.sister number two'	second elder sister	20.	<i>hia nombor dua</i> 'elder.brother number two' <i>ji-hia</i> 'two-elder.brother'	second elder brother
21.	<i>tachi nombor tiga</i> 'elder.sister number three'	third elder sister	22.	<i>hia nombor tiga</i> 'elder.brother number three' <i>sa-hia</i> 'three-elder.brother'	third elder brother
23.	<i>adék</i>	sibling, younger sibling	24.	<i>adék-beradék</i> sibling-PL	siblings, younger siblings
25.	<i>adék perempuan</i> 'younger.sibling female'	younger sister	26.	<i>adék jantan</i> 'younger.sibling male'	younger brother
27.	<i>adék nombor satu</i> 'younger.sibling number one'	eldest younger sibling	28.	<i>adék bongsu</i> 'sibling youngest'	youngest sibling
29.	<i>pangkat adék-beradék</i> 'rank sibling-PL'	cousins	30.	<i>kopiau</i>	paternal cousins
31.	<i>chau</i>	elder sister's husband	32.	<i>so</i> <i>eng-so</i> 'honorific-elder.brother's.wife'	elder brother's wife
33.	<i>yi</i>	mother's sister	34.	<i>ku</i> <i>eng-ku</i> 'honorific-mother's.brother'	mother's brother
35.	<i>tio</i> <i>eng-tio</i> 'honorific-mother's.sister's.husband'	mother's sister's husband	36.	<i>kim</i> <i>eng-kim</i> 'honorific-mother's.brother's.wife'	mother's brother's wife
37.	<i>yi-yi</i>	mother's eldest sister	38.	<i>ku-ku</i>	mother's eldest brother
39.	<i>tio-tio</i>	mother's eldest sister's husband	40.	<i>kim-kim</i>	mother's eldest brother's wife

41.	<i>ji-yi</i> 'two-mother's.sister'	mother's second sister	42.	<i>ji-tio</i> 'two-mother's. sister's.husband'	mother's second sister's husband
43.	<i>ji-ku</i> 'two-mother's.brother'	mother's second brother	44.	<i>ji-kim</i> 'two-mother's.brother's wife'	mother's second brother's wife
45.	<i>ko</i>	father's sister	46.	<i>pek</i> <i>eng-pek</i> 'honorific- father's.elder.brother'	father's elder brother
47.	<i>tio</i> <i>eng-tio</i> 'honorific-father's. sister's.husband'	father's sister's husband (also mother's sister's husband)	48.	<i>em</i>	father's elder brother's wife
49.	<i>mak-ko</i> mother-father's.sister	father's eldest sister'	50.	<i>pek-pek</i>	father's eldest brother
51.	<i>tio-tio</i>	father's eldest sister's husband (also mother's eldest sister's husband)	52.	<i>mak-em</i> 'mother- father's.elder.brother's. wife'	father's eldest brother's wife
53.	<i>chék</i> <i>eng-chék</i> honorific- father's.younger.brother	father's younger brother	54.	<i>chim</i> <i>eng-chim</i> 'honorific-father's.younger. brother's.wife	father's younger brother's wife
55.	<i>ji-ko</i> 'two-father's.elder.sister'	father's second sister	56.	<i>ji pek</i> 'two-father's.elder.brother'	father's second elder brother
57.	<i>ji-chék</i> 'two- father's.younger.brother'	father's second younger brother'	58.	<i>ji-chim</i> 'two- father's.younger.brother's wife'	father's second younger brother's wife.
59.	<i>gong</i>	grandparent, also grandfather	60.	<i>cho</i>	great- grandparent, also great- grandmother
61	<i>mamak</i>	grandmother	62.	<i>gong.gong</i>	grandfather

63.	<i>mak-cho</i> 'mother- great.grandparent' <i>cho-cho</i>	great- grandmother	64.	<i>cho-gong</i> 'great-grandfather'	great- grandfather
65.	<i>po</i> <i>em-po</i> 'honorific-grandaunt'	grand-aunt	66.	<i>chek-gong</i> 'father's.younger.brother- grandparent'	grandfather's younger brother
67.	<i>ku-gong-cho</i> 'mother's.brother- grandparent- great.grandparent'	great- grandmother's brother	68.	<i>cho-po</i> 'great.grandparent-grandaunt'	great- grandaunt
69.	<i>gong-cho</i> grandparent- great.grandparent	ancestors	70.	<i>chuchu</i>	grandchild
71.	<i>chichi</i>	great- grandchild	72.	<i>onéng.onéng</i>	great-great grandchild

Table 55: Kinship terms in Baba Malay

Expressions for day, month and time

The days of the week are indicated by the word for day (*ari* or *hari*), followed by a number that represents a specific day of the week. This is with the exception of Sunday—*ari minggu* literally means ‘day week’. This is similar to the Hokkien system, where a number is assigned to each day; Tuesday corresponds to ‘two’, and Sunday corresponds to ‘week’. For example, Tuesday in Hokkien is *lèpài jī* ‘week two’, while Sunday is *lèpài jīt* ‘week day’. This is different from Malay, where Tuesday is *Selasa* and Sunday is *Ahad*.

No.	Day	Gloss	No.	Day	Gloss
1.	<i>ari satu</i> ‘day one’	Monday	2.	<i>ari dua</i> ‘day two’	Tuesday
3.	<i>ari tiga</i> ‘day three’	Wednesday	4.	<i>ari empat</i> ‘day four’	Thursday
5.	<i>ari lima</i> ‘day five’	Friday	6.	<i>ari enam</i> ‘day six’	Saturday
7.	<i>ari minggu</i> ‘day week’	Sunday			

Table 56: Days of the week in Baba Malay

Similarly, the months of the year are formed by combining the word for month (*bulan*) and a number that represents a specific month. Again, this is closer to the Hokkien system, where specific numbers represent specific months, although the word for month occurs after the numeral in Hokkien, instead of before. For example, February is *jī guek*, literally ‘two month’. These expressions may also refer to the lunar month.

No.	Day	Gloss	No.	Day	Gloss
1.	<i>bulan satu</i> 'month one'	January	2.	<i>bulan dua</i> 'month two'	February
3.	<i>bulan tiga</i> 'month three'	March	4.	<i>bulan empat</i> 'month four'	April
5.	<i>bulan lima</i> 'month five'	May	6.	<i>bulan enam</i> 'month six'	June
7.	<i>bulan tujuh</i> 'month seven'	July	8.	<i>bulan lapan</i> 'month eight'	August
9.	<i>bulan sembilan</i> 'month nine'	September	10.	<i>bulan puloh</i> 'month ten'	October
11.	<i>bulan se-belas/belair</i> 'month one-ten'	November	12.	<i>bulan dua-belas/belair</i> 'month-two-ten'	December
13.	<i>bulan tujuh satu</i>	first of July	14.	<i>bulan empat lima</i>	fifth of May

Table 57: Months in Baba Malay

Although the above expressions can be used to denote lunar month, Peranakans also use Hokkien forms to indicate special occasions marked by the lunar calendar.

No.	Day	Gloss	No.	Day	Gloss
1.	<i>chay-it</i> 'beginning.of.lunar.calendar-one'	first day of the lunar month; new year's day	2.	<i>chay-gau</i> 'beginning.of.lunar.calendar-nine' <i>ti-gong séh</i> 'sky-deity birth'	ninth day of the lunar month, birth of Jade Emperor (ninth day of first lunar month)
3.	<i>chap-gor</i> 'ten-five'	fifteenth day of lunar month; lovers' day	4.	<i>chap-gor méh</i> 'ten-five night'	night of the fifteenth day of the lunar;
5.	<i>chit guék</i> 'seven month'	seventh month; Hungry Ghost month	6.	<i>puay guék chap-gor</i> 'eight month ten-five'	fifteenth day of the eighth lunar month; Mid-autumn festival
7.	<i>tangchék</i>	Winter Solstice	8.	<i>lun-guék</i> intercalary-month	intercalary month, in the year that an extra month occurs
9.	<i>lun-chit-guék</i> 'intercalary-seven-month'	intercalary seventh month, in the year that there are two seventh months	10.	<i>ji-gau méh</i> 'two-nine night'	night of the twenty-ninth day of the twelfth lunar month; night of new year's eve

Table 58: Special occasions on the lunar calendar in BM

General time expressions are as follows:

No.	Time expression	Gloss	No.	Time expression	Gloss
1.	<i>pagi</i>	morning	2.	<i>hari/ari</i>	day
3.	<i>tengah hari/ari</i> 'middle day'	midday	4.	<i>petang</i>	evening
5.	<i>sinjakala</i>	dusk	6.	<i>malam</i>	night
7.	<i>siang hari/ari</i> 'day'	daytime	8.	<i>siang malam</i> 'night day'	night time
9.	<i>se-malam</i> 'one-night'	yesterday	10.	<i>ini hari/ ini ari/ ni ari</i> 'this day'	today
11.	<i>besok</i>	tomorrow	12.	<i>rusak</i>	day after tomorrow

Table 59: General time expressions in Baba Malay

The expressions for specific time are similar to those in Malay. The words *pagi* 'morning' or *petang* 'evening' follow these expressions to indicate if the event takes place before midday or after midday.

No.	Time	Gloss	No.	Day	Gloss
1.	<i>pukol satu</i> 'strike one'	one o'clock	2.	<i>pukol dua</i> 'strike two'	two o'clock
3.	<i>pukol tiga</i> 'strike three'	three o'clock	4.	<i>pukol empat</i> 'strike four'	four o'clock
5.	<i>pukol lima</i> 'strike five'	five o'clock	6.	<i>pukol enam</i> 'strike six'	six o'clock
7.	<i>pukol tujuh</i> 'strike seven'	seven o'clock	8.	<i>pukol lapan</i> 'strike eight'	eight o'clock
9.	<i>pukol sembilan</i> 'strike nine'	nine o'clock	10.	<i>pukol puloh</i> 'strike ten'	ten o'clock
11.	<i>pukol se-belas/belair</i> 'strike-ten'	eleven o'clock	12.	<i>pukol dua-belas/belair</i> 'strike twelve'	twelve o'clock
13.	<i>pukol se-belas puloh pagi</i> 'strike one-ten ten morning'	ten minutes past eleven in the morning	14.	<i>pukol lima lapan petang</i> 'strike five eight evening'	eight minutes past five in the evening

Table 60: Specific time in Baba Malay

Appendix B: Texts

Transcription conventions for texts

The transcription conventions used in the texts that follow are modelled after Du Bois et al's (2012) Discourse Transcription conventions. Each line on the transcript represents one intonation unit (see section 2.2.5). The interpretation of transitions between intonation units are based on the findings regarding BM intonation patterns reported in section 3.6.3. The following table shows the conventions used in the texts. These conventions are indicated in the lines representing intonation units.

No.	Symbol	Interpretation
1.	.	Final intonation boundary, usually associated with fall in overall pitch contour in BM
2.	,	Continuing intonation boundary, usually associated with a rise at the end of the intonation unit in BM
3.	?	Appeal intonation boundary, associated with overall rise-fall for WH-questions, and rise at the end of the intonation unit for tag questions in BM
4.	...	Noticeable pause between 0.3 to 0.6 seconds
5.	=	Lengthening on preceding segment
6.	@	Pulse of laughter
7.	#	Indecipherable syllable
8.	(2.1)	Pause of 2.1 seconds. Numbers in brackets represent duration in seconds. Pauses which are longer than 0.6 seconds are indicated by this convention.

Table 61: Transcription conventions for texts

Pear story¹⁰⁶ with Peter Wee

((NL1-022:Pear Story with Peter Wee: 0.00-379.00))

PETER WEE;

1.

Baba Malay IU	Ah ini dekat mana ini?
Baba Malay words	ah ini dekat mana
English gloss	filler this PREP where
English free translation	Ah what country (is) this?

2.

(2.7)

3.

Baba Malay IU	Ini tempat pokok,
Baba Malay words	ini tempat pokok
English gloss	this place tree
English free translation	This is an orchard,

4.

(1.1)

5.

Baba Malay IU	Ini buah pear,
Baba Malay words	Ini buah pear
English gloss	this fruit
English free translation	These are pears,

6.

Baba Malay IU	Gua rasa orang ni orang ada peték buah pear.
Baba Malay words	gua rasa ni orang ada peték buah pear.
English gloss	1.SG think this person PROG pluck fruit
English free translation	I think the person is plucking pears.

7.

Baba Malay IU	Masok ini kebun,
Baba Malay words	Masok ini kebun
English gloss	enter this garden
English free translation	Enter this garden,

8.

Baba Malay IU	Kebun dekat Europe ini.
Baba Malay words	Kebun dekat Europe ini
English gloss	garden PREP this
English free translation	This garden (is) in Europe.

¹⁰⁶ Chafe, Wallace (ed.), *The Pear Stories: Cognitive, Cultural, and Linguistic Aspects of Narrative Production*. Norwood, New Jersey: Ablex (1980)

9.
 Baba Malay IU Ada pear.
 Baba Malay words Ada pear
 English gloss EXIST
 English free translation There (are) pears.
10.
 Baba Malay IU Kita sini tak-a tak-a pokok.
 Baba Malay words Kita sini tak a tak a pokok
 English gloss 1.PL here NEG have NEG have tree
 English free translation We here, do not have trees.
11.
 Baba Malay IU Pear tau?
 Baba Malay words Pear tau
 English gloss know
 English free translation Pear (trees you) know?
12.
 Baba Malay IU Oh betol betol...oh satu sudah jatoh.
 Baba Malay words Oh betol betol oh satu sudah jatoh
 English gloss Filler true true Filler one already fall
 English free translation Oh true enough, one fell.
13.
 (2.6)
14.
 Baba Malay IU Buah pear bukan?
 Baba Malay words Buah pear bukan
 English gloss fruit no
 English free translation It (is) pear, no?
15.
 (1.0)
16.
 Baba Malay IU Mata gua tak gitu betol gua nampak begi buah paya hijau.
 Baba Malay words Mata gua tak gitu betol gua nampak begi buah paya hijau
 English gloss eye 1.SG NEG like.this correct 1.SG see like.that fruit papaya green
 English free translation My eyes (are) not accurate, (they) look like green papayas.
17.
 Baba Malay IU Bukan lah... ini semua pear=
 Baba Malay words Bukan lah ini semua pear
 English gloss no EMP this all
 English free translation No these are all pears.
18.
 Baba Malay IU Dia sudah dia sudah punggot punggot dia mia pear ... dia salin-kan dekat bakol.
 Baba Malay words Dia sudah dia sudah punggot punggot dia mia pear dia salin-kan dekat bakol
 English gloss 3.SG already 3.SG already pick.up pick.up 3.SG POSS 3.SG transfer-TR PREP basket
 English free translation He picked up his pears (and) transferred (them) to a basket.

19.
(2.4)

20.
Baba Malay IU Tu ini dia mia buah pear ini,
Baba Malay words Tu ini dia mia buah pear ini
English gloss that this 3.SG POSS fruit pear this
English free translation Those pears of his,

21.
Baba Malay IU Kalu mesti mo angkat tangan kalu tak-a jatoh kan dia,
Baba Malay words Kalu mesti mo angkat tangan tarok kalu tak a tak a jatoh-kan dia
English gloss if must want hold hand put if NEG EXIST NEG EXIST drop-TR 3.SG
English free translation If you want (you) put (your) hand (there), if not (you will) drop it.

22.
Baba Malay IU Dia sumua pichah.
Baba Malay words Dia sumua pichah
English gloss 3.PL all break
English free translation They all break.

23.
(1.9)

24.
Baba Malay IU Mesti angkat baik baik,
Baba Malay words Mesti angkat baik baik
English gloss must hold good good
English free translation Must hold them well,

25.
(1.9)

26.
Baba Malay IU Itu ada tangga...dia naik tangga atas pokok.
Baba Malay words Itu ada tangga dia naik tangga atas pokok
English gloss that COP ladder 3.SG climb ladder top tree
English free translation That is a ladder, he climbs the ladder up the tree.

27.
Baba Malay IU Sekarang dia mo apa?
Baba Malay words Sekarang dia mo apa
English gloss now 3.SG want what
English free translation Now what (does) he want (to do)?

28.
Baba Malay IU Satu satu dia mo apa bikin tu?
Baba Malay words Satu satu dia mo apa bikin tu
English gloss one one 3.SG want what make that
English free translation One by one, what (is) he doing?

29.
(2.3)

30.
Baba Malay IU Dia mo piléh gua rasa.
Baba Malay words Dia mo piléh gua rasa
English gloss 3.SG want choose 1.SG think
English free translation He wants to choose, I think.

31.
(2.1)

32.
Baba Malay IU Dia mo piléh mana baik...mana mana tak baik.
Baba Malay words Dia mo piléh mana baik mana satu tak baik
English gloss 3.SG want choose which good which one Neg good
English free translation He wants to choose which (is) good, and which one (is) not good.

33.
Baba Malay IU Mana sudah terbantot,
Baba Malay words Mana sudah terbantot
English gloss which already unripe
English free translation which has not ripened.

34.
Baba Malay IU Mana sudah masak,
Baba Malay words mana sudah masak
English gloss which already ripe
English free translation which (is) already ripe,

35.
Baba Malay IU Mana belum masak,
Baba Malay words Mana belum masak
English gloss which not.yet ripe
English free translation which (is) not yet ripe.

36.
Baba Malay IU Aye dia naik,
Baba Malay words Aye dia naik
English gloss EXCLAM.surprise 3. SG climb
English free translation Aye he is going up,

37.
Baba Malay IU Naik tangga pi pétek lagik.
Baba Malay words Naik tangga pi pétek lagik
English gloss climb ladder go pluck more
English free translation Climbing the ladder to go pluck more.

38.
Baba Malay IU Ini kawan dia datang sama apa?
Baba Malay words Ini kawan dia datang sama apa
English gloss this friend 3.SG come with what
English free translation These friends are coming with what?

39.								
Baba Malay IU	Sama	donkey	eh	#sama	goat.			
Baba Malay words	Sama	donkey	eh	sama				
English gloss	with		filler	with	goat			
English free translation	With a donkey, (no) with a goat.							
40								
(1.2)								
41.								
Baba Malay IU	Kambing=,							
Baba Malay words	Kambing							
English gloss	goat							
English free translation	Goat,							
42.								
Baba Malay IU	Kawan	dia	datang	angkat	satu	kambing.		
Baba Malay words	Kawan	dia	datang	angkat	satu	kambing		
English gloss	friend	3.SG	come	hold	one	goat		
English free translation	His friend comes holding a goat.							
43.								
Baba Malay IU	Kambing	dia	ada	locheng.				
Baba Malay words	Kambing	dia	ada	locheng				
English gloss	goat	3.SG	have	bell				
English free translation	His goat has a bell.							
44.								
Baba Malay IU	Itu	kambing	gua	tak	tau	perompuan	ka	jantan.
Baba Malay words	itu	kambing	gua	tak	tau	perompuan	ka	jantan
English gloss	that	goat	1.SG	Neg	know	female	or	male
English free translation	I don't know if that goat (is) male or female.							
45								
(3.2)								
46.								
Baba Malay IU	Aye	mampus...	kambing	dia	kuat	sekali.		
Baba Malay words	Aye	mampus	kambing	dia	kuat	sekali		
English gloss	EXCLAM.	surprise	dead	goat	3.SG	strong	very	
English free translation	Oh this goat is so strong!							
47.								
Baba Malay IU	Taré	dia.						
Baba Malay words	Taré	dia						
English gloss	pull	3.SG						
English free translation	Pulling it.							

48.
 Baba Malay IU Aye ini dia pétek lagik lah.
 Baba Malay words Aye ini dia pétek lagik lah
 English gloss EXCLAM.surprise this 3.SG pluck again Emp
 English free translation Ah he plucks again.
49.
 (2.4)
50.
 Baba Malay IU Buah pear...pear tree.
 Baba Malay words Buah pear pear tree
 English gloss fruit
 English free translation Pear tree.
51.
 (2.3)
52.
 Baba Malay IU Ala orang dia mia mungka bukan main garang sekali,
 Baba Malay words Ala orang dia mia mungka bukan main garang sekali
 English gloss EXCLAM.regret people 3.SG POSS face NEG play fierce very
 English free translation This person's face, (he is) not to be trifled with, (he) is very fierce,
53.
 Baba Malay IU Kus semangat.
 Baba Malay words Kus semangat
 English gloss EXCLAM.cry.to.a.dead.spirit
 English free translation Oh my goodness.
54.
 (3.0)
55.
 Baba Malay IU Ini ada bicycle=
 Baba Malay words Ini ada bicycle
 English gloss this EXIST
 English free translation There is a bicycle.
56.
 Baba Malay IU Gua rasa ini anak dia ka chuchu dia,
 Baba Malay words Gua rasa ini anak dia ka chuchu dia
 English gloss 1.SG feel this child 3.SG or grandchild 3.SG
 English free translation I think this (is) his child or his grandchild,
57.
 Baba Malay IU pi charék dia mia gonggong...gua rasa.
 Baba Malay words pi charék dia mia gonggong gua rasa
 English gloss go find 3.SG POSS grandfather 1.SG think
 English free translation goes to find his grandfather, I think.

58.
 Baba Malay IU Ini mia tempat bukan Singapore lah...bukan this region...gua rasa ini sumua
 Mediterranean.
 Baba Malay words Ini mia tempat bukan lah bukan gua rasa ini sumua Mediterranean
 English gloss this POSS place NEG EMP NEG 1.SG feel this all
 English free translation This place (is) neither Singapore, nor this region. It (is) all Mediterranean.
59.
 (3.3)
60.
 Baba Malay IU Ala panas sair pakay kopiah dia.
 Baba Malay words Ala panas sair pakay kopiah dia
 English gloss EXCLAM.regret hot indeed wear hat 3.SG
 English free translation Goodness, it (is) hot indeed, (the child) is wearing his hat.
61.
 (3.8)
62.
 Baba Malay IU Ah ini anak dia ka chuchu dia panggay dia mia gonggong.
 Baba Malay words Ah ini anak dia ka chuchu dia panggay dia mia gonggong.
 English gloss filler this child 3.SG or grandchild 3.SG call 3.SG POSS grandfather
 English free translation Ah this child or his grandchild is calling his grandfather.
63.
 (5.0)
64.
 Baba Malay IU Apa dia buat?
 Baba Malay words Apa dia buat
 English gloss what 3.SG do
 English free translation What (is) he doing?
65.
 (2.7)
66.
 Baba Malay IU Dia chakap sama dia.
 Baba Malay words Dia chakap sama dia
 English gloss 3.SG speak with 3.SG
 English free translation He is speaking with him.
67.
 (2.5)
68.
 Baba Malay IU Ah dia mo angkat.
 Baba Malay words Ah dia mo angkat
 English gloss filler 3.SG want lift
 English free translation Oh he wants to lift,

69.		
Baba Malay IU	Oh dia mo angkat tarok mana tarok dia mia,	
Baba Malay words	Oh dia mo angkat tarok mana tarok dia mia	
English gloss	Filler 3.SG want carry put where put 3.SG Poss	
English free translation	Oh he wants to carry and put where, put on his	
70.		
Baba Malay IU	Bicycle.	
Baba Malay words	Bicycle	
English gloss		
English free translation		
71.		
(1.0)		
72.		
Baba Malay IU	Eh tak jadi.	
Baba Malay words	Eh tak jadi	
English gloss	filler NEG happen	
English free translation	Ey it didn't happen.	
73.		
(5.5)		
74.		
Baba Malay IU	Ala...mia berat.	
Baba Malay words	Ala mia berat	
English gloss	EXCLAM.regret REL heavy	
English free translation	Goodness, (this) heavy.	
75.		
Baba Malay IU	Jatoh lah=.	
Baba Malay words	Jatoh lah	
English gloss	fall Emp	
English free translation	(He will) fall (for certain).	
76.		
Baba Malay IU	Amcam...oh dia tarok kat depan.	
Baba Malay words	Amcam oh dia tarok kat depan	
English gloss	how Filler 3.SG put PREP front	
English free translation	How? Oh he is putting (it) in front.	
77.		
(2.9)		
78.		
Baba Malay IU	Téngok budak anak itu kechik,	
Baba Malay words	téngok budak anak itu kechik	
English gloss	see child child that small	
English free translation	See that small child,	

79.
 Baba Malay IU Sudah boléh pakay bicycle.
 Baba Malay words Sudah boléh pakay bicycle
 English gloss already can use
 English free translation Already can use a bicycle.
80.
 (2.5)
81.
 Baba Malay IU Aye jatoh lagik.
 Baba Malay words Aye jatoh lagik
 English gloss EXCLAM.surprise fall again
 English free translation Ah (he) fell again.
82.
 (1.0)
83.
 Baba Malay IU Gua sudah kata nanti jatoh... lu téngok.
 Baba Malay words Gua sudah kata nanti jatoh lu téngok
 English gloss 1.SG already say later fall 2.SG see
 English free translation I told you (he) would fall later you see.
84.
 (0.9)
85.
 Baba Malay IU Mesti jatoh.
 Baba Malay words Mesti jatoh
 English gloss must fall
 English free translation Must fall.
86.
 Baba Malay IU Budak ini... sangat kechik lah.
 Baba Malay words Budak ini sangat kechik lah
 English gloss child this very small Emp
 English free translation This boy (is) so small.
87.
 (3.4)
88.
 Baba Malay IU Ah jumpa kawan dia.
 Baba Malay words Ah jumpa kawan dia
 English gloss filler meet friend 3.SG
 English free translation Ah (he) met his friends.
89.
 Baba Malay IU Gua rasa mesti ter-langgair tau?
 Baba Malay words Gua rasa mesti ter-langgair tau
 English gloss 1.SG think must ACD-crash know
 English free translation I think (there) must be a crash (you) know.

90
(1.5)

91.
Baba Malay IU Ada téngok ini dia mesti ter-langgair.
Baba Malay words Ada téngok ini dia mesti ter-langgair
English gloss PROG see this 3.SG must ACD-crash
English free translation Seeing this he must crash.

92.
(2.0)

93.
Baba Malay IU Ah dia téngok dia habis.
Baba Malay words Ah dia téngok dia habis
English gloss filler 3.PL see 3.SG finish
English free translation Ah they see him finish.

94.
Baba Malay IU Tu... gua sudah kata lu betol,
Baba Malay words Tu gua sudah kata lu betol,
English gloss that 1.SG already say 2.SG correct
English free translation That I already told you, correct,

95.
(1.0)

96.
Baba Malay IU Ala,
Baba Malay words Ala
English gloss EXCLAM.regret
English free translation Goodness,

97.
Baba Malay IU Habis... dia mia buah pear sumua,
Baba Malay words Habis dia mia buah pear sumua
English gloss finish 3.SG POSS fruit all
English free translation The end, his pears all

98.
Baba Malay IU Ter-lambong,
Baba Malay words Ter-lambong
English gloss ACD-toss
English free translation: were tossed,

99.
Baba Malay IU Bicycle jatoh,
Baba Malay words Bicycle jatoh
English gloss fall
English free translation The bicycle fell,

100.
 Baba Malay IU Budak jatuh,
 Baba Malay words Budah jatuh
 English gloss child fall
 English free translation The child fell,
101.
 Baba Malay IU Buah pear dia sumua sudah,
 Baba Malay words Buah pear dia sumua sudah
 English gloss fruit 3.PL all already
 English free translation The pears they all had
102.
 Baba Malay IU Sudah jatuh.
 Baba Malay words Sudah jatuh
 English gloss already fall
 English free translation Had fallen.
103.
 (1.0)
104.
 Baba Malay IU Ah kaki dia sudah kena.
 Baba Malay words Ah kaki dia sudah kena
 English gloss filler leg 3.SG already PASS
 English free translation Ah his foot was affected.
105.
 (3.0)
106.
 Baba Malay IU Ah siapa ni apa ni?
 Baba Malay words Ah siapa ni apa ni
 English gloss filler who this what this
 English free translation Ah who (is) this (and) what (is) this?
107.
 Baba Malay IU Ah kawan-kawan dia datang tolong.
 Baba Malay words Ah kawan kawan dia datang tolong
 English gloss filler friend friend 3.SG come help
 English free translation Ah his friends come to help.
108.
 (2.5)
109.
 Baba Malay IU Sumua tolong dia lah.
 Baba Malay words Sumua tolong dia lah
 English gloss all help 3.SG Emp
 English free translation All (are) helping him.

110.
 Baba Malay IU Punggot dia mia buah pear masok-kan bakol dia.
 Baba Malay words Punggot dia mia buah pear masok-kan bakol dia
 English gloss pick.up 3.SG POSS fruit enter-TR basket 3.SG
 English free translation Picking his pears putting them in his basket.
111.
 Baba Malay IU Aye dia ini lambong-kan.
 Baba Malay words Aye dia ini lambong-kan
 English gloss EXCLAM.surprise 3.SG this toss-TR
 English free translation He is throwing (it).
112.
 (1.0)
113.
 Baba Malay IU Si tua tu pelan-pelan,
 Baba Malay words Si tua tu pelan pelan
 English gloss PERSON old that slow slow
 English free translation The old man slowly,
114.
 Baba Malay IU Angkat satu-satu tak mo ini budak-budak amék-kan sumua sudah lambong.
 Baba Malay words Angkat satu satu tak mo ini budak budak amék-kan sumua sudah lambong
 English gloss hold one one NEG want this child child take-TR all already toss
 English free translation Holding (the pears) one by one, (if they) didn't want, these children would have taken them all (and) tossed them.
115.
 Baba Malay IU Gua rasa dia mia pear ini kerair.
 Baba Malay words Gua rasa dia mia pear ini kerair
 English gloss 1.SG think 3.PL POSS pear this hard
 English free translation I think their pears (are) hard.
116.
 Baba Malay IU Belom masak=
 Baba Malay words Belom masak
 English gloss not.yet ripe
 English free translation Not yet ripe.
117.
 Baba Malay IU Tu sudah masak dia bikin macam habis nua.
 Baba Malay words Tu sudah masak dia bikin macam habis nua
 English gloss that already ripe 3.SG make like finish smashed.up
 English free translation That (is) already ripe (and) he makes it smashed up.
118.
 (3.1)

119.		
Baba Malay IU	Ah ni apa ni?	
Baba Malay words	Ah ni apa ni	
English gloss	filler this what this	
English free translation	Ah what (is) this?	
120.		
(1.1)		
121.		
Baba Malay IU	Budak ini buat buat main main lah.	
Baba Malay words	Budak ini buat buat main main lah	
English gloss	child this do do play play Emp	
English free translation	The children are playing.	
122.		
(3.1)		
123.		
Baba Malay IU	Mmm=,	
Baba Malay words	Mmm	
English gloss		
English free translation		
124.		
Baba Malay IU	Sudah sudah habis ter-jalan balék.	
Baba Malay words	Sudah sudah habis ter-jalan balék	
English gloss	already already finish MVT-walk back	
English free translation	(when it is) already finished, he returns.	
125.		
Baba Malay IU	Ah sekarang dia tolak bicycle itu.	
Baba Malay words	Ah sekarang dia tolak bicycle itu	
English gloss	filler now 3.SG push that	
English free translation	Ah now he is pushing that bicycle.	
126.		
(1.5)		
127.		
Baba Malay IU	Dia mia dua... dia mia kawan.	
Baba Malay words	Dia mia dua dia mia kawan	
English gloss	3.SG POSS two 3.SG POSS friend	
English free translation	His two ... his friends.	
128.		
(0.8)		
129.		
Baba Malay IU	Dia jalan lain tempat dia jalan lain tempat.	
Baba Malay words	Dia jalan lain tempat dia jalan lain tempat	
English gloss	3.SG walk another place 3.PL walk another place	
English free translation	He walks to another place, they walk to another place.	

130.
(2.1)

131.

Baba Malay IU	Ah dia ni punggot apa ini.
Baba Malay words	Ah dia ni punggot apa ni
English gloss	filler 3.SG this pick.up what this
English free translation	Ah what (is) he picking up?

132.

Baba Malay IU	Oh dia punggot kopiah dia.
Baba Malay words	Oh dia punggot kopiah dia
English gloss	Filler 3.SG pick.up hat 3.SG
English free translation	Oh he is picking up his hat.

133.

Baba Malay IU	Kopiah dia sudah jatuh.
Baba Malay words	Kopiah ada sudah jatuh
English gloss	hat PFV already fall
English free translation	The hat had fallen.

134.

Baba Malay IU	Dia teriak dia,
Baba Malay words	Dia teriak dia
English gloss	3.SG call.out 3.SG
English free translation	He calls out to him.

135.

Baba Malay IU	Oi?
Baba Malay words	Oi
English gloss	
English free translation	

136.
(2.0)

137.

Baba Malay IU	Ah...lu sudah jatuh lu mia kopiah.
Baba Malay words	Ah lu sudah jatuh lu mia kopiah
English gloss	filler 2.SG already drop 2.SG POSS hat
English free translation	You dropped your hat.

138.
(1.6)

139.

Baba Malay IU	Ada angkat-kan kasi-kan dia balék,
Baba Malay words	Ada angkat-kan kasi-kan dia balék
English gloss	PFV pick.up-TR give-TR 3.SG back
English free translation	(He) picked (it) up (and) returned it back to him.

140.
 Baba Malay IU Habi dia kasi dia dua bijik pear.
 Baba Malay words Habi dia kasi dia dua bijik pear
 English gloss finish 3.SG give 3.SG two CLF.small.round
 English free translation Then he gave him two pears.
141.
 (1.0)
142.
 Baba Malay IU Amboi,
 Baba Malay words Amboi
 English gloss EXCLAM.surprise
 English free translation Suprising,
143.
 (1.5)
144.
 Baba Malay IU Dia kasi-kan kopiah,
 Baba Malay words Dia kasi-kan kopiah
 English gloss 3.SG give-TR hat
 English free translation He gave (him) the hat.
145.
 Baba Malay IU Dia dapat dua bijik pear balék huh,
 Baba Malay words Dia dapat dua bijik pear balék huh
 English gloss 3.SG receive two CLF.small.round back
 English free translation He received two pears back huh,
146.
 Baba Malay IU Happy dia.
 Baba Malay words Happy dia
 English gloss 3.SG
 English free translation He (is) happy.
147.
 Baba Malay IU Aye-
 Baba Malay words Aye-
 English gloss EXCLAM.surprise
 English free translation Surprising,
148.
 Baba Malay IU Ah dia kasi-kan kawan dia.
 Baba Malay words Ah dia kasi-kan kawan dia
 English gloss filler 3.SG give-TR friend 3.SG
 English free translation Ah he gave it to his friends.
149.
 (2.0)

150.
Baba Malay IU Ah tu lah kasi-kan kawan dia,
Baba Malay words Ah tu lah kasi-kan kawan dia
English gloss filler that EMP give-TR friend 3.SG
English free translation Ah that (he) gave to his friends.
151.
(0.9)
152.
Baba Malay IU Oh dia kasi dia tiga.
Baba Malay words Oh dia kasi dia tiga
English gloss Filler 3.SG give 3.PL three
English free translation Oh he gave them three.
153.
Baba Malay IU Oh sekarang budak-budak ini makan,
Baba Malay words Oh sekarangbudak budak ini makan
English gloss Filler now child child this eat
English free translation Oh now these children are eating,
154.
Baba Malay IU Gua rasa pear tu tak boléh makan lah,
Baba Malay words Gua rasa pear tu tak boléh makan lah
English gloss 1.SG think that NEG can eat EMP
English free translation I think those pears cannot be eaten.
155.
Baba Malay IU Kerair begi batu.
Baba Malay words Kerair begi batu
English gloss hard like.that rock
English free translation Hard like rocks.
- 156
(1.0)
157.
Baba Malay IU Tak masak mentah=
Baba Malay words Tak masak mentah
English gloss NEG ripe raw
English free translation Not ripe (still) raw.
158.
(2.0)
159.
Baba Malay IU Ah ni si tua dia.
Baba Malay words Ah ni si tua dia
English gloss Filler this PERSON old 3.SG
English free translation Ah this the old man he
160.
(0.8)

161.							
Baba Malay IU	Turun tangga.						
Baba Malay words	turun tangga						
English gloss	descend ladder						
English free translation	goes down the ladder.						
162.							
(1.0)							
163.							
Baba Malay IU	Téngok si tua pétek dia mia pear,						
Baba Malay words	Téngok si tua pétek dia mia pear						
English gloss	see PERSON old pluck 3.SG POSS						
English free translation	See the old man pluck his pear						
164.							
Baba Malay IU	Berapa sayang.						
Baba Malay words	Berapa sayang						
English gloss	how.many care						
English free translation	How much care.						
165.							
Baba Malay IU	Pelan-pelan=,						
Baba Malay words	Pelan pelan						
English gloss	slow slow						
English free translation	Slowly,						
166.							
Baba Malay IU	Satu-satu=,						
Baba Malay words	Satu satu						
English gloss	one one						
English free translation	One by one,						
167.							
Baba Malay IU	Budak masok-an.						
Baba Malay words	Budak masok-an						
English gloss	child enter TR						
English free translation	The boys enter.						
168.							
(3.0)							
169.							
Baba Malay IU	Aye mati sekair mana gua mia pear?						
Baba Malay words	Aye mati sekair mana gua mia pear						
English gloss	EXCLAM.surprise die indeed where 1.SG POSS						
English free translation	Oh goodness, die, where (are) my pears?						
170.							
Baba Malay IU	On satu bakol sudah hilang.						
Baba Malay words	Oh satu bakol sudah hilang						
English gloss	Filler one basket already lose						
English free translation	Oh one basket has gone missing.						

171.
(4.0)

172.

Baba Malay IU	Aye dia jumpa ini tiga ekor budak.
Baba Malay words	Aye dia jumpa ini tiga ekor budak
English gloss	EXCLAM.surprise 3.SG meet this three CLF.animals child
English free translation	Aye he meets these three children.

173.
(2.5)

174.

Baba Malay IU	Ada makan pear.
Baba Malay words	Ada makan pear
English gloss	PROG eat
English free translation	Eating pear.

175.
(6.0)

176.

Baba Malay IU	Oh sudah si tua ini heran,
Baba Malay words	Oh sudah si tua ini heran
English gloss	Filler already PERSON old this wonder
English free translation	Oh already this old man wonders,

177.

Baba Malay IU	Mana budak ini dapat pear #makan?
Baba Malay words	Mana budak ini dapat pear makan
English gloss	where child this get eat
English free translation	Where did these children get the pears to eat (from)?

178.
(2.0)

179.

Baba Malay IU	Ada lagik eh?
Baba Malay words	Ada lagik eh
English gloss	EXIST more filler
English free translation	Is there more?

180.

Baba Malay IU	Sumua habis.
Baba Malay words	Sumua habis
English gloss	all finish
English free translation	All finished.

181.
(3.0)

182.
Baba Malay IU Habis.
Baba Malay words Habis
English gloss finish
English free translation the end.

<T=379.00>

Turnip story¹⁰⁷ with Victor

((NL1-034:The Enormous Turnip with Victor: 0.00-183.00))

VICTOR;

1.

Baba Malay IU	Ada satu hari,
Baba Malay words	Ada satu hari
English gloss	EXIST one day
English free translation	One day,

2.

Baba Malay IU	Ada satu tukang kebun,
Baba Malay words	Ada satu tukang kebun
English gloss	EXIST one laborer garden
English free translation	There was a gardener,

3.

Baba Malay IU	Sama bini dia,
Baba Malay words	Sama bini dia
English gloss	and wife 3.SG
English free translation	And his wife,

4.

Baba Malay IU	Tanam,
Baba Malay words	Tanam
English gloss	plant
English free translation	(they) were planting,

5.

Baba Malay IU	Bangkuang.
Baba Malay words	Bangkuang
English gloss	turnip
English free translation	Turnip.

6.

Baba Malay IU	Dekat dia nia,
Baba Malay words	Dekat dia nia
English gloss	at 3.PL POSS
English free translation	At their,

7.

(1.8)

¹⁰⁷ The turnip story told here is a retelling of Alexander Afanasyev's The Enormous Turnip, published in 1863. See Афанасьев, А. Н. (1863)1984. Репка: Сказка N 89. Фундаментальная электронная библиотека: Русская литература и фольклор.

8.
 Baba Malay IU Kebun.
 Baba Malay words Kebun
 English gloss garden
 English free translation Garden.
9.
 Baba Malay IU Ada satu...bangkuang dia téngok betol besair.
 Baba Malay words Ada satu bangkuang dia téngok betol besair
 English gloss EXIST one turnip 3.PL see really big
 English free translation There was one turnip they saw (that was) really big.
10.
 (2.9)
11.
 Baba Malay IU Tukang kebun chakap,
 Baba Malay words Tukang kebun chakap
 English gloss laborer garden speak
 English free translation The gardener spoke,
12.
 Baba Malay IU Gua mo cha=bot itu bangkuang besair sekarang.
 Baba Malay words gua mo chabot itu bangkuang besair sekarang
 English gloss 1.SG want pull.out that turnip big now
 English free translation I want to pull out that big turnip now.
13.
 (1.4)
14.
 Baba Malay IU Lepas itu,
 Baba Malay words Lepas itu
 English gloss after that
 English free translation After that,
15.
 Baba Malay IU Kita boléh...buat,
 Baba Malay words Kita boléh buat
 English gloss 1.PL can make
 English free translation We can make,
16.
 (2.1)
17.
 Baba Malay IU Sup bangkuang.
 Baba Malay words Sup bangkuang
 English gloss soup turnip
 English free translation Turnip soup

18.
 Baba Malay IU Untok¹⁰⁸ makan malam.
 Baba Malay words untok makan malam
 English gloss for for eat night
 English free translation for dinner.
19.
 Baba Malay IU Dia nia bini chakap.
 Baba Malay words Dia nia bini chakap.
 English gloss 3.SG POSS wife speak
 English free translation His wife spoke.
20.
 (1.6)
21.
 Baba Malay IU Tukang kebun,
 Baba Malay words Tukang kebun
 English gloss laborer garden
 English free translation The gardener,
22.
 (1.1)
23.
 Baba Malay IU Tarék itu,
 Baba Malay words Tarék itu
 English gloss pull that
 English free translation Pulled that
24.
 Baba Malay IU Tarék itu,
 Baba Malay words Tarék itu
 English gloss pull that
 English free translation Pulled that
25.
 Baba Malay IU Bangkuang.
 Baba Malay words Bangkuang
 English gloss turnip
 English free translation Turnip.
26.
 Baba Malay IU Tapi tu bangkuang...tak bergerak.
 Baba Malay words Tapi tu bangkuang tak bergerak
 English gloss but that turnip NEG move
 English free translation But that turnip did not move.
27.
 (1.6)

¹⁰⁸ The usage of ‘untok’ is Malay. It can be used in BM by those who speak Malay.

28.
 Baba Malay IU Tukang kebun,
 Baba Malay words Tukang kebun
 English gloss laborer garden
 English free translation The gardener,
29.
 Baba Malay IU Ta=rék itu bangkuang lagik,
 Baba Malay words Tarék itu bangkuang lagik
 English gloss pull that turnip again
 English free translation Pulled that turnip again,
30.
 Baba Malay IU Tu bangkuang,
 Baba Malay words Tu bangkuang
 English gloss that turnip
 English free translation That turnip,
31.
 Baba Malay IU Pokok bangkuang...tak bergerak.
 Baba Malay words Pokok bangkuang tak bergerak
 English gloss tree turnip NEG move
 English free translation The turnip did not move.
32.
 (1.8)
33.
 Baba Malay IU Tukang kebun panggay dia nia bini tolong.
 Baba Malay words Tukang kebun panggay dia nia bini tolong
 English gloss laborer garden call 3.SG POSS wife help
 English free translation That gardener called his wife to help.
34.
 Baba Malay IU Tukang kebun sama bini dia,
 Baba Malay words Tukang kebun sama bini dia
 English gloss laborer garden and wife 3.SG
 English free translation The gardener and his wife,
35.
 Baba Malay IU Ta=rék itu bangkuang.
 Baba Malay words Tarék itu bangkuang
 English gloss pull that turnip
 English free translation Pulled that turnip.
36.
 Baba Malay IU Itu bangkuang pun...tak bergerak.
 Baba Malay words Itu bangkuang pun tak bergerak
 English gloss that turnip also NEG move
 English free translation That turnip also did not move.
37.
 (3.1)

38.
 Baba Malay IU Tukang kebun mia bini,
 Baba Malay words Tukang kebun mia bini
 English gloss laborer garden POSS wife
 English free translation The gardner' wife,
39.
 Baba Malay IU Panggay anjing dia tolong.
 Baba Malay words Panggay anjing dia tolong
 English gloss call dog 3.PL help
 English free translation Called their dog to help.
40.
 Baba Malay IU Tukang kebun,
 Baba Malay words Tukang kebun
 English gloss laborer garden
 English free translation The gardener,
41.
 Baba Malay IU Bini dia sama anjing,
 Baba Malay words Bini dia sama anjing
 English gloss wife 3.SG and dog
 English free translation his wife and the dog,
42.
 Baba Malay IU Ta=rék tu bangkuang.
 Baba Malay words Tarék tu bangkuang
 English gloss pull that turnip
 English free translation Pulled that turnip.
43.
 Baba Malay IU Tu bangkuang,
 Baba Malay words Tu bangkuang
 English gloss that turnip
 English free translation That turnip,
44.
 Baba Malay IU Pun= tak bergerak.
 Baba Malay words pun tak bergerak
 English gloss also NEG move
 English free translation also did not move.
45.
 (1.7)
46.
 Baba Malay IU Tu anjing,
 Baba Malay words Tu anjing
 English gloss that dog
 English free translation That dog,
47.
 (1.5)

48.		
Baba Malay IU	Menyalap.	
Baba Malay words	Menyalap	
English gloss	howl	
English free translation	howled.	
49.		
Baba Malay IU	Panggay kucing tolong.	
Baba Malay words	Panggay kucing tolong	
English gloss	call cat help	
English free translation	called the cat to help.	
50.		
Baba Malay IU	Gonggong ... #.	
Baba Malay words	Gonggong	
English gloss	ONOMATOPOEIA.bark	
English free translation	Barked.	
51.		
Baba Malay IU	Anjing tu gonggong.	
Baba Malay words	Anjing tu gonggong	
English gloss	dog that ONOMATOPOEIA.bark	
English free translation	That dog barked	
52.		
Baba Malay IU	Sama,	
Baba Malay words	Sama	
English gloss	with	
English free translation	with,	
53.		
Baba Malay IU	Si kucing.	
Baba Malay words	Si kucing	
English gloss	PERSON cat	
English free translation	the cat.	
54.		
Baba Malay IU	Mintak tolong.	
Baba Malay words	Mintak tolong	
English gloss	ask.sincerely help	
English free translation	Asking sincerely for help.	
55.		
Baba Malay IU	Tukang kebun,	
Baba Malay words	Tukang kebun	
English gloss	laborer garden	
English free translation	The gardener,	
56.		
Baba Malay IU	Bini dia,	
Baba Malay words	bini dia	
English gloss	wife 3.SG	
English free translation	his wife	

57.
 Baba Malay IU Anjing sama kucing,
 Baba Malay words anjing sama kucing
 English gloss dog and cat
 English free translation dog and cat,
58.
 Baba Malay IU Tarék tu bangkuang.
 Baba Malay words Tarék tu bangkuang
 English gloss pull that turnip
 English free translation Pulled that turnip.
59.
 Baba Malay IU Bangkuang tu pun...tak bergerak.
 Baba Malay words Bangkuang tu pun tak bergerak
 English gloss turnip that also NEG move
 English free translation That turnip also did not move.
60.
 (1.3)
61.
 Baba Malay IU Tukang kebun chakap,
 Baba Malay words Tukang kebun chakap
 English gloss laborer garden speak
 English free translation The gardener spoke,
62.
 Baba Malay IU Tak guna lah.
 Baba Malay words Tak guna lah
 English gloss NEG use EMP
 English free translation (It is) no use.
63.
 Baba Malay IU Ini=,
 Baba Malay words Ini
 English gloss this
 English free translation This,
64.
 Baba Malay IU Tak guna betol.
 Baba Malay words Tak guna betol
 English gloss Neg use really
 English free translation (It is) no use really.
65.
 Baba Malay IU Ini bangkuang,
 Baba Malay words Ini bangkuang
 English gloss this turnip
 English free translation This turnip,

66.
(2.0)

67.
Baba Malay IU Tak bergerak.
Baba Malay words Tak bergerak
English gloss NEG move
English free translation Does not move.

68.
(1.9)

69.
Baba Malay IU Si kucing,
Baba Malay words Si kucing
English gloss PERSON cat
English free translation The cat,

70.
Baba Malay IU Panggay,
Baba Malay words Panggay
English gloss call
English free translation called,

71.
Baba Malay IU Ah...si burong.
Baba Malay words Ah si burong
English gloss filler PERSON bird
English free translation the bird.

72.
Baba Malay IU Pi #tolong mo tolong.
Baba Malay words Pi tolong mo tolong
English gloss go help want help
English free translation Go help (we) want help.

73.
(1.7)

74.
Baba Malay IU Tukang kebun...bini dia,
Baba Malay words Tukang kebun bini dia
English gloss laborer garden wife 3.SG
English free translation The gardener, his wife,

75.
Baba Malay IU Anjing...kuching sama,
Baba Malay words Anjing kuching sama
English gloss dog cat and
English free translation dog, cat and

76.		
Baba Malay IU	Burong,	
Baba Malay words	Burong	
English gloss	bird	
English free translation	bird,	
77.		
Baba Malay IU	Ta=rék,	
Baba Malay words	Tarék	
English gloss	pull	
English free translation	Pulled	
78.		
(1.9)		
79.		
Baba Malay IU	Tarék itu bangkuang.	
Baba Malay words	Tarék itu bangkuang	
English gloss	pull that turnip	
English free translation	pulled that turnip.	
80.		
(1.5)		
81.		
Baba Malay IU	Sama sekejap tu bangkuang,	
Baba Malay words	sama sekejap tu bangkuang	
English gloss	with a.while that turnip	
English free translation	After a while that turnip	
82.		
Baba Malay IU	Pun= bergerak.	
Baba Malay words	Pun bergerak	
English gloss	also move	
English free translation	Also moved.	
83.		
(1.0)		
84.		
Baba Malay IU	Itu malam...sumua dapat,	
Baba Malay words	Itu malam sumua dapat	
English gloss	that night all get	
English free translation	That night all got,	
85.		
Baba Malay IU	Makan,	
Baba Malay words	Makan	
English gloss	eat	
English free translation	to eat,	

86.
 Baba Malay IU Sup bangkuang.
 Baba Malay words Sup bangkuang
 English gloss soup turnip
 English free translation turnip soup.
87.
 Baba Malay IU Untok makan malam,
 Baba Malay words for makan malam
 English gloss for eat night
 English free translation for dinner,
88.
 Baba Malay IU Untok makan besok pagi,
 Baba Malay words Untok makan besok pagi
 English gloss for eat tomorrow morning
 English free translation for breakfast tomorrow,
89.
 Baba Malay IU Untok makan tengah hari,
 Baba Malay words Untok makan tengah hari
 English gloss for eat middle day
 English free translation for lunch,
90.
 Baba Malay IU Untok makan,
 Baba Malay words Untok makan
 English gloss for eat
 English free translation to eat,
91.
 Baba Malay IU Minum téh nia jam.
 Baba Malay words Minum téh nia jam.
 English gloss drink tea REL time
 English free translation (for) teatime.

<T=183.00>

Conversation with Aunty Jane

((NL1-142: Conversation 2 with Jane Quek: 0.00-372.70))

1.

JQ;

IU ###,

2.

JQ;

Baba Malay IU	Ah tachi...apa khabair=?
Baba Malay words	Ah tachi apa khabair
English gloss	filler elder.sister what news
English free translation	Ah elder sister how are you?

3.

JQ;

Baba Malay IU	Lama tak jum=pa,
Baba Malay words	Lama tak jumpa
English gloss	long.time NEG meet
English free translation	Haven't met you in a long time,

4.

JQ;

Baba Malay IU	Ah...yah lah,
Baba Malay words	Ah yah lah
English gloss	filler yes EMP
English free translation	Ah yes,

5.

JQ;

Baba Malay IU	Lama tak...jalan sini,
Baba Malay words	Lama tak jalan sini
English gloss	long.time NEG walk here
English free translation	Haven't walked here in a long while,

6.

JQ;

Baba Malay IU	Tak jumpa lu lu apa macham?
Baba Malay words	Tak jumpa lu lu apa macham
English gloss	NEG meet 2.SG 2.SG what like that
English free translation	Haven't met you, how are you?

7.

JQ;

Baba Malay IU	Ada baik tak baik.
Baba Malay words	ada baik tak baik
English gloss	COP good NEG good
English free translation	Are you well or not.

8.

JQ;

Baba Malay IU	Se - lama ada pi main judi tak?				
Baba Malay words	se-lama ada pi main judi tak				
English gloss	one-long.time HAB go play gamble NEG				
English free translation	Do you still gamble regularly?				

9.

JQ;

Baba Malay IU	Mahjong se-ka se-kali lah=.				
Baba Malay words	Mahjong se-ka se-kali lah				
English gloss	one-time one-time EMP				
English free translation	(I play mahjong) once in a while.				

10.

JQ;

Baba Malay IU	Ada kaki...pi main,				
Baba Malay words	Ada kaki pi main				
English gloss	EXIST friend go play				
English free translation	If there are friends (I) go play,				

11.

JQ;

Baba Malay IU	Kalu tak-a orang panggay tak main,				
Baba Malay words	Kalu tak-a orang panggay tak main				
English gloss	if NEG-EXIST people call NEG play				
English free translation	If there is no calling (I) do not play,				

12.

JQ;

Baba Malay IU	Jalan jalan.				
Baba Malay words	Jalan jalan				
English gloss	walk walk				
English free translation	(I) take walks.				

13.

JQ;

Baba Malay IU	Abi main cherki tak?				
Baba Malay words	Abi main cherki tak				
English gloss	then play Peranakan.card.game NEG				
English free translation	Then do you play 'cherki'(card game)?				

14.

JQ;

Baba Malay IU	Cherki pun sama lah.				
Baba Malay words	Cherki pun sama lah				
English gloss	Peranakan.card.game also same EMP				
English free translation	'Cherki' (is) also the same.				

15.
 JQ;
 Baba Malay IU Se-lama kun...tak banyak orang main cherki.
 Baba Malay words Se-lama kun tak banyak orang main cherki
 English gloss one-long.time also NEG many people play Peranakan.card.game
 English free translation For a long time (now) also, there are not many people playing 'cherki'.

16.
 JQ;
 Baba Malay IU Tak # gitu main cherki lah.
 Baba Malay words Tak gitu main cherki lah
 English gloss NEG like.this play Peranakan.card.game EMP
 English free translation Not like this (I do not) play 'cherki'.

17.
 JQ;
 Baba Malay IU Tak apa pi jalan jalan,
 Baba Malay words Tak apa pi jalan jalan
 English gloss NEG what go walk walk
 English free translation (If there is) nothing (I) go for a walk,

18.
 JQ;
 Baba Malay IU Makan makan=,
 Baba Malay words Makan makan
 English gloss eat eat
 English free translation Eat (here and there),

19.
 JQ;
 Baba Malay IU Macham lah.
 Baba Malay words Macham lah
 English gloss like.that EMP
 English free translation Like that.

20.
 JQ;
 Baba Malay IU Pass...Pass time lah.
 Baba Malay words pass pass time lah
 English gloss EMP
 Free translation Pass time lah.

21.
 JQ;
 Baba Malay IU Day by day pass macham lah.
 Baba Malay words day by day pass macham lah
 English gloss like.that EMP
 English free translation Day by day pass like that.

22.
 (1.9)

23.

JQ;

Baba Malay IU	Ada pegi,
Baba Malay words	Ada pegi
English gloss	HAB go
English free translation	(I do) go,

24.

JQ;

Baba Malay IU	Mana lagik,
Baba Malay words	Mana lagik
English gloss	where more
English free translation	Where else,

25.

JQ;

Baba Malay IU	Apa macham lu ada berapa banyak chuchu?
Baba Malay words	Apa macham lu ada berapa banyak chuchu
English gloss	what like 2.SG have how.many many grandchild
English free translation	How are you, how many grandchildren do you have?

26.

JQ;

Baba Malay IU	Ada lah tiga.
Baba Malay words	Ada lah tiga
English gloss	have EMP three
English free translation	I have three.

27.

JQ;

Baba Malay IU	Anak prompuan nia satu=,
Baba Malay words	Anak prompuan nia satu
English gloss	Child female REL one
English free translation	One that is a girl,

28.

JQ;

Baba Malay IU	Anak jantan mia dua=,
Baba Malay words	Anak jantan mia dua
English gloss	child male REL two
English free translation	Two that are boys,

29.

JQ;

Baba Malay IU	Ah sudah tiga chuchu lah.
Baba Malay words	Ah sudah tiga chuchu lah
English gloss	filler already three grandchild EMP
English free translation	There (are) already three grandchildren.

30.

JQ;

Baba Malay IU	Lu jaga chuchu tak?
Baba Malay words	Lu jaga chuchu tak
English gloss	2.SG take.care grandchild NEG
English free translation	Do you take care of your grandchildren?

31.

JQ;

Baba Malay IU	Tak a lah tak gitu lah.
Baba Malay words	Tak-a lah tak gitu lah
English gloss	NEG-EXIST EMP NEG like.this EMP
English free translation	No not like this.

32.

JQ;

Baba Malay IU	Se-ka se-kali dia orang mo kuluar,			
Baba Malay words	Se-ka se-kali dia-orang	mo	kuluar	
English gloss	one-time one-time	3-PL	want	go.out
English free translation	Once in a while they want to go out.			

33.

JQ;

Baba Malay IU	Dia panggay...mata- mata-kan.
Baba Malay words	Dia panggay mata-mata -kan
English gloss	3.PL call eye-eye-TR
English free translation	They call me to watch over (the grandchildren).

34.

JQ;

Baba Malay IU	Ah...mata-mata-kan lah.
Baba Malay words	Ah mata-mata-kan lah
English gloss filler	eye-eye-TR EMP
English free translation	(I) watch (them).

35.

JQ;

Baba Malay IU	Satu dua jam macham lah.
Baba Malay words	Satu dua jam macham lah
English gloss	one two hour like.that EMP
English free translation	One (or) two hours like that.

36.

JQ;

Baba Malay IU	Macham sahja orang tua gitu lah.
Baba Malay words	Macham sahja orang tua gitu lah
English gloss	like.this only people old like.this EMP
English free translation	Like this only, old people (do it) like this.

37.						
JQ;						
Baba Malay IU	Boléh makan makan=,					
Baba Malay words	Boléh makan makan					
English gloss	can eat eat					
English free translation	(If you) can eat, eat.					
38.						
JQ;						
Baba Malay IU	Boléh jalan jalan=.					
Baba Malay words	Boléh jalan jalan					
English gloss	can walk walk					
English free translation	(If you) can walk, walk.					
39.						
JQ;						
IU	Be happy=,					
40.						
JQ;						
Baba Malay IU	Toksa sikit sikit marah.					
Baba Malay words	Toksa sikit sikit marah					
English gloss	do.not.need little little angry					
English free translation	(You) don't need to be angry over small things.					
41.						
JQ;						
Baba Malay IU	Apa sair lu siap-siap maki maid?					
Baba Malay words	Apa sair lu siap-siap maki maid					
English gloss	what reason 2.SG perpetually scold					
English free translation	Why do you perpetually scold the maid?					
42.						
JQ;						
Baba Malay IU	Maid kun orang jugak.					
Baba Malay words	Maid kun orang jugak					
English gloss	also people also					
English free translation	Maids (are) people also.					
43.						
JQ;						
Baba Malay IU	Abi gua téngok banyak= sair pi pasar.					
Baba Malay words	Abi gua téngok banyak sair pi pasar					
English gloss	then 1.SG see many indeed go market					
English free translation	Then I see many going to the market.					
44.						
JQ;						
Baba Malay IU	Sumua= complain pasair maid.					
Baba Malay words	Sumua complain pasair maid					
English gloss	all matter					
English free translation	All complaining about maid matters.					

45.

JQ;

Baba Malay IU	Maid kun orang ini lah maid ini lah.
Baba Malay words	Maid kun orang ini lah maid itu lah
English gloss	also people this EMP that EMP
English free translation	Maid (are) also people, (maid) this and (maid) that.

46.

JQ;

Baba Malay IU	Maid itu lah.
Baba Malay words	Maid itu lah
English gloss	that EMP
English free translation	The maid (does) that.

47.

JQ;

Baba Malay IU	Tak-a satu chakap maid baik.
Baba Malay words	Tak-a satu chakap maid baik
English gloss	NEG-EXIST one speak good
English free translation	Not one says the maid (is) good.

48.

JQ;

Baba Malay IU	Tak tau apa sair ni orang orang sumua.
Baba Malay words	Tak tau apa sair ni orang orang sumua
English gloss	NEG know what reason this people people all
English free translation	Don't know why these people all.

49.

JQ;

Baba Malay IU	Orang orang sumua.
Baba Malay words	orang orang sumua
English gloss	people people all
English free translation	People all.

50.

JQ;

IU	Huh?
----	------

51.

JQ;

Baba Malay IU	Orang maid pun orang human being,
Baba Malay words	Orang maid pun orang human being
English gloss	people also people
English free translation	The maids (are) also human beings,

52.

JQ;

Baba Malay IU	Kalu tak-a maid datang sama lu krejar,
Baba Malay words	Kalu tak-a maid datang sama lu krejar
English gloss	if NEG-EXIST come with 2.SG work
English free translation	If there are no maids to come work for you,

53.
JQ;
Baba Malay IU Orang chakap lu,
Baba Malay words Orang chakap lu
English gloss people speak 2.SG
English free translation People speak about you,
54.
JQ;
Baba Malay IU Misti mo ##,
Baba Malay words Misti mo
English gloss must want
English free translation Must ##,
55.
JQ;
Baba Malay IU Lu mia duit banyak besar bayar orang,
Baba Malay words Lu mia duit banyak besar bayar orang
English gloss 2.SG POSS money many big pay people
English free translation Your money (is) very big (it) pays people,
56.
JQ;
Baba Malay IU Su=mua bising= pasair maid.
Baba Malay words Sumua bising pasair maid
English gloss all noisy matter
English free translation All make a lot of noise about maid matters.
57.
JQ;
Baba Malay IU Dengar pun boring.
Baba Malay words Dengar pun boring
English gloss listen also
English free translation It is boring to even hear about this.
58.
(13.4)
59.
JQ;
Baba Malay IU Bila gua senang gua jalan jalan lah.
Baba Malay words Bila gua senang gua jalan jalan lah
English gloss when 1.SG free 1.SG walk walk EMP
English free translation When I (am) free I take walks.
60.
JQ;
Baba Malay IU Jalan pi Parkway Parade,
Baba Malay words Jalan pi Parkway Parade
English gloss walk go
English free translation Walk to Parkway Parade (shopping mall),

61.				
JQ				
Baba Malay IU	Jalan jalan makan,			
Baba Malay words	Jalan jalan makan			
English gloss	walk walk eat			
English free translation	Walk a bit and eat,			
62.				
JQ;				
Baba Malay IU	Téngok baju=,			
Baba Malay words	Téngok baju			
English gloss	see clothes			
English free translation	See clothes,			
63.				
JQ;				
Baba Malay IU	Ah chanték kita beli satu lay=,			
Baba Malay words	Ah chanték kita beli satu	lay		
English gloss	filler beautiful 1.PL buy one	CLF.piece of fabric or paper		
English free translation	(If it is) pretty we buy one piece,			
64.				
JQ;				
Baba Malay IU	Tak chanték jalan jalan=,			
Baba Malay words	Tak chanték jalan jalan			
English gloss	NEG beautiful walk walk			
English free translation	(If it is) not pretty (I) take a walk,			
65.				
JQ;				
Baba Malay IU	Minum kopi=,			
Baba Malay words	Minum kopi			
English gloss	drink coffee			
English free translation	Drink coffee,			
66.				
JQ;				
Baba Malay IU	Sama kawan kuluar=,			
Baba Malay words	Sama kawan kuluar			
English gloss	with friend go.out			
English free translation	Go out with friends,			
67.				
JQ;				
Baba Malay IU	Ada kali...gua mia friend kuluar,			
Baba Malay words	Ada kali gua mia friend kuluar			
English gloss	EXIST time 1.SG POSS friend kuluar			
English free translation	There are times my friend go out,			

68.
 JQ;
 Baba Malay IU From Australia panggay,
 Baba Malay words From Australia panggay
 English gloss call
 English free translation From Australia (she) calls ,
69.
 JQ;
 Baba Malay IU Aunty kita pi,
 Baba Malay words Aunty kita pi
 English gloss 1.PL go
 English free translation Aunty we go ,
70.
 JQ;
 Baba Malay IU Pi Takashimaya= Orchard Road side,
 Baba Malay words Pi Takashimaya Orchard Road side
 English gloss go
 English free translation Go to Takashimaya (shopping mall) at Orchard Road,
71.
 JQ;
 Baba Malay IU Jalan jalan=,
 Baba Malay words Jalan jalan
 English gloss walk walk
 English free translation Take walk,
72.
 JQ;
 Baba Malay IU Téngok téngok,
 Baba Malay words Téngok téngok
 English gloss see see
 English free translation Take a look-see,
73.
 JQ;
 Baba Malay IU Kalu téngok barang hak,
 Baba Malay words Kalu téngok barang hak
 English gloss if see thing suitable
 English free translation If (we) see a suitable thing,
74.
 JQ;
 Baba Malay IU Kita beli lah.
 Baba Malay words Kita beli lah
 English gloss 1.PL buy EMP
 English free translation We buy.

75.							
JQ;							
Baba Malay IU		Barang tak hak tak beli lah.					
Baba Malay words		Barang tak hak tak beli lah					
English gloss		thing NEG suitable NEG buy EMP					
English free translation		(If) the thing (is) not suitable we won't buy.					
76.							
JQ;							
Baba Malay IU		Pegi Isetan,					
Baba Malay words		Pegi Isetan					
English gloss		go					
English free translation		Go to Isetan (departmental store),					
77.							
JQ;							
Baba Malay IU		Pegi Cold Storage,					
Baba Malay words		Pegi Cold Storage					
English gloss		go					
English free translation		Go to Cold Storage (supermarket),					
78.							
JQ;							
Baba Malay IU		Gua mia kawan suka pi Isetan nia,					
Baba Malay words		Gua mia kawan suka pi Isetan nia					
English gloss		1.SG POSS friend like go POSS					
English free translation		My friend likes to go to Isetan's,					
79.							
JQ;							
IU		Shopping centre,					
80.							
JQ;							
Baba Malay IU		Pasair dia mia meat very fresh.					
Baba Malay words		pasair dia mia meat very fresh					
English gloss		because 3.SG Poss					
English free translation		Because its meat (is) very fresh.					
81.							
JQ;							
Baba Malay IU		Tapi mahal lah.					
Baba Malay words		Tapi mahal lah					
English gloss		but expensive EMP					
English free translation		But it (is) expensive.					
82.							
JQ;							
Baba Malay IU		Dia mia barang mahal,					
Baba Malay words		Dia mia barang mahal					
English gloss		3.SG POSS thing expensive					
English free translation		Its things (are) expensive,					

83.						
JQ;						
Baba Malay IU	Tapi	dia	mia	meat	betol...fresh.	
Baba Malay words	Tapi	dia	mia	meat	betol	fresh
English gloss	but	3.SG	POSS		really	
English free translation	But its meat (is) really fresh.					
84.						
JQ;						
Baba Malay IU	Chanték sekali.					
Baba Malay words	Chanték		sekali			
English gloss	beautiful		very			
English free translation	Very beautiful.					
85.						
JQ;						
Baba Malay IU	Dia bikin ## steamboat,					
Baba Malay words	Dia	bikin	steamboat			
English gloss	3.SG	make				
English free translation	She makes steamboat (hot pot),					
86.						
JQ;						
Baba Malay IU	Dia selalu pi sana.					
Baba Malay words	Dia	selalu	pi	sana		
English gloss	3.SG	always	go	there		
English free translation	She always goes there.					
87.						
JQ;						
Baba Malay IU	Dia tak beli kita market nia.					
Baba Malay words	Dia	tak	beli	kita	market	nia
English gloss	3.SG	NEG	buy	1.PL		REL
English free translation	She doesn't buy the ones that are at our markets.					
88.						
JQ;						
Baba Malay IU	Dia sudah used to it lah.					
Baba Malay words	Dia	sudah	used	to	it	lah
English gloss	3.SG	already				EMP
English free translation	She (is) already used to it.					
89.						
JQ;						
IU	My Australia friend is like that,					
90.						
JQ;						
Baba Malay IU	Tapi ya lah betul betul fresh.					
Baba Malay words	Tapi	ya	lah	betul	betul	fresh
English gloss	but	yes	EMP	really	really	
English free translation	But yes it is really really fresh.					

91.				
JQ;				
Baba Malay IU	Mahal	sikit	lah.	
Baba Malay words	Mahal	sikit	lah	
English gloss	expensive	little	EMP	
English free translation	A little expensive.			
92.				
JQ;				
Baba Malay IU	Jalan	jalan	abi	tu
Baba Malay words	Jalan	jalan	abi	tu
English gloss	walk	walk	finish	that
English free translation	Take a walk and after that eat,			
93.				
JQ:				
Baba Malay IU	Lunch	habi,		
Baba Malay words	Lunch	habi		
English gloss		finish		
English free translation	Lunch (is) over,			
94.				
JQ;				
Baba Malay IU	Jalan	jalan	habi,	
Baba Malay words	Jalan	jalan	habi	
English gloss	walk	walk	finish	
English free translation	After taking a walk,			
95.				
JQ;				
Baba Malay IU	Dudok	kopi	house,	
Baba Malay words	Dudok	kopi	house	
English gloss	sit	coffee		
English free translation	Sit at the coffee house,			
96.				
JQ;				
Baba Malay IU	Minum	kopi	lah.	
Baba Malay words	Minum	kopi	lah	
English gloss	drink	coffee	EMP	
English free translation	Drink coffee.			
97.				
JQ;				
Baba Malay IU	Cappucino	ka=,		
Baba Malay words	Cappucino	ka		
English gloss		or		
English free translation	Cappucino or,			

98.
 JQ;
 Baba Malay IU Latte ka=
 Baba Malay words Latte ka
 English gloss or
 English free translation Latte or,
99.
 JQ;
 Baba Malay IU Makan ## itu pancake share share,
 Baba Malay words Makan itu pancake share share
 English gloss eat that
 English free translation Share some pancake,
100.
 JQ;
 Baba Malay IU Macham lah.
 Baba Malay words Macham lah
 English gloss like.that EMP
 English free translation Like that.
101.
 JQ;
 Baba Malay IU Dudok= sampay pukol lima,
 Baba Malay words Dudok sampay pukol lima
 English gloss sit until strike five
 English free translation Sit until five o'clock,
102.
 JQ;
 Baba Malay IU ## Charék makan dinner,
 Baba Malay words Charék makan dinner
 English gloss find eat
 English free translation Find dinner to eat,
103.
 JQ;
 Baba Malay IU Abi itu pulang,
 Baba Malay words Abi itu pulang
 English gloss finish that return
 English free translation After that (we) return,
104.
 JQ;
 Baba Malay IU Pulang sampay rumah,
 Baba Malay words Pulang sampay rumah
 English gloss return until house
 English free translation (We) return home,
105.
 JQ;
 IU Seven eight or nine,

106.
JQ;
Baba Malay IU Pukol lapan pukol semilan
Baba Malay words Pukol lapan pukol semilan
English gloss strike eight strike nine
English free translation Eight or night o'clock.
107.
JQ;
Baba Malay IU macham lah.
Baba Malay words macham lah
English gloss like.that EMP
108.
JQ;
Baba Malay IU Balék rumah mandi=
Baba Malay words Balék rumah mandi
English gloss return house bathe
English free translation Return home to bathe,
109.
JQ;
Baba Malay IU Dudok= téngok tv,
Baba Malay words Dudok téngok tv
English gloss sit see
English free translation Sit and watch tv,
110.
JQ;
Baba Malay IU Tv bagus=
Baba Malay words Tv bagus
English gloss good
English free translation If the television show is good,
111.
JQ;
Baba Malay IU Téngok satu pukol dua belas.
Baba Malay words Téngok sampay pukol dua belas
English gloss see until strike two ten
English free translation Watch until twelve o'clock.
112.
JQ;
Baba Malay IU Tv tak bagus,
Baba Malay words Tv tak bagus
English gloss NEG good
English free translation (if) the television show (is) not good,

113.

JQ;

Baba Malay IU	Se-belas	tutup.
Baba Malay words	Se-belas	tutup
English gloss	one-ten	close
English free translation	Turn off at eleven.	

114.

JQ;

Baba Malay IU	Pi	tidor.
Baba Malay words	Pi	tidor
English gloss	go	sleep
English free translation	Go sleep.	

115.

JQ;

Baba Malay IU	Then...	pagi	sudah	mangun.
Baba Malay words	Then	pagi	sudah	mangun
English gloss		morning	already	wake
English free translation	Then in the morning I already wake up.			

116.

JQ;

Baba Malay IU	Sua	tua=	tak	boléh	tidor	tau?
Baba Malay words	Sudah	tua	tak	boléh	tidor	tau
English gloss	already	old	NEG	can	sleep	know
English free translation	Already old (I am) cannot sleep (you) know?					

117.

JQ;

Baba Malay IU	Kalu	apa	macham	lambat,
Baba Malay words	Kalu	apa	macham	lambat
English gloss	if	what	like	late
English free translation	Even if (I sleep) late,			

118.

JQ;

Baba Malay IU	Tidor	pun	tak	boléh	tidor.
Baba Malay words	Tidor	pun	tak	boléh	tidor
English gloss	sleep	also	NEG	can	sleep
English free translation	(I still) cannot sleep.				

119.

JQ;

Baba Malay IU	Apa	macham	pukol	lima	lebéh	sudah	mangun.	
Baba Malay words	Apa	macham	pukol	lima	lebéh	sudah	mangun	
English gloss	what	like	that	strike	five	more	already	wake
English free translation	However (late I go to bed, I) wake up slightly after five.							

120.					
JQ;					
Baba Malay IU	Mata sudah ter-bukak.				
Baba Malay words	Mata sudah ter-bukak				
English gloss	eye already MVT-open				
English free translation	Eyes already opened.				
121.					
JQ;					
Baba Malay IU	Sudah tak boléh tidor balék.				
Baba Malay words	Sudah tak boléh tidor	balék			
English gloss	already NEG can sleep	return			
English free translation	Already cannot return to sleep.				
122.					
JQ;					
Baba Malay IU	Orang kata,				
Baba Malay words	Orang kata				
English gloss	people say				
English free translation	People say,				
123.					
JQ;					
Baba Malay IU	boléh tidor pukol se-puloh lah.				
Baba Malay words	boléh tidor pukol se	puloh lah			
English gloss	can sleep strike one	ten EMP			
English free translation	(they) can sleep until ten o'clock.				
124.					
JQ;					
Baba Malay IU	Aye=,				
Baba Malay words	Aye				
English gloss	EXCLAM.surprise,				
English free translation	Goodness,				
125.					
JQ;					
Baba Malay IU	Ho mia,				
Baba Malay words	Ho mia				
English gloss	good life				
English free translation	Good life,				
126.					
JQ;					
Baba Malay IU	Gua tak boléh...pukol se-puloh,				
Baba Malay words	Gua tak boléh pukol	se-puloh			
English gloss	1.SG NEG can strike	one-ten			
English free translation	I cannot (sleep until) ten o'clock,				

127.

JQ;

Baba Malay IU	Amcham pun,
Baba Malay words	Amcham pun
English gloss	how also
English free translation	However (much I try),

128.

JQ;

Baba Malay IU	Pukul lima lebéh gua sudah mangun.
Baba Malay words	Pukul lima lebéh gua sudah mangun
English gloss	strike five more 1.Sg already wake
English free translation	After five I already wake up.

129.

JQ;

Baba Malay IU	Dia orang boléh tidor pukul se-puloh.
Baba Malay words	Dia-orang boléh tidor pukul se-puloh
English gloss	3-PL can sleep strike one-ten
English free translation	They can sleep (until) ten o'clock.

130.

JQ;

Baba Malay IU	Mata ari gemoh pantat pun tak boléh mangun.
Baba Malay words	Mata ari gemoh pantat pun tak boléh mangun
English gloss	eye day sun.dry buttocks also NEG can wake
English free translation	The sun shines on your backside (they) also cannot get up.

131.

(6.0)

132.

NL;

Baba Malay IU	Australia friend bila dia mo balék?
Baba Malay words	Australia friend bila dia mo balék
English gloss	when 3.SG want return
English free translation	When does your friend from Australia want to return?

133.

JQ;

Baba Malay IU	Bila dia pulang=,
Baba Malay words	Bila dia pulang
English gloss	when 3.SG return
English free translation	When she returns,

134.

JQ;

Baba Malay IU	Tiga minggu...macham dia balék lah.
Baba Malay words	Tiga minggu macham dia balék lah
English gloss	three week like.that 3.SG return EMP
English free translation	Three weeks or so she returns.

135.

JQ;

Baba Malay IU	Because dia Singaporean=.
Baba Malay words	Because dia Singaporean
English gloss	3.SG
English free translation	Because she (is) Singaporean.

136.

JQ;

Baba Malay IU	Pasair dia tinggal Singapore,
Baba Malay words	Pasair dia tinggal Singapore
English gloss	because 3.SG live
English free translation	Because she lives in Singapore,

137.

JQ;

Baba Malay IU	Dia kena balék for,
Baba Malay words	Dia kena balék for
English gloss	3.SG subjected.to return
English free translation	She has to return for,

138.

JQ;

Baba Malay IU	Ini apa?
Baba Malay words	Ini apa
English gloss	this what
English free translation	This what?

139.

JQ;

Baba Malay IU	Chop kan?
Baba Malay words	Chop kan
English gloss	NEG
English free translation	The stamp (on the passport) no?

140.

JQ;

Baba Malay IU	Ah...balék kat Australia lama,
Baba Malay words	Ah balék kat Australia lama
English gloss	filler return PREP long.time
English free translation	She returned to Australia for a long time,

141.

JQ;

IU	Seven eight years.
----	--------------------

142.

JQ;

Baba Malay IU	Macham lah se-ka se-kali dia kuluair,
Baba Malay words	Macham lah se-ka se-kali dia kuluair
English gloss	like.that EMP one-time one-time 3.SG go.out
English free translation	Like that once in a while she comes out,

143.

JQ;

Baba Malay IU	Jalan jalan=,
Baba Malay words	jalan jalan
English gloss	walk walk
English free translation	Take a walk,

144.

JQ;

Baba Malay IU	Dua tiga minggu=,
Baba Malay words	Dua tiga minggu
English gloss	two three week
English free translation	Two three weeks,

145.

JQ;

Baba Malay IU	Satu bulan dia pulang lah.
Baba Malay words	Satu bulan dia pulang lah
English gloss	one month 3.SG return EMP
English free translation	(After) one month she returns.

146.

JQ;

Baba Malay IU	Macham lah.
Baba Malay words	Macham lah
English gloss	like.that EMP
English free translation	Like that.

147.

JQ;

Baba Malay IU	Tapi dia orang tak suka Singapore.
Baba Malay words	Tapi dia-orang tak suka Singapore
English gloss	but 3-PL NEG like
English free translation	But she doesn't like Singapore.

148.

JQ;

Baba Malay IU	Manyak panas dia tak boléh tahan.
Baba Malay words	Manyak panas dia tak boléh tahan
English gloss	many hot 3.SG NEG can withstand
English free translation	She cannot stand that it is very hot.

149.

NL;

Baba Malay IU	Dia toksa jaga chuchu ah?
Baba Malay words	Dia toksa jaga chuchu ah
English gloss	3.SG do.not.need take.care grandchild filler
English free translation	She doesn't have to take care of her grandchildren?

150.

JQ;

Baba Malay IU

Dia tak boléh tahan panas.

Baba Malay words

Dia tak boléh tahan panas

English gloss

3.SG NEG can withstand hot

English free translation

She cannot stand that it is hot.

151.

JQ;

Baba Malay IU

Dia sendiri pi jalan suka pi aircon.

Baba Malay words

Dia sendiri pi jalan suka pi aircon

English gloss

3.SG self go walk like go

English free translation

She herself likes to go to airconditioned places when she walks.

152.

JQ;

Baba Malay IU

Makan itu foodcourt,

Baba Malay words

Makan itu foodcourt

English gloss

eat that

English free translation

Eat at that foodcourt,

153.

JQ;

Baba Malay IU

Dekat tengah jalan foodcourt dia tak masok.

Baba Malay words

Dekat tengah jalan foodcourt dia tak masok

English gloss

PREP middle walk 3.SG NEG enter

English free translation

The foodcourt in the middle of the road she doesn't enter.

154.

JQ;

Baba Malay IU

Dia pegi # aircon mia foodcourt.

Baba Malay words

Dia pegi aircon mia foodcourt

English gloss

3.SG go REL

English free translation

She goes to foodcourts that are air-conditioned.

155.

JQ;

Baba Malay IU

Foodcourt foodcourt pun dia seldom,

Baba Malay words

Foodcourt foodcourt pun dia seldom

English gloss

also 3.Sg

English free translation

She also seldom (even goes to) the foodcourt,

156.

JQ;

Baba Malay IU

Dia téngok ma=na nia foodcourt.

Baba Malay words

Dia téngok mana nia foodcourt

English gloss

3.SG see which REL

English free translation

She sees which foodcourt it is.

157.

JQ;

Baba Malay IU	Takashimaya nia foodcourt ka apa.				
Baba Malay words	Takashimaya nia foodcourt	ka	apa		
English gloss	POSS	or	what		
English free translation	Takashimaya's foodcourt or what.				

158.

JQ;

Baba Malay IU	Lain nia dia tak pi.				
Baba Malay words	Lain	nia	dia	tak	pi
English gloss	other	REL	3.SG	Neg	go
English free translation	Others she does not go.				

159.

NL;

Baba Malay IU	Ada apa tak sama?			
Baba Malay words	Ada	apa	tak	sama
English gloss	EXIST	what	NEG	same
English free translation	What is different?			

160.

JQ;

Baba Malay IU	Hot panas dirty lah.			
Baba Malay words	Hot	panas	dirty	lah
English gloss		hot		EMP
English free translation	Hot, dirty.			

161.

JQ;

Baba Malay IU	Tempat tu ko=tor=.		
Baba Malay words	Tempat	tu	kotor
English gloss	place	that	dirty
English free translation	That place (is) dirty.		

162.

JQ;

Baba Malay IU	Kuluar nia tak bagus kotor.				
Baba Malay words	kuluar	nia	tak	bagus	kotor
English gloss	outside	REL	NEG	good	dirty
English free translation	(Those) that are outside are not good, dirty.				

163.

JQ;

Baba Malay IU	Dia orang sudah biasa macham.				
Baba Malay words	Dia-orang	sudah	biasa		macham
English gloss	3-PL		already used.to.it		like.that
English free translation	She is already used to it like that.				

164.
 JQ;
 Baba Malay IU Kita apa court pun ma=sok lah.
 Baba Malay words Kita apa court pun masok lah
 English gloss I.PL what also enter EMP
 English free translation We enter whatever court it is.

165.
 JQ;
 IU Small court big court,

166.
 JQ;
 Baba Malay IU Sumua ma=sok lah.
 Baba Malay words Sumua masok lah
 English gloss all enter EMP
 English free translation All (we) enter.

167.
 JQ;
 Baba Malay IU Ada radat,
 Baba Malay words Ada radat
 English gloss Exist greed.vulgar
 English free translation (We) have greed (vulgar),

168.
 JQ;
 Baba Malay IU Cher=kek sudah cher=kek lah.
 Baba Malay words Cherkek sudah cherkek lah
 English gloss eat.vulgar already eat.coarse EMP
 English free translation Already eaten (we) eat more.

169.
 JQ;
 IU @@@@.

170.
 JQ;
 Baba Malay IU Macham lah apa mo bikin,
 Baba Malay words Macham lah apa mo bikin
 English gloss like.that EMP what want make
 English free translation Like that what (do we) want to do,

171.
 JQ;
 Baba Malay IU Human being macham lah.
 Baba Malay words Human being macham lah
 English gloss like.that EMP
 English free translation Human beings are like that.

172.
JQ;
Baba Malay IU Dia sudah biasa,
Baba Malay words Dia sudah biasa
English gloss 3.SG already used.to.it
English free translation She (is) used to it,
173.
JQ;
Baba Malay IU From young macham,
Baba Malay words From young macham
English gloss like.that
English free translation From young like that,
174.
JQ;
Baba Malay IU Kita ini tak ah kita ini chin=chai lah.
Baba Malay words Kita ini tak ah kita ini chinchai lah
English gloss 1.PL this NEG filler 1.PL this not.fussy EMP
English free translation We (are) not we (are) not fussy.
175.
JQ;
Baba Malay IU Ma=na pun dudok ma=kan lah.
Baba Malay words Mana pun dudok makan lah
English gloss where also sit eat EMP
English free translation Anywhere we sit and eat.
176.
JQ;
Baba Malay IU Orang panggay makan ma=kan lah.
Baba Malay words Orang panggay makan makan lah
English gloss people call eat eat EMP
English free translation (If) people call us to eat (we) eat.
177.
JQ;
Baba Malay IU Ada #...ini saya makan,
Baba Malay words Ada ini saya makan
English gloss EXIST this 1.SG eat
English free translation (If) there is this, I (will) eat,
178.
JQ;
Baba Malay IU Itu tak tau makan.
Baba Malay words Itu tak tau makan
English gloss that NEG know eat
English free translation That (you) do not know how to eat.

179.

JQ;

Baba Malay IU Bodoh mo mampus.
Baba Malay words Bodoh mo mampus
English gloss stupid want EXCLAM.dead
English free translation Stupid (until I) want to die. [humorous intent]

180.

JQ;

Baba Malay IU Stupid sekali tak tau makan.
Baba Malay words Stupid sekali tak tau makan
English gloss very NEG know eat
English free translation Very stupid (if you) do not know how to eat. [humorous intent]

181.

JQ;

Baba Malay IU Ini tak tau makan,
Baba Malay words Ini tak tau makan
English gloss this NEG know eat
English free translation This (you) do not know how to eat,

182.

JQ;

Baba Malay IU Ada kali kawan pun sama,
Baba Malay words Ada kali kawan pun sama
English gloss EXIST time friend also same
English free translation There are times friends (are) also the same,

183.

JQ;

Baba Malay IU Nasik ulam dia tak tau makan.
Baba Malay words Nasik ulam dia tak tau makan
English gloss cooked.rice mixed.herbs.anchovies 3.PL NEG know eat
English free translation Mixed herbs rice with shredded anchovies they do not know how to eat.

184.

JQ;

Baba Malay IU Nasik goreng pun tak tau makan.
Baba Malay words Nasik goreng pun tak tau makan
English gloss cooked.rice fry also Neg know eat
English free translation Fried rice too (they) do not know how to eat.

185.

JQ;

Baba Malay IU A=pa nia Peranakan?
Baba Malay words Apa nia Peranakan
English gloss what REL Peranakan
English free translation What kind of Peranakan? [humorous intent]

186.

JQ;

Baba Malay IU	Tak sepekah.
Baba Malay words	Tak sepekah
English gloss	Neg acceptable
English free translation	Not proper.

187.

NL;

Baba Malay IU	Aunty Jane suka masak tak?
Baba Malay words	Aunty Jane suka masak tak
English gloss	like cook NEG
English free translation	Aunty Jane likes to cook?

188.

JQ;

Baba Malay IU	Su=ka=.
Baba Malay words	Suka
English gloss	like
English free translation	(I) like.

189.

JQ;

Baba Malay IU	Tapi selalu gua chakap,
Baba Malay words	Tapi selalu gua chakap
English gloss	but always I.SG speak
English free translation	But I always say,

190.

JQ;

Baba Malay IU	Lain kali kalu gua mati ah,
Baba Malay words	Lain kali kalu gua mati ah
English gloss	other time if I.SG die filler
English free translation	Next time if I die,

191.

JQ;

Baba Malay IU	Kalu chutsi gua tak mo angkat pot.
Baba Malay words	Kalu chutsi gua tak mo angkat pot
English gloss	if born I.SG NEG want carry
English free translation	If reborn I do not want to carry the pot.

192.

JQ;

Baba Malay IU	Tak mo angkat pen.
Baba Malay words	Tak mo angkat pen
English gloss	NEG want carry
English free translation	I do not want to carry the pen.

193.

JQ;

Baba Malay IU	Hari hari tak bikin apa,
Baba Malay words	Hari hari tak bikin apa
English gloss	day day NEG do what
English free translation	Everyday (I) do not do anything ,

194.

JQ;

Baba Malay IU	Ma=sak sahja.
Baba Malay words	Masak sahja
English gloss	cook only
English free translation	Only cook.

195.

NL;

Baba Malay IU	Angkat pen?
Baba Malay words	Angkat pen
English gloss	carry
English free translation	Carry the pen?

196.

JQ;

Baba Malay IU	Angkat pot pen sumua tak mo.
Baba Malay words	Angkat pot pen sumua tak mo
English gloss	carry all Neg want
English free translation	Carry the pot pen I do not want all.

197.

JQ;

Baba Malay IU	Mo angkat duit kuluar.
Baba Malay words	Mo angkat duit kuluar
English gloss	want carry money out
English free translation	(I) want to carry money out (when reborn). [humorous intent]

198.

NL;

IU @@@@@.

199.

JQ;

Baba Malay IU	Tony gua chakap,
Baba Malay words	Tony gua chakap
English gloss	1.SG speak
English free translation	My Tony said,

200.

JQ;

Baba Malay IU	Tony gua chakap,
Baba Malay words	Tony gua chakap
English gloss	1.SG speak
English free translation	My Tony said,

201.

JQ;

Baba Malay IU

Oh...lu Catholic ada chutsi eh=?

Baba Malay words

Oh lu Catholic ada chutsi eh

English gloss

Filler 2.SG have born filler

English free translation

Oh you Catholic have the concept of rebirth? [humorous intent]

202.

JQ;

Baba Malay IU

@@@@@,

<T=372.70>

Pantuns by Albert Ku

The following are *pantuns* composed by Albert Ku from Malacca. The *pantun* is a traditional Malay verse with an *abab* rhyme scheme. The main message of the *pantun* is embedded in the third and fourth lines. Peranakan *pantuns* often relate to Peranakan subject matters, although it is interesting that the Malay language is followed closely. Pantuns can be composed on the spot, exchanged between performers (see verses in C), and sung to a tune called the *dondang sayang* ‘melody (of) love’. The art of singing *pantun* to the *dondang sayang* is endangered, but there are still cultural associations such as the Gunong Sayang Association in Singapore that preserve this tradition. The following pantuns were written, and have no recordings associated with them.

A.	<p>Apa ada di gunong api</p> <p>Nampak api merah menjulang</p> <p>Sudah lama gua tabor budi</p> <p>Emas juga di pandang orang</p>	<p>1. Apa ada di gunong api what EXIST PREP mountain fire ‘What exists on top of fire mountain?’</p> <p>2. Nampak api merah menjulang see fire red tower ‘See the towering red fire’</p> <p>3. Sudah lama gua tabor budi already long.time 1.SG sow character ‘I have given my respect for a long time.’</p> <p>4. Emas juga di-pandang orang gold also PASS-view people ‘Gold is still looked at by people.’</p>
B.	<p>Dari Melaka ka Pulau Daik</p> <p>Mau beli sakati kerang</p> <p>Kalau baka pokok itu baik</p> <p>Ranting jatuh di pungot orang</p>	<p>5. Dari Melaka ka Pulau Daik from Malacca to Island ‘From Malacca to Daik Island’</p> <p>6. Mau beli sa-kati kerang want buy one-catty clam ‘Wanted to buy one catty of clams’</p> <p>7. Kalau baka pokok itu baik if lineage tree that good ‘If the lineage of that tree is good’</p>

		8. Ranting jatuh di pungot orang branch fall PASS pick.up people 'A fallen branch will be picked up by people.'
C.	Diberi hendak tak hendak Walau pun si ayer susu Tahun ini kawinkan anak Tahun depan timang chuchu Walau pun si ayer susu Susu dibeli rumah chek Ah Tan Tahun depan dapat chuchu Gua harap chuchu jantan Susu dibeli rumah chek Ah Tan Beli labu buat kua	9. Di-beri hendak tak hendak PASS-give want NEG want 'Given something whether you want it or not' 10. Walau pun si ayer susu although also PERSON water milk 'Even though it is a person with milk' 11. Tahun ini kawin-kan anak year this marry-TR child 'This year (you) marry off your child' 12. Tahun depan timbang chuchu year front balance grandchild 'Next year (you) balance your grandchild (on your lap).'
		13. Walau pun si ayer susu although also Person water milk 'Even though it is a person with milk' 14. Susu di-beli rumah chek Ah Tan Milk PASS-buy house uncle 'Milk bought from the house of Uncle Ah Tan' 15. Tahun depan dapat chuchu year front get grandchild 'Next year (you) get a grandchild.' 16. Gua harap chuchu jantan 1.SG hope grandchild male 'I hope for a grandson.'
		17. Susu di-beli rumah chek Ah Tan milk PASS-buy house uncle 'Milk bought from the house of Uncle Ah Tan'

	<p>Kalau dapat chuchu jantan Jangan lupa chiah gua</p>	<p>18. Beli labu buat kua buy pumpkin make gravy 'Buy pumpkin to make gravy'</p> <p>19. Kalau dapat chuchu jantan if get grandchild male 'If you get a grandson'</p> <p>20. Jangan lupa chiah gua do.not forget invite 1.SG 'Do not forget to invite me.'</p>
D.	<p>Apa ada di Gunong Daik Nampak api merah menjulang Jika ada budi yang baik Sampai mati dikenang orang</p>	<p>21. Apa ada di Gunong Daik what EXIST PREP mountain 'What exists on Daik Mountain?'</p> <p>22. Nampak api merah menjulang see fire red tower 'See the towering red fire'</p> <p>23. Jika ada budi yang baik if have character REL good 'If (you) have a good character'</p> <p>24. Sampai mati di-kenang orang until die PASS-remisces people 'Until (you) die you will be remembered by people.'</p>
E.	<p>Bibik Tengah sudah pulang Kena hujan baju basah Apa guna banyak wang Kalau tidak berbudi bahasa</p>	<p>25. Bibik Tengah sudah pulang Bibik middle already return 'Middle Bibik has already returned'</p> <p>26. Kena hujan baju basah PASS rain clothes wet '(She) was rained on (so her) clothes are wet'</p> <p>27. Apa guna banyak wang what use many money 'What is the use of a lot of money?'</p>

		28. Kalau tidak ber-budi bahasa if NEG POSS-character respect 'If (one) does not have a respectable character'
F.	Kain di ikat dengan benang Letak mari di kayu Jati Sembilan laut-an dah gua berenang Belum dapat kehendak hati	29. Kain di-ikat dengan benang cloth PASS-tie with thread 'Cloth tied with thread' 30. Letak mari di kayu jati place come PREP wood teak 'Come to the place with the teak wood' 31. Sembilan laut-an dah gua berenang Nine sea-NMZ already 1.SG swim 'I have swam nine oceans' 32. Belum dapat ke-hendak hati not.yet get NMZ-want heart 'I have not gotten what the heart wants'
G.	Pergi pasar beli manga Mangga di beli dari chek Dara Kalau lu dah berumah tangga Jangan pulak lupa saudara mara	33. Pergi pasar beli manga go market buy mango 'Went to the market to buy mangoes' 34. Mangga di-beli dari Chek Dara mango PASS-buy from uncle 'Mangoes bought from Uncle Dara' 35. Kalau lu dah ber-rumah tangga if 2.SG already POSS-house ladder' 'If you get married' 36. Jangan pulak lupa saudara mara do.not instead forget relations relatives 'Do not forget (your) relatives instead.'
H.	Pergi kedai beli arang Beli juga sa ikat buah Kalau ada di negeri orang	37. Pergi kedai beli arang go shop buy charcoal 'Went to a shop to buy charcoal'

	<p>Jangan pulak lupa kan gua</p>	<p>38. Beli juga sa-ikat buah buy also one-tie fruit 'Also bought a bundle of fruit.'</p> <p>39. Kalau ada di negeri orang if EXIST PREP country people 'If (you) are in (other) people's country'</p> <p>40. Jangan pulak lupa-kan gua do.not instead forget-TR 1.SG 'Do not forget me.'</p>
I.	<p>Sakit jari kena sembilu</p> <p>Sembilu kena di Tiang Dua</p> <p>Emak Bapak besar kan lu</p> <p>Jangan berdusta bila besar</p>	<p>41. Sakit jari kena sembilu pain finger PASS prick 'Pain in the finger from a prick'</p> <p>42. Sembilu kena di Tiang Dua prick PASS PREP pillar two 'Pricked at Pillar Two'</p> <p>43. Emak Bapak besar-kan lu mother father big-TR 2.SG '(Your) parents raised you.'</p> <p>44. Jangan berdusta bila besar do.not betray when big 'Do not betray them when you are old'</p>
J.	<p>Dari Dali ka Tiong Baru</p> <p>Mau beli ikan keli</p> <p>Sekarang ini dunia baru</p> <p>Cinta di hati mulai sekali</p>	<p>45. Dari Dali ka Tiong Baru From to Center New 'From Dali to New Center'</p> <p>46. Mau beli ikan keli want buy fish catfish 'Wanted to buy catfish'</p> <p>47. Sekarang ini dunia baru now this world new 'Now this is a new world'</p> <p>48. Cinta di hati mulai sekali love PREP heart from very 'Love from the heart is most important.'</p>

Appendix C: Vocabulary

Singapore Baba Malay-English

A

- a** [a] var. of **ada**
- a'us** [a.ʔus] var. of **ha'us**
- abi** [a.bi] var. of **habis**
- abis** [a.bis] var. of **habis**
- abu** [a.bu] (Malay) *n* ash
- ada** [a.da] (var. **a**) (Malay) *v* **1**) have **2**) PROG **3**) EXIST **4**) PFV **5**) COP **6**) HAB
- adék** [a.deʔ] (Malay) *n* **1**) sibling **2**) younger sibling der.
adék beradék
- adék beradék** [a.deʔ bæ.ra.deʔ] (Malay) (der. of **adék**)
n **1**) siblings **2**) younger siblings **pangkat adek beradek**
- agak** [a.gaʔ] (Malay) *v* estimate **agak agak**
- agak agak** [a.gaʔ a.gaʔ] (Malay) (**agak**) *adv* roughly
- aiya** [aj.ja] (Hokkien) EXCLAM annoyance
- aiyo** [aj.jo] (Hokkien) EXCLAM irritation
- ajair** [a.dʒɛ] refined var. of **ajar**
- ajar** [a.dʒar] (var. **ajair**) (Malay) *v* teach (coarse)
- akair** [a.ke] refined var. of **akar**
- akar** [a.kar] (var. **akair**) (Malay) *n* root (coarse)
- ala** [a.la] (English) EXCLAM regret
- alamak** [a.la.maʔ] (English), (Malay) EXCLAM dismay
- aloji** [a.lo.dʒi] (Portuguese) *n* **1**) small clock **2**) wrist watch
- amboi** [am.boj] source uncertain EXCLAM surprise
- ambun** [am.bun] (Malay) *n* tapioca
- amcham** [am.tʃam] var. of **apa macham**
- amék** [a.meʔ] (Malay) *v* take
- ampun** [am.pun] (Malay) **1**) *v* forgive **2**) *n* forgiveness
- an** [-an] (Malay) Nom
- anak** [a.naʔ] (Malay) *n* child
- ancheng** [an.tʃəŋ] (Hokkien) *n* blessing of the marital bed
- angin** [a.ŋin] (Malay) *n* wind
- angkat** [aŋ.kat] (Malay) *v* **1**) lift **2**) hold **3**) pick up **4**) carry
- angpau** [aŋ.paw] (Hokkien) *n* red packet of monetary gift
- anjing** [an.dʒing] (Malay) *n* dog
- ano** [a.no] (English) *adj* anonymous
- anting-anting** [an.tiŋ-an.tiŋ] (Malay) *n* earrings
- apa** [a.pa] (Malay) *interrog* what **amcham** , comp. **apa cham** , comp. **apa macham** , comp. **apa pasal** , contr. **apa sair**
- apa cham** [a.pa.tʃam] var. of **apa macham**
- apa macham** [a.pa ma.tʃam] (var. **apa cham**; var. **amcham**) (Malay) (comp. of **apa**, **macham**) *interrog* how
- apa pasair** [a.pa pa.sɛ] refined var. of **apa pasal**
- apa pasal** [a.pa pa.sal] (var. **apa pasair**) (Malay) (comp. of **apa**, **pasal**) *interrog* why (coarse)
- apa sair** [a.pa sɛ] (Malay) (contr. of **apa**, **sair 1**) *interrog* why (refined)
- api** [a.pi] (Malay) *n* fire
- arap** [a.rap] var. of **harap**
- arat** [a.rat] (Malay) *n* alcohol
- arga** [ar.ga] var. of **harga**
- ari** [a.ri] var. of **hari** comp. **mata ari** sun
- arimo** [a.ri.mo] var. of **harimo**
- asam** [a.sam] (Malay) **1**) *n* tamarind **2**) *adj* sour
- asap** [a.sap] (Malay) *n* smoke
- asién asin** [a.sjen asin] (Malay) (der. of **asin**) *adj*

somewhat salty
asin [a.sin] (Malay) *adj* salty comp. **mulot asin** , der.
asién asin atas [a.tas] (Malay) *prep* top
ati [a.ti] var. of **hati** comp. **ati baik** kind , comp. **ati busok** cruel , comp. **jantong ati** beloved , id. **ati it gor it chap** nervous
ati baik [a.ti baj?] (Malay) (comp. of **ati**, **baik**) *adj* kind
ati busok [a.ti bu.so?] (Malay) (comp. of **ati**, **busok**) *adj* cruel

ati it gor it chap [a.ti it gɔ it tʃap] (Malay), (Hokkien) (id. of **ati**, **it**, **gor**, **chap**) *adj* nervous
auban [aw.ban] (Hokkien) *adj* selfish
awan [a.wan] (Malay) *n* cloud
ayam [a.jam] (Malay) *n* chicken comp. **mulot pantat ayam**
ayé [a.je] (Malay) *n* water **mata ayé** , comp. **buang ayé** , comp. **buang ayé besair** , comp. **buang ayé kechik**
ayi [a.ji] source uncertain EXCLAM surprise

B

ba [ba] (Hokkien) (contr. of **Baba**) *n* son
ba'u [ba.ʔu] (Malay) *n* smell
baba [ba.ba] var. of **bapak**
Baba [ba.ba] (Hokkien) *n* **1** Peranakan male **2**) son contr.
ba
babi [ba.bi] (Malay) *n* pig
bacha [ba.tʃa] (Malay) *v* read
badan [ba.dan] (Malay) *n* body
bagi [ba.gi] (Malay) *adv* similar to
bagus [ba.gus] (Malay) *interj* good
bahasa [ba.ha.sa] (Malay) *n* language
bahu [ba.hu] (var. **bau**) (Malay) *n* shoulder
baik [baj?] (Malay) *adj* good comp. **ati baik** , comp. **naik baik**
baju [ba.dʒu] (Malay) *n* clothes
bakair [ba.kɛ] refined var. of **bakar**
bakar [ba.kar] (var. **bakair**) (Malay) *v* burn (coarse)
bakol [ba.kol] (Malay) *n* basket
baku [ba.ku] (Malay) *adj* standard
balay-balay [ba.le-ba.le] Indonesian *n* resting platform
baldi [bal.di] (Malay) *n* pail
balék [ba.le?] (Malay) **1**) *v* return **2**) *v* turn over **3**) *adv* back **4**) *adv* again
bangkit [baŋ.kit] (Malay) *n* coconut cookie

bangkuang [baŋ.kwaŋ] (Hokkien) *n* turnip
bapak [ba.pa?] (var. **baba**) (Malay) *n* father **mak bapak**
barang [ba.raŋ] (Malay) *n* thing
baring [ba.riŋ] (Malay) *v* lie down
baru [ba.ru] (Malay) **1**) *adj* new **2**) *adv* just
basi [ba.si] (Malay) *adj* stale
batang [ba.taŋ] (Malay) *clf* long and thin
batu [ba.tu] (Malay) **1**) *n* rock **2**) *n* stone **3**) *n* cave
bau [baw] var. of **bahu**
bawah [ba.wah] (Malay) *prep* under
bawak [ba.wa?] (Malay) *v* bring
bawang [ba.waŋ] (Malay) *n* onion
bayair [ba.jɛ] refined var. of **bayar**
bayang [ba.jaŋ] (Malay) *n* shadow
bayar [ba.jaŋ] (var. **bayair**) (Malay) *v* pay (coarse)
bedék [bɛ.de?] (Malay) *v* pummel
bédék [be.de?] (Malay) *v* tell a lie
bég [be?] (English) *n* bag
begi [bɛ.gi] (Malay) (contr. of **begitu**) *adv* like that
begitu [bɛ.gi.tu] (var. **gitu**) (Malay) *adv* like that contr.
begi
belachu [bɛ.la.tʃu] (Malay) *n* unbleached cotton outfit used for mourning
belah [bɛ.lah] (Malay) *prep* side

belair [bə.lɛ] refined var. of **belas**
belajair [bə.la.dʒɛ] var. of **belajar**
belajar [bə.la.dʒar] (var. **belajair**) (Malay) *v* learn
 (coarse)
belakang [bə.la.karŋ] (Malay) *prep* behind
belanda [bə.lan.da] (Malay) **1** *n* (Dutch) **2** *adj* (Dutch)
 comp. **kuching belanda**
belangkat [bə.larŋ.kat] (Malay) *v* crawl
belanja [bə.lan.dʒa] (Malay) *v* **1** spend **2** treat
belas [bə.las] (refined var. **belair**) (Malay) *cardnum* ten
 (coarse)
beli [bə.li] (Malay) *v* buy
beliak [bə.ljaʔ] (Malay) *v* glare **mata beliak**
belién [bə.ljen] (Malay) *n* diamond
bélok [be.loʔ] (Malay) *v* turn
belum [bə.lom] (Malay) *adv* **1** not yet **2** before
benair [bə.nɛ] refined var. of **benal**
benal [bə.nal] (var. **benair**) (Malay) *adj* rational (coarse)
bengis [bə.ŋis] (Malay) **1** *adj* fierce **2** *adj* serious
béngok [be.ŋoʔ] (Malay) *adj* twisted comp. **mulot**
béngok
bérak [be.raʔ] (Malay) *v* defecate
beramay [bə.ra.me] (Malay) *adj* crowded
berapa [bə.ra.pa] (Malay) **1** *interrog* how many **2** *adv*
 some
berat [bə.rat] (Malay) *adj* heavy comp. **mulot berat**
berenang [bə.rə.narŋ] (Malay) *v* swim
bergerak [bə.r.gə.raʔ] (Malay) *v* move
bergetair [bə.r.gə.tɛ] refined var. of **bergetar**
bergetar [bə.r.gə.tar] (var. **bergetair**) (Malay) *v* tremble
 (coarse)
berkelay [bə.r.kə.le] (Malay) *v* quarrel
berléiléi [bə.r.lej.lej] (Malay) *v* drip
bersi [bə.r.si] (Malay) *v* clean
bertengkair [bə.r.təŋ.kɛ] refined var. of **bertengkar**
bertengkar [bə.r.təŋ.kar] (var. **bertengkair**) (Malay) *v*
 argue (coarse)
beruang [bə.r.warŋ] (Malay) *n* bear

beruba [bə.ru.ba] (Malay) *v* repent
berus [bə.rus] (English) *v* brush
besair [bə.sɛ] refined var. of **besar** comp. **buang ayé**
besair defecate
besar [bə.sar] (var. **besair**) (Malay) *adj* big
besi [bə.si] (Malay) *n* iron
bésok [be.soʔ] (Malay) *adv* tomorrow
betol [bə.tol] (Malay) **1** *adj* correct **2** *adj* true **3** *adv*
 really
biasa [bja.sa] (Malay) *adj* used to it
bibik [bi.biʔ] (Malay) *n* older Peranakan woman
bijik [bi.dʒiʔ] (Malay) **1** *clf* small and round **2** *n* seed
 comp. **bijik mata**
bijik mata [bi.dʒiʔ ma.ta] (Malay) (comp. of **bijik**,
mata) *n* favourite child
bikin [bi.kin] Indonesian *v* **1** make **2** do
bila [bi.la] (Malay) *adv* when
bilang [bi.larŋ] (Malay) *v* tell
bilék [bi.leʔ] (Malay) *n* room
bilis [bi.lis] (Malay) *n* anchovy
bini [bi.ni] (Malay) *n* wife
binpo [bin.po] (Hokkien) *n* handkerchief
bintang [bin.taŋ] (Malay) *n* star
bising [bi.siŋ] (Malay) *adj* noisy
bocho [bo.tʃo] (Malay) *adj* leaky comp. **mulot bocho**
bodoh [bo.doh] (Malay) *adj* stupid comp. **buat bodoh**
bohong [bo.hoŋ] (Malay) *n* lie
boik [bojk] (Hokkien) *n* socks
bok-bok [boʔ-boʔ] (Hokkien) (contr. of **embok-embok**)
n traditional Peranakan elder
boléh [bo.leh] (Malay) *v* can **bongsu** [boŋ.su] (Malay)
adj youngest (only for familial relations)
botak [bo.taʔ] (Malay) *adj* bald
botol [bo.tol] (English) *n* bottle
bresi [bre.si] var. of **bersi**
buah [bwah] (Malay) *n* fruit
buang [bwarŋ] (Malay) *v* throw comp. **buang ayé**, comp.
buang ayé besair, comp. **buang ayé kechik**, comp.
buang buang, comp. **buang mata**, comp. **buang**

mulot , comp. **buang sebelah** , comp. **buang segan** ,
comp. **buang tebiét**

buang ayé [bwaŋ a.je] (Malay) (comp. of **buang, ayé**) *v*
urinate

buang ayé besair [bwaŋ a.je bæ.sɛ] (Malay) (comp. of
buang, ayé, besair) *v* defecate

buang ayé kechik [bwaŋ a.je kə.chiʔ] (Malay) (comp.
of **buang, ayé, kechik**) *v* urinate

buang buang [bwaŋ bwaŋ] (Malay) (comp. of **buang**) *v*
exorcise

buang mata [bwaŋ ma.ta] (Malay) (comp. of **buang,**
mata) *v* keep an eye on someone or something

buang mulot [bwaŋ mu.lot] (Malay) (comp. of **buang,**
mulot) *v* convey by speaking (as opposed to writing)

buang sebelah [bwaŋ sə.bə.lah] (Malay) (comp. of
buang) *adv* aside from

buang segan [bwaŋ sə.gan] (Malay) (comp. of **buang,**
segan) *v* stretch upon waking up

buang tebiét [bwaŋ tə.bjet] (Malay) (comp. of **buang,**
terbiat) *v* throw tantrum

buat [bwat] (Malay) *v* **1** do **2** make comp. **buat bodoh** ,
comp. **buat mahal** , comp. **buat malu** , comp. **buat**
mungka , comp. **buat suay** , comp. **buat tak dengair**

buat bodoh [bwat bo.doh] (Malay) (comp. of **buat,**
bodoh) *v* feign ignorance

buat mahal [bwat ma.hal] (Malay) (comp. of **buat,**
mahal) *v* play hard to get

buat malu [bwat ma.lu] (Malay) (comp. of **buat, malu**)
v cause embarrassment

buat mungka [bwat muŋ.ka] (Malay) (comp. of **buat,**
mungka) *v* be sour-faced

buat suay [bwat swe] (Malay), (Hokkien) (comp. of
buat, suay) *v* cause misfortune

buat tak dengair [bwat taʔ dəŋ.ɛ] (Malay) (comp. of
buat, tak, dengair) *v* pretend to not hear

bubor [bu.bɔ] (Hokkien) *n* porridge

budak [bu.daʔ] (Malay) *n* child

budi [bu.di] (Malay) *n* character

bukak [bu.kaʔ] (Malay) *v* open

bukak kun [bu.kaʔ kun] (Malay), (Hokkien) *n* bride's
assistant for kneeling and general help

bukan [bu.kan] (var. **kan**) (Malay) **1** *interj* no **2**) Neg
(noun)

bukit [bu.kit] (Malay) *n* hill

buku [bu.ku] (Malay) *n* book

bulan [bu.lan] (Malay) *n* **1** moon **2**) month

bulat [bu.lat] (Malay) *adj* round

bulu [bu.lu] (Malay) *n* **1** feather **2**) fine hair

bunbong [bun.boŋ] (Malay) *n* roof

bunchit [bun.tʃit] (Malay) *n* distended .stomach

bunga [bu.ŋa] (Malay) *n* flower

bungkus [buŋ.kus] (Malay) *n* bundle

bunuh [bu.noh] (Malay) *v* kill

buntot [bun.tot] (Malay) *n* **1** tail **2**) backside

bunyi [bu.ɲi] (Malay) *n* sound

burok [bu.roʔ] (Malay) *adj* ugly

burong [bu.roŋ] (Malay) *n* bird

busok [bu.soʔ] (Malay) *adj* smelly comp. **ati busok** ,
comp. **mulot busok**

buta [bu.ta] (Malay) *adj* blind

C

chabot [tʃa.bot] (Malay) *v* pull out

chaiki [tʃaj.ki] (Hokkien) *n* banner

chaiteng [tʃaj.təŋ] (Hokkien) *n* nunnery that serves
vegetarian food

chaiyen [tʃaj.jen] (Hokkien) *n* jelly

chakap [tʃaj.kap] (Malay) *v* speak

chakiak [tʃa.kjaʔ] (Hokkien) *n* clogs

cham 1 [tʃam] (Hokkien) *v* observe

cham 2 [tʃam] (Malay) (contr. of **macham**) *adv* like
champor [tʃam.por] (Malay) *v* mix
chanab [tʃa.nab] (Hokkien) *n* decorative altar stand
changkay [tʃaŋ.ke] (Malay) *n* cup
chanték [tʃan.teʔ] (Malay) *adj* beautiful
chap [tʃap] (Hokkien) *cardnum* ten id. **ati it gor it chap**
charék [tʃa.reʔ] (Malay) *v* find
chau [tʃaw] (Hokkien) *n* elder sister's husband
chauhah [tʃaw.kah] (Hokkien) *n* bad sport
chauhuan [tʃaw.kwan] (Hokkien) *n* bad type of person
chay [tʃe] (Hokkien) *n* beginning of lunar month
chay-em [tʃe.əm] (Hokkien) *n* mother-in-law (indirect address)
chék [tʃeʔ] (Hokkien) *n* father's younger brother
chék-gong [tʃeʔ-goŋ] (Hokkien) *n* grandfather's younger brother
chekék [tʃə.keʔ] (Malay) *v* strangle
chelop [tʃə.lɔp] (Malay) *v* dip in dye
cepat [tʃə.pat] (Malay) *adj* quick
chepuat [tʃə.pwat] (Hokkien) *n* family
cherdék [tʃər.deʔ] (Malay) *adj* streetwise
cherita [tʃə.ri.ta] (Malay) *n* story
cherkek [tʃər.keʔ] (Malay) *v* eat (vulgar)
cherki [tʃər.ki] (Hokkien) *n* Peranakan card game
chia [tʃja] (Hokkien) *n* car
chiam [tʃjam] (Hokkien) *n* divination stick
chichi [tʃi.tʃi] (Malay) *n* great-grandchild

chilaka [tʃi.la.ka] Javanese EXCLAM cursed one
chim [tʃim] (Hokkien) *n* father's younger brother's wife
chin [tʃin] (Hokkien) *adj* close to one another
China [tʃi.na] (Malay) *adj* Chinese
chinchai [tʃin.tʃaj] (Hokkien) *adj* not fussy
chinchang [tʃin.tʃaŋ] (Malay) *v* mince
chingkay [tʃiŋ.ke] (Hokkien) *n* **1** parent-in-law **2** father-in-law (indirect address)
chit [tʃit] (Hokkien) *cardnum* seven
chiwan [tʃi.wan] source unknown *n* toilet
cho [tʃo] (Hokkien) *n* **1** great-grandparent **2** great-grandmother **mak-cho**, comp. **gong-cho**
cho-cho [tʃo-tʃo] (Hokkien) *n* great-grandmother
cho-gong [tʃo-goŋ] (Hokkien) *n* great-grandfather
cho-po [tʃo-po] (Hokkien) *n* great-grandaunt
chobak [tʃo.baʔ] (Malay) *v* taste
chobék [tʃo.beʔ] (Malay) *adj* long-jawed
chonténg [tʃon.teŋ] (Malay) *v* draw
chu-lang [tʃu.laŋ] (Hokkien) *n* bridegroom's family's host
chuchi [tʃu.tʃi] (Malay) *v* wash
chuchu [tʃu.tʃu] (Malay) *n* grandchild
chuikuéh [tʃwi.kweh] (Hokkien) *n* rice cake
chukop [tʃu.kop] (Malay) *adv* enough
chupak [tʃu.paʔ] (Malay) *clf* cylindrical measure of quarter gallon of rice
chutsi [tʃut.si] (Hokkien) *n* birth

D

dadair [da.dɛ] refined var. of **dadar**
dadar [da.dar] (var. **dadair**) (Malay) *n* omelette (coarse)
dagin [da.gin] (Malay) *n* meat
dahi [da.hi] (var. **dai**) (Malay) *n* forehead
dai [daj] var. of **dahi**

dalam [da.lam] (Malay) *prep* inside
daoun [da.own] (Malay) *n* leaves
dapat [da.pat] (Malay) *v* **1** get **2** receive
dapor [da.pɔ] (Malay) *n* kitchen
darah [da.rah] (Malay) *n* blood comp. **naik darah**
dari [da.ri] (Malay) *prep* from

datang [da.taŋ] (Malay) *v* come
degil [də.gil] (Malay) *adj* stubborn
dekat [də.kat] (var. **kat**) (Malay) **1** *prep* at **2** *prep* near **3** *prep* Prep
dengair [də.ŋɛ] refined var. of **dengar** comp. **buat tak dengair** pretend to not hear
dengar [də.ŋaŋ] (var. **dengair**) (Malay) *v* listen (coarse)
depan [də.pan] (Malay) *prep* front
di [di] (Malay) *prep* Prep comp. **diluar**
dia [dja] (Malay) **1** *pro* 3.SG; **2** *pro* 3.PL
diam [djam] (Hokkien) *adj* quiet
diam diam [djam djam] (Hokkien) *adv* quietly
dia-orang [dja-o.raŋ] (Malay) 3. PL *pro*

diluar [di.lwaŋ] (var. **duluar**) (Malay) (comp. of **di**, **keluar**) *prep* outside
diri [di.ri] (Malay) *v* stand
dit [dit] var. of **duit**
dondang [don.daŋ] *n* melody
dua [dwa] (Malay) *cardnum* two
dudok [du.doʔ] (Malay) *v* **1** sit **2** stay
duit [dwit] (var. **dit**) (Dutch) *n* money
dulu [du.lu] (Malay) *adv* **1** before **2** long ago **3** first
duluar [du.lwaŋ] var. of **diluar**
dunya [du.ja] (Arabic) *n* world comp. **mulot dunya**

E

éh [eh] source unknown EXCLAM jibe
ékor [e.kor] (Malay) *clf* animal
éla [e.la] var. of **helay**
em [əm] (Malay) *n* **1** father's elder brother's wife **2** old aunty **sangkék-em**
embair-embair [əmbɛ-əmbɛ] (Malay) *adj* half-cooked (refined)

embar-embar [əmbaŋ-əmbaŋ] (Malay) *adj* half-cooked (coarse)
embok-embok [əm.boʔ-əm.boʔ] (Hokkien) *n* traditional Peranakan elder contr. **bok-bok**
empat [əm.pat] (Malay) *cardnum* four
enam [e.nam] (Malay) *cardnum* six
eng- [əŋ] (Hokkien) honorific prefix for familial relations

G

gaga [ga.ga] (Hokkien) *adj* daring
gambair [gam.bɛ] refined var. of **gambar gambar** [gam.baŋ] (var. **gambair**) (Malay) *n* picture (coarse)
ganjil [gan.jil] (Malay) *adj* odd-numbered
gantang [gan.taŋ] (Malay) *clf* cylindrical measure of one gallon of rice
gantong [gan.toŋ] (Malay) *v* hang
gaou [ga.ow] (Malay) *v* mix
garam [ga.ram] (Malay) *n* salt

garang [ga.raŋ] (Malay) *adj* fierce
gasak [ga.saʔ] (Malay) *v* guess
gatair [ga.te] refined var. of **gatal** comp. **mulot gatair** uncontrollable mouth
gatal [ga.tal] (var. **gatair**) (Malay) *adj* itchy (coarse)
gau [gaw] (Hokkien) *cardnum* nine
gaya [ga.ja] (Malay) *v* splurge
gedebak-gedebuk [gə.də.baʔ-gə.də.buʔ] (Malay) onomatopoeia thudding of the heart
gedebang-gedebong [gə.də.baŋ-gə.də.boŋ] (Malay)

onomatopoeia loud noises

gelang [gə.laŋ] (Malay) *n* bracelet
gelap [gə.lap] (Malay) *adj* dark comp. **mata gelap**
geléték [gə.le.teʔ] (Malay) *v* tickle comp. **naik geléték**
gemoh [gə.moh] (Malay) *v* sun dry
gemok [gə.moʔ] (Malay) *adj* fat
geram [gə.ram] (Malay) *adj* furious comp. **naik geram**
gerang [gə.raŋ] (Malay) *n* enthusiasm
gerja [gər.dʒa] (var. **greja**) (Portuguese) *n* church
gerong [gə.roŋ] (Malay) *v* scratch
gertak [gər.taʔ] (var. **gretak**) (Malay) *v* threat
gigi [gi.gi] (Malay) *n* tooth
gigit [gi.git] (Malay) *v* bite
gila [gi.la] (Malay) *adj* mad comp. **naik gila**
gitu [gi.tu] var. of **begitu**
gonchang [gon.tʃaŋ] (Malay) *v* shake
gong [goŋ] (Hokkien) **1** *n* grandparent **2** grandfather
comp. **gong-cho**

gong-cho [goŋ-tʃo] (Hokkien) (comp. of **gong, cho**) *n*
ancestors
gonggong [goŋ-goŋ] (Malay) *v* bark
gor [gɔ] (Hokkien) *cardnum* five id. **ati it gor it chap**
gorblok [gɔ.bloʔ] Indonesian *adj* stupid
goréng [go.reŋ] (Malay) *v* fry
gostan [go.stan] (English) *v* reverse
goyang [go.jaŋ] (Malay) **1** *vi* rock **2** *vt* shake
greja [grə.dʒa] var. of **gerja**
gretak [grə.taʔ] var. of **gertak**
gua [gwa (coarse)] (Hokkien) *pro* 1.SG
guék [gweʔ] (Hokkien) *n* lunar month
gugol [gu.gol] (Malay) *n* dried fruit
gui [gwi] (Hokkien) *v* kneel
gula [gu.la] (Malay) *n* sugar
guna [gu.na] (Malay) **1** *v* use **2** *n* use
gunong [gu.noŋ] (Malay) *n* mountain

H

ha'us [ha.ʔus] (var. **a'us**) (Malay) *adj* **1** thirsty **2** worn out
habis [ha.bis] (var. **abis**; var. **abi**) (Malay) **1** *v* finish 1.1)
adv finish **2** *conn* finish
hak [haʔ] (Hokkien) *adj* suitable
halo [ha.lo] (Malay) *v* chase away
hantok [han.toʔ] (Malay) *v* bang
harap [ha.rap] (var. **arap**) (Malay) *v* hope
harga [har.ga] (var. **arga**) (Sanskrit) *n* price
hari [ha.ri] (var. **ari**) (Malay) *n* day **tengah hari**
harimo [ha.ri.mo] (var. **arimo**) (Malay) *n* tiger
hati [ha.ti] (var. **ati**) (Malay) *n* **1** heart **2** liver
haulam [haw.lam] (Hokkien) *n* male mourner
hauli [haw.li] (Hokkien) *n* female mourner
helay [hə.le] (var. **lay**; var. **éla**) (Malay) *clf* sheet
héran [he.ran] (Malay) *v* wonder

hia [hja] (Hokkien) *n* elder brother
hidong [hi.doŋ] (var. **idong**) (Malay) *n* nose
hijo [hi.dʒo] (var. **ijo**) (Malay) *adj* green
hilang [hi.laŋ] (Malay) *v* lose
hitam [hi.tam] (var. **itam**) (Malay) *adj* black
ho [ho] (Hokkien) *adj* good
homia [ho.mja] (Hokkien) *interj* good life (comp. of **ho, mia**)
hormat [hor.mat] (var. **ormat**) (Malay) *v* respect
horpau [hor.paw] (var. **orpau**) (Hokkien) *n* purse
hu [hu] (Hokkien) *n* amulet
huahi [hwa.hi] (Hokkien) *adj* happy
huantiok [hwan.tjoʔ] (Hokkien) *v* meet ill spiritual forces
huat [hwat] (Hokkien) *v* **1** expand **2** prosper
huésio [hwe.sjo] (Hokkien) *n* monk

hujan [hu.dʒan] (var. **ujan**) (Malay) *n* rain
hutan [hu.tan] (var. **utan**) (Malay) *n* jungle

huyi [hu.ji] (Hokkien) *n* fishball

I

idong [i.doŋ] var. of **hidong**
ijo [i.dʒo] var. of **hijo**
ikan [i.kan] (Malay) *n* fish comp. **mata ikan**
ikat [i.kat] (Malay) *v* tie
ikot [i.kot] (Malay) *v* follow
ilang [i.laŋ] (Malay) *v* lose
ingat [i.ŋat] (Malay) *v* **1** remember **2** think
ini [i.ni] (Malay) *dem* this contr. **ni**

intan [in.tan] (Malay) *n* diamond chip
iris [i.ris] (Malay) *v* slice
isi [i.si] (Malay) *n* flesh
it [it] (Hokkien) *cardnum* one id. **ati it gor it chap**
itam [i.tam] var. of **hitam**
iték [i.teʔ] (Malay) *n* duck
itu [i.tu] (Malay) *dem* that contr. **tu**

J

ja.ou [dʒa.ow] (Malay) *adj* far
jadi [dʒa.di] (Malay) *v* **1** happen **2** become
jaga [dʒa.ga] (Malay) *v* **1** guard **2** take care of someone or something
jahat [dʒa.hat] (Malay) *adj* evil comp. **mulot jahat**
jait [dʒajt] (Malay) *v* sew
jalan [dʒa.lan] (Malay) *v* walk **tengah jalan**
jam [dʒam] (Malay) *n* **1** time **2** hour
jambu [dʒam.bu] (Malay) *n* Syzygium fruit (pink)
jangan [dʒa.ŋan] (Malay) *adv* do not
jantan [dʒan.tan] (Malay) *n* male
jantong [dʒan.toŋ] (Malay) *n* heart comp. **jantong ati**
jantong ati [dʒan.toŋ a.ti] (Malay) (comp. of **jantong, ati**) *n* beloved
jarang [dʒa.raŋ] (Malay) *adv* seldom
jari [dʒa.ri] (Malay) *n* finger
jati [dʒa.ti] (Malay) **1** *n* teak **2** *adj* pure

jatoh [dʒa.toh] (Malay) *v* **1** fall **2** drop
jawab [dʒa.wab] (Malay) *v* answer
ji [dʒi] (Hokkien) *cardnum* two
jin [dʒin] (Hokkien) *v* recognize
jinak [dʒi.naʔ] (Malay) *adj* tame
jodoh [dʒo.doh] (Malay) *n* match of love
jogét [dʒo.get] (Malay) *v* stylised dance
juair [dʒwɛ] refined var. of **jual**
jual [dʒwal] (var. **juair**) (Malay) *v* sell (coarse)
judi [dʒu.di] (Malay) *v* gamble
jugak [dʒu.gaʔ] (Malay) *adv* also
juling [dʒu.liŋ] (Malay) *adj* squinty comp. **mata juling**
jumoh [dʒu.moh] (Malay) *v* dry in the sun
jumpa [dʒum.pa] (Malay) *v* meet

K

ka [ka] (Hokkien) *conn* or

kabair [ka.be] refined var. of **kabar**

kabar [ka.bar] (var. **kabair**) (Arabic) *n* news (coarse)

kacho [ka.tʃo] (Malay) *v* disturb

kachua [ka.tʃwa] (Hokkien) *n* cockroach

kain [kajn] (Malay) *n* cloth

kaki [ka.ki] (Malay) *n* **1** leg **2** friend

kala [ka.la] (Malay) *n* time **sinjakala**

kalah [ka.lah] (Malay) *v* lose (opposed to win)

kalang-kabot [ka.laŋ-ka.bot] (Malay) *v* fumble

kali [ka.li] (Malay) *n* time

kalo [ka.lo] var. of **kalu**

kalu [ka.lu] (var. **kalo**) (Malay) *conn* if

kambing [kam.biŋ] (Malay) *n* **1** goat **2** sheep

kampong [kam.poŋ] (Malay) *n* village

kamsiah [kam.sjah] (Hokkien) *interj* thank you

kamtiok [kam.tjoʔ] (Hokkien) *v* develop an illness

kamwan [kam.wan] (Hokkien) *adj* satisfied

kan [kan] var. of **bukan**

-kan [kan] (Malay) *Tr*

kanan [ka.nan] (Malay) *adv* right

kangkong [kaŋ.koŋ] (Malay) *n* water spinach

kapair [ka.pɛ] refined var. of **kapal**

kapal [ka.pal] (var. **kapair**) (Malay) *n* ship (coarse)

kasair [ka.sɛ] refined var. of **kasar**

kasar [ka.sar] (var. **kasair**) (Malay) *adj* coarse (coarse)

kasi [ka.si] Indonesian *v* **1** give **2** let **3** cause **4** PASS

kasot [ka.sot] (Malay) *n* shoe

kat [kat] var. of **dekat**

kata [ka.ta] (Malay) *v* say

katék [ka.teʔ] (Hokkien) *adj* short

kawan [ka.wan] (Malay) *n* friend

kawin [ka.win] (Hokkien) *v* marry

kaya 1 [ka.ja] (Malay) *adj* rich

kaya 2 [ka.ja] (Hokkien) *n* coconut jam

kayu [ka.ju] (Malay) *n* wood

kebaya [kə.ba.ja] (Malay) *n* traditional blouse

kebun [kə.bun] (Malay) *n* garden

kechik [kə.tʃiʔ] (Malay) *adj* small comp. **buang ayé**

kechik keday [kə.de] (Malay) *n* shop

kejair [kə.dʒɛ] refined var. of **kejar**

kejar [kə.dʒar] (var. **kejair**) (Malay) *v* chase (coarse)

kejut [kə.dʒut] (Malay) *adj* shock

kék [keʔ] (English) *n* cake

kéksim [keʔ.sim] (Hokkien) *adj* unhappy

kelaka [kə.la.ka] (Malay) *v* joke

kelék.kelé [kəleʔ.kəleʔ] (Malay) *v* blink

keluak [kə.lwaʔ] (Malay) *n* *Pangium edule* (black fruit with hard shell)

keluar [kə.lwar] (refined var. **kuluair**; var. **kuluar**)

(Malay) *v* go out (coarse) comp. **diluar**

kemair [kə.mɛ] refined var. of **kemas**

kemantin [kə.man.tin] (Malay) *n* bride

kemas [kə.mas] (var. **kemair**) (Malay) *v* tidy (coarse)

kena [kə.na] (Malay) **1** *v* Pass **2** *v* subjected to

kenair [kə.nɛ] (Malay) *v* (refined var. **kenal**)

kenal [kə.nal] (Malay) *v* know a person (coarse)

kenang [kə.naŋ] (Malay) *v* reminisce

kenonay [kə.no.ne] source unclear *adj* proper

kenyang [kə.ŋaŋ] (Malay) *adj* full from eating

kepala [kə.pa.la] (Malay) *n* head

kepék [kə.peʔ] (Malay) *v* pinch

kepéng [kə.pɛŋ] (Malay) *clf* piece
kepok [kə.pɔʔ] (Malay) *n* box
kerair [kərəɛ] refined var. of **keras**
keras [kərəs] (var. **kerair**) (Malay) *adj* hard (coarse)
kerekot [kə.re.kot] (Malay) *adj* bent and curled up
keréta [kə.re.ta] (Malay) *n* car
kering [kə.rɪŋ] (Malay) **1** *n* hay **2** *adj* dry comp. **mulot kering**
kerja [kə.r.dʒa] (var. **kreja**) (Malay) **1** *v* work **2** *n* work
kerosi [kə.ro.si] (Malay) *n* chair
kertair [kə.r.tɛ] refined var. of **kertas**
kertas [kə.r.tas] (var. **kertair**) (Malay) *n* paper (coarse)
ketawa [kə.ta.wa] (Malay) *v* laugh
ketok [kə.toʔ] (Malay) *v* knock
kéwat [ke.wat] source unknown *adj* fussy
kiam [kjam] (Hokkien) *adj* stingy
kiamchai [kjam.chaj] (Hokkien) *n* preserved vegetables
kiamsiap [kjam.sjap] (Hokkien) *adj* miserly
kiasai [kjasaj] (Hokkien) *n* **1** son-in-law **2** groom
kim [kim] (Hokkien) *n* mother's brother's wife
kipas [ki.pas] (Malay) *n* fan
kita [ki.ta] (Malay) *pro* I.PL
ko 1 [ko] (Hokkien) *n* father's sister **mak-ko**
ko 2 [ko] (Hokkien) *n* elder brother (non-familial)
kochék [ko.tʃeʔ] (Malay) *n* pocket
kodok [ko.doʔ] (Malay) *n* frog
koktok [kɔʔ.tɔʔ] (Hokkien) *v* ill-treat
kolomi [ko.lo.mi] (Hokkien) *n* type of noodles
konchi [kon.tʃi] (Malay) *n* key
kopék [ko.peʔ] (Hokkien) *v* peel
kopi [ko.pi] (Hokkien) *n* coffee
kopiah [ko.pjah] source unknown *n* hat
kopiau [ko.pjaw] (Hokkien) *n* paternal cousins
kotor [ko.tor] (Malay) *adj* dirty
koya [ko.ja] (Malay) *n* green bean cookies
koyak [ko.jaʔ] (Malay) *adj* torn

kreja [krə.dʒa] var. of **kerja**
ku 1 [ku] (Hokkien) *n* tortoise
ku 2 [ku] (Hokkien) *n* mother's brother comp. **ku-ku**
ku-ku [ku-ku] (Hokkien) (comp. of **ku 2**) *n* mother's eldest brother
kua 1 [kwa] (Malay) *n* gravy
kua 2 [kwa] (Hokkien) *n* father-in-law
kuat [kwat] (Hokkien) *adj* strong
kuatay [kwa.te] (Malay) *v* worry **kuching** [ku.tʃɪŋ] (Malay) *n* cat comp. **kuching belanda**, comp. **mata kuching**
kuching belanda [ku.tʃɪŋ bælanda] (Malay) (comp. of **kuching belanda**) *n* rabbit
kuda [ku.da] (Malay) *n* horse
kudut [ku.dut] (Malay) *adj* crumpled
kuéh [kweh] (Hokkien) *n* traditional cakes comp. **kuéh-yi**
kuéh-yi [kweh-ji] (Hokkien) (comp. of **kuéh, yi 1**) *n* glutinous rice balls
kuku [ku.ku] (Malay) *n* fingernail
kukus [ku.kus] (Malay) *v* steam
kulit [ku.lit] (Malay) *n* skin
kuluair [ku.lwɛ] refined var. of **keluar**
kuluar [ku.lwar] var. of **keluar**
kun 1 [kun] (Hokkien) *n* skirt
kun 2 [kun] var. of **pun**
kunang [ku.naŋ] source unclear *adj* under the influence of black magic
kuning [ku.niŋ] (Malay) *adj* yellow
kuntum [kun.tum] (Malay) *clf* bloom
kupas [ku.pas] (Malay) *v* peel
kuping [ku.piŋ] (Portuguese) *n* ear
kupu-kupu [ku.pu-ku.pu] (Malay) *n* butterfly
kura-kura [ku.ra-ku.ra] (Malay) *n* tortoise
kurang [ku.raŋ] (Malay) *adv* less
kuro [ku.ro] (Malay) *n* threadfin fish
kurus [ku.rus] (Malay) *adj* thin
kus semangat [kus sə.ma.ŋat] (Malay) EXCLAM cry to a

dead spirit

kuya [ku.ja] (Hokkien) *n* groom's boy assistant

L

labi-labi [la.bi-la.bi] (Malay) *n* turtle

lagik [la.giʔ] (Malay) **1** *adv* more **2** *adv* still **3** *adv* again

lagu [la.gu] (Malay) *n* song

lah [lah] (Hokkien) *prt* EMP

lain [lajn] (Malay) *adv* other

lak [laʔ] (Hokkien) *cardnum* six

laki [la.ki] (Malay) (contr. of **lelaki**) *n* husband

lalu [la.lu] (Malay) *v* pass

lama [la.ma] (Malay) *adj* long time

lambat [lam.bat] (Malay) *adj* late

lambong [lam.boŋ] (Malay) *v* toss

lampu [lam.pu] (Dutch) *n* lamp

lanchang [lan.tʃaŋ] (Malay) *adj* fluent

langgair [laŋ.ge] refined var. of **langgar**

langgar [laŋ.gar] (var. **langgair**) (Malay) *v* crash (coarse)

langit [la.ŋit] (Malay) *n* sky

langkék [laŋ.keʔ] (Hokkien) *n* guest

langsong [laŋ.soŋ] (Malay) *adv* straightaway

lantay [lan.te] (Malay) *n* floor

lapair [la.pe] refined var. of **lapar**

lapan [la.pan] (Malay) *cardnum* eight

lapar [la.par] (var. **lapair**) (Malay) *adj* hungry (coarse)

lapchai [lap.tʃaj] (Hokkien) *n* wedding gift exchange ceremony

lari [la.ri] (Malay) *v* run

lauk [lawʔ] (Malay) *n* cooked food

laut [lawt] (Malay) *n* **1** sea **2** pond

lawa [la.wa] (Malay) *adj* stylish **lawan** [la.wan] (Malay) *v* race

lay [le] var. of **helay**

lebat [lə.bat] (Malay) *adj* heavy

lebéh [lə.beh] (Malay) *adv* more

léihéi [lej.hej] (Malay) *n* neck

lekair [lə.ke] refined var. of **lekas**

lekas [lə.kas] (var. **lekair**) (Malay) *adj* quick (coarse)

lelaki [lə.la.ki] (Malay) *n* boy contr. **laki**

lemak [lə.maʔ] (Malay) *adj* cooked in coconut milk comp.
naik lemak

lembéh [ləm.beh] (Malay) *adj* soft

lembu [ləm.bu] (Malay) *n* cow comp. **mata lembu**

lembut [ləm.but] (Malay) *adj* supple

lémo [le.mo] (Malay) *n* lemon

lentang [lən.taŋ] (Malay) *v* fall backwards

lepas [lə.pas] (Malay) *prep* after

letak [lə.taʔ] (Malay) *v* park

letay [lə.te] (Malay) *adj* exhausted

liar [ljaɾ] (Malay) *n* wildness

lichin [li.tʃin] (Malay) *adj* smooth

lidah [li.dah] (Malay) *n* tongue

lihai [li.haj] (Hokkien) *adj* cunning

lima [li.ma] (Malay) *cardnum* five

lio [ljo] (Malay) *n* saliva

lipat [li.pat] (Malay) *v* fold

lobang [lo.baŋ] (Malay) *n* hole

lochéng [lo.tʃeŋ] (Hokkien) *n* bell

logong [lo.goŋ] (Malay) *v* carry baby

lompat [lom.pat] (Malay) *v* jump

longkang [loŋ.kaŋ] (Malay) *n* drain

lontong [lon.toŋ] (Malay) *n* rice cake served in spicy
gravy

lor [lor] (Hokkien) EMP

loténg [lo.teŋ] (Hokkien) *adv* upstairs

lu [lu] (Hokkien) *pro* 2.SG

luan [lwan] (Hokkien) *adj* messy
luka [luka] (Malay) **1** *adj* wounded **2** *n* wound
lun [lun] (Hokkien) *adj* intercalary

lun-guék [lun gweʔ] (Hokkien) *n* intercalary month
lupa [lu.pa] (Malay) *v* forget

M

mabok [ma.boʔ] (Malay) *adj* giddy
macham [ma.tʃam] (Malay) **1** *adv* like **2** *adv* seems **3**
adv like that **amcham**, comp. **apa cham**, comp. **apa macham**, contr. **cham 2**
magam [ma.gam] (Malay) *adj* overripe
mahal [ma.hal] (Malay) *adj* expensive comp. **buat mahal**
main [majn] (Malay) **1** *v* play **2** *v* perform
mair [mɛ] refined var. of **mas**
mak [maʔ] (Hokkien) *n* mother **mak bapak**, **mak-cho**, **mak-ko**, **mak-yi**
mak bapak [maʔ ba.paʔ] (Hokkien), (Malay) (**mak**, **bapak**) *n* parents
mak-cho [maʔ-tʃo] (Hokkien) (**mak**, **cho**) *n* great-grandmother
mak-ko [maʔ-ko] (Hokkien) (**mak**, **ko 1**) *n* father's eldest sister
mak-yi [maʔ-ji] (Hokkien) (**mak**, **yi 2**) *n* second mother
makan [ma.kan] (Malay) **1** *v* eat **2** *n* food
maki [ma.ki] (Malay) *v* scold
malair [ma.lɛ] refined var. of **malas**
malam [ma.lam] (Malay) *n* night **malas** [ma.las] (var. **malair**) (Malay) *adj* lazy (coarse)
malu [ma.lu] (Malay) *adj* embarrassed comp. **buat malu**
mamak [ma.maʔ] (Hokkien) *n* grandmother
mampus [mam.pus] (Malay) EXCLAM dead
mana [ma.na] (Malay) *interrog 1* where **2** which
mandi [man.di] (Malay) *v* bathe
manék [ma.neʔ] (Malay) *n* bead
mangkok [maŋ.koʔ] (Malay) *n* bowl
mangun [ma.ŋun] (Malay) *v* wake

manis [ma.nis] (Malay) *adj* sweet comp. **mulot manis**
manyak [ma.naʔ] (Malay) *adv* many
marah [ma.rah] (Malay) *adj* angry
mari [ma.ri] (Malay) *v* come let us
mas [mas] (var. **mair**) (Malay) *n* gold (coarse)
masak [ma.saʔ] (Malay) **1** *v* cook **2** *adj* ripe
masok [ma.soʔ] (Malay) **1** *v* enter **2** *v* put in
mata [ma.ta] (Malay) *n* eye **mata ayé**, **mata beliak**, comp. **bijik mata**, comp. **buang mata**, comp. **mata ari**, comp. **mata gelap**, comp. **mata ikan**, comp. **mata juling**, comp. **mata kuching**, comp. **mata lembu**, comp. **mata piso**, comp. **tanda mata**
mata ari [ma.ta a.ri] (Malay) (comp. of **mata**, **ari**) *n* sun
mata ayé [ma.ta a.je] (Malay) (**mata**, **ayé**) *n* sweetheart
mata beliak [ma.ta bæ.ljaʔ] (Malay) (**mata**, **beliak**) *n* protruding eyes
mata gelap [ma.ta gə.lap] (Malay) (comp. of **mata**, **gelap**) *n* detective
mata ikan [ma.ta i.kan] (Malay) (comp. of **mata**, **ikan**) *n* wart
mata juling [ma.ta dʒu.liŋ] (Malay) (comp. of **mata**, **juling**) *n* cockeyed
mata kuching [ma.ta ku.tʃiŋ] (Malay) (comp. of **mata**, **kuching**) *n* longan fruit
mata lembu [ma.ta ləm.bu] (Malay) (comp. of **mata**, **lembu**) *n* fried egg with yolk intact
mata piso [ma.ta pi.so] (Malay) (comp. of **mata**, **piso**) *n* blade of knife
mati [ma.ti] (Malay) **1** *v* die **2** *adj* dead
mayang [ma.jaŋ] (Malay) *n* palm blossom
méh [meh] (Hokkien) *n* night (when referring to lunar dates)

méja [me.dʒa] (Malay) *n* table
melekat [mə.lə.kat] (Malay) *adj* sticky
mémang [me.marŋ] (Malay) *adv* indeed
memisék [mə.mi.seʔ] (Malay) *v* whisper
menang [mə.naŋ] (Malay) *v* win
menangis [mə.na.ŋis] (Malay) *v* cry
mengantok [mə.ŋan.toʔ] (Malay) **1** *v* yawn **2** *adj* sleepy
mengintar [mə.ŋin.tar] (Malay) *v* shiver
menidi [mə.ni.di] (Malay) *adj* boiling
menjéla [mə.n.dʒe.la] (Portuguese) *n* window
mentah [mə.n.tah] (Malay) *adj* raw
mentéga [mə.n.te.ga] (Portuguese) *n* butter
menyalap [mə.ŋa.lap] (Malay) *v* howl
menyanyi [mə.ŋa.ŋi] (var. **nyanyi**) (Malay) *v* sing
mérah [me.rah] (Malay) *adj* red
mesti [mə.s.ti] (Malay) *v* must
mia [mja] (Hokkien) *n* life
1
mia [mja] (Malay) (contr. of **punya**) **1** *v* POSS **2** REL
2
mimpi [mim.pi] (Malay) *v* dream
mincharok [min.tʃa.roʔ] (Malay) *v* curse
minggu [miŋ.gu] (Malay) *n* week
minit [mi.nit] (English) *n* minute
mintak [min.taʔ] (Malay) *v* ask sincerely
minum [mi.num] (Malay) *v* drink
minyak [mi.ŋak] (Malay) *n* **1** grease **2** oil
miskin [mis.kin] (Malay) *adj* poor **mo** [mo] (Malay) *v* want
mudah [mu.dah] (Malay) *adj* young
muilang [mwi.laŋ] (Hokkien) *n* matchmaker
mula [mu.la] (Malay) *adj* original
mula mula [mu.la mu.la] (Malay) *adv* originally
mulot [mu.lot] (Malay) *n* mouth comp. **buang mulot** ,

comp. **mulot asin** , comp. **mulot béngok** , comp. **mulot berat** , comp. **mulot bocho** , comp. **mulot busok** , comp. **mulot dunya** , comp. **mulot gatair** , comp. **mulot jahat** , comp. **mulot kering** , comp. **mulot manis** , comp. **mulot pantat ayam** , comp. **mulot ringan**
mulot asin [mu.lot a.sin] (Malay) (comp. of **mulot, asin**) *n* ability to make accurate predictions
mulot béngok [mu.lot be.ŋoʔ] (Malay) (comp. of **mulot, béngok**) *n* twisted mouth
mulot berat [mu.lot bæ.rat] (Malay) (comp. of **mulot, berat**) *n* inability to express oneself
mulot bocho [mu.lot bo.tʃo] (Malay) (comp. of **mulot, bocho**) *n* inability to keep a secret
mulot busok [mu.lot bu.soʔ] (Malay) (comp. of **mulot, busok**) *n* bad breath
mulot dunya [mu.lot du.ŋia] (Malay), (Arabic) (comp. of **mulot, dunya**) *n* public opinions
mulot gatair [mu.lot ga.tɛ] (Malay) (comp. of **mulot, gatair**) *n* uncontrollable mouth
mulot jahat [mu.lot dʒa.hat] (Malay) (comp. of **mulot, jahat**) *n* caustic mouth
mulot kering [mu.lot kə.riŋ] (Malay) (comp. of **mulot, kering**) *n* inability to say more
mulot manis [mu.lot ma.nis] (Malay) (comp. of **mulot, manis**) *n* ability to speak sweetly
mulot pantat ayam [mu.lot pan.tat a.jam] (Malay) (comp. of **mulot, pantat, ayam**) *n* deceitful mouth
mulot ringan [mu.lot ri.ŋan] (Malay) (comp. of **mulot, ringan**) *n* polite character
mungka [muŋ.ka] (Malay) *n* face comp. **buat mungka**
musim [mu.sim] (Arabic) *n* season

N

na [na] (Hokkien) *interj* here you go

naik [najʔ] (Malay) *v* **1** climb **2** rise **3** ascend comp. **naik baik** , comp. **naik darah** , comp. **naik geléték** ,

comp. **naik geram** , comp. **naik gila** , comp. **naik lemak** , comp. **naik pangkat** , comp. **naik sedap** , comp. **naik seram**

naik baik [naj? baj?] (Malay) (comp. of **naik**, **baik**) *v* change for good

naik darah [naj? da.rah] (Malay) (comp. of **naik**, **darah**) *v* be angry

naik geléték [naj? gə.le.te?] (Malay) (comp. of **naik**, **geléték**) *v* be up to mischief

naik geram [naj? gə.ram] (Malay) (comp. of **naik**, **geram**) *v* be furious

naik gila [naj? gi.la] (Malay) (comp. of **naik**, **gila**) *v* be mad

naik lemak [naj? lə.ma?] (Malay) (comp. of **naik**, **lemak**) *v* be up to mischief

naik pangkat [naj? paŋ.kat] (Malay) (comp. of **naik**, **pangkat**) *v* be promoted

naik sedap [naj? sə.dap] (Malay) (comp. of **naik**, **sedap**) *v* be satisfied

naik seram [naj? sə.ram] (Malay) (comp. of **naik**, **seram**) *v* be frightened

nama [na.ma] (Malay) *n* name

nampak [nam.pa?] (Malay) *v* see

nang [naŋ] var. of **yang**

nangis [na.ŋis] (Malay) *v* cry

nanti [nan.ti] (Malay) *adv* later

napas [na.pas] (Malay) *n* breath

napsu [nap.su] (Malay) **1** *v* desire **2** *n* desire

nasik [na.si?] (Malay) *n* cooked rice

nganga [ŋa.ŋa] (Malay) *adj* mouth agap

ngéngé [ŋe.ŋe] (Hokkien) *adj* obstinate

ni [ni] (Malay) (contr. of **ini**) *dem* this

nia [ɲa] (Malay) (**punya** **1**) *v* POSS **2**) REL

nio [njo] (Hokkien) *n* mother-in-law

nua [nwa] (Hokkien) *adj* smashed up

number [num.bor] (English) *n* number

nya [ɲa] (Portuguese) (contr. of **Nyonya**) *n* **1**) mother **2**) daughter

nyanyi [ɲa.ɲi] var. of **menyanyi**

Nyonya [ɲo.ɲa] (Portuguese) *n* Peranakan female contr. **nya**

O

onéng.onéng [o.neŋ.o.neŋ] (Malay) *n* great-great grandchild

opan [o.pan] (Hokkien) *n* freckles

orang [o.raŋ] (Malay) *n* **1**) people **2**) person

ormat [or.mat] var. of **hormat**

orna [or.na] (Malay) *n* colour

orpau [or.paw] var. of **horpau**

otak [o.ta?] (Malay) *n* brain

P

paderi [pa.də.ri] (Portuguese) *n* priest

pagi [pa.gi] (Malay) *n* morning

paiséh [paj.seh] (Hokkien) *adj* embarrassed

pait [pajt] (Malay) *adj* bitter

pak [pa?] (Malay) *n* male person

pak-chindék [pa?-tʃin.de?] (Malay) *n* groom's ceremonial

assistant
pakay [pa.ke] (Malay) *v* **1** wear **2** use
paksa [paʔ.sa] (Malay) *v* force
paku [pa.ku] (Malay) *n* nail
panair [pa.nɛ] refined var. of **panas**
panas [pa.nas] (var. **panair**) (Malay) *adj* hot (coarse)
panday [pan.de] (Malay) *adj* clever
panggang [paŋ.gaŋ] (Malay) *v* roast
panggay [paŋ.ge] (Malay) *v* call
panggung [paŋ.goŋ] (Malay) *n* stage
pangkat [paŋ.kat] (Malay) *n* rank **pangkat adek beradek**, comp. **naik pangkat**
pangkat adek beradek [paŋ.kat a.deʔ bə.ra.deʔ] (Malay) (**pangkat, adék beradék**) *n* cousins
panjang [pan.dʒaŋ] (Malay) *adj* long
panjat [pan.dʒat] (Malay) *v* climb
pantang [pan.taŋ] (Malay) *adj* superstitious **pantang ayam** [pan.tat] (Malay) *n* buttocks comp. **mulot pantang ayam**
parang [pa.raŋ] (Malay) *n* chopper
pasair 1 [pa.se] refined var. of **pasar**
pasair 2 [pa.se] refined var. of **pasal** contr. **sair 1** matter reason
pasal [pa.sal] (var. **pasair 2**) (Malay) *n* **1** matter (coarse) **2**) reason comp. **apa pasal**
pasang 1 [pa.saŋ] (Malay) *v* use
pasang 2 [pa.saŋ] (Malay) *clf* pair
pasar [pa.sar] (var. **pasair 1**) (Malay) *n* market (coarse)
paser [pa.ser] (Malay) *n* sand
patah [pa.tah] (Malay) *v* snap
pati [pa.ti] (Malay) *n* first cream of coconut
paya [pa.ja] (English) *n* papaya
payong [pa.joŋ] (Malay) *n* umbrella
pé'él [peʔe] (Arabic) *n* good character
pechah [pə.tʃah] (Malay) *v* break
pégang [pe.gaŋ] (Malay) *v* hold
pegi [pə.gi] (Malay) *v* go contr. **pi**

pék [peʔ] (Hokkien) *n* father's elder brother
pék-pék [peʔ-peʔ] (Hokkien) *n* father's eldest brother
pekak [pə.kaʔ] (Malay) *adj* deaf
pekara [pə.ka.ra] (Malay) *n* matter
pekat [pə.kat] (Malay) *adj* thick (describing liquid)
pekék [pə.keʔ] (Malay) *v* shout
pékji [peʔ.dʒi] (Hokkien) (**puay**) *n* eight characters of Chinese horoscope
pelan [pə.lan] (Malay) *adj* slow comp. **pelan pelan**
pelan pelan [pə.lan pə.lan] (Malay) (comp. of **pelan**) *adv* slowly
peléchok [pə.le.tʃoʔ] (Malay) *v* to twist one's foot
penat [pə.nat] (Malay) *adj* tired
péndék [pen.deʔ] (Malay) *adj* short
pengapék [pəŋ.ga.peʔ] (Malay) *n* bride's girl assistant
pengat [pə.ŋat] (Malay) *n* sweet dessert of thick coconut milk with banana chunks
péngsan [peŋ.san] (Malay) *v* faint
penoh [pə.noh] (Malay) *adj* full
perah [pə.rah] (Malay) *v* wring
perama [pə.ra.ma] (Malay) *adj* polite
Peranakan [pə.ra.na.kan] (Malay) *n* Straits-born Chinese
peranjat [pə.ra.n.dʒat] (Malay) *adj* surprised comp. **terperanjat**
perchaya [pəɾ.tʃa.ja] (Malay) *v* trust
periok [pə.rjoʔ] (Malay) *n* cooking pot
pernah [pəɾ.nah] (Malay) *adv* ever
perompuan [pə.rom.pwan] (Malay) *n* female
perot [pə.rot] (Malay) *n* **1** belly **2**) stomach
pertama [pəɾ.ta..a] (Malay) *adj* first
pésan [pe.san] (Malay) *v* instruct
petang [pə.taŋ] (Malay) *n* evening
petay [pə.te] (Malay) *n* flat bean
peték [pə.teʔ] (Malay) *v* pluck
pi [pi] (Malay) (contr. of **pegi**) *v* go
piara [pja.ra] (Malay) *v* raise children
pikay [pi.ke] (Arabic) *v* think
piléh [pi.leh] (Malay) *v* choose

pilu [pi.lu] (Malay) *n* sorrow
pinggan [piŋ.gan] Persian *n* plate
pinggang [piŋ.gan] (Malay) *n* waist
pinjak [pin.dʒaʔ] (Malay) *v* step on
pinjam [pin.dʒam] (Malay) *v* lend
pintu [pin.tu] (Malay) *n* door
pisang [pi.saŋ] (Malay) *n* banana
pisso [pi.so] (Malay) *n* knife comp. **mata pisso**
po [po] (Hokkien) *n* grand-aunt
pokok [po.koʔ] (Malay) *n* tree
popiah [po.pjah] (Hokkien) *n* spring roll with turnip
potong [po.ton] (Malay) *v* cut
puay [pwe] (Hokkien) *cardnum* eight **pékji**

pukol [pu.kol] (Malay) *v* **1** hit **2** strike
pulak [pu.laʔ] (Malay) *adv* instead
pulang [pu.laŋ] (Malay) *v* return
puloh [pu.loh] (Malay) *cardnum* ten
pulot [pu.lot] (Malay) *n* glutinous rice
pun [pun] (var. **kun 2**) (Malay) *adv* also
punggot [puŋ.got] (Malay) *v* pick up
punya [pu.ɲa] (Malay) **1** *v* POSS **2** REL **nia**, contr. **mia**
2
purot [pu.rot] (Malay) *n* stomach
pusing [pu.siŋ] (Malay) *v* **1** whirl **2** turn
putéh [pu.teh] (Malay) *adj* white

R

raba [ra.ba] (Malay) *v* touch
radat [ra.dat] (Malay) *adj* greed (vulgar)
raja [ra.dʒa] (Sanskrit) *n* king
rambat [ram.bat] (Malay) *n* hallway
rambot [ram.bot] (Malay) *n* **1** hair **2** coarse hair
ranjang [ran.dʒaŋ] (Malay) *n* bed
rasa [ra.sa] (Sanskrit) **1** *v* feel **2** *v* think
ratus [ra.tus] (Malay) *cardnum* hundred
rebus [rə.bus] (Malay) *v* boil
rendah [rən.dah] (Malay) *adj* low
renjis [rən.dʒis] (Malay) *v* to water plants
reti [rə.ti] (Malay) *v* understand

ribu [ri.bu] (Malay) *cardnum* thousand
rindu [rin.du] (Malay) *v* miss
ringan [ri.ŋan] (Malay) *adj* slim comp. **mulot ringan**
riyang [ri.jaŋ] (Malay) *adj* lively
ronggég [roŋ.geŋ] (Malay) *n* faster version of jogét
rosak [ro.saʔ] (Malay) *adj* spoilt
roti [ro.ti] (Sanskrit) *n* bread
ruboh [ru.boh] (Malay) *v* collapse
rumah [ru.mah] (Malay) *n* house
rumpot [rum.pot] (Malay) *n* grass
rupa [ru.pa] (Sanskrit) *n* appearance
rusak [ru.saʔ] (Malay) *adv* day after tomorrow

S

sa [sa] (Hokkien) *cardnum* three
sahja [sa.dʒa] (Malay) *adv* only
sair 1 [sɛ] (Malay) (contr. of **pasair 2**) **1** *n* matter **2**) *n*
 reason contr. **apa sair**

sair 2 [sɛ] (Malay) *pvt* Confirmative
sakit [sa.kit] (Malay) *adj* sick
salah [sa.lah] (Malay) **1** *n* mistake **2**) *adj* wrong
salat [sa.lat] (Malay) *n* custard rice

salin [sa.lin] (Malay) *v* transfer
sama [sa.ma] (Sanskrit) **1** *adj* same **2** *conn* and **3** *prep* with
sambal [sam.bal] (Malay) *n* chilli paste made with shrimps or anchovies
sampay [sam.pe] (Malay) **1** *prep* until **2** *v* reach
samplang [sam.plaŋ] source uncertain *adj* promiscuous
samséng [sam.serŋ] (Hokkien) *n* offering of three meats
sana [sa.na] (Malay) *adv* there
sandah [san.dah] (Malay) *v* lean
sangat [sa.ŋat] (Malay) *adv* very
sangka [saŋ.ka] (Malay) *v* expect
sangkék-em [saŋ.keʔ-əm] (Hokkien) (**em**) *n* bride's ceremonial assistant
sapu [sa.pu] (Malay) *v* sweep
satu [sa.tu] (Malay) *cardnum* one *contr.* **se**
saya (Sanskrit) *pers* 1.SG (refined)
sayang [sa.jaŋ] (Malay) **1** *v* care **2** *n* care **3** *interj* pity
sayor [sa.jor] (Malay) *n* vegetables
sayup [sa.jup] (Malay) *adj* faint
se [sə] (Malay) (*contr.* of **satu**) *cardnum* one
sebelair [sə.bə.lɛ] refined *var.* of **sebelas**
sebelas [sə.bə.las] (*var.* **sebelair**) (Malay) *cardnum* eleven (*coarse*)
seberang [sə.bə.raŋ] (Malay) *v* cross
sébok [se.boʔ] (Malay) *adj* busy
sedang [sə.daŋ] (Malay) *adj* medium
sedap [sə.dap] (Malay) *adj* delicious *comp.* **naik sedap**
sedéh [sə.deh] (Malay) *adj* sad
segan [sə.gan] (Malay) *adj* shy *comp.* **buang segan**
séhjit [seh.dʒit] (Hokkien) *n* birthday
séisema [sej.sə.ma] (Malay) *n* cold
sejok [sə.dʒoʔ] (Malay) *adj* cold
sekair [sə.ke] (Malay) *adv* indeed (refined) (*unsystematic variant* of *sekali*)
sekali [sə.ka.li] (Malay) *adv* **1** very **2** indeed
sekarang [sə.ka.raŋ] (Malay) *adv* now
sekejab [sə.kə.dʒab] (Malay) *adv* awhile

sekolah [sə.ko.lah] (Dutch) *n* school
selalu [sə.la.lu] (Malay) *adv* always
selamat [sə.la.mat] (Arabic) *adj* safe
seluar [sə.lwar] (*var.* **suluar**) (Malay) *n* pants (*coarse*)
semayang [sə.ma.jaŋ] (Malay) *v* pray
semilan [sə.mi.lan] (Malay) *cardnum* nine
semonyet [sə.mo.jɛt] (Malay) *v* hide
sempang [səm.paŋ] (Malay) *n* traffic junction
senang [sə.naŋ] (Malay) **1** *adj* easy **2** *adj* free
sendiri [sən.di.ri] (Malay) *n* self
séndok [sen.doʔ] (Malay) *n* spoon
sengbuay [səŋ.bwe] (Hokkien) *n* sour plums
séngét [se.ŋeʔ] (Malay) *adj* crooked
séngkang [seŋ.kaŋ] (Malay) *v* block
senyap [səŋap] (Malay) *adj* silent
senyum [sə.ɲum] (Malay) **1** *v* smile **2** *n* smile
sepekah [sə.pə.kah] (Malay) *adj* acceptable
seram [sə.ram] (Malay) *adj* frightening *comp.* **naik seram**
seronoh [sə.ro.noh] (Malay) *adj* proper
setuju [sə.tu.dʒu] (Malay) *v* agree
shiok [ʃjoʔ] (Punjabi) **1** *adj* satisfied feeling **2** EXCLAM satisfaction
si 1 [si] Indonesian *dem* Person
si 2 [si] (Hokkien) *cardnum* four
siang [sjaŋ] (Malay) **1** *adj* early **2** *n* day
siap [sjap] (Malay) *v* prepared
siap-siap [sjap-sjap] (Malay) *adv* perpetually
siapa [sja.pa] (Malay) *interrog* who
sién [sjen] (Hokkien) *n* Chinese immortal
sihat [si.hat] (Malay) *adj* healthy
siki [si.ki] (Hokkien) *n* death anniversary
sikit [si.kit] (Malay) *adv* little
simpan [sim.pan] (Malay) *v* keep something safe
sini [si.ni] (Malay) *adv* here
sinjakala [sin.dʒa.ka.la] (Malay) (**kala**) *n* dusk
sinkék [sin.keʔ] (Hokkien) *n* newcomer from China

siram [si.ram] (Malay) *v* flush with water
siséh [si.seh] Indonesian *n* comb **sisék** [si.seʔ] (Hokkien)
n four colours cardgame
siut [sjut] (Hokkien) *v* singe
so 1 [so] (Hokkien) *v* burn **so.hio**
so 2 [so] (Hokkien) *n* elder brother's wife
so.hio [sohjo] (Hokkien) (**so 1**) *v* burn incense
sua [swa] (Malay) (contr. of **sudah**) *adv* already
suap [swap] source unclear *v* feed
suay [swe] (Hokkien) *adj* unlucky comp. **buat suay**
sudah [su.dah] (Malay) *adv* already contr. **sua**

suka [su.ka] (Malay) *v* like
suluair [su.lwɛ] refined var. of **suluar**
suluar [su.lwar (coarse)] (var. **suluair**) var. of **seluar**
sumua [su.mwa] (Arabic) *adv* all
sunggay [suŋ.ge] (Malay) *n* river
sunggu [suŋ.gu] (Malay) *adv* really
sup [sup] (English) *n* soup
surat [su.rat] (Malay) *n* letters
susah [su.sah] (Malay) *adj* difficult
susu [su.su] (Malay) *n* milk

T

tachi [ta.tʃi] (Hokkien) *n* elder sister
tadi [ta.di] (Malay) *adv* just
tahan [ta.han] (Malay) *v* withstand
tait [tajt] (Malay) *n* excrement
tak [taʔ] (Malay) *adv* Neg verb comp. **buat tak dengair**
takot [ta.kot] (Malay) *v* afraid
taman [ta.man] (Malay) *n* garden
tambah [tam.bah] (Malay) *v* add on
tampair [tam.pe] refined var. of **tampal**
tampal [tam.pal] (var. **tampair**) (Malay) *v* mend (coarse)
tanah [ta.nah] (Malay) *n* **1** soil **2** ground **3** earth
tanam [ta.nam] (Malay) *v* plant
tanda [tan.da] (Malay) *n* sign comp. **tanda mata**
tanda mata [tan.da ma.ta] (Malay) (comp. of **tanda**,
mata) *n* heirloom
tandok [tan.doʔ] (Malay) *n* horn
tangan [ta.ŋan] (Malay) *n* hand
tangchék [taŋ.tʃeʔ] (Hokkien) *n* Winter Solstice
tangga [taŋ.ga] (Malay) *n* **1** ladder **2** stairs
tangkap [taŋ.kap] (Malay) *v* capture
tanya [tan.ja] (Malay) *v* ask
taoun [ta.own] (Malay) *n* year

tapay [ta.pe] (Malay) *n* fermented rice dessert
tapi [ta.pi] (Sanskrit) *conn* but
tapis [ta.pis] (Malay) *v* filter
tarék [ta.reʔ] (Malay) *v* pull
tarok [ta.roʔ] (Malay) *v* put
tau [taw] (Malay) *v* know
taukay [taw.ke] (Hokkien) *n* boss
tauyy [taw.ju] (Malay) *n* dark soya sauce
tawair [ta.wɛ] refined var. of **tawar**
tawar [ta.war] (var. **tawair**) (Malay) *adj* tasteless (coarse)
téh [teh] (Hokkien) *n* tea
teloh [tə.loh] (Malay) *n* egg
téloh [te.loh] (Malay) *n* accent
témbok [tem.boʔ] (Malay) *n* wall
tempat [təm.pat] (Malay) *n* place
tendang [tən.dan] (Malay) *v* kick
tengah [tə.ŋah] (Malay) *adj* middle **tengah hari** ,
tengah jalan tengah hari [tə.ŋah ha.ri] (Malay)
(tengah, hari) *n* midday
tengah jalan [tə.ŋah dʒa.lan] (Malay) (**tengah, jalan**)
n road
tenggelam [təŋ.gə.lam] (Malay) *v* sink

téngok [te.ŋoʔ] (Malay) *v* see
tentu [tən.tu] (Malay) *adj* definite
tepi [tə.pi] (Malay) **1** *n* side **2** *adv* side
tepok [tə.poʔ] (Malay) *v* clap
ter- [tər] (Malay) **1** accidental **2** movement comp.
terperanjat
terang [tə.raŋ] (Malay) *adj* **1** bright **2** clear
terbang [tər.baŋ] (var. **trebang**) (Malay) *v* fly
terbantot [tər.ban.tot] (Malay) *adj* unripe
terbiat [tər.bjat] (Malay) *n* attitude comp. **buang tebiét**
teriak [tə.rjaʔ] (Malay) *v* call out
terima [tə.ri.ma] (Malay) *v* accept
terlalu [tə.la.lu] (Malay) *adv* too
terok [tə.roʔ] (Malay) *adj* terrible
terompak [tə.rom.paʔ] (Malay) *n* slippers
terperanjat [tər.pə.ran.dʒat] (Malay) (comp. of **ter-**,
peranjat) *adj* be shocked
terus [tə.rus] (Malay) *adj* straight
téték [te.teʔ] (Malay) *n* breast
ti-gong [ti-goŋ] (Hokkien) *n* sky-deity (Jade Emperor)
tiap [tjap] (Malay) *adv* every
tiarap [tja.rap] (Malay) *v* fall on one's front
tidor [ti.dor] (Malay) *v* sleep
tiga [ti.ga] (Malay) *cardnum* three
tikam [ti.kam] (Malay) *v* **1** stab **2** bet
tim [tim] (Hokkien) *v* double-boil
timbang [tim.baŋ] (Malay) *v* weigh
timbol [tim.bol] (Malay) *v* float
timun [ti.mun] (Malay) *n* cucumber
tinggair [tiŋ.gɛ] refined var. of **tinggal**
tinggal [tiŋ.gal] (var. **tinggair**) (Malay) *v* **1** live (coarse)
2 stay (coarse)
tinggi [tiŋ.gi] (Malay) *adj* tall
tio [tjo] (Hokkien) *n* parent's sister's husband
tio-tio [tjo-tjo] (Hokkien) *n* **1** parents' eldest sister's

husband **2** second father
tiop [tjop] (Malay) *v* blow
tipu [ti.pu] (Malay) *v* cheat
titék [ti.teʔ] source unclear *n* spice paste (for papaya soup)
tok [toʔ] (Hokkien) *n* table
toksa [toʔ.sa] (Malay) *v* do not need
tolak [to.laʔ] (Malay) *v* push
tolong [to.loŋ] (Malay) *v* help
tompang [tom.paŋ] (Malay) *v* hitch (a ride)
tongchit [toŋ.tʃit] (Cantonese) *v* eat all (mahjong term)
tongkat [toŋ.kat] (Malay) *n* walking stick
tork [tɔʔ] (Hokkien) *adj* **1** poisonous **2** evil (used in
particular for describing stepmothers)
trebang [trə.baŋ] var. of **terbang**
tu [tu] (Malay) (contr. of **itu**) *dem* that
tua [twa] (Hokkien) *adj* **1** old **2** big
tuaban [twa.ban] (Hokkien) *adj* handsome
tuakang [twa.kan] (Hokkien) *adj* generous
tuan [twan] (Malay) *n* boss
tuang [twaŋ] (Malay) *v* pour and serve
tujoh [tu.dʒoh] (Malay) *cardnum* seven
tukair [tu.ke] refined var. of **tukar**
tukang [tu.kan] (Malay) *n* laborer
tukar [tu.kar] (var. **tukair**) (Malay) *v* change (coarse)
tulang [tu.laŋ] (Malay) *n* bone
tulis [tu.lis] (Malay) *v* write
tumbok [tum.boʔ] (Malay) *v* pound
tumis [tu.mis] (Malay) *v* saute spices
tumpus [tum.pus] (Malay) *adj* blunt **tunggang**
[tuŋ.gaŋ] (Malay) *v* ride
tunggu [tuŋ.gu] (Malay) *v* wait
tunjok [tun.dʒoʔ] (Malay) *v* show
turun [tu.run] (Malay) *v* descend
tutup [tu.top] (Malay) *v* close

U

udang [u.daŋ] (Malay) *n* prawn

ujan [u.dʒan] var. of **hujan**

ulam [u.lam] (Malay) *n* mixed herbs rice with shredded anchovies

ular [u.lar] (var. **uluair**) (Malay) *n* snake (coarse)

uluair [u.lɛ] refined var. of **ular**

umor [u.mor] (Arabic) *n* age

untok [un.toʔ] (Malay) *adv* for

utan [u.tan] var. of **hutan**

W

wak [waʔ] (Malay) *n* elderly Malay person

wangi [wa.ŋi] (Malay) *adj* fragrant

wayang [wa.jaŋ] (Malay) *n* play

wilo [wi.lo] (Hokkien) *n* hot pot for lunar new year reunion dinner

Y

ya [ja] (Malay) *interj* yes

yang [jaŋ] (var. **nang**) (Malay) REL

yauguai [jaw.gwaj] (Hokkien) *n* demon

yaukin [jau.kin] (Hokkien) *adj* important

yénchi [jen.tʃi] (Hokkien) *n* rouge

yi 1 [ji] (Hokkien) *n* food in the shape of a ball comp. **kuéh-yi**

yi 2 [ji] (Hokkien) *n* mother's sister **mak-yi**

yi-yi [ji-ji] (Hokkien) *n* mother's eldest sister

yok.hun [joʔ.hun] (Hokkien) *n* Chinese medicinal powder

English-Singapore Baba Malay

1

1.PL kita [ki.ta] *pro*
1.SG (coarse) gua [gwa] *pro*

1.SG (refined) saya [saja] *pro*

2

2.SG lu [lu] *pro*

3

3.SG dia [dja] *pro*

3.PL dia [dja] *pro*; dia-orang [dja-o.raŋ] *pro*

A

ability to make accurate predictions mulot asin [mu.lot a.sin] (comp. of mulot, asin) *n*
ability to speak sweetly mulot manis [mu.lot ma.nis] (comp. of mulot, manis) *n*
accent téloh [te.loh] *n*
accept terima [tə.ri.ma] *v*
acceptable sepekah [sə.pə.kah] *adj*
accidental ter- 1 [tər] (comp. terperanjat)
add on tambah [tam.bah] *v*
afraid takot [ta.kot] *v*
after lepas [lə.pas] *prep*
again balék 4 [ba.leʔ] *adv*; lagik 3 [la.giʔ] *adv*
age umur [u.mor] *n*
agree setuju [sə.tu.dʒu] *v*
alcohol arat [a.rat] *n*
all sumua [su.mwa] *adv*

already sua [swa] (contr. of sudah) *adv*; sudah [su.dah] *adv*
also jugak [dʒu.gaʔ] *adv*; kun 2 [kun] (var. of pun) *adv*; pun [pun] *adv*
always selalu [sə.la.lu] *adv*
amulet hu [hu] *n*
ancestors gong-cho [goŋ-tʃo] (comp. of gong, cho) *n*
anchovy bilis [bi.lis] *n*
and sama 2 [sa.ma] *conn*;
angry marah [ma.rah] *adj*
animal ékor [e.kor] *clf*
anonymous ano [a.no] *adj*
answer jawab [dʒa.wab] *v*
appearance rupa [ru.pa] *n*
argue (coarse) bertengkar [bər.təŋ.kar] *v*
argue (refined) bertengkair [bər.təŋ.kɛ] *v*

ascend naik 3 [najʔ] *v* (comp. naik baik comp. naik darah comp. naik geléték comp. naik geram comp. naik gila comp. naik lemak comp. naik pangkat comp. naik sedap comp. naik seram)
ash abu [a.bu] *n*
aside from buang sebelah [bwaŋ sə.bə.lah] (comp. of

buang, sebelah) *adv*
ask tanya [tan.ja] *v*
ask sincerely mintak [min.taʔ] *v*
at dekat 1 [də.kat] (var. kat) *prep*
attitude terbiat [tər.bjat] *n* (comp. buang tebiét)
awhile sekejab [sə.kə.dʒab] *adv*

B

back balék 3 [ba.leʔ] *adv*
backside buntut 2 [bun.tot] *n*
bad breath mulot busok [mu.lot bu.soʔ] (comp. of mulot, busok) *n*
bad sport chaukah [tʃaw.kah] *n*
bad type of person chaukuan [tʃaw.kwan] *n*
bag bég [beʔ] *n*
bald botak [bo.taʔ] *adj*
banana pisang [pi.saŋ] *n*
bang hantok [han.toʔ] *v*
banner chaiki [tʃaj.ki] *n*
bark gonggong [goŋ-goŋ] *v*
basket bakol [ba.kol] *n*
bathe mandi [man.di] *v*
be angry naik darah [najʔ da.rah] (comp. of naik, darah) *v*
be frightened naik seram [najʔ sə.ram] (comp. of naik, seram) *v*
be furious naik geram [najʔ gə.ram] (comp. of naik, geram) *v*
be mad naik gila [najʔ gi.la] (comp. of naik, gila) *v*
be promoted naik pangkat [najʔ paŋ.kat] (comp. of naik, pangkat) *v*
be satisfied naik sedap [najʔ sə.dap] (comp. of naik, sedap) *v*
be shocked terperanjat [tər.pə.ran.dʒat] (comp. of ter-, peranjat) *adj*
be sour-faced buat mungka [bwat muŋ.ka] (comp. of

buat, mungka) *v*
be up to mischief naik geléték [najʔ gə.le.teʔ] (comp. of naik, geléték) *v*; naik lemak [najʔ lə.maʔ] (comp. of naik, lemak) *v*
bead manék [ma.neʔ] *n*
bear beruang [bə.rwaŋ] *n*
beautiful chanték [tʃan.teʔ] *adj*
become jadi 2 [dʒa.di] *v*
bed ranjang [ran.dʒaŋ] *n*
before belom 2 [bə.lom] *adv*; dulu 1 [du.lu] *adv*
beginning of lunar month chay [tʃe] *n*
behind belakang [bə.la.kəŋ] *prep*
bell lochéng [lo.tʃeŋ] *n*
belly perot 1 [pə.rot] *n*
beloved jantung atí [dʒan.toŋ a.ti] (comp. of jantung, atí) *n*
bent and curled up kerekot [kə.re.kot] *adj*
bet tikam 2 [ti.kam] *v*
big besar [bə.sar] (var. besair) *adj*; tua 2 [twa] *adj*
big (refined) besair [bə.sə] (var. of besar) *adj* (comp. buang ayé besair)
bird burong [bu.roŋ] *n*
birth chutsi [tʃut.si] *n*
birthday séhjit [seh.dʒit] *n*
bite gigit [gi.git] *v*
bitter pait [pajt] *adj*
black hitam [hi.tam] (var. itam) *adj*; itam [i.tam] *adj*

blade of knife mata piso [ma.ta pi.so] (comp. of mata, piso) *n*
blending of the marital bed ancheng [an.tʃəŋ] *n*
blind buta [bu.ta] *adj*
blink kelék.keléék [kələʔ.kələʔ] *v*
block séngkang [seŋ.karŋ] *v*
blood darah [da.rah] *n* (comp. naik darah)
bloom kuntum [kun.tum] *clf*
blow tiop [tjop] *v*
blunt tumpus [tum.pus] *adj*
body badan [ba.dan] *n*
boil rebus [rə.bus] *v*
boiling menidi [mə.ni.di] *adj*
bone tulang [tu.larŋ] *n*
book buku [bu.ku] *n*
boss taukay [taw.ke] *n*; tuan [twan] *n*
bottle botol [bo.tol] *n*
bowl mangkok [maŋ.koʔ] *n*
box kepok [kə.poʔ] *n*
boy lelaki [lə.la.ki] *n* (contr. laki)
bracelet gelang [gə.larŋ] *n*
brain otak [o.taʔ] *n*
bread roti [ro.ti] *n*

break pechah [pə.tʃah] *v*
breast tétéék [te.teʔ] *n*
breath napas [na.pas] *n*
bride kemantin [kə.man.tin] *n*
bridegroom's family's host chu-lang [tʃu.larŋ] *n*
bride's assistant for kneeling and general help bukak kun [bu.kaʔ kun] *n* (comp. of kun 2)
bride's ceremonial assistant sangkéék-em [sarŋ.keʔ-əm] *n* (comp of em)
bride's girl assistant pengapék [pəŋ.ga.peʔ] *n*
bright terang 1 [tə.raŋ] *adj*
bring bawak [ba.waʔ] *v*
brush berus [bə.rus] *v*
bundle bungkus [buŋ.kus] *n*
burn so 1 [so] *v* (comp. sohio)
burn (coarse) bakar [ba.kar] *v*
burn (refined) bakair [ba.kə] *v*
burn incense sohio [sohjo] (comp of so 1)
busy sébok [se.boʔ] *adj*
but tapi [ta.pi] *conn*
butter mentéga [mən.te.ga] *n*
butterfly kupu-kupu [ku.pu-ku.pu] *n*
buttocks pantat [pan.tat] *n* (comp. mulot pantat ayam)
buy beli [bə.li] *v*

C

cake kéék [keʔ] *n*
call panggay [paŋ.ge] *v*
call out teriak [tə.rjaʔ] *v*
can boléh [bo.leh] *v*
capture tangkap [taŋ.kap] *v*
car chia [tʃja] *n*; keréta [kə.re.ta] *n*
care sayang 1 [sa.jarŋ] *v*; sayang 2 [sa.jarŋ] *n*
carry angkat 4 [aŋ.kat] *v*
carry baby logong [lo.goŋ] *v*
cat kucing [ku.tʃiŋ] *n* (comp. kucing belanda comp.

mata kucing)
cause kasi 3 [ka.si]
cause embarrassment buat malu [bwat ma.lu] (comp. of buat, malu) *v*
cause misfortune buat suay [bwat swe] (comp. of buat, suay) *v* **caustic mouth** mulot jahat [mu.lot dʒa.hat] (comp. of mulot, jahat) *n*
cave batu 3 [ba.tu] *n*
chair kerosi [kə.ro.si] *n*
change (coarse) tukar [tu.kar] *v*

change (refined) tukair [tu.kɛ] *v*
change for good naik baik [naj? baj?] (comp. of naik, baik) *v*
character budi [bu.di] *n*
chase (coarse) kejar [kə.dʒar] *v*
chase (refined) kejair [kə.dʒɛ] *v*
chase away halo [ha.lo] *v*
cheat tipu [ti.pu] *v*
chicken ayam [a.jam] *n* (comp. mulot pantat ayam)
child anak [a.na?] *n*; budak [bu.da?] *n*
chilli paste made with shrimps or anchovies sambal [sam.bal] *n*
Chinese China [tʃi.na] *adj*
Chinese medicinal powder yok.hun [jo?.hun] *n*
Chinese immortal sién [sjen] *n*
choose piléh [pi.leh] *v*
chopper parang [pa.raŋ] *n*
church gerja [gɛr.dʒa] (var. greja) *n*; greja [grɛ.dʒa] *n*
clap tepok [tə.po?] *v*
clean bersi [bɛr.si] (var. bresi) *v*; bersi [bɛr.si] *n*
clear terang 2 [tɛ.raŋ] *adj*
clever panday [pan.de] *adj* ever
climb naik 1 [naj?] *v* (comp. naik baik comp. naik darah comp. naik geléték comp. naik geram comp. naik gila comp. naik lemak comp. naik pangkat comp. naik sedap comp. naik seram); panjat [pan.dʒat] *v*
clogs chakiak [tʃa.kja?] *n*
close tutop [tu.top] *v*
close to one another chin [tʃin] *adj*
cloth kain [kajn] *n*
clothes baju [ba.dʒu] *n*
cloud awan [a.wan] *n*
coarse (coarse) kasar [ka.sar] *adj*
coarse (refined) kasair [ka.sɛ] *adj*
coarse hair rambot 2 [ram.bot] *n*
cockeyed mata juling [ma.ta dʒu.liŋ] (comp. of mata, juling) *n*

cockroach kachua [ka.tʃwa] *n*
coconut cookie bangkit [ban.kit] *n*
coconut jam kaya 2 [ka.ja] *n*
coffee kopi [ko.pi] *n*
cold séisema [sej.sə.ma] *n*; sejok [sə.dʒo?] *adj*
collapse ruboh [ru.boh] *v*
colour orna [or.na] *n*
comb siséh [si.seh] *n*
come datang [da.taŋ] *v*
come let us mari [ma.ri] *v*
Confirmative sair 2 [sɛ] *prt*
convey by speaking (as opposed to writing) buang mulot [bwaŋ mu.lot] (comp. of buang, mulot) *v*
cook masak 1 [ma.sa?] *v*
cooked food lauk [law?] *n*
cooked in coconut milk lemak [lə.ma?] *adj*
cooked rice nasik [na.si?] *n*
cooking pot periok [pɛ.rjo?] *n*
COP ada 5 [a.da] (var. a) *v*
correct betul 1 [bɛ.tol] *adj*
cousins pangkat adek beradek [paŋ.kat a.de? bɛ.ra.de?] (comp. of pangkat, adék beradék) *n*
cow lembu [ləm.bu] *n* (comp. mata lembu)
crash (coarse) langgar [laŋ.gar] *v*
crash (refined) langgair [laŋ.gɛ] *v*
crawl belangkat [bə.laŋ.kat] *v*
crooked séngét [se.ŋe?] *adj*
cross seberang [sɛ.bɛ.raŋ] *v*
crowded beramay [bɛ.ra.me] *adj*
cruel ati busok [a.ti bu.so?] (comp. of ati, busok) *adj*
crumpled kudut [ku.dut] *adj*
cry menangis [mə.na.ŋis] *v*; nangis [na.ŋis] *v*
cucumber timun [ti.mun] *n*
cunning lihai [li.haj] *adj*
cup changkay [tʃaŋ.ke] *n*
curse mincharok [min.tʃa.ro?] *v*
custard rice salad [sa.lat] *n*
cut potong [po.toŋ] *v*

cylindrical measure of one gallon of rice gantang
[gan.taŋ] *clf*

cylindrical measure of quarter gallon of rice
chupak[tʃu.paʔ] *clf*

D

daring gaga [ga.ga] *adj*
dark gelap [gə.lap] *adj* (comp. mata gelap)
dark soya sauce tauyu [taw.ju] *n*
daughter nya 2 [ɲa] (contr. of Nyonya) *n*
day ari [a.ri] (var. of hari) *n* (comp. mata ari); hari [ha.ri]
(var. ari) *n* (comp. tengah hari); siang 2 [sjaŋ] *n*
day after tomorrow rusak [ru.saʔ] *adv*
dead mati 2 [ma.ti] *adj*
deaf pekak [pə.kaʔ] *adj*
death anniversary siki [si.ki] *n*
deceitful mouth mulot pantat ayam [mu.lot pan.tat
a.jam] (comp. of mulot, pantat, ayam) *n*
decorative altar stand chanab [tʃa.nab] *n*
defecate bérak [be.raʔ] *v*; buang ayé besair [bwaŋ a.je
bə.sɛ] (comp. of buang, ayé, besair) *v*
definite tentu [tən.tu] *adj*
delicious sedap [sə.dap] *adj* (comp. naik sedap)
demon yauguai [jaw.gwaj] *n*
descend turun [tu.run] *v*
desire napsu 1 [nap.su] *v*; napsu 2 [nap.su] *n*
detective mata gelap [ma.ta gə.lap] (comp. of mata,
gelap) *n*
develop an illness kamtiok [kam.tjoʔ] *v*
diamond belién [bə.ljen] *n*
diamond chip intan [in.tan] *n*
die mati 1 [ma.ti] *v*
difficult susah [su.sah] *adj*

dip in dye chelop [tʃə.lɔp] *v*
dirty kotor [ko.tor] *adj*
distended stomach bunchit [bun.tʃit] *n*
disturb kacho [ka.tʃo] *v*
divination stick chiam [tʃjam] *n*
do bikin 2 [bi.kin] *v*; buat 1 [bwat] *v* (comp. buat bodoh
comp. buat mahal comp. buat malu comp. buat
mungka comp. buat suay comp. buat tak dengair)
do not jangan [dʒa.ŋan] *adv*
do not need toksa [toʔ.sa] *v*
dog anjing [an.dʒing] *n*
door pintu [pin.tu] *n*
double-boil tim [tim] *v*
drain longkang [loŋ.kan] *n*
draw chonténg [tʃon.teŋ] *v*
dream mimpi [mim.pi] *v*
dried fruit gugol [gu.gol] *n*
drink minum [mi.num] *v*
drip berléliléi [bər.lej.lej] *v*
drop jatoh 2 [dʒa.toh] *v*
dry kering 2 [kə.riŋ] *adj* (comp. mulot kering)
dry in the sun jumoh [dʒu.moh] *v*
duck iték [i.teʔ] *n*
durian flesh uluair 2 [u.lɛ] (var. of ular) *clf*
dusk sinjakala [sin.dʒa.ka.la] (comp. of kala) *n*
Dutch belanda [bə.lan.da] (comp. kuching belanda)

E

ear kuping [ku.piŋ] <i>n</i>	ever pernah [pər.nah] <i>adv</i>
early siang [sjaŋ] <i>adv</i>	every tiap [tjap] <i>adv</i>
earrings anting-anting [an.tiŋ-an.tiŋ] <i>n</i>	evil jahat [dʒa.hat] <i>adj</i> (comp. mulot jahat)
earth tanah 3 [ta.nah] <i>n</i>	evil (used in particular for describing stepmothers) tork 2 [tɔʔ] <i>adj</i>
easy senang 1 [sə.naŋ] <i>adj</i>	EXCLAM annoyance aiya [aj.ja]
eat makan 1 [ma.kan] <i>v</i>	EXCLAM cry to a dead spirit kus semangat [kus sə.ma.ŋat]
eat (vulgar) cherkek [tʃər.keʔ] <i>v</i>	EXCLAM cursed one chilaka [tʃi.la.ka]
eat all (mahjong term) tongchit [toŋ.tʃit] <i>v</i>	EXCLAM dead mampus [mam.pus]
egg teloh [tə.loh] <i>n</i>	EXCLAM dismay alamak [a.la.maʔ]
eight lapan [la.pan] <i>cardnum</i> ; puay [pwe] <i>cardnum</i> (comp. pékji)	EXCLAM irritation aiyo [aj.jo]
eight characters of Chinese horoscope pékji [peʔ.dʒi] (comp. of puay) <i>n</i>	EXCLAM jibe éh [eh]
elder brother hia [hja] <i>n</i>	EXCLAM regret ala [a.la]
elder brother (non-familial) ko 2 [ko] <i>n</i>	EXCLAM satisfaction shiok 2 [ʃjoʔ]
elder brother's wife so 2 [so] <i>n</i>	EXCLAM surprise amboi [am.boj]; ayi [a.ji]
elder sister tachi [ta.tʃi] <i>n</i>	excrement tait [tajt] <i>n</i>
elder sister's husband chau [tʃaw] <i>n</i>	exhausted letay [lə.te] <i>adj</i>
elderly Malay person wak [waʔ] <i>n</i>	EXIST ada 3 [a.da] (var. a) <i>v</i>
eleven (coarse) sebelas [sə.bə.las] <i>cardnum</i>	exorcise buang buang [bwaŋ bwaŋ] (comp. of buang) <i>v</i>
eleven (refined) sebelair [sə.bə.lɛ] <i>cardnum</i>	expand huat 1 [hwat] <i>v</i>
embarrassed malu [ma.lu] <i>adj</i> (comp. buat malu); paiséh [paj.seh] <i>adj</i>	expect sangka [saŋ.ka] <i>v</i>
EMP lah [lah] <i>prt</i> ; lor [lor] <i>prt</i>	expensive mahal [ma.hal] <i>adj</i> (comp. buat mahal)
enough chukop [tʃu.kop] <i>adv</i>	eye mata [ma.ta] <i>n</i> (comp. bijak mata comp. buang mata comp. mata ari mata ayé mata beliak comp. mata gelap comp. mata ikan comp. mata juling comp. mata kuching comp. mata lembu comp. mata piso comp. tanda mata)
enter masok 1 [ma.soʔ] <i>v</i>	
enthusiasm gerang [gə.raŋ] <i>n</i>	
estimate agak [a.gaʔ] <i>v</i> (comp. agak agak)	
evening petang [pə.taŋ] <i>n</i>	

F

- face mungka** [muŋ.ka] *n* (comp. buat mungka)
faint péngsan [peŋ.san] *v*; sayup [sa.jup] *adj*
fall jatuh 1 [dʒa.toh] *v*
fall backwards lentang [lən.taŋ] *v*
fall on one's front tiarap [tja.rap] *v*
family chepuat [tʃə.pwat] *n*
fan kipas [ki.pas] *n*
far ja.ou [dʒa.ow] *adj*
faster version of jogét ronggéng [roŋ.geŋ] *n*
fat gemok [gə.moʔ] *adj*
father baba [ba.ba] (var. of bapak) *n*; bapak [ba.paʔ] *n*
 (comp. mak bapak)
father-in-law kua 2 [kwa] *n*
father-in-law (indirect address) chingkey 2 [tʃiŋ.ke] *n*
father's elder brother pék [peʔ] *n*
father's elder brother's wife em 1 [əm] *n* (comp. sangkék-em)
father's eldest brother pék-pék [peʔ-peʔ]
father's eldest sister mak-ko [maʔ-ko] (comp. of mak, ko 1) *n*
father's sister ko 1 [ko] *n* (comp. mak-ko)
father's younger brother chék [tʃeʔ] *n*
father's younger brother's wife chim [tʃim] *n*
favourite child bijak mata [bi.dʒiʔ ma.ta] (comp. of bijak, mata) *n*
feather bulu 1 [bu.lu] *n*
feed suap [swap] *v*
feel rasa 1 [ra.sa] *v*
feign ignorance buat bodoh [bwat bo.doh] (comp. of buat, bodoh) *v*
female perompuan [pə.rom.pwan] *n*
female mourner hauli [haw.li] *n*
fermented rice dessert tapay [ta.pe] *n*
fierce bengis 1 [bə.ŋis] *adj*; garang [ga.raŋ] *adj*
filter tapis [ta.pis] *v*
find charék [tʃa.reʔ] *v*
fine hair bulu 2 [bu.lu] *n*
finger jari [dʒa.ri] *n*
finger nail kuku [ku.ku] *n*
finish abi [a.bi] (var. of habis) *conn*; abis [a.bis] (var. of habis) *v*; habis [ha.bis] 1 *v*; habis 2 [ha.bis] *adv*
fire api [a.pi] *n*
first dulu 3 [du.lu] *adv*; pertama [pəɾ.ta.ma] *adj*
first cream of coconut pati [pa.ti] *n*
fish ikan [i.kan] *n* comp. mata ikan
fishball huyi [hu.ji] *n*
five gor [gɔ] *cardnum* id. ati it gor it chap; lima [li.ma] *cardnum*
flat bean petay [pə.te] *n*
flesh isi [i.si] *n*
float timbol [tim.bol] *v*
floor lantay [lan.te] *n*
flower bunga [bu.ŋa] *n*
fluent lanchang [lan.tʃaŋ] *adj*
flush with water siram [si.ram] *v*
fly terbang [təɾ.baŋ] (var. trebang) *v*; trebang [trə.baŋ] *v*
fold lipat [li.pat] *v*
follow ikot [i.kot] *v*
food makan 2 [ma.kan] *n*
food in the shape of a ball yi 1 [ji] *n* (comp. kuéh-yi)
force paksa [paʔ.sa] *v*
for untok [un.toʔ] *adv*
forehead dahi [da.hi] (var. dai) *n*; dai [daj] *n*
forget lupa [lu.pa] *v*
forgive ampun 1 [am.pun] *v*
forgiveness ampun 2 [am.pun] *n*
four empat [əm.pat] *cardnum*; si 2 [si] *cardnum*

four colours cardgame sisék [si.seʔ] *n*
fragrant wangi [wa.ŋi] *adj*
freckles opan [o.pan] *n*
free senang 2 [sə.naŋ] *adj*
fried egg with yolk intact mata lembu [ma.ta læm.bu]
 (comp. of mata, lembu) *n*
friend kaki 2 [ka.ki] *n*; kawan [ka.wan] *n*
frightening seram [sə.ram]
frog kodok [ko.doʔ] *n*

from dari [da.ri] *prep*
front depan [də.pan] *prep*
fruit buah [bwah] *n*
fry goréng [go.reŋ] *v*
full penoh [pə.noh] *adj*
full from eating kenyang [kə.ŋaŋ] *adj*
fumble kalang-kabot [ka.laŋ-ka.bot] *v*
furious geram [gə.ram] *adj* (comp. naik geram)
fussy kéwat [ke.wat] *adj*

G

gamble judi [dʒu.di] *v*
garden kebun [kə.bun] *n*; taman [ta.man] *n*
generous tuakang [twa.kaŋ] *adj*
get dapat 1 [da.pat] *v*
giddy mabok [ma.boʔ] *adj*
give kasi 1 [ka.si] *v*
glare beliak [bə.ljaʔ] *v* (comp. mata beliak)
glutinous rice pulot [pu.lot] *n*
glutinous rice balls kuéh-yi [kweh-ji] (comp. of kuéh, yi
 1) *n*
go pegi [pə.gi] *v* contr. pi; pi [pi] *v*
go out (coarse) keluar [kə.lwar] (var. kular) (comp.
 diluar); kular [ku.lwar] *v*
go out (refined) kuluar [ku.lwɛ] *v*
goat kambing 1 [kam.biŋ] *n*
gold (coarse) mas [mas] *n*
gold (refined) mair [mɛ] *n*
good bagus [ba.gus] *interj*; baik [bajʔ] *adj* (comp. ati
 baik, comp. naik baik); ho [ho] *adj* (comp. homia)
good character pé'él [peʔe] *n*
good life homia [ho.mja] *interj* (comp. of ho, mia)
grand-aunt po [po] *n*
grandchild chuchu [tʃu.tʃu] *n*
grandfather gong 2 [goŋ] (comp. gong-cho)

grandfather's younger brother chék-gong [tʃeʔ-goŋ] *n*
grandmother mamak [ma.maʔ] *n*
grandparent gong 1 [goŋ] *n* comp. gong-cho
grass rumput [rum.pot] *n*
gravy kua 1 [kwa] *n*
grease minyak 1 [mi.ŋak] *n*
great-grandaunt cho-po [tʃo-po] *n*
great-grandchild chichi [tʃi.tʃi] *n*
great-grandfather cho-gong [tʃo-goŋ] *n*
great-grandmother cho 2 [tʃo] *n* (comp. gong-cho, mak-
 cho); cho-cho [tʃo-tʃo] *n*; mak-cho [maʔ-tʃo] (comp. of
 mak, cho) *n*
great-grandparent cho 1 [tʃo] *n* (comp. gong-cho mak-
 cho)
great-great grandchild onéng.onéng [o.neŋ.o.neŋ] *n*
greed (vulgar) radat [ra.dat] *adj*
green hijo [hi.dʒo] (var. ijo) *adj*; ijo [i.dʒo] *adj*
green bean cookies koya [ko.ja] *n*
groom kiasai 2 [kjasaj] *n*
groom's ceremonial assistant pak-chindék [paʔ-tʃin.deʔ]
n
groom's boy assistant kuya [ku.ja] *n*
ground tanah 2 [ta.nah] *n*
guard jaga 1 [dʒa.ga] *v*
guess gasak [ga.saʔ] *v*

guest langkék [laŋ.keʔ] *n*

H

HAB ada 6 [a.da] (var. a) *v*

hair rambot 1 [ram.bot] *n*

half-cooked (coarse) embar-embar [əmbaɾ-əmbaɾ] *adj*

half-cooked (refined) embair-embair [əmbɛ-əmbɛ] *adj*

hallway rambat [ram.bat] *n*

hand tangan [ta.ŋan] *n*

handkerchief binpo [bin.po] *n*

handsome tuaban [twa.ban] *adj*

hang gantong [gan.ton] *v*

happen jadi 1 [dʒa.di] *v*

happy huahi [hwa.hi] *adj*

hard (coarse) keras [kəras] *adj*

hard (refined) kerair [kəɾɛ] *adj*

hat kopiah [ko.pjah] *n*

have a [a] (var. of ada) *v*; ada 1 [a.da] (var. a) *v*

hay kering 1 [kə.riŋ] *n* comp. mulot kering

head kepala [kə.pa.la] *n*

healthy sihat [si.hat] *adj*

heart ati 1 [a.ti] (var. of hati) *n* (comp. ati baik, comp. ati busok, comp. jantung ati (id. ati it gor it chap); hati 1 [ha.ti] *n*; jantung [dʒan.ton] *n* (comp. jantung ati)

heavy berat [bə.rat] *adj* (comp. mulot berat); **lebat** [lə.bat] *adj*

heirloom tanda mata [tan.da ma.ta] (comp. of tanda, mata) *n*

help tolong [to.lon] *v*

here sini [si.ni] *adv*

here you go na [na] *interj*

hide semonyet [sə.mo.ɲet] *v*

hill bukit [bu.kit] *n*

hit pukol 1 [pu.kol] *v*

hitch (a ride) tompang [tom.paŋ] *v*

hold angkat 2 [aŋ.kat] *v*; pégang [pe.gaŋ] *v*

hole lobang [lo.baŋ] *n*

honorific prefix for familial relations eng- [əŋ]

hope arap [a.rap] (var. of harap); harap [ha.rap] *v*

horn tandok [tan.doʔ] *n*

horse kuda [ku.da] *n*

hot (coarse) panas [pa.nas] *adj*

hot (refined) panair [pa.nɛ] *adj* hot

hot pot for lunar new year reunion dinner wilo [wi.lo] *n*

hour jam 2 [dʒam] *n*

house rumah [ru.mah] *n*

how (coarse) amcham [am.tʃam] (var. of apa cham [a.pa.tʃam]; apa macham [a.pa ma.tʃam], comp. of apa, macham) *interrog*

how (refined) apa pasair [a.pa pa.sɛ] *interrog*

how many berapa 1 [bə.ra.pa] *interrog*

howl menyalap [mə.pa.lap] *v*

hundred ratus [ra.tus] *cardnum*

hungry (coarse) lapar [la.paɾ] *adj*

hungry (refined) lapair [la.pɛ] *adj*

husband laki [la.ki] (contr. of lelaki) *n*

I

if kalo [ka.lo] (var. of kalu) *conn*; kalu [ka.lu] *conn*

ill-treat koktok [kɔʔ.tɔʔ] *v*

important *yaukin* [jau.kin] *adj*
inability to express oneself *mulot berat* [mu.lot bæ.rat]
(comp. of *mulot*, *berat*) *n*
inability to keep a secret *mulot bocho* [mu.lot bo.tʃo]
(comp. of *mulot*, *bocho*) *n*
inability to say more *mulot kering* [mu.lot kə.rin] (comp.
of *mulot*, *kering*) *n*
indeed *mémang* [me.man] *adv*; *sekali 2* [sə.ka.li] *adv*
indeed (refined) (unsystematic variant of *sekali*) *sekair*
[sə.kɛ] *adv*

inside *dalam* [da.lam] *prep*
instead *pulak* [pu.laʔ] *adv* instead instead
instruct *pésan* [pe.san] *v* instruct instruct
intercalary *lun* [lun] *adj* intercalary intercalary
intercalary month *lun-guék* [lun gweʔ] *n* (comp. of *lun*,
guék)
iron *besi* [bə.si] *n* iron iron
itchy (coarse) *gatal* [ga.tal] *adj* itchy
itchy (refined) *gatair* [ga.te] *adj* (comp. *mulot gatair*)

J

jelly *chaiyen* [tʃaj.jen] *n*
joke *kelaka* [kə.la.ka] *v*
jump *lompat* [lom.pat] *v*

jungle *hutan* [hu.tan] (var. *utan*) *n*; *utan* [u.tan] *n*
just *baru 2* [ba.ru] *adv*; *tadi* [ta.di] *adv*

K

keep an eye on someone or something *buang mata* [bwaŋ
ma.ta] (comp. of *buang*, *mata*) *v*
keep something safe *simpan* [sim.pan] *v*
key *konchi* [kon.tʃi] *n*
kick *tendang* [tən.daŋ] *v*
kill *bunuh* [bu.noh] *v*
kind *ati baik* [a.ti bajʔ] (comp. of *ati*, *baik*) *adj*
king *raja* [ra.dʒa] *n*

kitchen *dapor* [da.pɔ] *n*
kneel *gui* [gwi] *v*
knife *piso* [pi.so] *n* (comp. *mata piso*)
knock *ketok* [kə.tɔʔ] *v*
know *tau* [taw] *v*
know a person (coarse) *kenal* [kə.nal] *v*
know a person (refined) *kenair* [kə.nɛ] *v*

L

laborer *tukang* [tu.kan] *n*
ladder *tangga 1* [taŋ.ga] *n*
lamp *lampu* [lam.pu] *n*
language *bahasa* [ba.ha.sa] *n*

late *lambat* [lam.bat] *adj*
later *nanti* [nan.ti] *adv*
laugh *ketawa* [kə.ta.wa] *v*
lazy (coarse) *malas* [ma.las] *adj*

lazy (refined) malair [ma.lɛ] *adj*
leaky bocho [bo.tʃo] *adj* (comp. mulot bocho)
lean sandah [san.dah] *v*
learn (coarse) belajar [bə.la.dʒar] *v*
learn (refined) belajair [bə.la.dʒɛ] *v*
leaves daoun [da.own] *n*
leg kaki 1 [ka.ki] *n*
lemon lémo [le.mo] *n*
lend pinjam [pin.dʒam] *v*
less kurang [ku.raŋ] *adv*
let kasi 2 [ka.si] *v*
letters surat [su.rat] *n*
lie bohong [bo.hoŋ] *n*
lie down baring [ba.riŋ] *v*
life mia 1 [mja] *n*
lift angkat 1 [aŋ.kat] *v*
like suka [su.ka] *v*
like cham 2 [tʃam] (contr. of macham) *adv*; macham 1 [ma.tʃam] *adv* (comp. apa cham, comp. apa macham)
like that macham 3 [ma.tʃam] *adv* (comp. apa cham comp. apa macham)
like that begi [bə.gi] (contr. of begitu), gitu [gi.tu] (contr.

of begitu) *adv*; begitu [bə.gi.tu] *adv*,
listen (coarse) dengar [də.ŋar] *v*
listen (refined) dengair [dəŋ.ɛ] *v* (comp. buat tak dengair)
little sikit [si.kit] *adv*
live (coarse) tinggal 1 [tiŋ.gal] *v*
live (refined) tinggair 1 [tiŋ.gɛ] *v*
lively riyang [ri.jaŋ] *adj*
liver ati 2 [a.ti] (var. of hati) *n* (comp. ati baik, comp. ati busok, id. ati it gor it chap, comp. jantung ati); hati 2 [ha.ti] *n*
long panjang [pan.dʒaŋ] *adj*
long ago dulu 2 [du.lu] *adv*
long and thin batang [ba.taŋ] *clf*
long time lama [la.ma] *adj*
longan fruit mata kuching [ma.ta ku.tʃiŋ] (comp. of mata, kuching) *n*
long-jawed chobék [tʃo.beʔ] *adj*
lose hilang [hi.laŋ]; ilang [i.laŋ] *v*
lose (opposed to win) kalah [ka.lah] *v*
low rendah [rən.dah] *adj*
lunar month guék [gweʔ] *n*

M

mad gila [gi.la] *adj* (comp. naik gila)
make bikin 1 [bi.kin] *v*; buat 2 [bwat] *v* (comp. buat bodoh, comp. buat mahal, comp. buat malu, comp. buat mungka, comp. buat suay, comp. buat tak dengair)
male jantan [dʒan.tan] *n*
male mourner haulam [haw.lam] *n*
male person pak [paʔ] *n*
many banyak [ma.naʔ] *adv*
market (coarse) pasar [pa.sar] *n*
market (refined) pasair 1 [pa.sɛ] *n*

marry kawin [ka.win] *v*
match of love jodoh [dʒo.doh] *n*
matchmaker muilang [mwi.laŋ] *n*
matter perkara [pə.ka.ra] *n*; sair 1 [sɛ] (contr. of pasair 2) *n* (comp. apa sair)
matter (coarse) pasal 1 [pa.sal] *n* (comp. apa pasal)
matter (refined) pasair 2 [pa.sɛ] *n* (contr. sair 1)
meat daging [da.gin] *n*
medium sedang [sə.daŋ] *adj*
meet jumpa [dʒum.pa] *v*
meet ill spiritual forces huantiok [hwan.tjoʔ] *v*

melody dondang [don.daŋ] *n*
mend (coarse) tampal [tam.pal] *v*
middle tengah [tə.ŋah] *adv* (comp tengah jalan)
milk susu [su.su] *n*
mince chinchang [tʃin.tʃaŋ] *v*
minute minit [mi.nit] *n*
miserly kiamsiap [kjam.sjap] *adj*
miss rindu [rin.du] *v*
mistake salah 1 [sa.lah] *n*
mix champor [tʃam.por] *v*; gaou [ga.ow] *v*
mixed herbs rice with shredded anchovies ulam [u.lam] *n*
money dit [dit] (var. of duit) *n*; duit [dwit] *n*
monk huésio [hwe.sjo] *n*
month bulan 2 [bu.lan] *n*
moon bulan 1 [bu.lan] *n*
more lagik 1 [la.giʔ] *adv*; lebéh [lə.beh] *adv*
morning pagi [pa.gi] *n*
mother mak [maʔ] *n* (comp. mak bapak, mak-cho, mak-ko, mak-yi); nya 1 [ŋa] (contr. of Nyonya) *n*

mother-in-law nio [ŋjo] *n*
mother-in-law (indirect address) chay-em [tʃe.əm] *n*
mother's brother ku 2 [ku] *n* (comp. ku-ku)
mother's brother's wife kim [kim] *n*
mother's eldest brother ku-ku [ku-ku] (comp. of ku 2) *n*
mother's eldest sister yi-yi [ji-ji] *n*
mother's sister yi 2 [ji] *n* (comp. mak-yi)
mountain gunung [gu.noŋ] *n*
mouth mulut [mu.lot] *n* (comp. buang mulut, comp. mulut asin, comp. mulut béngok, comp. mulut berat, comp. mulut bocho, comp. mulut busok, comp. mulut dunya, comp. mulut gatair, comp. mulut jahat, comp. mulut kering, comp. mulut manis, comp. mulut pantat ayam, comp. mulut ringan)
mouth agape nganga [ŋa.ŋa] *adj*
move bergerak [bər.gə.raʔ] *v*
movement ter- 2 [tər] (comp. terperanjat)
must mesti [məs.ti] *v*

N

nail paku [pa.ku] *n*
name nama [na.ma] *n*
near dekat 2 [də.kat] (var. kat) *prep*
neck léihéi [lej.hej] *n*
NEG kan [kan] (contr of bukan 1) *interj*; tak [taʔ] *adv* (comp. buat tak dengair)
NEG (noun) bukan 2 [bu.kan]
nervous ati it gor it chap [a.ti it gɔ it tʃap] (id. of ati, it, gor, chap) *adj*
new baru 1 [ba.ru] *adj*
newcomer from China sinkék [sin.keʔ] *n*
news (coarse) kabar [ka.bar] *n*
news (refined) kabair [ka.bɛ] *n*

night malam [ma.lam] *n*
night (when referring to lunar dates) méh [meh] *n*
nine gau [gaw] *cardnum*; semilan [sə.mi.lan] *cardnum*
no bukan 1 [bu.kan] (var. kan) *interj*
noisy bising [bi.siŋ] *adj*
NOM -an [-an]
nose hidong [hi.doŋ] (var. idong) *n*; idong [i.doŋ] *n*
not fussy chinchai [tʃin.tʃaj] *adj*
not yet belom 1 [bə.lom] *adv*
now sekarang [sə.ka.raŋ] *adv*
number nombor [num.boʔ] *n*
nunnery that serves vegetarian food chaiteng [tʃaj.təŋ] *n*

O

observe cham 1 [tʃam] *v*
obstinate ngéngé [ŋe.ŋe] *adj*
odd-numbered ganjil [gan.jil] *adj*
offering of three meats samséng [sam.seŋ] *n*
oil minyak 2 [mi.nak] *n*
old tua 1 [twa] *adj*
old aunty em 2 [əm] *n* (comp. sangkék-em)
older Peranakan woman bibik [bi.biʔ] *n*
omelette (coarse) dadar [da.dar] *n*
omelette (refined) dadair [da.dɛ] *n*
one it [it] *cardnum* (id. ati it gor it chap); satu [sa.tu]
cardnum (contr. se [sə])
onion bawang [ba.waŋ] *n*

only sahja [sa.dʒa] *adv*
onomatopoeia loud noises gedebang-gedebong
 [gə.də.baŋ-gə.də.boŋ]
onomatopoeia thudding of the heart gedebak-gedebuk
 [gə.də.baʔ-gə.də.buʔ]
open bukak [bu.kaʔ] *v*
or ka [ka] *conn*
original mula [mu.la] *adj*
originally mula mula [mu.la mu.la] *adv*
other lain [lajn] *adv*
outside diluar [di.lwar] (var. dular) (comp. of di, keluar)
prep; dular [du.lwar] (var. of diluar) *prep*
overripe magam [ma.gam] *adj*

P

pail baldi [bal.di] *n*
pair pasang 2 [pa.saŋ] *clf*
palm blossom mayang [ma.jaŋ] *n*
Pangium.edule (black fruit with hard shell) keluak
 [kə.lwaʔ] *n*
pants (coarse) suluar [su.lwar] (var. seluair) *n*
pants (coarse) seluar [sə.lwar] *n*
pants (refined) suluar [su.lwɛ] *n*
papaya paya [pa.ja] *n*
paper (coarse) kertas [kər.tas] *n*
paper (refined) kertair [kər.tɛ] *n*
parent-in-law chingkay 1 [tʃiŋ.ke] *n*
parents mak bapak [maʔ ba.paʔ] (comp. of mak, bapak)
n
parents' eldest sister's husband tio-tio 1 [tjo-tjo] *n*
parent's sister's husband tio [tjo] *n*
park letak [lə.taʔ] *v*

pass lalu [la.lu] *v*
PASS kena 1 [kə.na] *v*; kasi 4 [ka.si] *v*
paternal cousins kopiau [ko.pjaw] *n*
pay (coarse) bayar [ba.jar] *v*
pay (refined) bayair [ba.jɛ] *v*;
peel kopék [ko.peʔ] *v*; kupas [ku.pas] *v*
people orang 1 [o.raŋ] *n*
Peranakan female Nyonya [ŋo.ŋa] *n* (contr. nya)
Peranakan male Baba 1 [ba.ba] *n* (contr. ba)
Peranakan card game cherki [tʃɛr.ki] *n*
perform main 2 [majn] *v*
perpetually siap-siap [sjap-sjap] *adv*
person orang 2 [o.raŋ] *n*
PERSON si 1 [si] *dem*
PFV ada 4 [a.da] (var. a) *v*
pick up angkat 3 [aŋ.kat] *v*; punggot [puŋ.got] *v*

picture (coarse) gambar [gam.bar] *n*
picture (refined) gambair [gam.be] *n*
piece kepéng [kə.peŋ] *clf*
piece of fabric or paper lay [le] (var. of helay) *clf*
pig babi [ba.bi] *n*
pinch kepék [kə.peʔ] *v*
pity sayang 3 [sa.jaŋ] *interj*
place tempat [təm.pat] *n*
plant tanam [ta.nam] *v*
plate pinggan [piŋ.gan] *n*
play main 1 [majn] *v*; wayang [wa.jaŋ] *n*
play hard to get buat mahal [bwat ma.hal] (comp. of
 buat, mahal) *v*
pluck peték [pə.teʔ] *v*
pocket kochék [ko.tʃeʔ] *n*
poisonous tork 1 [tɔʔ] *adj*
polite perama [pə.ra.ma] *adj*
polite character mulot ringan [mu.lot ri.ŋan] (comp. of
 mulot, ringan) *n*
pond laut 2 [lawt] *n*
poor miskin [mis.kin] *adj*
porridge bubor [bu.bɔ] *n*
POSS mia 2 [mja] (contr. of punya) *v*; nia 1 [nja] *v*; punya
 1 [pu.ja] *v*
pound tumbok [tum.boʔ] *v*
pour and serve tuang [twaŋ] *v*

prawn udang [u.daŋ] *n*
pray semayang [sə.ma.jaŋ] *v*
PREP dekat 3 [də.kat] (var. kat) *prep*; di [di] *prep* (comp.
 diluar); kat [kat] *prep*
prepared siap [sjap] *v*
preserved vegetables kiamchai [kjam.chaj] *n*
pretend to not hear buat tak dengair [bwat taʔ dəŋ.ɛ]
 (comp. of buat, tak, dengair) *v*
price arga [ar.ga] (var. of harga) *n*; harga [har.ga] *n*
priest paderi [pa.də.ri] *n*
PROG ada 2 [a.da] (var. a) *v*
promiscuous samplang [sam.plaŋ] *adj*
proper kenonay [kə.no.ne] *adj*; seronoh [sə.ro.noh] *adj*
prosper huat 2 [hwat] *v*
protruding eyes mata beliak [ma.ta bæ.ljaʔ] (mata,
 beliak) *n*
public opinions mulot dunya [mu.lot du.ɲia] (comp. of
 mulot, dunya) *n*
pull tarék [ta.reʔ] *v*
pull out chabot [tʃa.bot] *v*
pummel bedék [bə.deʔ] *v*
pure jati 2 [dʒa.ti] *adj*
purse horpau [hor.paw] (var. orpau) *n*; orpau [or.paw] *n*
push tolak [to.laʔ] *v*
put tarok [ta.roʔ] *v*
put in masok 2 [ma.soʔ] *v*

Q

quarrel berkelay [bər.kə.le] *v*
quick chepat [tʃə.pat] *adj*
quick (coarse) lekas [lə.kas] *adj*

quick (refined) lekair [lə.kɛ] *adj*
quiet diam [djam] *adj*
quietly diam diam [djam djam] *adv*

R

rabbit kucing belanda [ku.tʃiŋ bələnda] (comp. of kucing, belanda) *n*
race lawan [la.wan] *v*
rain hujan [hu.dʒan] (var. ujan) *n*; ujan [u.dʒan] *n*
raise children piara [pja.ra] *v*
rank pangkat [paŋ.kat] *n* (comp. naik pangkat, pangkat adek beradek)
rational (coarse) benal [bə.nal] *adj*
rational (refined) benair [bə.nɛ] *adj*
raw mentah [mən.tah] *adj*
reach sampai 2 [sam.pe] *v*
read bacah [ba.tʃa] *v*
really betul 3 [bə.tol] *adv*; sungguh [suŋ.gu] *adv*
reason (refined) pasair 2 [pa.sɛ] *n* (contr. sair 1); sair 1 [sɛ] (contr. of pasair) *n* (comp. apa sair)
reason (coarse) pasal 2 [pa.sal] *n* (comp. apa pasal)
receive dapat 2 [da.pat] *v*
recognize jin [dʒin] *v*
red mérah [me.rah] *adj*
red packet of monetary gift angpau [aŋ.paw] *n*
REL mia 2 [mja] (contr. of punya), nia 2 [nja] (contr. of punya) *rel*, punya 2 [pu.nja] *rel*; nang [naŋ] (var. of yang) *rel*; yang [jaŋ] *rel*
remember ingat 1 [i.ŋat] *v*
reminisce kenang [kə.naŋ] *v*
repent beruba [bə.ru.ba] *v*
respect hormat [hor.mat] (var. ormat) *v*; ormat [or.mat] *v*

resting platform balay-balay [ba.le-ba.le] *n*
return balék 1 [ba.leʔ] *v*; pulang [pu.laŋ] *v*
reverse gostan [go.stan] *v*
rice cake chuikuéh [tʃwi.kweh] *n*
rice cake served in spicy gravy lontong [lon.ton] *n*
rich kaya 1 [ka.ja] *adj*
ride tunggang [tuŋ.gaŋ] *v*
right kanan [ka.nan] *adv*
ripe masak 2 [ma.saʔ] *adj*
rise naik 2 [najʔ] *v* (comp. naik baik, comp. naik darah, comp. naik geléték, comp. naik geram, comp. naik gila, comp. naik lemak, comp. naik pangkat, comp. naik sedap, comp. naik seram)
river sunggay [suŋ.ge] *n*
road tengah jalan [tə.ŋah dʒa.lan] (comp. of tengah, jalan) *n*
roast panggang [paŋ.gaŋ] *v*
rock batu 1 [ba.tu] *n*
rock goyang 1 [go.jaŋ] *v*
roof bunbong [bun.boŋ] *n*
room bilék [bi.leʔ] *n*
root (coarse) akar [a.kar] *n*
root (refined) akair [a.kɛ] *n*
rouge yénchi [jen.tʃi] *n*
roughly agak agak [a.gaʔ a.gaʔ] (agak) *adv*
round bulat [bu.lat] *adj*
run lari [la.ri] *v*

S

sad sedéh [sə.deh] *adj*
safe selamat [sə.la.mat] *adj*

saliva lio [ljo] *n*
salt garam [ga.ram] *n*

salty asin [a.sin] *adj* der. asién asin comp. mulot asin
same sama 1 [sa.ma] *adj*
sand paser [pa.ser] *n*
satisfied kamwan [kam.wan] *adj*
satisfied feeling shiok 1 [ʃjoʔ] *adj*
saute spices tumis [tu.mis] *v*
say kata [ka.ta] *v*
school sekolah [sə.ko.lah] *n*
scold maki [ma.ki] *v*
scratch gerong [gə.rɔŋ] *v*
sea laut 1 [lawt] *n*
season musim [mu.sim] *n*
second father tio-tio 2 [tjo-tjo] *n*
second mother mak-yi [maʔ-ji] (comp. of mak, yi 2) *n*
see nampak [nam.paʔ] *v*; téngok [te.ŋoʔ] *v*
seed bijik 2 [bi.dʒiʔ] *n* (comp. bijik mata)
seems macham 2 [ma.tʃam] *adv*; amcham (comp. of apa
 cham), (comp. apa macham, contr. cham 2)
seldom jarang [dʒa.raŋ] *adv*
self sendiri [sən.di.ri] *n*
selfish auban [aw.ban] *adj*
sell (coarse) jual [dʒwal] *v*
sell (refined) juair [dʒwɛ] *v*
serious bengis 2 [bə.ŋis] *adj*
seven chit [tʃit] *cardnum*; tujuh [tu.dʒoh] *cardnum*
sew jait [dʒajt] *v*
shadow bayang [ba.jaŋ] *n*
shake gonchang [gon.tʃaŋ] *v*; goyang 2 [go.jaŋ] *v*
sheep kambing 2 [kam.biŋ] *n*
sheet éla [e.la] (var. of helay) *clf*; lay [le] (var. of helay);
 helay [hə.le] *clf*
ship (coarse) kapal [ka.pal] *n*
ship (refined) kapair [ka.pɛ] *n* ship
shiver mengintar [mə.ŋin.tar] *v*
shock kejut [kə.dʒut] *adj*
shoe kasot [ka.sot] *n*
shop keday [kə.de] *n*

short katék [ka.teʔ] *adj*; péndék [pen.deʔ] *adj*
shoulder bahu [ba.hu] (var. bau) *n* shoulder shoulder; bau
 [baw] *n*
shout pekék [pə.keʔ] *v*
show tunjok [tun.dʒoʔ] *v*
shy segan [sə.gan] *adj* comp. buang segan
sibling adék 1 [a.deʔ] *n* (der. adék beradék)
siblings adék beradék 1 [a.deʔ bə.ra.deʔ] (der. of adék)
n (comp. pangkat adek beradek)
sick sakit [sa.kit] *adj*
side belah [bə.lah] *prep*; tepi 1 [tə.pi] *n*; tepi 2 [tə.pi]
adv
sign tanda [tan.da] *n* (comp. tanda mata)
silent senyap [səŋap] *adj*
similar to bagi [ba.gi] *adv*
sing menyanyi [mə.ŋa.ŋi] (var. nyanyi) *v*; nyanyi [ŋa.ŋi]
v
singe siut [sjut] *v*
sink tenggelam [təŋ.gə.lam] *v*
sit dudok 1 [du.doʔ] *v*
six enam [e.nam] *cardnum* six six; lak [laʔ] *cardnum*
skin kulit [ku.lit] *n*
skirt kun 1 [kun] *n*
sky langit [la.ŋit] *n*
sky-deity (Jade Emperor) ti-gong [ti-gon] *n*
sleep tidor [ti.dor] *v*
sleepy mengantok 2 [me.ŋan.toʔ] *v*
slice iris [i.ris] *v*
slim ringan [ri.ŋan] *adj* (comp. mulot ringan)
slippers terompak [tə.rom.paʔ] *n*
slow pelan [pə.lan] *adj* (comp. pelan pelan)
slowly pelan pelan [pə.lan pə.lan] (der. of pelan) *adv*
small kechik [kə.tʃiʔ] *adj* (comp. buang ayé kechik)
small and round bijik 1 [bi.dʒiʔ] *clf* (comp. bijik mata)
small clock aloji 1 [a.lo.dʒi] *n*
smashed up nua [nwa] *adj*
smell ba'u [ba.ʔu] *n*
smelly busok [bu.soʔ] *adj* (comp. ati busok, comp. mulot

busok)
smile senyum 1 [sə.ɲum] *v*; senyum 2 [sə.ɲum] *n*
smoke asap [a.sap] *n*
smooth lichin [li.tʃin] *adj*
snake (coarse) ular [u.lar] *n* snake
snake (refined) uluair 1 [u.ɛ] *n*
snap patah [pa.tah] *v*
socks boik [bojk] *n*
soft lembéh [ləm.bəh] *adj*
soil tanah 1 [ta.nah] *n*
some berapa 2 [bə.ra.pa] *adv*
somewhat salty asién asin [a.sjen asin] (der. of asin) *adj*
son ba [ba] (contr. of Baba) *n*; Baba 2 [ba.ba] *n*
song lagu [la.gu] *n*
son-in-law kiasai 1 [kjasaj] *n*
sorrow pilu [pi.lu] *n*
sound bunyi [bu.ɲi] *n*
soup sup [sup] *n*
sour asam 2 [a.sam] *adj*
sour plums sengbuay [səŋ.bwe] *n*
speak chakap [tʃaj.kap] *v*
spend belanja 1 [bə.lan.dʒa] *v*
spice paste (for papaya soup) titék [ti.teʔ] *n*
splurge gaya [ga.ja] *v*
spoilt rosak [ro.saʔ] *adj*
spoon séndok [sen.doʔ] *n*
spring roll with turnip popiah [po.pjah] *n*
squinty juling [dʒu.liŋ] *adj* (comp. mata juling)
stab tikam 1 [ti.kam] *v*
stage panggong [paŋ.gon] *n*
stairs tangga 2 [taŋ.ga] *n*
stale basi [ba.si] *adj*
stand diri [di.ri] *v*
standard baku [ba.ku] *adj*
star bintang [bin.taŋ] *n*
stay dudok 2 [du.doʔ] *v*
stay (coarse) tinggal 2 [tiŋ.gal] *v* stay

stay (refined) tinggair 2 [tiŋ.gɛ] *v*
steam kukus [ku.kus] *v*
step on pinjak [pin.dʒaʔ] *v*
sticky melekat [mə.lə.kat] *adj*
still lagik 2 [la.giʔ] *adv*
stingy kiam [kjam] *adj*
stomach perot 2 [pə.rot] *n*; purot [pu.rot] *n*
stone batu 2 [ba.tu] *n*
story cherita [tʃə.ri.ta] *n*
straight terus [tə.rus] *adj*
straightaway langsung [laŋ.son] *adv*
Straits-born Chinese Peranakan [pə.ra.na.kan] *n*
strangle chekék [tʃə.keʔ] *v*
streetwise cherdék [tʃər.deʔ] *adj*
stretch upon waking up buang segan [bwaŋ sə.gan]
 (comp. of buang, segan) *v*
strike pukol 2 [pu.kol] *v*
strong kuat [kwat] *adj*
stubborn degil [də.gil] *adj*
stupid bodoh [bo.doh] *adj* (comp. buat bodoh); gorblok
 [gɔ.bloʔ] *adj*
stylised dance jogét [dʒo.get] *v*
stylish lawa [la.wa] *adj*
subjected to kena 2 [kə.na] *v*
sugar gula [gu.la] *n*
suitable hak [haʔ] *adj*
sun mata ari [ma.ta a.ri] (comp. of mata, ari) *n*
sun dry gemoh [gə.moh] *v*
superstitious pantang [pan.taŋ] *adj*
supple lembut [ləm.but] *adj*
surprised peranjat [pə.ran.dʒat] *adj* (comp. terperanjat)
sweep sapu [sa.pu] *v*
sweet manis [ma.nis] *adj* (comp. mulot manis)
sweet dessert of thick coconut milk with banana chunks
 pengat [pə.ŋat] *n*
sweetheart mata ayé [ma.ta a.je] (comp. of mata, ayé) *n*
swim berenang [bə.rə.naŋ] *v*
Syzygium fruit (pink) jambu [dʒam.bu] *n*

T

- table** méja [me.dʒa] *n*; tok [toʔ] *n*
- tail** buntot 1 [bun.tot] *n*
- take** amék [a.meʔ] *v*
- take care of someone or something** jaga 2 [dʒa.ga] *v*
- tall** tinggi [tiŋ.gi] *adj*
- tamarind** asam 1 [a.sam] *n*
- tame** jinak [dʒi.naʔ] *adj*
- tapioca** ambun [am.bun] *n*
- taste** chobak [tʃo.baʔ] *v*
- tasteless (coarse)** tawar [ta.war] *adj*
- tasteless (refined)** tawair [ta.wɛ] *adj*
- tea** téh [teh] *n*
- teach (coarse)** ajar [a.dʒar] *v*
- teach (refined)** ajair [a.dʒɛ] *v*
- teak** jati 1 [dʒa.ti] *n*
- tell** bilang [bi.laŋ] *v*
- tell a lie** bédék [be.deʔ] *v*
- ten** chap [tʃap] *cardnum* (id. ati it gor it chap); puloh [pu.loh] *cardnum*
- ten (coarse)** belas [bə.las] *cardnum*
- ten (refined)** belair [bə.lɛ] *cardnum*
- terrible** terok [tə.roʔ] *adj*
- thank you** kamsiah [kam.sjah] *interj*
- that** itu [i.tu] *dem* (contr. tu)
- there** [sa.na] *adv*
- thick (describing liquid)** [pə.kat] *adj*
- thin** kurus [ku.rus] *adj*
- thing** barang [ba.raŋ] *n*
- think** ingat 2 [i.ŋat] *v*; pikay [pi.ke] *v*; rasa 2 [ra.sa] *v*
- thirsty** a'us 1 [a.ʔus] (var. of ha'us) *adj*; ha'us 1 [ha.ʔus] *adj*
- this** ini [i.ni] *dem* (contr. ni)
- thousand** ribu [ri.bu] *cardnum*
- threadfin** fish kuro [ku.ro] *n*
- threat** gertak [gər.taʔ] (var. gretak) *v*; gretak [grə.taʔ] *v*
- three** sa [sa] *cardnum*; tiga [ti.ga]
- throw** buang [bwaŋ] *v* (comp. buang ayé, comp. buang ayé besair, comp. buang ayé kechik, comp. buang buang, comp. buang mata, comp. buang mulot, comp. buang sebelah, comp. buang segan, comp. buang tebiét)
- throw** tantrum buang tebiét [bwaŋ tə.bjet] (comp. of buang, terbiat) *v*
- tickle** geléték [gə.le.teʔ] *v* (comp. naik geléték)
- tidy (coarse)** kemas [kə.mas] *v* tidy
- tidy (refined)** kemair [kə.mɛ] *v*
- tie** ikat [i.kat] *v*
- tiger** arimo [a.ri.mo] (var. of harimo) *n*; harimo [ha.ri.mo] *n*
- time** jam 1 [dʒam] *n*; kala [ka.la] *n* (comp. sinjakala); kali [ka.li] *n*
- tired** penat [pə.nat] *adj* **to twist one's foot** peléchok [pə.le.tʃoʔ] *v*
- to water** plants renjis [rən.dʒis] *v*
- toilet** chiwan [tʃi.waŋ] *n*
- tomorrow** bésok [be.soʔ] *adv*
- tongue** lidah [li.dah] *n*
- too** terlalu [tə.la.lu] *adv*
- tooth** gigi [gi.gi] *n*
- top** atas [a.tas] *prep*
- torn** koyak [ko.jaʔ] *adj*
- tortoise** ku 1 [ku] *n*; kura-kura [ku.ra-ku.ra] *n*
- toss** lambong [lam.boŋ] *v*
- touch** raba [ra.ba] *v*
- TR** -kan [kan]
- traditional** blouse kebaya [kə.ba.ja] *n*
- traditional** cakes kuéh [kweh] *n*. (comp. kuéh-yi)
- traditional** Peranakan elder bok-bok [boʔ-boʔ] (contr. of

embok-embok *n* ; embok-embok [ə̃m.boʔ-ə̃m.boʔ] *n*
traffic junction sempang [sə̃m.paŋ] *n*
transfer salin [sa.lin] *v*
treat belanja 2 [bə̃.lan.dʒa] *v*
tree pokok [po.koʔ] *n*
tremble (coarse) bergetar [bər.gə̃.taɾ] *v*
tremble (refined) bergetair [bər.gə̃.tɛ] *v*
true betul 2 [bə̃.tol] *adj*
trust percaya [pər.tʃa.ja] *v*
turn b́elok [be.loʔ] *v*; **pusing** 2 [pu.sij] *v*

turn over balék 2 [ba.leʔ] *v*
turnip bangkuang [baŋ.kwaŋ] *n*
turtle labi-labi [la.bi-la.bi] *n*
twisted b́engok [be.ŋoʔ] *adj* (comp. mulot b́engok)
twisted mouth mulot b́engok [mu.lot be.ŋoʔ] (comp. of mulot, b́engok) *n*
two dua [dwa] *cardnum*; **ji** [dʒi] *cardnum*
type of noodles kolomi [ko.lo.mi] *n*

U

ugly burok [bu.roʔ] *adj*
umbrella payong [pa.joŋ] *n*
unbleached cotton outfit used for mourning belachu [bə̃.la.tʃu] *n*
uncontrollable mouth mulot gatair [mu.lot ga.tɛ] (comp. of mulot, gatair) *n*
under bawah [ba.wah] *prep*
under the influence of black magic kunang [ku.naŋ] *adj*
understand reti [rə̃.ti] *v*
unhappy kéksim [keʔ.sim] *adj*

unlucky suay [swe] *adj* (comp. buat suay)
unripe terbantot [tər.ban.tot] *adj*
until sampay 1 [sam.pe] *prep*
upstairs loténg [lo.teŋ] *adv*
urinate buang ayé kechik [bwaŋ a.je kə̃.chiʔ] (comp. of buang, ayé, kechik) *v*; buang ayé [bwaŋ a.je] (comp. of buang, ayé) *v*
use guna 1 [gu.na] *v*; guna 2 [gu.na] *n*; pakay 2 [pa.ke] *v*; pasang 1 [pa.saŋ] *v*
used to it biasa [bja.sa] *adj*

V

vegetables sayor [sa.jor] *n*
very sangat [sa.ŋat] *adv* *v*; sekali 1 [sə̃.ka.li] *adv*

village kampong [kam.poŋ] *n*

W

waist pinggang [piŋ.gaŋ] *n*
wait tunggu [tuŋ.gu] *v*
wake mangun [ma.ŋun] *v*

walk jalan [dʒa.lan] *v* (comp. tengah jalan)
walking stick tongkat [toŋ.kat] *n*
wall témbok [tem.boʔ] *n*

want mo [mo] *v*
wart mata ikan [ma.ta i.kan] (comp. of mata, ikan) *n*
wash chuchi [tʃu.tʃi] *v*
water ayé [a.je] *n* (comp. buang ayé, comp. buang ayé besair, comp. buang ayé kechik, mata ayé)
water spinach kangkong [kaŋ.koŋ] *n*
wear pakay 1 [pa.ke] *v*
wedding gift exchange ceremony lapchai [lap.tʃaj] *n*
week minggu [miŋ.gu] *n*
weigh timbang [tim.baŋ] *v*
what apa [a.pa] *interrog* (comp amcham, comp. apa cham, comp. apa macham, comp. apa pasal, contr. apa sair)
when bila [bi.la] *adv*
where mana 1 [ma.na] *interrog*
which mana 2 [ma.na] *interrog*
whirl pusing 1 [pu.siŋ] *v*
whisper memisék [mə.mi.seʔ] *v*
white putéh [pu.teh] *adj*
who siapa [sja.pa] *interrog*
why (coarse) apa pasal [a.pa pa.sal] (comp. of apa, pasal) *interrog*
why (refined) apa sair [a.pa se] (contr. of **apa**, **sair** 1)

interrog why
wife bini [bi.ni] *n*
wildness liar [ljaɾ] *n*
win menang [mə.naŋ] *v*
wind angin [a.ŋin] *n*
window menjéla [mən.dʒe.la] *n*
Winter Solstice tangchék [taŋ.tʃeʔ] *n*
with sama 3 [sa.ma] *prep*
withstand tahan [ta.han] *v*
wonder héran [he.ran] *v*
wood kayu [ka.ju] *n*
work kerja 1 [kəɾ.dʒa] (var. kreja) *v*; kerja 2 [kəɾ.dʒa] (var. kreja) *n*; kreja 1 [krə.dʒa] *v*; kreja 2 [krə.dʒa] *n*
world dunya [du.ja] *n* (comp. mulot dunya)
worn out ha'us 2 [ha.ʔus] (var. a'us) *adj*; a'us 2 [a.ʔus] *adj*
worry kuatay [kwa.te] *v*
wound luka 2 [luka] *n*
wounded luka 1 [luka] *adj*
wring perah [pə.rah] *v*
wrist watch aloji 2 [a.lo.dʒi] *n*
write tulis [tu.lis] *v*
wrong salah 2 [sa.lah] *adj*

Y

yawn mengantok 1 [mə.ŋan.toʔ] *v*
year taoun [ta.own] *n*
yellow kuning [ku.niŋ] *adj*
yes ya [ja] *interj*
young mudah [mu.dah] *adj*
younger sibling adék 2 [a.deʔ] *n* (der. adék beradék)
younger siblings adék beradék 2 [a.deʔ bə.ra.deʔ] (der. of adék) *n*
youngest (only for familial relations) bongsu [boŋ.su] *adj*

Appendix D: Sociophonetics

Post-listening questionnaire

1. Do you have any known auditory impairment? YES NO
2. What is your age? _____
3. What is your gender? FEMALE MALE
4. Rate your proficiency in Baba Malay from 1 to 5, 1 extremely weak, 5 being extremely proficient.
1 2 3 4 5
5. What language(s) did you speak from birth? (Circle the appropriate response(s))
 - a. Baba Malay
 - b. English
 - c. Mandarin
 - d. Bahasa Melayu
 - e. Tamil
 - f. Hokkien
 - g. Teochew
 - h. Cantonese
 - i. Hakka
 - j. Other(s): _____
6. What are other languages that you speak? (Circle the appropriate response(s))
 - a. Baba Malay
 - b. English
 - c. Mandarin
 - d. Bahasa Melayu
 - e. Tamil
 - f. Hokkien
 - g. Teochew
 - h. Cantonese
 - i. Hakka
 - j. Other(s): _____

7. What are the other languages that you understand? (Circle the appropriate response(s))
- a. Baba Malay
 - b. English
 - c. Mandarin
 - d. Bahasa Melayu
 - e. Tamil
 - f. Hokkien
 - g. Teochew
 - h. Cantonese
 - i. Hakka
 - j. Other(s): _____
8. If you have children, would you want your children to speak Baba Malay?
YES NO
9. If a person does not speak Baba Malay, can that person be Peranakan?
YES NO
10. Do you feel that Baba Malay is endangered?
YES NO
11. Are you worried that Baba Malay is endangered?
YES NO

END OF POST-LISTENING QUESTIONNAIRE

Responses to fillers

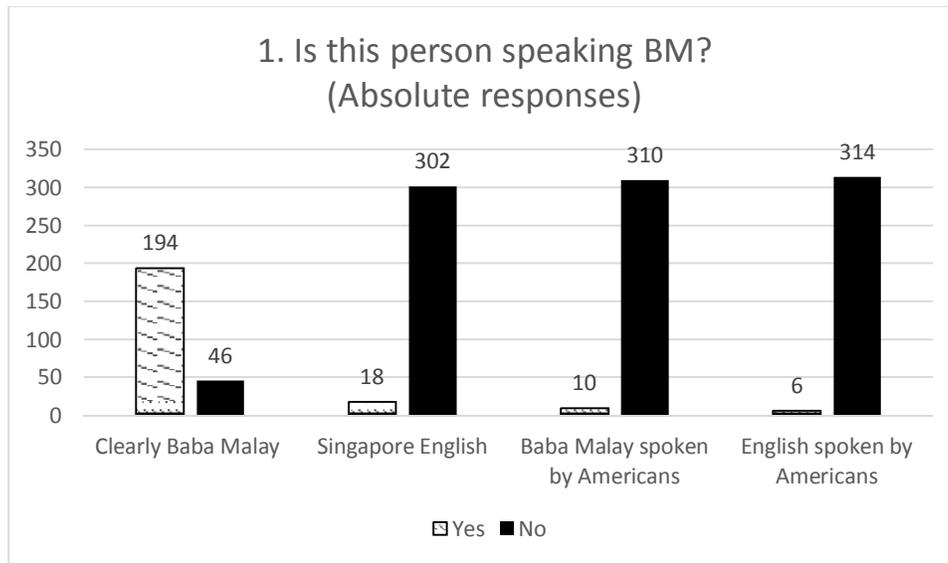


Figure 34: Responses to fillers for the question ‘Is this person speaking Baba Malay?’

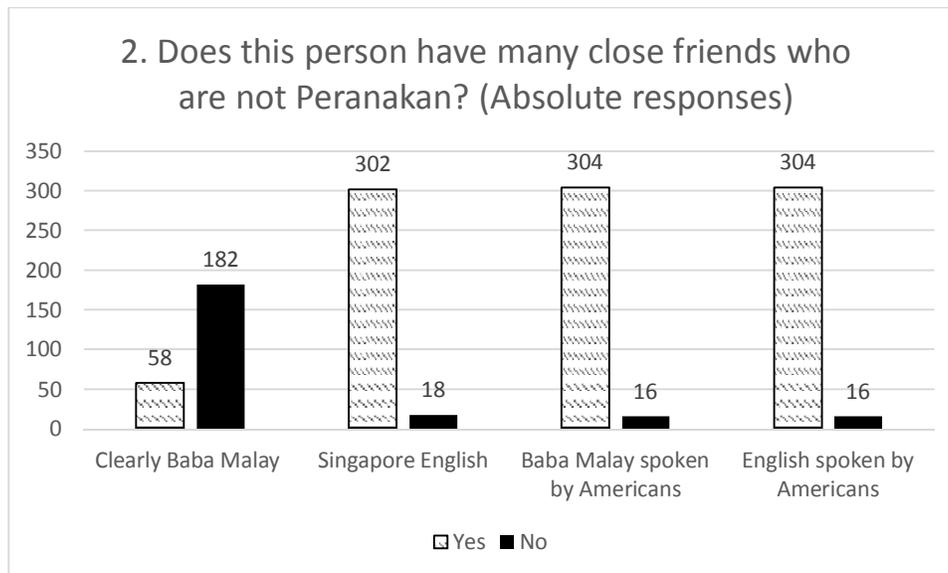


Figure 35: Responses to fillers for the question ‘Does this person have many close friends who are not Peranakan?’

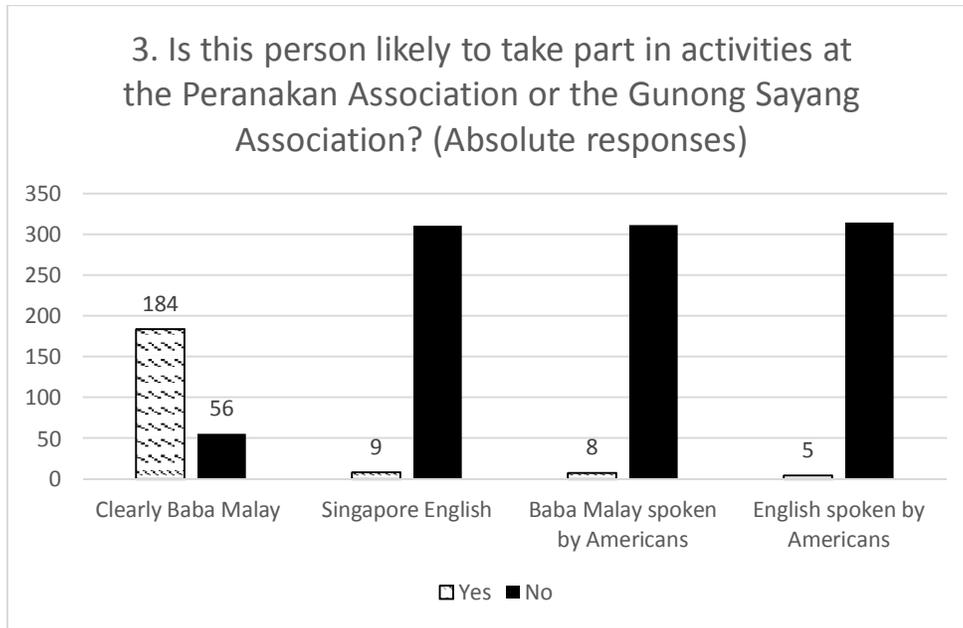


Figure 36: Responses to fillers for the question ‘Is this person likely to take part in activities at the Peranakan Association or the Gunong Sayang Association?’

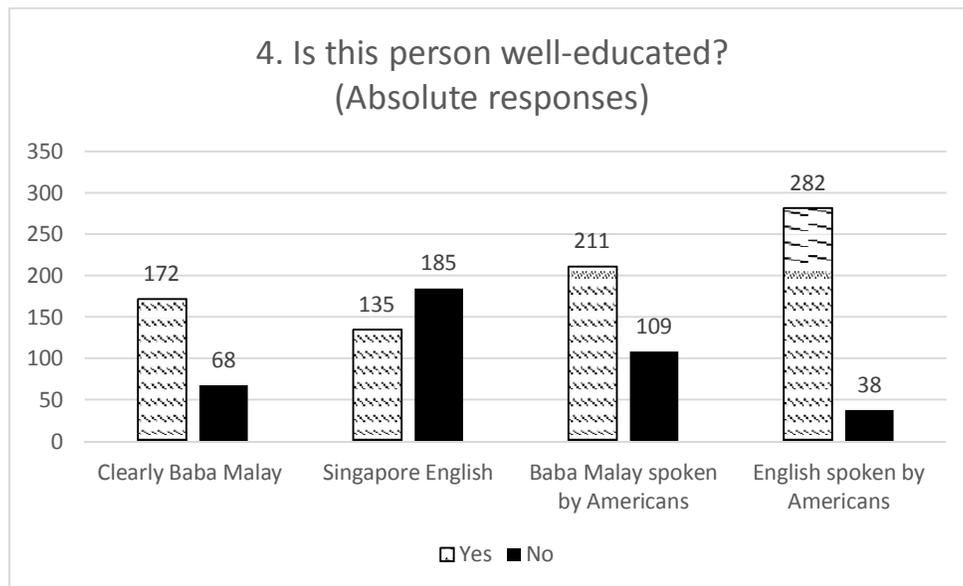


Figure 37: Responses to fillers for the question ‘Is this person well-educated?’